МАТЕМАТИЧЕСКИЙ АНАЛИЗ-I

Конспект лекций

I семетр

Лекция 1

Элементы теории множеств

На первых порах мы будем использовать следующие множества: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \varnothing$. Они знакомы вам со школьных времён. Также считаются известными операции включения элемента во множество $a \in A$ и множества во множество $A \subset B$, пересечения и объединения множеств $A \cap B, A \cup B$, разности множеств $A \setminus B$ и декартова произведения множеств $A \times B$.

В школе вы также изучали отображения множеств $f\colon A \to B$. Напомним, что можно выделить несколько видов отображений множеств: *инъективное, сюръективное и биективное.* Отображение $f\colon A \to B$ называется инъективными или отображением "в", если $f(x)$ не совпадает с $f(y)$ для всяких несовпадающих $x, y \in A$. Отображение $f\colon A \to B$ называется сюръективным или отображением "на", если для всякого $b \in B$ существует $a \in A$ такой, что $f(a) = b$. Отображение называется биективным, или взаимно-однозначным, если оно одновременно инъективно и сюръективно. Конечно же, есть отображения, которые не являются ни инъективными, ни сюръективными.

В матанализе активно используется язык кванторов. Нам понадобятся кванторы *всеобщности* $\forall$ и *существования* $\exists$. Их пока следует понимать как символьное обозначение слов "для всякого, для любого" в случает квантора $\forall$ и "существует" в случае квантора $\exists$. Например определение сюръективного отображения перепишется следующим образом:

$$\forall b \in B\ \exists a \in A\colon f(a) = b.$$

Двоеточие читается как "такой что".

Напомним также , что *мощностью конечного множества* называется число элементов в нём. При этом мощность множества $\varnothing$ равна 0. Для бесконечных множеств понятие мощности не определяется, но определяется понятие *равномощных множеств*: *два множества называются равномощными, если между ними существует биекция.*

Нам понадобятся две следующие теоремы, которые будут доказаны в курсе **"Введение в дискретную математику и топологию"**.

Теорема. *Для любых двух множеств A и B верно следующее: они либо равномощны, либо нет. Причём в последнем случае верно, что либо A равномощно подмножеству B, либо B равномощно подмножеству A.*

Теорема Кантора-Бернштейна. *Если A равномощно подмножеству B, а B равномощно подмножеству A, то A и B равномощны.*

Счётным мы будем называть всякое множество, равномощное $\mathbb{N}$. Например, $\mathbb{Z}$ и $\mathbb{Q}$ счётны. Докажем следующую теорему

Теорема. *Множество бесконечных последовательностей из 0 и 1 несчётно.*

Доказательство. Предположим противное. Тогда последовательности из 0 и 1 можно занумеровать следующим образом: в первой строке выпишем первую последовательность, во второй строке - вторую и т.д. Предъявим последовательность из 0 и 1, которую нельзя вписать ни в какую строку. Её первый элемент отличен от первого элемента первой строки, второй элемент отличен от второго элемента второй строки и т.д. Такая последовательность не может стоять в $n-$ой строке, т.к. она отличается от этой строки $n-$м элементом. $\square$

Помимо счётных множеств нам будут очень часто встречаться *континуальные множества*: *множество называется континуальным, если оно равномощно множеству последовательностей из 0 и 1.* Примеры континуальных множеств: $\mathbb{R}, [0,1], (-2.5, 7)$.

Пример. Покажем, что всякое *бесконечное множество* (т.е. оно содержит бесконечно много элементов) A равномощно его объединению со счётным (или конечным) множеством B. Действительно, найдём в A счётную часть S (почему она есть?). Рассмотрим разбиение $A = (A \setminus S) \cup S$. Тогда биекцию с $A \cup B$ построим таким образом: на $A \setminus S$ — это тождественное отображение, а на S есть биекция с $S \cup B$.

Упражнение. Покажите, что произведение континуальных множеств континуально, а счётных счётно.

Теория вещественных чисел

Есть различные подходы к определению вещественного числа. Мы будем следовать аксиоматическому подходу, а именно предъявим такой набор аксиом, что множество, удовлетворяющее этому набору, будет называется множеством вещественных чисел $\mathbb{R}$. Корректность этого определения (т.е. что такое множество действительно существует и оно в некотором смысле единственно) будет доказана в начале второго модуля.

Чтобы предъявить вышеупомянутый набор аксиом нам понадобиться дать несколько определений. Мы начнём с определения *группы.*

Определение. *Множество G с бинарной операцией $*\colon G \times G \to G$ называется коммутативной (или абелевой) группой, если выполнены следующие условия:*

(i) Существует нейтральный элемент $e \in G$ такой, что $\forall x \in G$ справедливо $x * e = e * x = x$.
(ii) Существует обратный элемент: $\forall x \in G\ \exists y \in G$ такой, что $x * y = y * x = e$.
(iii) Операция $*$ ассоциативна: всегда $(a * b) * c = a * (b * c)$.
(iv) Операция $*$ коммутативна: всегда $a * b = b * a$.

При этом вместо $*$ часто используют либо знак $+$ сложения, либо знак $\times$ (он же $\cdot$) умножения. В случае использования $+$, группу называют *аддитивной группой* или *группой по сложению.* Если используется $\times$, то группу называют *мультипликативной* или *группой по умножению.* В мультипликативной группе очень часто знак операции не пишут вообще. В аддитивной группе нейтральный элемент называется 0, обратный элемент к a обозначается $-a$. В мультипликативной группе нейтральный элемент называется 1, обратный элемент к a обозначается a^{-1}. Конечно же, есть и некоммутативные группы (они нам пока не понадобятся). Подробнее группы будут изучаться в курсе алгебры.

Аксиомы операций. На множестве вещественных чисел заданы 2 коммутативные бинарные операции, сложения и умножения. Вот аксиомы, которым они удовлетворяют:

(i) $(\mathbb{R}, +)$ — аддитивная коммутативная группа.
(ii) $(\mathbb{R} \setminus \{0\}, \cdot)$ — мультипликативная коммутативная группа.
(iii) Умножение дистрибутивно относительно сложения: $\forall x, y, z \in \mathbb{R}$ верно, что

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

Если выполнены такие аксиомы, то множество называется *полем* (т.е. заданы две коммутативных операции, связанные между собой законом дистрибутивности). Таким образом, *первая аксиома вещественных чисел говорит, что на $\mathbb{R}$ заданы 2 операции, относительно которых $\mathbb{R}$ является полем.* Примеры полей: $\mathbb{R}, \mathbb{Q}$. На $\mathbb{Z}$ же со стандартными операциями $+$ и $\times$ нет структуры поля. Как и множества, поля бывают конечными и бесконечными. Например, поле $\mathbb{F}_2$ состоит из двух элементов 0 и 1.

Аксиомы порядка. Начнём с определений.

Определение. *Говорят, что на множестве A определено отношение $\mathcal{R}$, если в множестве пар (a, b), $a \in A$, $b \in A$ выбрано некоторое подмножество. Если пара (a, b) принадлежит подмножеству, то будем говорить, что $a\mathcal{R}b$.*

Определение. *Отношение $\leqslant$ называется отношением порядка, если выполнены следующие свойства:*

(i) $a \leqslant a$;
(ii) $(a \leqslant b) \wedge (b \leqslant a) \Rightarrow a = b$ *(символ "$\wedge$" читается как "и")*;
(iii) $(a \leqslant b) \wedge (b \leqslant c) \Rightarrow a \leqslant c$.

Если дополнительно выполнено свойство $\forall a, b \in A$ либо $a \leqslant b$, либо $b \leqslant a$ (то есть все элементы множества A *сравнимы*), то множество называется *линейно упорядоченным.* В противном случае (т.е. не все элементы сравнимы) — *частично упорядоченным.* Наряду со знаком $\leqslant$ мы будем использовать знак $\geqslant$, который определим так $x \leqslant y \Leftrightarrow y \geqslant x$. Пример отношения на множестве натуральных чисел (помимо очевидного): m делится на n. Проверка аксиом отношения порядка очевидна (но всё равно проверьте!).

Структуры одного лишь множества (или поля) с отношением порядка нам будет недостаточно. Поэтому дадим следующее определение.

Определение. *Поле F называется упорядоченным полем, если на этом поле задано отношение порядка, F линейно упорядочено, причём*

$$a \leqslant b \Rightarrow a + c \leqslant b + c \; \forall c \in F \quad \text{и} \quad (0 \leqslant a) \wedge (0 \leqslant b) \Rightarrow 0 \leqslant ab.$$

Примеры упорядоченных полей:
1) $\mathbb{Q}, \mathbb{R}$;
2) *множество рациональных функций над полем* $\mathbb{R}$, т.е. функций вида $R = P/Q$, где P, Q – многочлены, причём Q не тождественно 0 (при этом полагается, что $R > 0$, если коэффициенты при старших степенях в P и Q имеют один и тот же знак, а соотношение порядка таково: $R_1 \geqslant R_2 \Leftrightarrow R_1 - R_2 \geqslant 0$);
3) множество $\mathbb{Q}(\sqrt{3})$ чисел вида $a + b\sqrt{3}$, где $a, b \in \mathbb{Q}$;
4) множество вещественных алгебраических чисел (т.е. корней многочленов с целыми коэффициентами).

Аксиома непрерывности (полноты). *Пусть $A, B \subset \mathbb{R}, A, B \neq \varnothing$, причём $\forall a \in A, b \in B$ справедливо $a \leqslant b$. Тогда существует такое $c \in \mathbb{R}$, что $\forall a \in A$ справедливо $a \leqslant c$, и $\forall b \in B$ справедливо $c \leqslant b$.*

Несмотря на кажущуюся очевидность аксиомы непрерывности, она позволяет отделить поле $\mathbb{R}$ от всех остальных упорядоченных полей.

Итак, мы готовы сформулировать определение вещественных чисел.

Определение. *Полем вещественных чисел называется упорядоченное поле с аксиомой непрерывности.*

Для всех приведённых выше упорядоченных полей за исключением $\mathbb{R}$ аксиома непрерывности не выполняется. Действительно,

1) для $\mathbb{Q}$ и $\mathbb{Q}(\sqrt{3})$: между множествами $A = \{x \in \mathbb{Q} \mid x < \sqrt{2}\}$ и $B = \{x \in \mathbb{Q} \mid x > \sqrt{2}\}$ нельзя "вставить рациональное число" (предварительно нужно показать, что число $\sqrt{2}$ не является рациональным). Аналогичным образом для $\mathbb{Q}(\sqrt{3})$ между множествами $A = \{x \in \mathbb{Q}(\sqrt{3}) \mid x < \sqrt{2}\}$ и $B = \{x \in \mathbb{Q}(\sqrt{3}) \mid x > \sqrt{2}\}$ нельзя "вставить число" из $\mathbb{Q}(\sqrt{3})$ (почему?).

2) для поля рациональных функций над $\mathbb{R}$: рассмотрим множество чисел и множество многочленов вида $ax + b, a > 0$. Любой такой многочлен больше любого числа, однако нет рациональной функции, которая больше всех чисел и меньше всех таких многочленов (почему?).

Модели множества вещественных чисел. Под моделью мы будем понимать какое-то упорядоченное поле, на котором действует аксиома непрерывности. Рассмотрим некоторые из них.

(i) Множество десятичных чисел. Это конструкция изучалась в школе.

(ii) Множество чисел по заданному основанию (двоичные, троичные и т.д.). Также изучалась в школе.

(iii) Числовая прямая. Также изучалась в школе.

(iv) Дедекиндовы сечения. Ограничимся лишь определением: *сечением Дедекинда* называется разбиение множества $\mathbb{Q}$ на два подмножества A и A' таким образом, чтобы всякое число из A было бы меньше всякого числа из A'. На множестве таких сечений можно ввести структуру упорядоченного множества и проверить аксиому непрерывности (см. [2]).

(v) Множество классов эквивалентности фундаментальных последовательностей рациональных чисел. Здесь две фундаментальные последовательности называются эквивалентными, если их "объединение" — снова фундаментальная последовательность. Об этом речь пойдёт позже.

Упражнение. Проверьте аксиомы упорядоченного поля, а также аксиому непрерывности для моделей десятичных, двоичных чисел, а также для модели числовой прямой (как определить умножение?).

Лекция 2

Ограниченные множества

Определение. *Множеством $A \subset \mathbb{R}$ называется ограниченным сверху, если существует такое число $C \in \mathbb{R}$, что для всякого $a \in A$ верно, что $a \leqslant C$. Число C при этом называется верхней границей или гранью.*

На языке кванторов это определение запишется следующим образом: $A \subset \mathbb{R}$ называется ограниченным сверху $\Leftrightarrow \exists C \in \mathbb{R}\ \forall a \in A : a \leqslant C$. Отрицание этого определения даёт определение *неограниченного сверху множества.*

Определение. *Множество $A \subset \mathbb{R}$ называется неограниченным сверху $\Leftrightarrow \forall C \in \mathbb{R}\ \exists a \in A : a > C$.*

Аналогичным образом определяются ограниченные снизу (существует *нижняя граница $C \in \mathbb{R}$*) и неограниченные снизу множества (нижней границы не существует).

Определение. *Множеством $A \subset \mathbb{R}$ называется ограниченным, если оно ограниченно и сверху и снизу. Можно также сказать, что $A \subset \mathbb{R}$ называется ограниченным $\Leftrightarrow \exists C > 0 \in \mathbb{R}\ \forall a \in A : |a| \leqslant C$.*

Определение. *Число $M \in A$ называется максимальным или наибольшим элементом множества A, если оно является верхней границей множества A, то есть для любого $x \in A$ справедливо $x \leqslant M$. Число $m \in A$ называется минимальным или наименьшим элементом множества A, если оно является нижней границей множества A, то есть для любого $x \in A$ справедливо $x \geqslant m$.*

Обозначения. $M = \max A = \max\{x \mid x \in A\} = \max_{x \in A} x$; $m = \min A = \min\{x \mid x \in A\} = \min_{x \in A} x$.

Упражнение. Из аксиомы порядка выведите единственность наибольшего и наименьшего элементов ограниченного множества (при условии, что они есть).

Примеры: наибольшим элементов множества *неположительных вещественных чисел* является 0, наименьшего элемента нет (множество не ограниченно снизу); множество же *отрицательных чисел* не имеет ни наибольшего, ни наименьшего элементов (наибольшим элементом должен был бы быть 0, но он не является отрицательным числом), наименьшего элемента нет, т.к. множество не ограниченно снизу.

Предыдущий пример показывает, что ограниченные множества не очень удобно описывать при помощи их наибольших и наименьших элементов (т.к. их может не быть). Однако, если множество ограничено, то всегда есть границы. Дадим следующие определения.

Определение. *Число $M \in \mathbb{R}$ называется супремумом или точной верхней гранью множества $A \Leftrightarrow (\forall x \in A : x \leqslant M) \wedge (\forall y < M : \exists x \in A : x > y)$. Число $m \in \mathbb{R}$ называется инфимумом или точной нижней гранью множества $A \Leftrightarrow (\forall x \in A : x \geqslant m) \wedge (\forall y > m : \exists x \in A : x < y)$.*

Обозначения. Они аналогичны обозначениям для max и min: $M = \sup A$; $m = \inf A$.

Следующая теорема демонстрирует удобство этого определения.

Теорема. *У любого непустого ограниченного сверху множества есть супремум. У любого непустого ограниченного снизу множества есть инфимум.*

Доказательство. Доказательство проведём для случая ограниченного сверху множества. Для ограниченного снизу множества доказательство аналогично.

Пусть A^+ — множество верхних границ множества A. Оба множества A и A^+ не пустые и при этом выполнено условие аксиомы непрерывности: все элементы множества A не больше всех элементов множества A^+. Значит, в силу аксиомы непрерывности существует число M такое, что $\forall x \in A, y \in A^+ : x \leqslant M \leqslant y$. По построению, $M = \sup A$, т.к. $\forall x \in A : x \leqslant M$ и это, действительно, наименьшая верхняя граница, т.к. $\forall y \in A^+ : y \geqslant M$ $\square$.

На самом деле верна следующая

Теорема. *Теорема о существовании верхней грани эквивалентна аксиоме непрерывности. Иными словами, упорядоченное множество, у которого каждое ограниченное сверху множество имеет точную верхнюю грань есть* $\mathbb{R}$.

Лекция 3

Аксиома Архимеда

Несмотря на своё название, аксиома Архимеда в современном изложении является теоремой, выводящейся из аксиом вещественных чисел.

Теорема (аксиома Архимеда). *Для любых положительных вещественных чисел a и b верно, что среди чисел*

$$a,\ 2a(=a+a),\ 3a(=a+a+a),\dots,\ na,\dots$$

найдутся числа, большие b.

Доказательство. Рассмотрим множество A чисел вида na. Предположим, что среди них нет ни одного числа, большего b. Рассмотрим множество B вещественных чисел, больших всех чисел из A. Множества A и B не пустые, т.к. $b \in B$. Все элементы множества A меньше всех элементов B. Значит, по аксиоме непрерывности между A и B найдется промежуточное число c такое, что $\forall x \in A, y \in B : x \leqslant c \leqslant y$. Число c — точная верхняя грань A по построению. Однако, $d = c - a \in B$ также по построению. Действительно, $c > a + \ldots + a$ для любого количества слагаемых. Это противоречит тому, что $c = \sup A$ $\square$.

Определение.*Упорядоченное поле, удовлетворяющее аксиоме Архимеда, называется архимедовым.*

Примеры: архимедовы поля: $\mathbb{R}, \mathbb{Q}, \mathbb{Q}(\sqrt{3})$; неархимедово поле: множество рациональных функций над $\mathbb{R}$.

Упражнение. Убедитесь в том, что утверждения из предыдущих примеров верны (случай $\mathbb{R}$ уже разобран).

Другие следствия из аксиом вещественных чисел

Сформулируем и докажем некоторые следствия из аксиом вещественных чисел.

(i) Следующие элементы единственны: 0, 1, противоположный и обратный к данному элементы.

$\rhd$ Доказательство проведем лишь для 0 и противоположного элемента. Доказательство для 1 и обратного элемента аналогично. От противного. Пусть есть два нулевых элемента 0_1 и 0_2 и для данного $x \in \mathbb{R}$ пусть есть два противоположных элемента $(-x)_1$ и $(-x)_2$. Тогда

$$0_1 = 0_1 + 0_2 = 0_2.$$

$$(-x)_1 = (-x)_1 + 0 = (-x)_1 + (x + (-x)_2) = ((-x)_1 + x) + (-x)_2 = 0 + (-x)_2 = (-x)_2. \lhd$$

(ii) К обеим частям равенства можно добавлять число, сохраняя равенство. Обе части равенства можно умножать на число, сохраняя равенство.

(iii) Уравнение $a + x = b$ имеет решение $x = b - a$, причём единственное. Уравнение $a \cdot x = b$ имеет при $a \neq 0$ решение $x = b \cdot a^{-1}$, причём единственное.

$$\rhd\ a + x = b \Leftrightarrow a + x + (-a) = b + (-a) \Leftrightarrow a + (-a) + x = b - a \Leftrightarrow x = b - a. \lhd$$

(iv) Для любого $x \in \mathbb{R}$ справедливо $x \cdot 0 = 0 \cdot x = 0$. Если $x \cdot y = 0$,то либо $x = 0$, либо $y = 0$.

$$\rhd\ x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0 \Rightarrow x \cdot 0 + (-(x \cdot 0)) = x \cdot 0 + x \cdot 0 + (-(x \cdot 0)) \Rightarrow x \cdot 0 = 0$$

Если $x \cdot y = 0$ и $y \neq 0$,то $x = 0 \cdot y^{-1} = 0$. $\lhd$

(v) Справедливы равенства $-x = (-1) \cdot x; (-1) \cdot (-1) = 1; (-x) \cdot (-y) = x \cdot y$.

$\triangleright$ Докажем лишь первое равенство

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0. \triangleleft$$

(vi) $\forall x, y \in \mathbb{R} : (x < y) \vee (x = y) \vee (x > y)$; $(x < y) \wedge (y \leqslant z) \Rightarrow x < z$; $(x \leqslant y) \wedge (y < z) \Rightarrow x < z$.

(vii) $x < y \Rightarrow x + z < y + z$; $x > 0 \Rightarrow -x < 0$; $(a \geqslant b) \wedge (c \geqslant d) \Rightarrow a + c \geqslant b + d$; $(a \geqslant b) \wedge (c > d) \Rightarrow a + c > b + d$; $(a > b) \wedge (c > d) \Rightarrow a + c > b + d$;

(viii) $(x > 0) \wedge (y > 0) \Rightarrow xy > 0$; $(x > 0) \wedge (y < 0) \Rightarrow xy < 0$; $(x > y) \wedge (z > 0) \Rightarrow xz > yz$; $(x > y) \wedge (z < 0) \Rightarrow xz < yz$;

(ix) $1 > 0$; $x > 0 \Rightarrow x^{-1} > 0$; $x < 0 \Rightarrow x^{-1} < 0$; $(x > 0) \wedge (y > x) \Rightarrow x^{-1} > y^{-1}$.

$\triangleright$ Если $1 < 0$, то $1 \cdot 1 > 0$ и $1 > 0$. Если $1 = 0$, то $x \cdot 1 = x$ и $x \cdot 0 = 0$, отсюда $x = 0$. Значит, $1 > 0$. Из $x > 0$ и $x^{-1} < 0$ следует $x \cdot x^{-1} < 0$, что противоречит $1 > 0$. $\triangleleft$

Натуральные числа

Нам понадобится более строгое определение натуральных чисел, чем то, которым вы пользовались в школе. Вообще говоря, натуральные числа определяются с помощью *аксиом Пеано*, о которых пойдёт речь в курсе по дискретной математике и топологии. Мы не будем вовсе касаться аксиоматики Пеано натуральных чисел. Корректность определения, которое последует ниже (т.е. то, что определённые нами натуральные числа действительно удовлетворяют аксиомам Пеано), будет доказана в курсе “Введение в дискретную математику и топологию”.

Начнём с определения индуктивных множеств.

Определение.*Множество $A \subset \mathbb{R}$ называется индуктивным, если $\forall a \in A \colon a + 1 \in A$.*

Лемма. *Пересечение любого числа индуктивных множеств или пусто или само является индуктивным множеством.*

Доказательство. Пусть $A = \cap A_\alpha$, где α пробегает некоторое множество. Тогда

$$a \in A \Rightarrow \forall \alpha \colon a \in A_\alpha \Rightarrow \forall \alpha \colon (a + 1) \in A_\alpha \Rightarrow (a + 1) \in A. \square$$

Определение. *Множество натуральных чисел $\mathbb{N}$ — это пересечение всех индуктивных подмножеств $\mathbb{R}$, содержащих 1.*

Отметим, что определение немедленно влечёт, что $\min \mathbb{N} = 1$ (почему?). Натуральные числа обладают например таким свойствами: сумма и произведение двух натуральных чисел — натуральное число, между двумя натуральными числами n и $n + 1$ нет натуральных чисел (см. Глава 1, §2b в [1]). Подробно эти свойства будут обсуждаться в курсе “Введение в дискретную математику и топологию”. Здесь же мы отметим, что доказательство этих свойств опирается на следующее следствие:

Следствие из определения (принцип математической индукции). *Если $E \subseteq \mathbb{N}, 1 \in E$ и E — индуктивное множество, то $E = \mathbb{N}$.*

Напомним, что множество называется линейно упорядоченным, если любые два его элемента сравнимы.

Определение. *Назовём линейно упорядоченное множество вполне упорядоченным, если в каждом его непустом подмножестве есть наименьший элемент.*

Теорема. *Множество $\mathbb{N}$ вполне упорядочено.*

Доказательство. Пусть $A \subset \mathbb{N}$. Если $1 \in A$, то $\min A = 1$. Если $1 \notin A$, то $1 \in \mathbb{N} \setminus A$. рассмотрим множество $B = \mathbb{N} \setminus A$. Покажем, что в B найдётся такой элемент n, что $1, \ldots, n \in B$, $n + 1 \in A$. В самом деле, если такого n не существует, то множество B индуктивно и, следовательно, совпадает с $\mathbb{N}$. Но тогда $A = \mathbb{N} \setminus B = \emptyset$. Противоречие. Тогда n существует и значит $\min A = n+1$ $\square$.

Замечание. Множество $\mathbb{Q}$ не является вполне упорядоченным.

Теорема. *В любом непустом ограниченном сверху подмножестве $\mathbb{N}$ есть наибольший элемент.*

Доказательство. Пусть $E \subset \mathbb{N}$ — такое подмножество. По теореме о верхней грани существует $c = \sup E$. По определению точной верхней грани, в E найдётся число n такое, что $c - 1 < n \leqslant c$. Тогда $n = \max E$. В самом деле: прибавляя 1 к обеим частям неравенства, получаем $c < n+1$. Тогда $n \leqslant c < n+1$ и $c \in \mathbb{N}$. Значит, $c = n$ т.к. мы отмечали выше, что между двумя последующими натуральными числами нет натуральных чисел. $\square$.

Следствие. *Множество натуральных чисел не ограничено сверху.*

Доказательство. Иначе существовало бы наибольшее натуральное число. Но $n < n+1$. $\square$.

Корректно определив $\mathbb{N}$, мы теперь сможем определить $\mathbb{Z}, \mathbb{Q}$. В первом случае это натуральные числа, 0, и противоположные к натуральным, во втором случае добавим к $\mathbb{Z}$ ещё все обратные числа (кроме нуля) к числам из $\mathbb{Z}$ и добавим все произведения обратных на натуральные.

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

Доказательство. Аналогично теореме о наибольшем элементе непустого ограниченного сверху подмножества $\mathbb{N}$. $\square$.

Аксиома Архимеда уточнённая (принцип Архимеда). Для любого $h > 0$ и для любого $a \in \mathbb{R}$ найдется такое $n \in \mathbb{Z}$, что справедливы неравенства $(n-1)h \leqslant a < nh$.

Доказательство. Рассмотрим множество $E = \{n \in \mathbb{Z} \colon a/h < n\}$. Оно ограничено снизу. Оно не пусто. В самом деле, если $a > 0$, то не пусто по аксиоме Архимеда. Если $a < 0$, то оно не пусто так как содержит ноль. Если $a = 0$, то $E = \mathbb{N}$. Пусть $m = \min E$ (такой элемент обязательно есть по следствию об ограниченном снизу подмножестве $\mathbb{Z}$). Тогда $m - 1 \leqslant a/h < m$ $\square$.

Вот несколько следствий из принципа Архимеда.

Следствие 1. *Для любого $\varepsilon > 0$ существует $n \in \mathbb{N}$ такое, что $n^{-1} < \varepsilon \Leftrightarrow n^{-1} \in (0, \varepsilon)$.*

Доказательство. Применим принцип Архимеда к $a = 1, h = \varepsilon$. Тогда найдется n такое, что $1 < \varepsilon n$, отсюда $0 < 1/n < \varepsilon$. $\square$

Следствие 2. *Между двумя любыми вещественными числами есть рациональное число, т.е. для любых $a, b \in \mathbb{R}, a < b$ существует такое $r \in \mathbb{Q}$, что $a < r < b$.*

Доказательство. Пусть $a, b \in \mathbb{R}, b > a$. Выберем m так, чтобы $0 < h = m^{-1} < \varepsilon = b - a$. Найдется такое $n \in \mathbb{Z}$, что
$$\frac{n-1}{m} \leqslant a < \frac{n}{m}.$$
Теперь $\frac{n}{m} < b$, иначе получилось бы, что $\frac{n-1}{m} \leqslant a < b \leqslant \frac{n}{m}$, а это противоречит $0 < 1/m < b - a$. Таким образом, $r = n/m$ — искомое рациональное число. $\square$.

Следствие 3. *Для любого $x \in \mathbb{R}$ существует $k \in \mathbb{Z}$ такое, что $k \leqslant x < k+1$.*

Непосредственно следует из принципа Архимеда. Число k называется целой частью, число $x - k$ называется дробной частью.

Последовательности

Определение. *Последовательность — функция натурального аргумента, т.е.* $f\colon \mathbb{N} \to X$, *где* X *— некоторое множество.*

Обозначение. Когда говорят о последовательностях, то аргумент обычно пишут не в скобках, а в виде индекса: x_n. Часто саму последовательность обозначают таким образом $(x_n)_{n\in\mathbb{N}}$.

Рассмотрим два вида последовательностей: последовательности вещественных чисел и последовательности множеств.

(i) Последовательности вещественных чисел: $a\colon \mathbb{N} \to \mathbb{R}$.

(ii) последовательности множеств: $A\colon \mathbb{N} \to 2^X$ (здесь 2^X — обозначение множества всех подмножеств множества X). Будем также говорить, что A_n — последовательность вложенных множеств, если $\forall k \in \mathbb{N}\colon A_k \supset A_{k+1}$.

Определение. *Будем говорить, что последовательность вещественных чисел* x_n, *стремится или сходится к нулю, если*

$$\forall \varepsilon > 0\ \exists N = N(\varepsilon) \in \mathbb{N}\colon\ \forall n > N\colon |x_n| < \varepsilon.$$

Обозначение. $x_n \to 0$

Примеры. $x_n = 0$, $x_n = 1/n$, $x_n = q^n$, где $|q| < 1$.

Лекция 4

Теорема о вложенных отрезках. *В $\mathbb{R}$ любая последовательность вложенных отрезков $\Delta_n = [a_n, b_n]$ имеет непустое пересечение. Если их длины $b_n - a_n$ стремятся к нулю, то это пересечение $\bigcap_{n=1}^{\infty}\Delta_n$ состоит из единственной точки.*

Доказательство. Рассмотрим два множества: множество A, состоящее из точек a_n (левых концов отрезков) и множество B, состоящее из точек b_n (правых концов отрезков). Заметим, что какое-то из этих множеств вполне может оказаться конечным. Множества A и B удовлетворяют условиям аксиомы непрерывности (они не пустые и каждое a_m не превосходит каждого b_n), поэтому существует число c: $a_m \leqslant c \leqslant b_n$ при всех $m, n \in \mathbb{N}$. Отсюда следует искомое включение $c \in \bigcap_{n=1}^{\infty}\Delta_n$. Единственность точки c следует из $b_n - a_n \to 0$. Если бы было две таких различных точки $c_1 < c_2$, то из неравенств $a_n \leqslant c_1 < c_2 \leqslant b_n$ следовало бы неравенство $b_n - a_n > c_2 - c_1 > 0$, которое противоречит $b_n - a_n \to 0$. $\square$

Замечание. Последовательность вложенных интервалов не обязательно имеет общую точку. Например, пересечение вложенных интервалов

$$(0,1) \supset (0,1/2) \supset (0,1/3) \supset \ldots \supset (0,1/n) \supset \ldots$$

является пустым множеством.

Открытые и замкнутые множества

Определение. *Окрестностью точки называется любой интервал, содержащий эту точку. Пусть $\varepsilon > 0$, тогда ε-окрестность точки x_0 — это интервал $U_\varepsilon(x_0) = (x_0 - \varepsilon, x_0 + \varepsilon)$.*

Заметим, что пересечение двух окрестностей точки x также является окрестностью этой точки.

Определение. *Пусть дано множество $A \subset \mathbb{R}$. Точка $x \in \mathbb{R}$ называется предельной точкой A, если в любой окрестности точки x найдется $y \in A : y \neq x$.*

Эквивалентно можно сказать, что точка $x \in \mathbb{R}$ называется предельной точкой множества A, если в любой окрестности точки x найдется бесконечное множество точек из A. Действительно, пусть в любой окрестности точки x найдется $y \in A : y \neq x$. Возьмём какую-то окрестность U' точки x, зафиксируем $y' \in A \cap U', y' \neq x$, потом возьмём окрестность $U'' \subset U'$ точки x, не содержащую точку y' (почему так можно сделать?), в U'' по предположению есть точка $y'' \in A \cap U'', y'' \neq x$. По такой схеме мы найдём бесконечное множество точек, принадлежащих $A \cap U'$.

Определение. *Точки множества, не являющиеся предельными, называются изолированными. Изолированная точка A — точка, в некоторой окрестности которой кроме неё нет других точек из A.*

Лемма (о предельной точке). *Всякое ограниченное бесконечное множество имеет по крайней мере одну предельную точку.*

Доказательство. Дано ограниченное бесконечное множество E . Пусть отрезок $[a_1, b_1], b_1 > a_1$ содержит E . Разделим отрезок $[a_1, b_1]$ пополам точкой $(a_1 + b_1)/2$ на два отрезка: $L_1 = [a_1, (a_1 + b_1)/2]$ и $R_1 = [(a_1 + b_1)/2, b_1]$. По крайней мере одно из множеств $L_1 \cap E$ или $R_1 \cap E$ содержит бесконечное множество точек. Пусть например, это будет множество $R_1 \cap E$. Обозначим отрезок R_1 через $[a_2, b_2]$. Разделим его пополам на отрезки L_2 и R_2, снова по крайней мере одно из множеств $L_2 \cap E$ или $R_2 \cap E$ содержит бесконечное множество точек, выберем его и т.д. Мы получим последовательность вложенных отрезков $[a_n, b_n]$, длины этих отрезков равны $(b_n - a_n) = (b_1 - a_1)2^{-n+1} \to 0$. Итак, пересечение всех этих отрезков состоит из единственной точки c. Это есть предельная точка множества E: любая окрестность точки c содержит все отрезки $[a_n, b_n]$ с достаточно большими номерами (почему?), на каждом таком отрезке есть бесконечное количество точек множества E. $\square$.

Определение. *Множество U называется открытым, если для любой точки $x \in U$ существует окрестность точки x целиком лежащая в U. Множество называется замкнутым, если оно содержит все свои предельные точки. Точка $x \in U$ называется внутренней точкой множества U, если найдётся окрестность U' точки x, целиком лежащая в U. Таким образом, множество U открытое, если все его точки являются внутренними.*

Примеры: $\varnothing, \mathbb{R}$ одновременно и замкнутые, и открытые; $[a, b], [a, \infty), \mathbb{N}, \mathbb{Z}$ — замкнутые множества, $(a, b), (-\infty, b)$ — открытые множества. Промежутки $[a, b)$ и $(a, b]$ не замкнуты и не открыты.

Теорема. *Пусть A — открытое множество, B — замкнутое множество. Тогда*
1) $A \setminus B$ *— открытое множество.*
2) $B \setminus A$ *— замкнутое множество.*

На самом деле верна следующее следствие, которое часто даётся в качестве определения открытого и замкнутого множества в курсе топологии.

Следствие. *Пусть A — открытое множество, тогда $\mathbb{R} \setminus A$ — замкнутое. Пусть B — замкнутое множество, тогда $\mathbb{R} \setminus B$ — открытое.*

Иными словами, дополнение замкнутое множество - это дополнение до открытого и наоборот.

Упражнение. Докажите это следствие.

Определение. *Подмножество называется $A \subset \mathbb{R}$ связным, если не существует таких открытых непересекающихся множеств U и V в $\mathbb{R}$, что $A \cap U \neq \varnothing$ и $A \cap V \neq \varnothing$ и $A \subseteq U \cup V$.*

Примеры. $\mathbb{R}, (-\pi, \pi)$ связные, множество $(0, 1) \cup (1, 2)$ не связно.

Вернёмся к связным множествам позднее, а сейчас рассмотрим множество предельных точек некоторого подмножества $\mathbb{R}$. Обозначим через A' множество предельных точек множества A. Заметим, что в общем случае *не верны* включения $A' \subset A$, $A \subset A'$ (почему?).

Определение. *Замыкание множества A есть объединение A с множеством его предельных точек A'.*

Обозначение. Замыкание будем обозначать символом $\bar{A}$, т.е. $\bar{A} = A \cup A'$.

Лекция 5

Замыкание множеств

На прошлой лекции были определены понятия множества предельных точек данного множества и замыкания множеств. Рассмотрим их основные свойства.

Утверждение 1. *Множество A' предельных точек множества A замкнуто.*

Доказательство. В самом деле, пусть A' — множество предельных точек множества A. Пусть x — предельная точка множества A'. Тогда в любой окрестности $U(x)$ точки x лежит точка $y \in A'$, предельная точка множества A. Возьмём окрестность $U(y)$ точки y, целиком лежащую в $U(x)$ и не содержащую x. По определению предельной точки там лежит точка $z \in A$. Эта точка лежит в $U(x)$ и не совпадает с x. Мы доказали, что в любой окрестности точки x (предельной для A') лежит точка $z \neq x$, $z \in A$. Значит, $x \in A'$. Итак, каждая предельная точка множества A' ему принадлежит, множество A' замкнуто. $\square$

Утверждение 2. *Множество $\bar{A}$ замкнуто.*

Доказательство. Предельная точка множества $A \cup B$ — либо предельная точка для A, либо для B, либо для обоих множеств (если не так, то она изолированная и для A и для B, поэтому и для $A \cup B$). Поэтому предельная точка для $\bar{A} = A \cup A'$ — это либо предельная точка x для A, тогда $x \in A'$, либо $x \in (A')'$, тогда $x \in A'$ по утверждению 1. $\square$

Утверждение 3. *Объединение любого набора открытых множеств — открытое множество. Пересечение конечного набора открытых множеств — открытое множество. Объединение конечного набора замкнутых множеств — замкнутое множество. Пересечение любого набора замкнутых множеств — замкнутое множество.*

Упражнение. Докажите утверждение 3.

Пример. $\cap_{n=1}^{\infty}(-n^{-1}, n^{-1}) = \{0\}$, т.е. пересечение счётного набора открытых множеств не обязательно открытое множество.

На семинарах мы докажем

Утверждение 4. *Каждое непустое открытое множество U на прямой является объединением не более чем счётного набора непересекающихся открытых промежутков.*

Всюду плотные и нигде не плотные множества

Определение. *Подмножество A некоторого множества $X \subset \mathbb{R}$ называется всюду плотным в X, если всякая окрестность любой точки x из X содержит элемент из A. Подмножество A некоторого множества $X \subset \mathbb{R}$ называется нигде не плотным в X, если для любого интервала $U \subset \mathbb{R}$, имеющего непустое пересечение с X, найдётся интервал $V \subset U$, также имеющий непустое пересечение с X, такой, что $A \cap V = \varnothing$.*

Примеры.

1. Множества рациональных и иррациональных чисел всюду плотны в $\mathbb{R}$.
2. Существует множество обладающее следующими свойствами: дополнение к нему всюду плотно, само оно нигде не плотно. Бывают также замкнутые нигде не плотные множества и открытые всюду плотные множества. Не бывает открытых нигде не плотных множеств.
3. Дополнение к открытому всюду плотному множеству нигде не плотно.

Определение. *Пусть дано множество A и дана система множеств U_α, причём $A \subset \cup_\alpha U_\alpha$. Тогда говорят, что система этих множеств образует покрытие множества A:* $\forall x \in A\ \exists U_\alpha : x \in U_\alpha$.

Теорема Бэра о категориях. *Отрезок не может быть покрыт объединением счётного числа нигде не плотных множеств.*

Доказательство. Пусть $[0,1] \subset \cup_{n=1}^{\infty} U_n$, где все множества U_n нигде не плотны в $\mathbb{R}$. Тогда выберем отрезок $\Delta_1 \subset [0,1]$, на котором нет точек из U_1, потом выберем отрезок $\Delta_2 \subset \Delta_1$, на котором нет точек из U_2 и так далее. Как это сделать? В определении нигде не плотного множества стоят интервалы, поэтому будем выбирать интервалы, а потом в каждом из них выберем отрезок, чуть поменьше. Последовательность вложенных отрезков Δ_n имеет общую точку, эта точка не принадлежит ни одному из множеств U_n, следовательно предположение $[0,1] \subset \cup_{n=1}^{\infty} U_n$ не верное. $\square$.

Определение. *Множества, которые могут быть покрыты объединением счётного числа нигде не плотных множеств, называются множествами 1-й категории по Бэру, или тощими множествами. Остальные множества — множества 2й категории, или тучными.*

Таким образом, мы доказали, что отрезок — множество второй категории по Бэру.

Канторво множество

Важную роль в разнообразных конструкциях играют множества, которые называют канторовыми множествами. Опишем их общую конструкцию. Возьмём отрезок Δ_1 и выберем интервал U_1, так, чтобы его концы не совпадали с концами отрезка Δ_1. Множество $\Delta_1 \setminus U_1$ — это два отрезка, Δ_2 и Δ_3. Затем на каждом из этих отрезков выберем по интервалу $U_2 \subset \Delta_2$ и $U_3 \subset \Delta_3$ так, чтобы их концы не совпадали с концами отрезков Δ_3 и Δ_1. Множество $\Delta_1 \setminus (U_1 \cup U_2 \cup U_3)$ — объединение четырёх непересекающихся отрезков $\Delta_4, \ldots, \Delta_7$. Снова на каждом из них выберем по интервалу U_k и так далее. Предположим ещё, что мы так будем выбирать интервалы, что длины оставшихся отрезков будут стремиться к нулю. В результате получится последовательность непересекающихся интервалов U_k, их объединение — открытое множество U. Множество $\Delta_1 \setminus U_1$ — замкнутое. Оно содержит концы всех отрезков, мы их ни на каком шаге не выбросим. Эти концы отрезков образуют счётное множество. Кроме того, $\Delta_1 \setminus U_1$ содержит ещё континуальное множество точек. Проделав эту процедуру бесконечное число раз (как описать строго?), мы получим некоторое множество. Это и есть *канторово множество.* Самым знаменитым канторовым множеством является случай, когда $\Delta_1 = [0,1]$, а в качестве каждого интервала берётся «средняя треть». Часто именно это множество называется канторовым, на английском языке оно называется *Middle Third Cantor Set.*

На семинарах и в ДЗ2 мы докажем, что любое канторово множество континуальное, замкнутое, нигде не плотное, не содержит изолированных точек.

Лекция 6

Компактные множества (компакты)

Нам понадобятся следующие определения.

Определение. *Любое подмножество покрытия, если оно также является покрытием, называется подпокрытием. Покрытие называют открытым, если все множества U_α открытые. Покрытие называют конечным, если оно содержит конечное множество множеств U_α.*

Докажем следующую фундаментальную лемму.

Лемма Гейне - Бореля. *Любое открытое покрытие $\cup_\alpha U_\alpha$ отрезка Δ имеет конечное подпокрытие.*

Доказательство. Пусть отрезок $\Delta_1 = [a_1, b_1], b_1 > a_1$ покрыт бесконечной системой открытых множеств U_α. Пусть у этой систему не существует конечного подпокрытия, т.е. нельзя выбрать конечный набор $U_k, k = 1, 2, \ldots, N$ открытых множеств, также покрывающее отрезок Δ_1. Разделим отрезок $[a_1, b_1]$ пополам точкой $(a_1 + b_1)/2$ на два отрезка: $L_1 = [a_1, (a_1 + b_1)/2]$ и $R_1 = [(a_1 + b_1)/2, b_1]$. По крайней мере для одного из отрезков L_1 или R_1 не существует конечного подпокрытия из изначального покрытия. Пусть например, это R_1. Обозначим отрезок R_1 через $[a_2, b_2]$. Разделим его пополам на отрезки L_2 и R_2, снова для одного из отрезков L_2 или R_2 не существует конечного подпокрытие из изначального покрытия. Выберем его и повторим эту процедуру бесконечно много раз. Мы получим последовательность вложенных отрезков $[a_n, b_n]$, длины этих отрезков равны $(b_n - a_n) = (b_1 - a_1)2^{-n+1} \to 0$ при $n \to \infty$. Эти отрезки строились так, чтобы ни один из них не мог быть покрыт конечным набором множеств из изначального покрытия. По лемме о вложенных отрезках пересечение всех этих отрезков состоит из единственной точки c. Эта точка c принадлежит какому-то открытому множеству U из изначального покрытия. Вместе с точкой c открытое множество U содержит и некоторую окрестность точки c. Теперь, при достаточно больших n отрезки $[a_n, b_n]$ все покрыты одним единственным открытым множеством U из покрытия. Полученное противоречие доказывает лемму Гейне–Бореля. $\square$

Определение. *Подмножество $A \subset \mathbb{R}$ называется компактным (или компактом), если из любого его покрытия можно извлечь конечное подпокрытие.*

Таким образом, мы доказали, что отрезок компактен. Также из леммы Гейне-Бореля следует, что любое ограниченное замкнутое множество является компактом. Любой компакт на прямой — обязательно ограниченное замкнутое множество. Это будет доказано позднее.

Предел последовательности

Определение. *Число A называется пределом последовательности $(x_n)_{n\in\mathbb{N}}$, если вне любой окрестности точки A лежит лишь конечное множество элементов последовательности $(x_n)_{n\in\mathbb{N}}$, или, что то же самое, все элементы $(x_n)_{n\in\mathbb{N}}$ с достаточно большими номерами n лежат внутри этой окрестности.*

Обозначение. $\lim_{n\to\infty} x_n = A$. Также часто пишут $x_n \to A$ при $n \to \infty$.

На языке кванторов: $\lim_{n\to\infty} x_n = A \Leftrightarrow \forall\varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N : |x_n - A| < \varepsilon$.

Определение. *Будем говорить, что последовательность $(x_n)_{n\in\mathbb{N}}$ сходится к $+\infty$, если $\forall C\ \exists N \in \mathbb{N}\ \forall n > N : x_n > C$. Аналогично определяется сходимость к $-\infty$: $x_n \to \infty \Leftrightarrow \forall C\ \exists N \in \mathbb{N}\ \forall n > N : x_n < C$.*

Заметим, что натуральное число N зависит от числа ε и от самой последовательности.

Последовательность $(x_n)_{n\in\mathbb{N}}$ может не сходиться ни к какому числу A. На языке кванторов это означает

$$\forall A\ \exists\varepsilon>0\ \forall N\in\mathbb{N}\ \exists n>N: |x_n-A|>\varepsilon$$

Определение. *Последовательность $(x_n)_{n\in\mathbb{N}}$ называется сходящейся, если её предел равен какому-то числу. В остальных случаях последовательность называется расходящейся. Причём в случае расходящейся последовательности есть дихотомия: либо такого числа, которому последовательность сходилась бы не существует, либо предел последовательности равен $\pm\infty$.*

Утверждение. *Всякая сходящаяся последовательность ограничена.*

Доказательство. Пусть $x_n\to A$ при $n\to\infty$. Выберем $\varepsilon=1$, построим по этому ε число N по определению сходимости. Теперь все элементы последовательности с номерами больше N лежат в 1-окрестности точки A, конечное множество точек $\{x_1,\dots,x_N\}$ ограничено, значит, вся последовательность ограничена.$\square$

Утверждение. *Если последовательность сходится, то её предел единственный.*

Доказательство. В самом деле, пусть у последовательности $(x_n)_{n\in\mathbb{N}}$ есть 2 разных предела, A и B. Без потери общности пусть, например, $A<B$. Положим $\varepsilon=(B-A)/3$ и рассмотрим ε-окрестности $U_\varepsilon(A)$ и $U_\varepsilon(B)$ точек A и B. По построению, эти окрестности не пересекаются. По определению предела вне окрестности $U_\varepsilon(A)$ лежит лишь конечное множество элементов последовательности. В частности в окрестности $U_\varepsilon(B)$ лежит конечное число элементов последовательности. Тогда B — не предел. $\square$

Теорема. *Пусть заданы последовательности x_n и y_n, причём $x_n\geqslant y_n$ при каждом $n\in\mathbb{N}$. Пусть обе последовательности x_n и y_n сходятся:* $\lim_{n\to\infty}x_n=A, \lim_{n\to\infty}y_n=B$. *Тогда $A\geqslant B$.*

Доказательство. Пусть это не так. Тогда $B>A$. Положим $\varepsilon=(B-A)/3$ и рассмотрим ε-окрестности точек A и B. По построению, эти окрестности не пересекаются, причём все точки из окрестности $U_\varepsilon(B)$ строго больше (лежат правее) всех точек из окрестности $U_\varepsilon(A)$. Но $x_n\to A, y_n\to B$ при $n\to\infty$, поэтому при достаточно больших значениях n справедливы включения $x_n\in U_\varepsilon(A), y_n\in U_\varepsilon(B)$ (опишите это в строгих математических терминах). Отсюда $x_n<y_n$ при достаточно больших n, что противоречит предположению теоремы. $\square$

Замечание 1. Для справедливости утверждения теоремы достаточно, чтобы неравенство $x_n\geqslant y_n$ было справедливо лишь для "достаточно больших значений" n.

Замечание 2. Если вместо $x_n\geqslant y_n$ справедливо строгое неравенство $x_n>y_n$, то отсюда также следует $A\geqslant B$ (из теоремы), но может не следовать $A>B$. Контрпример: $x_n=n^{-1}, y_n=-n^{-1}$. Здесь $x_n>y_n$, однако $\lim_{n\to\infty}x_n=\lim_{n\to\infty}y_n=0$.

Лекция 7

Следующая теорема является одним из важнейших средств нахождения пределов.

Теорема о двух милиционерах. *Пусть даны три последовательности, причём* $x_n \geqslant y_n \geqslant z_n$ *и пусть существуют пределы* $\lim_{n\to\infty} x_n, \lim_{n\to\infty} z_n$ *и* $\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n = A$. *Тогда существует предел* $\lim_{n\to\infty} y_n$ *и он также равен* A.

Доказательство. Запишем определения $\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n = A$

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N : (|x_n - A| < \varepsilon) \wedge (|z_n - A| < \varepsilon).$$

Но тогда для этих ε и N имеем: $A-\varepsilon < x_n \leqslant y_n \leqslant z_n < A+\varepsilon$. Это и означает, что $\lim_{n\to\infty} y_n = A$. $\square$

Арифметические операции с переделами

Пусть заданы последовательности x_n и y_n. Тогда определены также следующие последовательности

$$(\alpha x_n)_{n\in\mathbb{N}}, \text{где } \alpha \in \mathbb{R}, \quad (x_n + y_n)_{n\in\mathbb{N}}, \quad (x_n \cdot y_n)_{n\in\mathbb{N}}, \quad \left(\frac{x_n}{y_n}\right)_{n\in\mathbb{N}}.$$

Последняя последовательность определена, если $y_n \neq 0$.

Теорема. *Пусть* $\lim_{n\to\infty} x_n = A, \lim_{n\to\infty} y_n = B$. *Тогда*

$$1)\ \lim_{n\to\infty} \alpha x_n = \alpha A, \quad 2)\ \lim_{n\to\infty} (x_n + y_n) = A + B, \quad 3)\ \lim_{n\to\infty} (x_n \cdot y_n) = AB,$$

и если $y_n, B \neq 0$, *то* 4) $\lim_{n\to\infty} \left(\frac{x_n}{y_n}\right) = \frac{A}{B}$.

Следствие. *Операция взятия предела линейна, т.е.* $\lim_{n\to\infty}(\alpha x_n + \beta y_n) = \alpha A + \beta B$.

Доказательство. Докажем лишь 1) и 3) свойства.

1) По определению $\lim_{n\to\infty} x_n = A$ имеем $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n > N : |x_n - A| < \varepsilon$. Нужно показать, что $\forall \varepsilon_0 > 0\ \exists N_0 \in \mathbb{N}\ \forall n > N_0 : |\alpha x_n - \alpha A| < \varepsilon_0$. Если $\alpha = 0$, то утверждение очевидно. Пусть $\alpha \neq 0$. Выберем какое-то $\varepsilon_0 > 0$. Рассмотрим число $\varepsilon_0/|\alpha|$. Так как число ε было выбрано произвольно, возьмем в качестве него число $\varepsilon = \varepsilon_0/|\alpha|$ и построим по нему соответствующее число $N_0 = N(\varepsilon_0/|\alpha|)$. При всех $n > N_0$ будет справедливо неравенство $|x_n - A| < \varepsilon_0/|\alpha|$, эквивалентное $|\alpha x_n - \alpha A| < \varepsilon_0$.

3) Имеем $\forall \varepsilon > 0\ \exists N_1 \in \mathbb{N}\ \forall n > N_1 : |x_n - A| < \varepsilon$ и $\forall \varepsilon > 0\ \exists N_2 \in \mathbb{N}\ \forall n > N_2 : |y_n - B| < \varepsilon$. Т.к. обе последовательности сходятся, то они ограничены и значит найдется такое $C > 0$, что $|x_n|, |y_n| \leqslant C$ при всех $n \in \mathbb{N}$. Отсюда следует, что также $|A|, |B| \leqslant C$. Выберем какое-то $\varepsilon_0 > 0$ и рассмотрим число $\varepsilon = \frac{\varepsilon_0}{2C}$. Построим по нему соответствующие числа $N_1(\frac{\varepsilon_0}{2C})$ и $N_2(\frac{\varepsilon_0}{2C})$. Положим $N_0 = \max\{N_1, N_2\}$. Заметим, что при $n > N_0$ верно, что $|x_n - A| < \frac{\varepsilon_0}{2C}$ и $|y_n - B| < \frac{\varepsilon_0}{2C}$. Тогда

$$|x_n y_n - AB| = |x_n y_n - A y_n + A y_n - AB| \leqslant |x_n y_n - A y_n| + |A y_n - AB| =$$

$$= |y_n||x_n - A| + |A||y_n - B| \leqslant C|x_n - A| + C|y_n - B| < \frac{C\varepsilon_0}{2C} + \frac{C\varepsilon_0}{2C} < \varepsilon_0.$$

Таким образом, $|x_n y_n - AB| < \varepsilon_0$ при $n > N_0$. $\square$

Упражнение. Докажите свойства 2) и 4).

Подпоследовательности

Определение. *Последовательность строго возрастает, если $x_{n+1} > x_n$ при $n \in \mathbb{N}$.*

Пусть у нас есть последовательность $(x_n)_{n\in\mathbb{N}}$. Предположим, задана строго возрастающая последовательность натуральных чисел $(n_k)_{k\in\mathbb{N}}$. Тогда можно рассмотреть композицию отображений $k \mapsto n_k$ и $n \mapsto x_n$, ставящую в соответствие каждому натуральному числу k элемент последовательности $(x_{n_k})_{k\in\mathbb{N}}$. Такая последовательность называется *подпоследовательностью* последовательности $(x_n)_{n\in\mathbb{N}}$. Выбирая различные n_k: получаем различные подпоследовательности $(x_{n_k})_{k\in\mathbb{N}}$.

Теорема. *Если последовательность $(x_n)_{n\in\mathbb{N}}$ сходится, то и каждая подпоследовательность $(x_{n_k})_{k\in\mathbb{N}}$ сходится, причём к тому же самому пределу.*

Доказательство. $(x_n)_{n\in\mathbb{N}}$ сходится, значит, $\forall\varepsilon > 0\ \exists N \in \mathbb{N}:\ \forall n > N:\ |x_n - A| < \varepsilon$. В частности, для $n_k > N$ верно, что $|x_{n_k} - A| < \varepsilon$. Тогда берём $K = \min\{k:\ n_k > N\}$. Имеем $\forall\varepsilon > 0\ \exists K \in \mathbb{N}:\ \forall k > K:\ |x_{n_k} - A| < \varepsilon$. $\square$

Теорема. *Если подпоследовательность $(x_{n_k})_{k\in\mathbb{N}}$ содержит все элементы последовательности $(x_n)_{n\in\mathbb{N}}$ кроме конечного их множества, то из сходимости подпоследовательности следует сходимость самой последовательности.*

Доказательство. Пусть $(x_{n_k})_{k\in\mathbb{N}} \subset (x_n)_{n\in\mathbb{N}} \setminus \{x_{i_1}, \ldots, x_{i_l}\}$. Последовательность $(x_{n_k})_{k\in\mathbb{N}}$ сходится, значит, $\forall\varepsilon > 0\ \exists N \in \mathbb{N}:\ \forall k > N:\ |x_{n_k} - A| < \varepsilon$. Возьмём $\tilde{N} = \max\{n_N, i_1, \ldots, i_l\}$. Тогда для $n > \tilde{N}$ последовательности $(x_{n_k})_{k\in\mathbb{N}}$ и $(x_n)_{n\in\mathbb{N}}$ совпадают как множества (т.к. мы рассматриваем номера членов послеовательностей большие любого из $i_1, \ldots, i_l$). Более того последовательность $(x_{n_k})_{k\in\mathbb{N}}$ сходится. Значит, $\forall\varepsilon > 0\ \exists\tilde{N} \in \mathbb{N}:\ \forall n > \tilde{N}:\ |x_n - A| < \varepsilon$, т.е. $(x_n)_{n\in\mathbb{N}}$ сходится. $\square$

Теорема (лемма Больцано - Вейерштрасса). *Всякая ограниченная последовательность $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ имеет сходящуюся подпоследовательность.*

Доказательство. Так как элементов последовательности бесконечное (счётное) множество, то либо множество значений, которые принимают элементы последовательности, бесконечно, либо какое-то из значений последовательность принимает бесконечное количество раз (возможно, таких значений много). В этом случае можно выбрать постоянную подпоследовательность, которая, очевидно, сходится. Если же множество значений, которые принимают элементы последовательности, бесконечно, то воспользуемся теоремой о том, что каждое ограниченное бесконечное множество имеет по крайней мере одну предельную точку A. Эта предельная точка и есть искомый предел подпоследовательности. Для указания конкретной подпоследовательности, сходящейся к этой предельной точке A, надо положить $\varepsilon_1 = 1$, выбрать ε_1-окрестность $U_{\varepsilon_1}(A)$, в ней найти элемент x_{n_1} (там лежит бесконечно много элементов последовательности), потом надо положить $\varepsilon_2 = 1/2$, выбрать меньшую ε_1-окрестность $U_{\varepsilon_1}(A)$, в ней найти элемент x_{n_2}, проследив, чтобы было $n_2 > n_1$, и так далее. Полученная подпоследовательность $(x_{n_k})_{k\in\mathbb{N}}$ сходится к A. $\square$

Фундаментальные последовательности

Определение. *Последовательность $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ называется фундаментальной последовательностью или последовательность Коши если*

$$\forall\varepsilon > 0\ \exists N \in \mathbb{N}\colon \forall n, m > N\ |x_n - x_m| < \varepsilon.$$

Перечислим основные свойства фундаментальных последовательностей

Лемма 1. *Всякая фундаментальная последовательность ограничена.*

Доказательство. Пусть $\varepsilon = 1$, по определению фундаментальности найдется такое N, что при $n, m > N$ расстояние между элементами последовательности меньше 1. Таким образом, все x_n, $n > N + 1$ лежат в 1-окрестности числа x_{N+1}. Теперь конечное множество $\{x_1, \ldots, x_N\}$ ограничено, поэтому и вся последовательность $(x_n)_{n\in\mathbb{N}}$ ограничена. □

Лемма 2. *Если подпоследовательность фундаментальной последовательности сходится к числу A, то сама последовательность сходится к A.*

Доказательство. Зададимся $\varepsilon > 0$. По числу $\varepsilon/2$ построим N из условия фундаментальности: при $n, m > N$ справедливо соотношение $|x_m - x_n| < \varepsilon/2$. По условию сходимости подпоследовательности $(x_{n_k})_{k\in\mathbb{N}}$ к пределу A, существует такое k_0, что при всех $k > k_0$ справедливы соотношения $|A - x_{n_k}| < \varepsilon/2$. Теперь пусть число $k_1 > k_0$ такое, что $n_{k_1} > N$. Тогда $|A - x_{n_{k_1}}| < \varepsilon/2$ и при всех $n > k_1$ справедливы неравенства $|x_{n_{k_1}} - x_n| < \varepsilon/2$, следовательно, справедливо неравенство

$$|A - x_n| \leqslant |A - x_{n_{k_1}}| + |x_{n_{k_1}} - x_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \square$$

Теорема (Критерий Коши). *Всякая фундаментальная последовательность сходится, всякая сходящаяся последовательность фундаментальна.*

Доказательство. В одну сторону доказательство простое. Пусть последовательность $(x_n)_{n\in\mathbb{N}}$ сходится к A. Тогда для любого $\varepsilon > 0$, начиная с некоторого N все элементы последовательности лежат в $\varepsilon/2$-окрестности точки A. Это значит, что начиная с этого N расстояние между элементами последовательности меньше ε. Фундаментальность сходящейся последовательности доказана.

Пусть $(x_n)_{n\in\mathbb{N}}$ — фундаментальная последовательность. Тогда по лемме 1 последовательность $(x_n)_{n\in\mathbb{N}}$ ограничена. Теперь, по теореме Больцано-Вейерштрасса из неё можно выбрать сходящуюся подпоследовательность $(x_{n_k})_{k\in\mathbb{N}}$. По лемме 2 вся последовательность $(x_n)_{n\in\mathbb{N}}$ сходится к тому же пределу. □

Лекция 8

Монотонные последовательности и теорема Вейерштрасса.

Определение. *Последовательность $(x_n)_{n\in\mathbb{N}}$ монотонно возрастает, если $x_{n+1} \geqslant x_n$ при $n \in \mathbb{N}$, последовательность строго возрастает, если $x_{n+1} > x_n$ при $n \in \mathbb{N}$. Про монотонно возрастающую последовательность говорят "неубывающая последовательность". Аналогично, последовательность монотонно убывает, если $x_{n+1} \leqslant x_n$ при $n \in \mathbb{N}$, последовательность строго убывает, если $x_{n+1} < x_n$ при $n \in \mathbb{N}$. Про монотонно убывающую последовательность говорят "невозрастащая последовательность". Последовательность $(x_n)_{n\in\mathbb{N}}$ называется монотонной, если она монотонно возрастает или монотонно убывает. Последовательность называется строго монотонной, если она строго монотонно возрастает или строго монотонно убывает.*

Теорема Вейерштрасса. *Ограниченная монотонная последовательность сходится. Более точно, невозрастающая ограниченная снизу последовательность сходится, неубывающая ограниченная сверху последовательность сходится.*

Доказательство. Докажем только одно утверждение, что неубывающая ограниченная сверху последовательность сходится. Случай ограниченной снизу последовательности аналогичен. Пусть $(x_n)_{n\in\mathbb{N}}$ — наша последовательность и пусть $x_{n+1} \geqslant x_n$. Множество значений этой последовательности ограничено сверху, значит, определено число $A = \sup_{n\in\mathbb{N}}\{x_n\}$. Покажем, что $\lim_{n\to\infty} x_n = A$. Зафиксируем $\varepsilon > 0$. Среди чисел $(x_n)_{n\in\mathbb{N}}$ есть числа, для которых $x_n > A - \varepsilon$ (иначе число A не было бы точной верхней гранью чисел x_n). Пусть $x_N > A - \varepsilon$, тогда при всех $n > N$, во-первых, $x_n > A - \varepsilon$, во-вторых, $x_n \leqslant A$ (иначе A не было бы верхней границей для множества значений x_n). Из $x_n > A - \varepsilon$ и $x_n \leqslant A$ следует $|x_n - A| < \varepsilon$. $\square$

Также нам будет полезна следующее утверждение

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

Доказательство. Пусть $x_n \leqslant x_{n+1}$ и A — точная верхняя грань последовательности $(x_n)_{n\in\mathbb{N}}$. Пусть $\varepsilon > 0$. Найдется такое число N, что $x_N > A - \varepsilon$ (иначе A — не точная верхняя грань). Тогда $x_n > A - \varepsilon$ при всех $n \geqslant N$. Кроме того, $x_n \leqslant A$ при всех n. $\square$

Частичные пределы.

Мотивация: разные подпоследовательности одной и той же последовательности могут сходиться к разным пределам (если последовательность не сходится). Пределы подпоследовательностей называются *частичными пределами.* Частичный предел — это величина, в любой окрестности которой находится бесконечное количество элементов последовательности. Частичные пределы называют также *предельными точками последовательности.* **Не путать с предельными точками множеств, в частности с предельными точками множества значений последовательности!**

Определение. *Нижним пределом последовательности $(x_n)_{n\in\mathbb{N}}$ называется называется предел $\lim_{n\to\infty}(\inf_{m\geqslant n} x_m)$. Он обозначается как $\liminf_{n\to\infty} x_n$ или $\underline{\lim}_{n\to\infty} x_n$. Верхним пределом последовательности $(x_n)_{n\in\mathbb{N}}$ называется называется предел $\lim_{n\to\infty}(\sup_{m\geqslant n} x_m)$. Он обозначается как $\limsup_{n\to\infty} x_n$ или $\overline{\lim}_{n\to\infty} x_n$.*

Замечание. *Не сложно понять, что из определения супремума множества следует, что последовательность $(s_n)_{n\in\mathbb{N}}$ с $s_n = \sup_{m\geqslant n} x_m$ является невозрастающей, т.е. $s_{n+1} \leqslant s_n$,* *а из определения инфимума множества следует, что последовательность* $(s_n)_{n\in\mathbb{N}}$ *с* $s_n = \sup_{m\geqslant n} x_m$ *является неубываеющей* $i_{n+1} \geqslant i_n$. *Если последовательность* $(s_n)_{n\in\mathbb{N}}$ *не ограничена снизу, то полагаем* $\limsup_{n\to\infty} x_n = -\infty$. *Если же последовательность* $(i_n)_{n\in\mathbb{N}}$ *не ограничена сверху, то полагаем* $\liminf_{n\to\infty} x_n = +\infty$.

Примеры. $\limsup_{n\to\infty} \frac{1}{n} = \liminf_{n\to\infty} \frac{1}{n} = \lim_{n\to\infty} \frac{1}{n} = 0$,
$\limsup_{n\to\infty}(-1)^n = 1, \liminf_{n\to\infty}(-1)^n = -1$, $\limsup_{n\to\infty} n = +\infty = \liminf_{n\to\infty} n$.

Очевидно также, что $\limsup_{n\to\infty} x_n \geqslant \liminf_{n\to\infty} x_n$.

Упражнение. Докажите это очевидное утверждение.

Обсуждение верхних и нижних пределов не даром происходит в разделе про частичные пределы: они также являются частичными пределами. Более того верна следующая теорема.

Теорема. *Нижний и верхний пределы ограниченной последовательности являются, соответственно, наименьшим и наибольшим из её частичных пределов.*

Доказательство. Доказательство проведём только для $\liminf_{n\to\infty} x_n$ (доказательство для $\limsup_{n\to\infty} x_n$ аналогично). Рассмотрим последовательность $i_n = \inf_{m\geqslant n} x_m$. Она монотонно не убывает, т.е. $i_n \leqslant i_{n+1}$. Т.к. по предположению последовательность $(x_n)_n$ ограничена, то, в частности, она ограничена сверху и по теореме Вейерштрасса существует предел $i = \lim_{n\to\infty} i_n$. Для чисел $n \in \mathbb{N}$ по определению инфимума подберём такие числа $k_n \in \mathbb{N}$, что $k_n < k_{n+1}$ и

$$i_n \leqslant x_{k_n} < i_n + \frac{1}{n}.$$

Т.к. $\lim_{n\to\infty} i_n = i = \lim_{n\to\infty} \left(i_n + \frac{1}{n}\right)$, то по теореме о двух милиционерах $\lim_{n\to\infty} x_{k_n} = i$,т.е. i — частиный предел последовательности $(x_n)_n$. Это наименьший частичный предел: $\forall \varepsilon > 0\ \exists n \in \mathbb{N}:\ i - \varepsilon < i_n$, т.е. $i - \varepsilon < \inf_{m\geqslant n} x_m \leqslant x_m$ для всякого $m \geqslant n$. Это означает, что ни один частичный предел не может быть меньше $i - \varepsilon$, т.е. для всякого частичного предела A верно, что $-\varepsilon \leqslant A$ (получено переходом к пределу в неравенстве). Тогда беря супремум по всем $\varepsilon > 0$, получаем $i = \sup_{\varepsilon>0}(i - \varepsilon) \leqslant A$, т.е. i — наименьший из частичных пределов. $\square$

Для того, чтобы понять что последовательность $(x_n)_{n\in\mathbb{N}}$ расходится достаточно найти два различных частичных предела. Действительно, если бы наша последовательность сходилась, тогда всякая бы её подпоследовательность также сходилась, причём к тому же самому числу, т.е. все частичные пределы были бы равны между собой (и равны пределу последовательности). Однако, как понять что наша последовательность сходится (если это сложно сделать “в лоб”)? По идее, нам бы помогло рассмотрение *всех* частичных предел. Если бы мы убедились, что все они равны, то это было бы аргументом в пользу сходимости самой последовательности. Разумеется, такой подход на практике сложно реализуем. Поэтому крайне удобной будет следующая теорема, которая вскрывает важность верхнего и нижнего пределов.

Теорема. *Последовательность* $(x_n)_{n\in\mathbb{N}}$ *сходится тогда и только тогда, когда* $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$.

Доказательство. Если последовательность $(x_n)_{n\in\mathbb{N}}$ сходится, то любая её подпоследовательность сходится причём к тому же числу. Т.е. все частичные пределы равны между собой и равны пределу $(x_n)_{n\in\mathbb{N}}$, в частности равны пределу $(x_n)_{n\in\mathbb{N}}$ и частичные пределы $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$.

Обратно. Рассмотрим последовательности $(i_n)_{n\in\mathbb{N}}$ и $(s_n)_{n\in\mathbb{N}}$ такие, что $i_n = \inf_{m\geqslant n} x_m, s_n = \sup_{m\geqslant n} x_m$. Имеем

$$i_n \leqslant x_n \leqslant s_n.$$

Применяя теорему о двух милиционерах получаем утверждение теоремы. $\square$

Числовые ряды

Важную роль играют последовательности, заданные в виде сумм элементов других последовательностей. Рассмотрим последовательность $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ и напишем «бесконечную сумму»:

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

Это выражение называется числовым рядом или просто рядом. Однако, такое определение врядли можно считать строгим. Чтобы дать строгое определение ряда рассмотрим величину $S_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + \ldots + a_n$. Она называется n-ой *частичной суммой.*

Определение. *(Числовым) рядом называется предел частичных сумм, т.е.* $\sum_{n=1}^{\infty} a_n := \lim_{n\to\infty} S_n$, *где* a_n *— элементы последовательности* $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$.

Определение. *Ряд* $\sum_{n=1}^{\infty} a_n$ *называется сходящимся, если сходится последовательность частичных сумм, иными словами, если существует предел* $S = \lim_{n\to\infty} S_n$. *Этот предел* S *(если он существует) называется суммой ряда* $\sum_{n=1}^{\infty} a_n$. *Если этот предел не существует, то ряд называется расходящимся, сумма ряда в этом случае не определена. Наконец, если предел равен* $\pm\infty$, *то ряд называется расходящимся к* $\pm\infty$ *соответственно.*

Теорема. *Сходимость ряда не зависит от выбрасывания (добавления) конечного множества начальных слагаемых, т.е.*

$$\sum_{n=1}^{\infty} a_n \text{ сходится} \Leftrightarrow \sum_{n=m+1}^{\infty} a_n \text{ сходится.}$$

Доказательство. Пусть $S_n = \sum_{k=1}^{n} a_k$ и $\tilde{S}_n = \sum_{k=m+1}^{n+m} a_k$. Заметим, что $\tilde{S}_n = S_{n+m} - S_m$. Утверждение следует т.к. $\lim_{n\to\infty} S_n = \lim_{n\to\infty} S_{n+m}$. Тогда $\lim \widetilde{S}_n = \lim_{n\to\infty} S_n - S_m$ $\square$

Замечание. Выбрасывание конечного множества начальных слагаемых меняет сумму ряда!

Примеры. 1) *Телескопический ряд*: путь дана последовательность $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$. Ряд вида $\sum_{n=1}^{\infty}(a_{n+1} - a_n)$ называется *телескопическим.* Найдем его сумму

$$\sum_{n=1}^{\infty}(a_{n+1} - a_n) = \lim_{n\to\infty} \sum_{k=1}^{n}(a_{n+1} - a_n) = \lim_{n\to\infty}((a_2 - a_1) + (a_3 - a_2) + \ldots + (a_{n+1} - a_n)) =$$
$$= \lim_{n\to\infty}(a_{n+1} - a_1) = \lim_{n\to\infty}(a_n) - a_1,$$

где в последнем равенстве мы воспользовались тем, что $\lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} a_n$, если $(a_n)_{n\in\mathbb{N}}$ сходится. Таким образом, телескопический ряд сходится тогда и только тогда, когда сходится образующая его последовательность $(a_n)_{n\in\mathbb{N}}$.

2) *Геометрический ряд*: он определяется как предел *геометрической прогрессии*, т.е.

$$\sum_{n=1}^{\infty} q^{n-1} = \lim_{n\to\infty} \sum_{k=1}^{n} q^{k-1} = \lim_{n\to\infty} \frac{1 - q^n}{1 - q},$$

где $q \neq 1$. Если $|q| < 1$, то геометрический ряд сходится и его сумма равна $\frac{1}{1-q}$.

3) *Гармонический ряд*: он определяется следующим образом

$$\sum_{n=1}^{\infty} \frac{1}{n} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k}.$$

Частичная сумма гармонического ряда имеет специальное обозначение H_n и называется *гармоническим числом.* Покажем, что гармонический ряд является расходящимся (к $+\infty$). Для этого используем неравенство, которое было доказано на семинарах:

$$\frac{1}{n+1} < \ln\left(1+\frac{1}{n}\right) < \frac{1}{n}. \tag{0.1}$$

(В этом неравенстве и при его выводе фигурирует пока не очень понятная функция ln. Пока что мы будем понимать её в "школьном смысле", т.е. все свойства логарифма считаем известными. В начале второго модуля мы строго обоснуем все эти свойства. Никаких логических циклов у нас при этом не возникнет.)

Это неравенство даёт:

$$\begin{gathered} \ln 2 - \ln 1 < 1, \\ \ln 3 - \ln 2 < \frac{1}{2}, \\ \dots \\ \ln(n+1) - \ln n < \frac{1}{n}. \end{gathered}$$

Складывая, получим:

$$\ln(n+1) < H_n. \tag{0.2}$$

Рассмотрим последовательность, чей n-ый член имеет вид $x_n = H_n - \ln n$. Мы утверждаем, что эта последовательность сходится. Действительно, $x_{n+1} > \frac{1}{n+1}$, как это следует из неравенства (0.2). Значит, последовательность $(x_n)_n$ ограничена снизу. С другой стороны, $x_n - x_{n+1} = -\frac{1}{n+1} + \ln(n+1) - \ln n = \ln(n+1) - \frac{1}{n+1} > 0$, как это следует из неравенства (0.1). Значит, $(x_n)_n$ монотонно убывает. Тогда по теореме Вейерштрасса существует передел $\gamma = \lim_{n\to\infty}(H_n - \ln n)$. Этот предел называется *постоянной Эйлера-Маскерони.* Её численное значение примерно равно $0,577216$. Это одна из замечательных констант в математике. Таким образом, $\lim_{n\to\infty}(H_n - \ln n - \gamma) = 0$. Перепишем это так

$$H_n - \ln n - \gamma = \varepsilon_n,$$

где $\varepsilon_n \to 0$ при $n \to \infty$. Тогда

$$H_n = \ln n + \gamma + \varepsilon_n.$$

Тогда переходя к пределу при $n \to \infty$, получим, что $\sum_{n=1}^{\infty} \frac{1}{n} = \lim_{n\to\infty} H_n = +\infty$, т.к. $\ln n$ не органичен сверху.

Сейчас мы займёмся общими вопросами сходимости рядов.

Теорема (необходимое условие сходимости ряда). *Если ряд* $\sum_{n=1}^{\infty} a_n$ *сходится, то последовательность* $(a_n)_{n\in\mathbb{N}}$ *инфинтезимальна, т.е.* $\lim_{n\to\infty} a_n = 0$.

Доказательство. Заметим, что $a_n = S_n - S_{n-1}$. Тогда

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}(S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = 0. \square$$

Замечание. Как показывает пример гармонического ряда, обратное к предыдущей теореме утверждение не верно.

Непосредственно из предыдущей теоремы получаем

Следствие (признак расходимости ряда). *Пусть последовательность* $(a_n)_{n\in\mathbb{N}}$ *не инфинтезимальна, т.е.* $\lim_{n\to\infty} a_n \neq 0$, *тогда ряд* $\sum_{n=1}^{\infty} a_n$ *расходится.*

Примеры. (i) Ряды $\sum_{n=1}^{\infty}(-1)^n, \sum_{n=1}^{\infty} \frac{1}{\sqrt[n]{n}}$ расходятся.

(ii) Рассмотрим *p-ряд* или *ряд Дирихле*: $\sum_{n=1}^{\infty}\frac{1}{n^p}$. При $p \leqslant 0$ имеем $\frac{1}{n^p} \nrightarrow 0$ при $n \to \infty$, тогда по признаку расходимости, ряд расходится.

Следующий критерий нужен для доказательства почти всех теорем о рядах, всех признаков сходимости и расходимости. Непосредственно к исследованию конкретных рядов критерий Коши, как правило, не применяется. Это просто перефразировка теоремы «последовательность сходится, если и только если она фундаментальна» для последовательности частичных сумм ряда.

Теорема (критерий Коши). *Для того, чтобы ряд $\sum_{n=1}^{\infty} a_n$ сходился, необходимо и достаточно, чтобы*

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N} \colon \forall m > 0, n > N\ |\sum_{k=n+1}^{n+m} a_k| < \varepsilon.$$

Доказательство. Достаточно заметить, $|S_{n+m} - S_n| = |\sum_{k=n+1}^{n+m} a_k|$ и воспользоваться критерием Коши для последовательности $(S_n)_{n\in\mathbb{N}}$. $\square$

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

Доказательство. Применим дважды признак Коши в разные стороны. Так как ряд $\sum_{n=1}^{\infty} |a_n|$ сходится, то

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N} \colon \forall m > 0, n > N\ \sum_{k=n+1}^{n+m} |a_k| < \varepsilon.$$

Поэтому, $\forall \varepsilon > 0$ можно взять то же самое N: $\forall m > 0, n > N$ справедливо $|\sum_{k=n+1}^{n+m} a_k| \leqslant \sum_{k=n+1}^{n+m} |a_k| < \varepsilon$. Отсюда следует сходимость ряда $\sum_{n=1}^{\infty} a_n$. $\square$

Определение. *Ряд $\sum_{n=1}^{\infty} a_n$ абсолютно сходится, если ряд $\sum_{n=1}^{\infty} |a_n|$ сходится. Ряд сходится условно, если он сходится, но не сходится абсолютно.*

Следствие произносится теперь по-другому: *абсолютно сходящийся ряд сходится.* Заметим также, что если ряд сходится условно, то оба ряда, составленные из его членов одного знака, расходятся. Если бы ровно один из этих рядов сходился, то весь ряд бы расходился; если бы сходились оба, исходный ряд сходился бы абсолютно.

Примеры. Ряд $\sum_{n=1}^{\infty}(-\frac{1}{2})^n$ сходится абсолютно; мы скоро покажем, что *знакопеременный гармонический ряд* $\sum_{n=1}^{\infty}(-1)^n\frac{1}{n}$ сходится условно; ряд $\sum_{n=1}^{\infty}(-1)^n$ не сходится ни абсолютно, ни условно (он расходится).

Теорема. *Гармонический ряд расходится.*

Доказательство. Воспользуемся критерием Коши. Положим $m = n$, тогда для каждого натурального n справедлива оценка

$$|\sum_{k=n+1}^{2n} \frac{1}{k}| = \sum_{k=n+1}^{2n} \frac{1}{k} \geqslant \sum_{k=n+1}^{2n} \frac{1}{2n} = n \cdot \frac{1}{2n} = \frac{1}{2}.$$

По критерию Коши гармонический ряд расходится. $\square$

Наконец, отметим такое свойство сходящихся рядов

Теорема (о линейной комбинации сходящихся рядов). *Пусть ряды $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} b_n$ сходятся. Тогда для любых $\alpha, \beta \in \mathbb{R}$ ряд $\sum_{n=1}^{\infty}(\alpha a_n + \beta b_n)$ тоже сходится.*

Доказательство. Достаточно воспользоваться линейностью предела, т.е.

$$\lim_{n\to\infty}(\alpha x_n + \beta y_n) = \alpha \lim_{n\to\infty} x_n + \beta \lim_{n\to\infty} y_n$$

и взять в качестве $(x_n)_{n\in\mathbb{N}}$ и $(y_n)_{n\in\mathbb{N}}$ последовательности частичных сумм для рядов $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} b_n$ соответственно. $\square$

Лекция 9

Ряды с положительными (неотрицательными) членами

Рассмотрим ряд $\sum_{n=1}^{\infty} a_n$, где $a_n \geqslant 0$. Заметим что последовательность его частичных сумм $(S_n)_{n\in\mathbb{N}}$ является неубывающей. Тогда верна следующая теорема

Теорема. *Для сходимости ряда с неотрицательными (положительными) членами необходимо и достаточно, чтобы последовательность частичных сумм была ограничена.*

Доказательство. Последовательность частичных сумм ряда с неотрицательными членами не убывает. Поэтому, если эта последовательность ограничена, то она сходится по теореме Вейерштрасса. Обратно, если эта последовательность сходится, то она ограничена. □

Замечание. (i) Сумма сходящегося ряда с положительными членами совпадает с точной верхней гранью частичных сумм.

(ii) Для ряда с неотрицательными (положительными) членами верна дихотомия: либо он сходится, либо он расходится к $+\infty$.

Признаки сходимости рядов

Теорема (признак сравнения). *Пусть даны два ряда $\sum_{n=1}^{\infty} a_n, \sum_{n=1}^{\infty} b_n, a_n, b_n \geqslant 0$. Пусть при некотором $C > 0$ для всех n справедливо неравенство $a_n \leqslant Cb_n$. Тогда из сходимости ряда $\sum_{n=1}^{\infty} b_n$ вытекает сходимость ряда $\sum_{n=1}^{\infty} a_n$, а из расходимости ряда $\sum_{n=1}^{\infty} a_n$ вытекает расходимость ряда $\sum_{n=1}^{\infty} b_n$.*

Доказательство. Пусть S_n — частичные суммы ряда $\sum_{n=1}^{\infty} a_n$, а $\tilde{S}_n$ — частичные суммы ряда $\sum_{n=1}^{\infty} b_n$. Имеем $0 \leqslant S_n \leqslant C\tilde{S}_n$ для всех n. Тогда если последовательность $(\tilde{S}_n)_{n\in\mathbb{N}}$ сходится, то сходится и последовательность $(S_n)_{n\in\mathbb{N}}$. Наоборот, если последовательность $(S_n)_{n\in\mathbb{N}}$ расходится, то расходится и последовательность $(\tilde{S}_n)_{n\in\mathbb{N}}$. □

Замечание. (i) Как всегда, в условиях теоремы достаточно чтобы неравенство $a_n \leqslant Cb_n$ выполнялось начиная с некоторого N.

(ii) В условиях теоремы ряд $\sum_{n=1}^{\infty} a_n$ называется *минорирующим рядом*, а ряд $\sum_{n=1}^{\infty} b_n$ называется *мажорирующим рядом.*

Теорема (предельный признак сравнения). *Пусть даны два ряда $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} b_n$ с положительными членами и пусть существует конечный положительный предел*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = C > 0.$$

Тогда ряды $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} b_n$ сходятся и расходятся одновременно.

Доказательство. В силу условия при достаточно больших n выполнены неравенства $\frac{1}{2}Cb_n \leqslant a_n \leqslant \frac{5}{2}Cb_n$. □

Следующая теорема — весьма практичный метод исследования сходимости рядов с положительными членами.

Теорема (конденсационный признак Коши). *Пусть последовательность $(a_n)_{n\in\mathbb{N}}, a_n > 0$ монотонна и убывает к нулю. Тогда ряды*

$$\sum_{n=1}^{\infty} a_n \text{ и } \sum_{n=1}^{\infty} 2^n a_{2^n}$$

сходятся или расходятся одновременно.

Доказательство. Запишем неравенства

$$\begin{gathered}
a_2 \leqslant a_2 \leqslant a_1, \\
2a_4 \leqslant a_3 + a_4 \leqslant 2a_2, \\
4a_8 \leqslant a_5 + a_6 + a_7 + a_8 \leqslant 4a_4, \\
8a_{16} \leqslant a_9 + a_{10} + \ldots + a_{15} + a_{16} \leqslant 8a_8, \\
\ldots \\
2^{n-1}a_{2^n} \leqslant a_{2^{n-1}+1} + \ldots + a_{2^n} \leqslant 2^{n-1}a_{2^{n-1}}
\end{gathered}$$

и сложим их. Положим $\tilde{S}_n = \sum_{k=1}^{n} 2^k a_{2^k}$, получим $\frac{1}{2}(\tilde{S}_n - a_1) \leqslant S_{2^n} - a_1 \leqslant \tilde{S}_n$. Тогда из сходимости последовательности $(S_n)_{n\in\mathbb{N}}$ следует сходимость последовательности $(\tilde{S}_n)_{n\in\mathbb{N}}$ и наоборот. $\square$

Применяя эту замечательную теорему можно доказать, что

(i) Гармонический ряд расходится: он удовлетворяет условию конденсационного признака Коши и потому

$$\sum_{n=1}^{\infty} \frac{1}{n} \sim \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = \sum_{n=1}^{\infty} 1,$$

что является расходящимся рядом.

(ii) Мы уже видели, что p-ряд расходится при $p \leqslant 0$. При $p > 0$ ряд удовлетворяет условию конденсационного признака Коши и потому

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \sim \sum_{n=1}^{\infty} 2^n \frac{1}{2^{pn}} = \sum_{n=1}^{\infty} \frac{1}{2^{(p-1)n}},$$

получили геометрический ряд; он сходится при $p > 1$.

(iii) Ряд $\sum_{n=1}^{\infty} \frac{1}{n \ln n}$ удовлетворяет условию конденсационного признака Коши (почему?) и потому

$$\sum_{n=1}^{\infty} \frac{1}{n \ln n} \sim \sum_{n=1}^{\infty} 2^n \frac{1}{2^n \ln 2^n} = \sum_{n=1}^{\infty} \frac{1}{n \ln 2} = \ln 2 \sum_{n=1}^{\infty} \frac{1}{n}$$

т.е. он расходится.

Лекция 10

Продолжим обсуждение признаков сходимости. Рассмотрим два следующих важных признака

Теорема (признак Даламбера). *Пусть дан ряд $\sum_{n=1}^{\infty} a_n$ с $a_n > 0$. Обозначим $q = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$. Тогда*
(i) если $q > 1$, то ряд расходится;
(ii) если $q < 1$, то ряд сходится.

Теорема (радикальный признак Коши). *Пусть дан ряд $\sum_{n=1}^{\infty} a_n$ с $a_n \geqslant 0$. Обозначим $q = \lim_{n\to\infty} \sqrt[n]{a_n}$. Тогда*
(i) если $q > 1$, то ряд расходится;
(ii) если $q < 1$, то ряд сходится.

Замечание. В обоих признаках в случае $q = 1$ вопрос о сходимости ряда остаётся открытым (например рассмотрите p-ряд при произвольном p).

Доказательство. Доказательство расходимости: если $q > 1$, то существует $N \in \mathbb{N}$ такое, что
1) для признака Даламбера: $\frac{a_{n+1}}{a_n} > 1 \Rightarrow a_{n+1} > a_n,\ \forall n > N$ (т.е. $a_n > a_N$);
2) для радикального Коши: $\sqrt[n]{a_n} > 1 \Rightarrow a_n > 1,\ \forall n > N$.
В обоих случая последовательность $(a_n)_n$ не инфинитезимальна. Значит, ряд расходится.
Доказательство сходимости: если $q < 1$, то выберем $r \in (q, 1)$. Тогда
1) для признака Даламбера: найдём N такое, что $\forall n \geqslant N$ имеем $a_{n+1} < ra_n$, тогда

$$a_{N+k} < ra_{N+k-1} < r^2 a_{N+k-2} < \ldots < r^k a_N,\ \forall k \in \mathbb{N};$$

2) для радикального Коши: найдём N такое, что $\forall n \geqslant N$ имеем $a_n < r^n$.
Такое N действительно найдётся: для этого достаточно положить $r = q+\varepsilon < 1$ (почему такое $\varepsilon > 0$ есть?) и воспользоваться условием сходимости $q = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ или $q = \lim_{n\to\infty} \sqrt[n]{a_n}$, соответственно.
Далее ряд $\sum_{n=1}^{\infty} r^n$ сходится. Тогда по признаку сравнения сходится и ряд $\sum_{n=1}^{\infty} a_n$. $\square$

Знакопеременные ряды

В заключении рассмотрим следующий сходимости знакопеременного ряда специального вида

Теорема (признак Лейбница). *Пусть дан знакопеременный ряд вида $\sum_{n=1}^{\infty}(-1)^{n-1}b_n$, где $b_n \to 0$ при $n \to \infty$ и $b_{n+1} \leqslant b_n$. Тогда этот ряд сходится.*

Доказательство. Рассмотрим последовательность чётных частичных сумм $(S_{2n})_n$. Она не убывает

$$S_{2n} = S_{2n-2} + b_{2n-1} - b_{2n} \geqslant S_{2n-2}.$$

Рассмотрим последовательность чётных частичных сумм $(S_{2n})_n$. Она не возрастает

$$S_{2n+1} = S_{2n-1} - b_{2n} + b_{2n+1} \leqslant S_{2n-1}.$$

Таким образом

$$S_2 \leqslant S_4 \leqslant \ldots \leqslant S_{2n} \leqslant S_{2n+1} \leqslant S_{2n-1} \leqslant \ldots \leqslant S_3 \leqslant S_1.$$

Следовательно, обе последовательности сходятся. Более того, т.к. $S_{2n+1} = S_{2n} + b_{2n+1}$, то переходя к пределу получаем, что предел обеих последовательностей один и тот же и равен S. Тогда и вся последовательность $(S_n)_n$ сходится к S т.к. начиная с некоторого N все члены этой последовательности лежат в ε-окрестности точки S. $\square$

Пример. Из признака Лейбница моментально получаем, что ряд $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ сходится. Заметим, что его абсолютный ряд расходится, т.е. ряд сходится условно.

Замечание. Условие монотонности в признаке Лейбница отбросить нельзя! чтобы в этом убедиться, рассмотрим такой ряд

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\sqrt{n}+(-1)^{n+1}}.$$

Легко убедиться в том, что для членов этого ряда нет никакой монотонности (правда есть условие инфинитезимальности). Предположим, что этот ряд сходится. Тогда вычтем его из ряда

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{\sqrt{n}},$$

который точно сходится по признаку Лейбница. Получим ряд

$$\sum_{n=1}^{\infty} \left(\frac{(-1)^{n+1}}{\sqrt{n}} - \frac{(-1)^{n+1}}{\sqrt{n}+(-1)^{n+1}} \right) = \sum_{n=1}^{\infty} \frac{1}{n\left(1+\frac{(-1)^{n+1}}{\sqrt{n}}\right)} \sim \sum_{n=1}^{\infty} \frac{1}{n},$$

где в последнем действии мы применили предельный признак сравнения. Итак, мы получили, что разность двух сходящихся рядов расходится. Это противоречит теореме о линейной комбинации сходящихся рядов. Значит, наш ряд расходится.

Представление числа e в виде ряда

На семинарах мы доказали, что

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

Рассмотрим ряд

$$\sum_{n=0}^{\infty} \frac{1}{n!}. \tag{0.3}$$

Он сходится по признаку Даламбера. Тогда сумма этого ряда — это некоторое число. Покажем, что оно равно e, т.е.

$$\sum_{n=0}^{\infty} \frac{1}{n!} = e.$$

Применим формулу бинома Ньютона

$$(1+b)^n = 1 + \frac{n}{1!}b + \frac{n(n-1)}{2!}b^2 + \ldots + \frac{n!}{(n-1)!}b^{n-1} + b^n$$

к $b = \frac{1}{n}$. Получим:

$$x_n = 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \ldots$$
$$+ \frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-1}{n}\right).$$

Отсюда следует, что

$$x_n < 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!} =: S_n,$$

где S_n — n-ая частичная сумма ряда (0.3). Покажем, что $\forall k \in \mathbb{N}$ $S_k \leqslant e$. Действительно, зафиксируем $k \in \mathbb{N}$. Получим, что

$$1+1+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)+\ldots+\frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k}{n}\right)<x_n$$

т.к. это урезанная формула бинома Ньютона для x_n. Выбранное k фиксированно, но n может меняться. Устремим его к бесконечности, получим:

$$S_k \leqslant e,$$

т.к. $x_n \to e$ при $n \to \infty$.

Таким образом, мы получили, что:

$$x_n < S_n \leqslant e.$$

Тогда по теореме о двух милиционерах, $\lim_{n\to\infty} S_n = e$. Это и есть сумма ряда (0.3).

Лекция 11

Компактность

Вернёмся к обсуждению компактов. Напомним, что множество называется компактным, если из любого его открытого покрытия можно извлечь конечное подпокрытие. Сейчас мы дадим ещё одну характеристику компактов в $\mathbb{R}$. Для этого сформулируем следующее определение.

Определение. *Множество E называется секвенциально компактным, если из любой последовательности $(x_n)_{n\in\mathbb{N}} \subset E$ можно извлечь подпоследовательность, сходящуюся к пределу, лежащему в E.*

Оказывается, что компактные подмножества в $\mathbb{R}$ являются также и секвеницально компактными и наоборот. Более точно, верна

Теорема. *(i) Множество $E \subset \mathbb{R}$ компактно, если и только если оно секвенциально компактно.*

(ii) Множество $E \subset \mathbb{R}$ компактно, если и только если оно замкнуто и ограничено.

Проведём доказательство теоремы по следующей схеме.
1) Если множество ограничено и замкнуто, то оно секвенциально компактно.
2) Если множество ограничено и замкнуто, то оно компактно.
3) Если множество не ограничено, то оно не компактно и не секвенциально компактно.
4) Если множество не замкнуто, то оно не компактно и не секвенциально компактно.

Доказательство. 1) Из ограниченной последовательности можно извлечь сходящуюся подпоследовательность, она также принадлежит множеству. Так как множество замкнуто, то и предел принадлежит множеству.

2) Ранее мы доказали, что отрезок компактен (лемма Гейне–Бореля). Доказательство в основном опиралось на вложенные отрезки. Для любого замкнутого ограниченного множества проходит похожая конструкция, но мы пойдем другим путём. Пусть есть ограниченное и замкнутое множество E и его открытое покрытие $\cup_\alpha U_\alpha$. Множество E принадлежит некоторому отрезку $[a, b]$. Множество $(a-1, b+1) \setminus E$ открыто. Рассмотрим покрытие отрезка $[a, b]$, состоящее из всех множеств U_α и множества $(a-1, b+1) \setminus E$. Каждая точка отрезка покрывается, точки из E покрываются множествами из U_α, точки не из E — точками из $(a-1, b+1) \setminus E$. Теперь выберем по лемме Гейне—Бореля из покрытия $\cup_\alpha U_\alpha \cup (a-1, b+1) \setminus E$ отрезка $[a, b]$ конечное подпокрытие. Оно состоит из конечного числа множеств U_α: $\{U_{\alpha_1}, \ldots, U_{\alpha_N}\}$ и множества $(a-1, b+1) \setminus E$, которое, заметим, с E не пересекается. Поэтому $\{U_{\alpha_1}, \ldots, U_{\alpha_N}\}$ — конечное покрытие E.

3) Пусть множество E неограничено например сверху. Тогда существует последовательность $(x_n)_{n\in\mathbb{N}} \subset E$, $x_n \to +\infty$. Без потери общности можно считать последовательность $(x_n)_{n\in\mathbb{N}}$ монотонной (почему?). Из этой последовательности нельзя извлечь сходящуюся. Поэтому, E не секвенциально компактно.

Рассмотрим счётную систему интервалов $U_n = (-n, n)$. По аксиоме Архимеда это не что иное как открытое покрытие $\mathbb{R}$, следовательно, любого подмножества $E \subset \mathbb{R}$. Если E не ограничено, то из этого покрытия нельзя извлечь конечное подпокрытие: объединение конечной системы ограниченных множеств —- ограниченное множество. Значит, E не секвенциально компактно.

4) Пусть множество E не замкнуто. Это значит, что у него есть предельная точка x^*, которая не принадлежит множеству E. По определению предельной точки выберем последовательность $(x_n)_{n\in\mathbb{N}} \subset E$, сходящуюся к x^*. Любая её подпоследовательность сходится к той же точке $x^* \notin E$. Поэтому множество E не секвенциально компактно.

Теперь построим открытое покрытие E, из которого нельзя выбрать конечное подпокрытие. Его легко выписать в явном виде, например, в него можно включить все лучи вида $(-\infty, x^* - 1/n)$ и все лучи $(x^* + 1/n, +\infty)$. Это покрытие множества $\mathbb{R} \setminus \{x^*\}$, следовательно, покрытие E. Любая конечная система множеств из такого покрытия не содержит некоторую окрестность точки x^*, но по определению предельной точки в ней обязательно есть точки множества E, то есть эта конечная система не является покрытием. Значит, E не компактно. $\square$

Предел функции

Определение. *Проколотая окрестность точки a — это окрестность точки a без самой точки a.*

Обозначение. $\overset{\circ}{U}(a)$

Пример. Множество $\{x : |x| \in (0, \delta)\}$ — проколотая окрестность точки 0 при любом $\delta > 0$, т.е. проколотая δ-окрестность $\overset{\circ}{U}_\delta(0)$ нуля.

Определение 1 (по Коши). *Пусть $E \subset \mathbb{R}$ и $a \in E'$ — предельная точка множества E. Пусть функция f определена на E. Говорят, что функция f стремится к $A \in \mathbb{R}$ при x, стремящемся к a, если и только если*

$$\forall \varepsilon > 0\ \exists \delta > 0 :\ \forall x \in \overset{\circ}{U}_\delta(a) \cap E :\ |f(x) - A| < \varepsilon.$$

Часто последняя запись записывается в более явном виде (расписывая определение $\overset{\circ}{U}_\delta(a)$)

$$\forall \varepsilon > 0\ \exists \delta > 0 :\ \forall x \in E, 0 < |x - a| < \delta :\ |f(x) - A| < \varepsilon.$$

Обозначение. $\lim\limits_{x \to a} f(x) = A, f(x) \to A$ при $x \to a$.

Определение 2 (переформулировка). $A = \lim_{x \to a} f(x)$, *если для любой окрестности $U(A)$ точки A найдётся проколотая окрестность $\overset{\circ}{U}(a)$ точки a такая, что $f(\overset{\circ}{U}(a) \cap E) \subset U(A)$.*

Упражнение. Объясните, почему определения 1 и 2 эквивалентны.

Сходимость к $\pm\infty$. Определения вполне естественные:

$$\lim_{x \to a} f(x) = +\infty \Leftrightarrow \forall C > 0\ \exists \delta > 0 :\ \forall x \in \overset{\circ}{U}_\delta(a) \cap E :\ f(x) > C.$$

$$\lim_{x \to a} f(x) = -\infty \Leftrightarrow \forall C > 0\ \exists \delta > 0 :\ \forall x \in \overset{\circ}{U}_\delta(a) \cap E :\ f(x) < C.$$

Определение 3 (по Гейне). *Число A является пределом функции f при $x \to a$, если для любой последовательности $(x_n)_n \subset E$ такой, что $x_n \to a, x_n \neq a$ справедливо соотношение $\lim\limits_{n \to \infty} f(x_n) = A$.*

Следующие теоремы показывают, что определение предела функции по Коши и по Гейне совпадают.

Теорема 1 (необходимое условие). *Если $A = \lim\limits_{x \to a} f(x)$, то для любой последовательности $(x_n)_n \subset E$ такой, что $x_n \to a, x_n \neq a$ справедливо соотношение $\lim\limits_{n \to \infty} f(x_n) = A$.*

Теорема 2 (достаточное условие). *Пусть для любой (строго монотонной последовательности) $(x_n)_n \subset E$ такой, что $x_n \to a$ справедливо соотношение $\lim\limits_{n \to \infty} f(x_n) = A$. Тогда $A = \lim\limits_{x \to a} f(x)$.*

Доказательство. 1) Пусть $A = \lim\limits_{x\to a} f(x)$ и $x_n \to a, x_n \neq a$. Покажем, что $\lim\limits_{n\to\infty} f(x_n) = A$. Выберем $\varepsilon > 0$ и построим по нему такое $\delta > 0$, что из $0 < |x - a| < \delta$ следует $|f(x) - A| < \varepsilon$. Теперь, в силу $x_n \to a$ по числу δ выберем такое N , что при всех $n > N$ выполняется неравенство $|x_n - a| < \delta$. Отсюда и по построению числа δ следует, что при всех $n > N$ выполняется неравенство $|f(x_n) - A| < \varepsilon$. Таким образом, $\lim\limits_{n\to\infty} f(x_n) = A$.

2) Пусть для любой строго монотонной последовательности $(x_n)_n \subset E$ такой, что $x_n \to a$ справедливо соотношение $\lim\limits_{n\to\infty} f(x_n) = A$. Покажем, что тогда $\lim\limits_{x\to a} f(x) = A$. Рассуждать будем от противного. Построим отрицание к

$$\forall\varepsilon > 0\ \exists\delta > 0 :\ \forall x \in E, 0 < |x - a| < \delta\ |f(x) - A| < \varepsilon.$$

Для этого поменяем местами кванторы и переверніём заключительное неравенство:

$$\exists\varepsilon > 0\ \forall\delta > 0 :\ \exists x \in E, 0 < |x - a| < \delta\ |f(x) - A| \geqslant \varepsilon.$$

Итак, пусть такое ε существует. В качестве произвольного δ будем выбирать числа $\delta_n = n^{-1}$, каждому такому δ_n соответствует $x_n \neq a$ такой, что $|x_n - a| < \delta_n$, но $|f(x_n) - A| \geqslant \varepsilon$. Очевидно, что $x_n \to a$ (почему?). Осталось воспользоваться леммой о том, что из любой последовательности $(x_n)_n$ такой, что $x_n \to a, x_n \neq a$ можно выбрать строго монотонную подпоследовательность $(x_{n_k})_k$. Эта строго монотонная подпоследовательность также стремится к a, однако $f(x_{n_k})$ не стремится к A по построению: $|f(x) - A| \geqslant \varepsilon$. $\square$

Свойства переделов функции

Из эквивалентности определений по Коши и по Гейне моментально следуют такие свойства:

Утверждение 1. *Пусть существуют пределы* $A = \lim\limits_{x\to a} f(x)$ *и* $B = \lim_{x\to a} g(x)$. *Тогда существует предел* $\lim\limits_{x\to a}(\alpha f(x) + \beta g(x))$ *и он равен* $\alpha A + \beta B$.

Утверждение 2. *Пусть существуют пределы* $A = \lim\limits_{x\to a} f(x)$ *и* $B = \lim\limits_{x\to a} g(x)$ *Тогда существует предел* $\lim\limits_{n\to a} f(x)g(x)$ *и он равен* AB. *Если* $B \neq 0$ *и* $g \neq 0$ *в некоторой окрестности точки* a, *то существует предел* $\lim\limits_{x\to a} f(x)/g(x)$ *и он равен* A/B.

Утверждение 3 (предельный переход в неравенствах). *Пусть существуют конечные пределы* $A = \lim\limits_{x\to a} f(x)$ *и* $B = \lim\limits_{x\to a} g(x)$. *Если* $f(x) \geqslant g(x)$ *для всех* x *из некоторой проколотой окрестности точки* a, *то* $A \geqslant B$. *Если* $f(x) \geqslant h(x) \geqslant g(x)$ *для всех* x *из некоторой проколотой окрестности точки* a *и* $A = B$, *то существует предел* $A = \lim\limits_{x\to a} h(x)$.

Теорема о двух милиционерах. *Пусть* $\lim\limits_{x\to a} f(x) = A = \lim\limits_{x\to a} g(x)$. *Если* $f(x) \geqslant h(x) \geqslant g(x)$ *для всех* x *из некоторой проколотой окрестности точки* a, *то существует предел* $\lim\limits_{x\to a} h(x)$ *и он равен* A.

Утверждение 3. *Если у функции* f *существует конечный предел* A *при* $x \to a$, *то она ограничена в некоторой проколотой окрестности точки* a. *Если* $A \neq 0$, *то в некоторой проколотой окрестности точки* a *выполняется оценка* $|f(x)| > |A|/2$.

Частичные пределы

Пусть последовательность $(x_n)_n, x_n \to a$ такова, что последовательность $(f(x_n))_n$ также сходится. Тогда $\lim\limits_{n\to\infty} f(x_n)$ называется *частичным пределом функции* f *при* $x \to a$.

Если есть две различные последовательности $(x_n)_n, (y_n)_n$ такие, что $x_n, y_n \to a$, то существуют и различные пределы $\lim\limits_{n\to\infty} x_n$, $\lim\limits_{n\to\infty} y_n$, и, в силу эквивалентности пределов по Коши и по Гейне, предел $\lim_{x\to a} f(x)$ не может существовать.

Теорема (критерий Коши). *Предел* $\lim\limits_{x\to a} f(x)$ *существует, если и только если*

$$\forall\varepsilon>0\ \exists\delta>0: \forall x, y \in \overset{\circ}{U}_\delta (a)\cap E\ |f(x)-f(y)|<\varepsilon.$$

Доказательство. Необходимость. Пусть существует $\lim\limits_{x\to a} f(x) = A$. Тогда, по определению Коши,

$$\forall\varepsilon>0\ \exists\delta>0:\ \forall x\in E, 0<|x-a|<\delta\ |f(x)-A|<\varepsilon/2.$$

Зафиксируем $\varepsilon>0$, построим по нему $\delta>0$. Теперь

$$\forall x, y \in \overset{\circ}{U}_\delta (a)\cap E\ |f(x)-A|<\varepsilon/2, |f(y)-A|<\varepsilon/2,$$

из чего следует, что $|f(x)-f(y)|<\varepsilon$. Необходимость доказана.

Достаточность. Пусть

$$\forall\varepsilon>0\ \exists\delta>0: \forall x, y \in \overset{\circ}{U}_\delta (a)\cap E\ |f(x)-f(y)|<\varepsilon.$$

Отсюда вытекает, что для каждой последовательности $(x_n)_n\subset E$ такой, что $x_n\to a$ последовательность образов $(f(x_n))_n$ фундаментальна (почему?). Поэтому для каждой $(x_n)_n\subset E$ такой, что $x_n\to a$ последовательность $f(x_n)$ сходится (почему?). Осталось увидеть, что для двух последовательностей $(x_n)_n, (y_n)_n\subset E$ таких, что $x_n, y_n\to a$ пределы последовательностей $(f(x_n))_n$ и $(f(y_n))_n$ совпадают. Для этого рассмотрим последовательность

$$x_1, y_1, x_2, y_2, \ldots, x_n, y_n, \ldots \subset E,$$

эта последовательность также сходится к a, поэтому последовательность образов

$$f(x_1), f(y_1), f(x_2), f(y_2), \ldots$$

также фундаментальна, следовательно, сходится. Значит, пределы $(f(x_n))_n$ и $(f(y_n))_n$ совпадают как частичные пределы сходящейся последовательности $\square$

Предел функции при $x\to\pm\infty$

Как и следует ожидать, определения предел функции при $x\to\infty$ таковы:

$$\lim_{x\to+\infty} f(x) = A \Leftrightarrow \forall\varepsilon>0\ \exists N>0: \forall x>N\ |f(x)-A|<\varepsilon,$$

$$\lim_{x\to+\infty} f(x) = +\infty \Leftrightarrow \forall C>0\ \exists N>0: \forall x>N\ f(x)>C.$$

$$\lim_{x\to+\infty} f(x) = -\infty \Leftrightarrow \forall C>0\ \exists N>0: \forall x>N\ f(x)<C.$$

Аналогично определения предела функции при $x\to-\infty$ таковы

$$\lim_{x\to-\infty} f(x) = A \Leftrightarrow \forall\varepsilon>0\ \exists N<0: \forall x<N\ |f(x)-A|<\varepsilon,$$

$$\lim_{x\to-\infty} f(x) = \infty \Leftrightarrow \forall C>0\ \exists N<0: \forall x<N\ f(x)>C.$$

$$\lim_{x\to-\infty} f(x) = -\infty \Leftrightarrow \forall C<0\ \exists N<0: \forall x>N\ f(x)<C.$$

Из определений видно, что из существования предела $\lim\limits_{x\to\infty} f(x)$ следует существование предела $\lim\limits_{n\to\infty} f(n)$. Обратное не верно. Пример: $f(x)=\sin(\pi x)$. В этом случае $f(n)=0$ при всех $n\in\mathbb{N}$, поэтому предел $\lim\limits_{n\to\infty} f(n)$ существует и равен нулю. В то же время, $\lim\limits_{x\to\infty}\sin(\pi x)$ не существует.

Упражнение. Докажите, что $\lim_{x\to\infty}\sin(\pi x)$ не существует.

Предел сложной функции. Замена переменных в пределах

Пусть даны две функции, f и g, причём их области определения и множества значений так согласованы, что определена функция $g(f(x))$. Пусть функция f определена в проколотой окрестности точки a и пусть существует предел $\lim\limits_{x\to a} f(x) = A$, причём $f(x) \neq A$ в этой проколотой окрестности. Пусть функция g определена в проколотой окрестности $\overset{\circ}{U}(A)$ точки A и существует предел $\lim\limits_{y\to A} g(y) = B$. Тогда в некоторой проколотой окрестности точки a определена функция $g(f(x))$ и существует предел $\lim\limits_{x\to a} g(f(x)) = B$, иными словами, если в формуле

$$\lim_{x\to a} g(f(x)) = \lim_{y\to A} g(y) = B$$

второй предел существует, то существует и первый, и он равен второму. **Докажем это:** дано:

$$\forall\varepsilon_0 > 0\ \exists\delta_0 > 0 : \forall x, 0 < |x-a| < \delta_0 \Rightarrow |f(x) - A| < \varepsilon_0,$$
$$\forall\varepsilon_1 > 0\ \exists\delta_1 > 0 : \forall x, 0 < |y-A| < \delta_1 \Rightarrow |g(y) - B| < \varepsilon_1.$$

Хотим доказать:

$$\forall\varepsilon > 0\ \exists\delta > 0 : \forall x, 0 < |x-a| < \delta_0 \Rightarrow |g(f(x)) - B| < \varepsilon.$$

Положим $\varepsilon = \varepsilon_1$ и построим по ε_1 величину δ_1 так, чтобы $\forall y : 0 < |y-A| < \delta_1 \Rightarrow |g(y)-B| < \varepsilon_1$. Теперь положим $\varepsilon_0 = \delta_1$ и построим по этому ε_0 величину δ_0. Покажем, что требуемая формула выполняется при $\delta = \delta_0$. По построению при $0 < |x-a| < \delta$ выполнено $|f(x) - A| < \varepsilon_0 = \delta_1$, при $|f(x) - A| < \delta_1$ выполнено $|g(f(x)) - B| < \varepsilon$. $\square$

Эту формулу удобно применять к вычислению пределов:

$$\lim_{x\to 0}\frac{\sin 4x}{x} = [y = 4x] = \lim_{y\to 0}\frac{\sin y}{y/4} = 4,$$
$$\lim_{x\to 0}\frac{\sin x^2}{x^2} = [y = x^2] = \lim_{y\to 0}\frac{\sin y}{y} = 1.$$

Лекция 12

Односторонние пределы

Важную роль играют так называемые односторонние пределы функций. Неформально говоря, это пределы, в которых $x \to a$ "с одной стороны", либо справа, либо слева от точки a. Соответственно, бывают два односторонних предела: левый (левосторонний, предел слева) и правый (правосторонний, предел справа).

Левый предел обозначается $\lim\limits_{x \to a-0} f(x)$, правый – $\lim\limits_{x \to a+0} f(x)$.

Определение. *Число A называется левым пределом пределом f при $x \to a$, $a \in E'$ если*

$$\forall \varepsilon > 0 \ \exists \delta > 0 \colon \forall x \in (a - \delta, a) \cap E \ |f(x) - A| < \varepsilon.$$

Число A называется правым пределом пределом f при $x \to a$, если

$$\forall \varepsilon > 0 \ \exists \delta > 0 \colon \forall x \in (a, a + \delta) \cap E \ |f(x) - A| < \varepsilon.$$

Замечание. Если существуют оба односторонних предела, и они равны $A \in \mathbb{R}$, то существует и обычный предел $\lim\limits_{x \to a} f(x)$ и он также равен A. Обратное так же верно. Это замечание часто оформляют в виду теоремы: *конечный предел A функции f в точке a существует тогда и только тогда, когда существуют конечные левые и правые пределы в a, равные A.*

Определения односторонних пределов могут также быть переформулированы в терминах Гейне. Ограничимся формулировками для левого предела.

Пусть функция f определена на интервале $(a - h, a)$.

Теорема о левом пределе по Гейне. *Если $A = \lim\limits_{x \to a-0} f(x)$, то для любой последовательности $(x_n)_n \subset (a - h, a)$, $x_n \to a$, справедливо соотношение $\lim\limits_{n \to \infty} f(x_n) = A$. Пусть для любой строго возрастающей последовательности $(x_n)_n \subset (a - h, a)$, $x_n \to a$ последовательность $(f(x_n))_n$ сходится. Тогда существует предел $\lim\limits_{x \to a-0} f(x)$.*

Замечание. Чтобы получить формулировку этой теоремы для правого предела, нужно заменить слова "строго возрастающей" на "строго убывающей", а интервал $(a - h, a)$ заменить на $(a, a + h)$.

Доказательство. Доказательство первой части теоремы опустим: оно ничем не отличается от доказательства утверждения про то, что определение по Коши влечёт определение по Гейне.

Во второй части сначала покажем, что в условиях теоремы все последовательности $(f(x_n))_n$ сходятся к одному и тому же пределу A. Пусть есть две строго возрастающих последовательности $(x_n)_n, (y_n)_n$, $x_n, y_n \to a$. Построим по ним объединяющую строго возрастающую последовательность $(z_n)_n, z_n \to a$, последовательности $(x_n)_n, (y_n)_n$ — подпоследовательности $(z_n)_n$. Теперь по предположению $(f(z_n))_n$ сходится к некоторому числу A, следовательно, обе подпоследовательности $(f(x_n))_n, (f(y_n))_n$ также сходятся к A. Дальнейшее доказательство можно провести от противного, аналогично доказательству утверждению, что определение по Гейне влечёт определение по Коши. $\square$

Из этой теоремы следует, например следующее важное утверждение. Мы скажем, что функция f является монотонно возрастающей, если $f(x) > f(y)$ при $x > y$, где x, y принадлежат области определения функции f. Аналогично формулируются понятия монотонно неубывающей, убывающей и невозрастающей функций. Во всех этих случая мы говорим, что функция f монотонная (в своей области определения).

Теорема об односторонних пределах монотонной функции. *Пусть на отрезке* $[c,d]$ *задана монотонно возрастающая функция* f. *Пусть* $a \in (c,d)$. *Тогда существуют пределы* $A = \lim\limits_{x\to a-0} f(x)$ *и* $B = \lim\limits_{x\to a+0} f(x)$.

Доказательство. Для доказательства достаточно отметить, что для любой строго возрастающей последовательности $(x_n)_n$, $x_n \to a$ последовательность $(f(x_n))_n$ не убывает и ограничена сверху числом $f(a)$ и что для любой строго убывающей последовательности $(x_n)_n$, $x_n \to a$ последовательность $(f(x_n))_n$ не возрастает и ограничена числом $f(a)$ снизу. Далее пользуемся теоремой Вейерштрасса. $\Box$

Замечательные пределы

Их два. **Первый замечательный предел**: $\lim\limits_{x\to 0}\dfrac{\sin x}{x} = 1$. **Второй замечательный предел**: $\lim\limits_{x\to+\infty}\left(1+\dfrac{1}{x}\right)^x = e$.

Доказательство первого замечательного предела: Докажем сначала, что

$$\cos^2 x < \frac{\sin x}{x} < 1,$$

при $0 < |x| < \pi/2$. Рассмотрим случай $x \in (0,\pi/2)$. Случай $x \in (-\pi/2, 0)$ рассматривается аналогично. Рассмотрим рис.1. Даны: тригонометрический круг, с центром в начале координат, луч OL, проведённый под углом $x \in (0,\pi/2)$ к положительному лучу оси абсцисс, тока K — точка пересечения окружности и луча OL, перпендикуляры к оси абсцисс проходят через K и точку с координатами $(1,0)$.

Теперь $S_{\Delta OAK} < S < S_{\Delta OAL}$, где S — площадь сектора OAK. Выразим эти площади через x: $|OA| = |OK| = 1, KH = \sin x, AL = \operatorname{tg} x$, тогда $S_{\Delta OAK} = \frac{1}{2}\sin x, S = \frac{1}{2}x, S_{\Delta OAL} = \frac{1}{2}\operatorname{tg} x$. Итак, $\sin x < x < \operatorname{tg} x$, откуда

$$\cos^2 x < \cos x < \frac{\sin x}{x} < 1.$$

Теперь из $\frac{\sin x}{x} < 1$ следует, что $0 < \sin x < x$. Тогда по теореме о двух милиционерах получаем, что $\lim\limits_{x\to 0}\sin x = 0$. Покажем, что $\lim\limits_{x\to 0}\cos^2 x = 1$. Действительно, $\lim\limits_{x\to 0}\cos^2 x = \lim\limits_{x\to 0}(1-\sin^2 x) = 1 - \lim\limits_{x\to 0}\sin^2 x = 1$. Тогда по теореме о двух милиционерах неравенство $\cos^2 x < \frac{\sin x}{x} < 1$ влечёт, что $\lim\limits_{x\to 0}\dfrac{\sin x}{x} = 1$. $\Box$

Следствие 1. $\lim\limits_{x\to 0}\dfrac{\operatorname{tg} x}{x} = 1$.

Следствие 2 *(из доказательства)*. $\lim\limits_{x\to 0}\sin x = 0$, $\lim\limits_{x\to 0}\cos x = 1$.

В первом следствии мы просто замечаем, что $\operatorname{tg} x = \frac{\sin x}{\cos x}$ и пользуемся арифметикой пределов. Во втором следствии из доказательства мы получаем, что $\lim\limits_{x\to 0+0}\sin x = 0$ и $\lim\limits_{x\to 0+0}\cos x = 1$ (т.к. в доказательстве мы рассматривали $x \in (0,\pi/2)$). Далее используя замечание в начале доказательства о том, что рассмотрение $x \in (-\pi/2, 0)$ происходит аналогично, мы получаем, что $\lim\limits_{x\to 0-0}\sin x = 0$ и $\lim\limits_{x\to 0-0}\cos x = 1$. Далее мы видим что как для косинуса так и для синуса левые и правые пределы в 0 равны между собой, а значит, существуют пределы этих функций в 0.

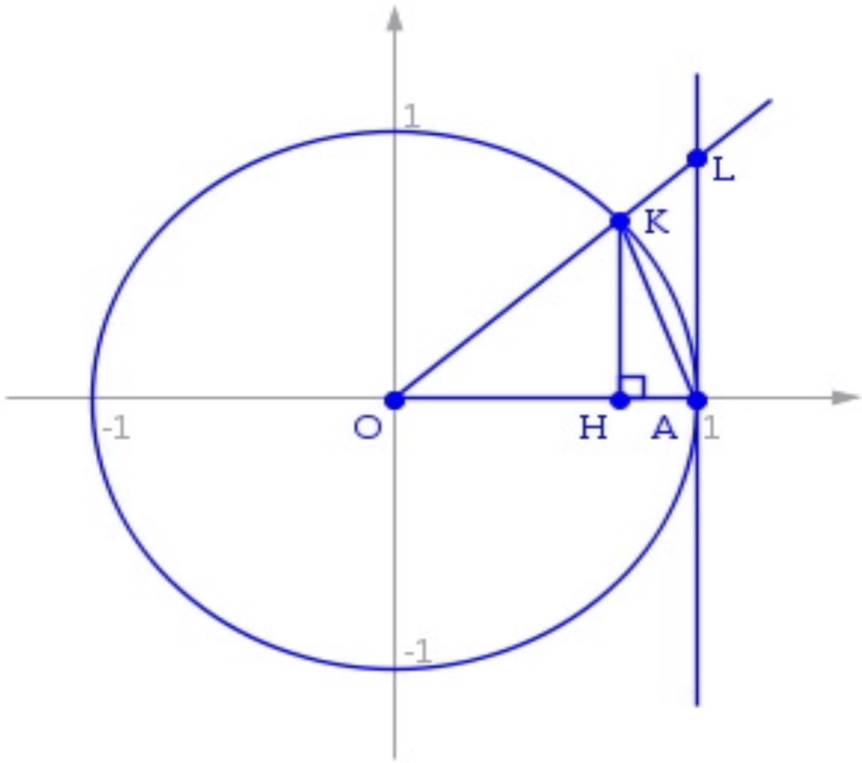


Рис. 1. К первому замечательному пределу

Доказательство второго замечательного предела: На семинарах мы доказали, что $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$. Пусть теперь x — вещественное число, а не натуральное, обозначим через $n=[x]$ его целую часть. Тогда

$$\left(1+\frac{1}{[x]+1}\right)^{[x]} \leqslant \left(1+\frac{1}{x}\right)^{x} \leqslant \left(1+\frac{1}{[x]}\right)^{[x]+1},$$

поэтому

$$\left(1+\frac{1}{n+1}\right)^{n+1}\frac{n+1}{n+2} \leqslant \left(1+\frac{1}{x}\right)^{x} \leqslant \left(1+\frac{1}{n}\right)^{n}\frac{n+1}{n}.$$

Легко видеть, что

$$\left(1+\frac{1}{n+1}\right)^{n+1}\frac{n+1}{n+2} \to e, \quad \left(1+\frac{1}{n}\right)^{n}\frac{n+1}{n} \to e, \text{ при } n\to\infty.$$

Значит, для любого $\varepsilon>0$ найдётся $N\in\mathbb{N}$, такое что при $n>N$

$$\left|\left(1+\frac{1}{n+1}\right)^{n+1}\frac{n+1}{n+2}-e\right|<\varepsilon, \quad \left|\left(1+\frac{1}{n}\right)^{n}\frac{n+1}{n}-e\right|<\varepsilon,$$

откуда получаем

$$\left|\left(1+\frac{1}{x}\right)^{x}-e\right|<\varepsilon$$

при $x>N$. $\square$.

Упражнение. Почему в доказательстве второго замечательного предела мы не воспользовались теоремой о двух милиционерах?

Непрерывные функции

Определение. *Функция $f\colon E \to \mathbb{R}$ называется непрерывной в точке $x_0 \in E$, если*

$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall x \in E \cap U_\delta(x_0) \Rightarrow f(x) \in U_\varepsilon(f(x_0)).$$

Иначе можно сказать, что $f\colon E \to \mathbb{R}$ непрерывна в точке $x_0 \in E$, если для любой окрестности $U(f(x_0))$ найдётся окрестность $U(x_0)$ такая, что $f(E \cap U(x_0)) \subset U(f(x_0))$.

Упражнение. Почему в предыдущем определении действительно можно так сказать?

Заметим, что всякая функция f непрерывна в изолированной точке области определения: существует окрестность $U(x_0)$ такая, что $f(E \cap U(x_0)) = f(x_0) \subset U(f(x_0))$. Сейчас мы покажем, что непрерывность функции f в точке $x_0 \in E$, где x_0 — предельная точка E, эквивалента $\lim\limits_{x\to x_0} f(x) = f(x_0)$ или более эффектно $\lim\limits_{x\to x_0} f(x) = f(\lim\limits_{x\to x_0} x)$. Используя затем эквивалентность определений по Гейне и по Коши, можно также сказать, что для любой последовательности $(x_n)_n$ сходящейся к x_0 справедливо, что $\lim\limits_{n\to\infty} f(x_n) = f(x_0)$.

Теорема. *Функция $f\colon E \to \mathbb{R}$ непрерывна в точке $x_0 \in E$ тогда и только тогда, когда $\lim\limits_{x\to x_0} f(x) = f(x_0)$.*

Доказательство. В одну сторону утверждение очевидно: если функция непрерывна в точке $x_0 \in E$, тогда $\lim\limits_{x\to x_0} f(x) = f(x_0)$. Для это нужно лишь заметить, что в определении непрерывной в точке $x_0 \in E$ функции речь идёт о $E \cap U_\delta(x_0)$ тогда как в определении предела мы имеем $E \cap \overset{\circ}{U}_\delta(x_0)$.

Обратно, пусть $\lim\limits_{x\to x_0} f(x) = f(x_0)$. Значит, по окрестности $U(f(x_0))$ найдём проколотую окрестность $\overset{\circ}{U}(x_0)$ такую, что $f(E \cap \overset{\circ}{U}(x_0)) \subset U(f(x_0))$. Т.к. $f(x_0) \in U(f(x_0))$, то $f(E \cap U(x_0)) \subset U(f(x_0))$. $\square$

Из теорем о пределах функции следуют “обычные” свойства непрерывных в некоторой точке x_0 функций: сумма и произведение непрерывных в x_0 функций — непрерывная в x_0 функция, частное непрерывных в x_0 функций f/g тоже непрерывна (если $g(x_0) \neq 0$).

Определение. *Функция называется непрерывной на множестве E, если она непрерывна в каждой точке этого множества.*

Соответственно, сумма и произведение на E непрерывных функций — непрерывная на E функция, частное f/g непрерывных функций тоже (если $g(x_0) \neq 0$ при $x \in E$). Более точно

Теорема. *Пусть функция $f\colon (a,b) \to \mathbb{R}$ непрерывна в точке $x_0 \in (a,b)$. Тогда функция f ограничена в некоторой окрестности точки x_0, а если $f(x_0) \neq 0$, то в некоторой окрестности точки x_0 функция f ненулевая и более того $|f| > |f(x_0)|/2 > 0$, и её знак совпадает со знаком $f(x_0)$.*

Теорема. *Пусть функции $f, g\colon (a,b) \to \mathbb{R}$ непрерывны в точке $x_0 \in (a,b)$. Тогда функции $\alpha f + \beta g, fg$ и (если, дополнительно, $g(x_0) \neq 0$) f/g непрерывны в x_0.*

Теорема о сложной функции. *Пусть функция $f\colon (a,b) \to \mathbb{R}$ непрерывна в точке $x_0 \in (a,b)$, а функция g определена в окрестности точки $f(x_0)$ и непрерывна в этой точке. Тогда функция $g(f(x))$ непрерывна в точке x_0.*

Следующая теорема содержит важное свойство непрерывных функций, которое часто принимают за определение непрерывной функции.

Теорема. *Функция f непрерывна на интервале, если прообраз любого открытого множества открыт.*

Односторонняя непрерывность

Пусть функция f определена на множестве $[x_0, x_0 + \delta)$, $\delta > 0$. Будем говорить, что f непрерывна справа в точке x_0, если существует односторонний предел f при $x \to x_0 + 0$ и $\lim\limits_{x \to x_0+0} f(x) = f(x_0)$. Аналогично, если функция f определена на множестве $(x_0 - \delta, x_0]$, $\delta > 0$, то f непрерывна слева в точке x_0, если существует односторонний предел f при $x \to x_0 - 0$ и $\lim\limits_{x \to x_0-0} f(x) = f(x_0)$.

O-большое и o-малое

Часто на практике и для наших дальнейших целей будет удобны понятия o-малого и O-большого.

Определение: *Функция α называется бесконечно малой по сравнению с функцией f при $x \to a$, если существует такая проколотая окрестность $\overset{\circ}{U}(a)$ точки a, на которой $\alpha(x) = \varepsilon(x) f(x)$ при $x \in \overset{\circ}{U}(a)$, причём $\lim\limits_{x \to a} \varepsilon(x) = 0$. При этом пишут, что $\alpha(x) = o(f(x))$ при $x \to a$.*

Определение: *Функция f называется ограниченной относительно функции g при $x \to a$, если функции f, g определены в проколотой окрестности $\overset{\circ}{U}(a)$ точки a и существует такое число $C > 0$, что $|f(x)| \leqslant C|g(x)|$ при $x \in \overset{\circ}{U}(a)$. При этом пишут, что $f(x) = O(g(x))$ при $x \to a$.*

Относительно последнего определения: эквивалентно можно сказать, что $f(x) = O(g(x))$ при $x \to a$, если существует такая ограниченная в $\overset{\circ}{U}(a)$ функция h, что $f(x) = h(x)g(x)$, $\forall x \in \overset{\circ}{U}(a)$.

Примеры: 1) $x^2 = o(x)$ при $x \to 0$;
2) $x = o(x^2)$ при $x \to +\infty$;
3) $(x+1)^2 = O(x^2)$ при $x \to +\infty$;
4) $o(x) = O(x)$ при $x \to 0$ однако $O(x) \neq o(x)$ при $x \to 0$.

Также полезным оказывается понятие *асимптотически равных или эквивалентных* функций.

Определение. *Функции f и g называются эквивалентными или асимптотически равными при $x \to a$, если существует такая проколотая окрестность $\overset{\circ}{U}(a)$ точки a, на которой $f(x) = \lambda(x) g(x)$ при $x \in \overset{\circ}{U}(a)$, причём $\lim\limits_{x \to a} \lambda(x) = 1$. При этом пишут, что $f(x) \sim g(x)$ при $x \to a$.*

Примеры: $\sin x \sim x$ при $x \to 0$, $\operatorname{tg} x \sim x$ при $x \to 0$ как следствие из первого замечательного предела

Несложно показать, что асимптотическое равенство или эквивалентность является отношением эквивалентности на функциях (при условии что $x \to a$ для всех функций).

Упражнение: проверьте это.

Лекция 13

Непрерывные функции на отрезке

Определение. *Будем говорить, что функция $f\colon [a,b] \to \mathbb{R}$ непрерывна на отрезке $[a,b]$, если f непрерывна на интервале (a,b), непрерывна слева в точке b и непрерывна справа в точке a.*

Всюду через $C([a,b])$ или $C^0([a,b])$ обозначается множество непрерывных функций на отрезке $[a,b]$, это множество является векторным пространством.

Теорема о промежуточном значении (Больцано–Коши). *Пусть $f \in C([a,b])$, причём $f(a)f(b) < 0$. Тогда существует такое $\theta \in (a,b)$, что $f(\theta) = 0$.*

Доказательство. Рассмотрим случай $f(a) < 0, f(b) > 0$. Определим множество $E = \{y \mid \forall x \in (y,b) : f(y) > 0\}$. Это множество не пусто: из $f(b) > 0$ следует, что $f(x) > 0$ в окрестности точки b. Это множество ограничено снизу: a — нижняя граница этого множества. Значит, у этого множества есть точная нижняя грань $\theta = \inf E$. Покажем, что $f(\theta) = 0$. Если это не так, то либо $f(\theta) > 0$, либо $f(\theta) < 0$.

Если $f(\theta) > 0$, то $f > 0$ в окрестности $U_\delta(\theta)$, таким образом $\theta - \delta/2 \in E$, это противоречит тому, что θ — нижняя грань E.

Если $f(\theta) < 0$, то $f < 0$ в окрестности $U_\delta(\theta)$, таким образом $\theta + \delta/2 \notin E$, это противоречит тому, что θ — точная нижняя грань для E. $\square$

Следствие. *Пусть $f \in C([a,b])$, причём $f(a) < f(b)$. Тогда для любого $A \in (f(a), f(b))$, найдётся такое $\theta \in (a,b)$, что $f(\theta) = A$.*

Доказательство. Для доказательства следствия достаточно рассмотреть функцию $f(x) - A$, для неё выполнены все условия теоремы Больцано–Коши. $\square$

Теорема Вейерштрасса о максимальном значении. *Пусть $f \in C([a,b])$. Тогда f ограничена сверху и снизу, причём найдутся такие точки x_1 и x_2 на $[a,b]$, что*

$$f(x_1) = \min_{[a,b]} f(x),\ f(x_2) = \max_{[a,b]} f(x).$$

Доказательство. Будем рассуждать от противного. Пусть функция f не ограничена, это значит, что найдётся последовательность $(x_n)_n \subset [a,b]$, для которой $\lim\limits_{n\to\infty} |f(x_n)| = +\infty$. Выберем из последовательности $(x_n)_n$ (она ограничена!) сходящуюся подпоследовательность $(x_{n_k})_k$, $\lim\limits_{k\to\infty} x_{n_k} = x_0$. Теперь $\lim\limits_{k\to\infty} x_{n_k} = x_0$, а $\lim\limits_{k\to\infty} f(x_{n_k}) = \pm\infty \neq f(x_0)$, что противоречит непрерывности f на $[a,b]$. Итак, функция f ограничена и сверху, и снизу. Пусть $M = \sup f(x)$, но f не принимает значение M ни в какой точке. Тогда функция $\varphi(x) = 1/(M - f(x))$ определена при всех $x \in [a,b]$, по доказанным ранее теоремам, она непрерывна. Значит, эта функция ограничена сверху: $\varphi(x) < K$, отсюда $M - f(x) > 1/K$, или $f(x) \leqslant M - 1/K$, что противоречит $M = \sup f(x)$. Аналогичная конструкция показывает, что достигается точка минимума. $\square$

Теорема. *Образ отрезка при непрерывном функции — отрезок или точка.*

Доказательство. Если $f = const$, то образ — единственная точка. Пусть $f \neq const$. Тогда образ $[a,b]$ не просто отрезок — это отрезок $\Delta = [\min\limits_{[a,b]} f, \max\limits_{[a,b]} f]$. Действительно, мы показали, что значения $\min\limits_{[a,b]} f$ и $\max\limits_{[a,b]} f$ достигаются. Более того, то, что любые значения из Δ достигаются, следует из теоремы о промежуточном значении. Наконец, очевидно, что значения вне Δ не принадлежат образу (поскольку эти значения находятся вне $\min\limits_{[a,b]} f$ и $\max\limits_{[a,b]} f$). $\square$

Разрывы, классификация точек разрыва

Непрерывные функции не имеют разрывов. Если функция не непрерывна, тогда мы будем интересоваться какого типа разрывы у неё могут быть. Одни разрывы в некотором смысле лучше, другие хуже.

Определение. *Если существуют конечные односторонние пределы* $\lim\limits_{x\to x_0-0} f(x)$, $\lim\limits_{x\to x_0+0} f(x)$ *и они равны между собой (но не равны* $f(x_0)$*), то* x_0 *называется устранимой точкой разрыва. Если существуют конечные односторонние пределы* $\lim\limits_{x\to x_0-0} f(x)$, $\lim\limits_{x\to x_0+0} f(x)$ *и они не равны между собой, то* x_0 *называется точкой разрыва 1-го рода. Если не существует (или бесконечен) хотя бы один односторонний предел* $\lim\limits_{x\to x_0-0} f(x)$, $\lim\limits_{x\to x_0+0} f(x)$*, то точка* x_0 *называется точкой разрыва 2-го рода.*

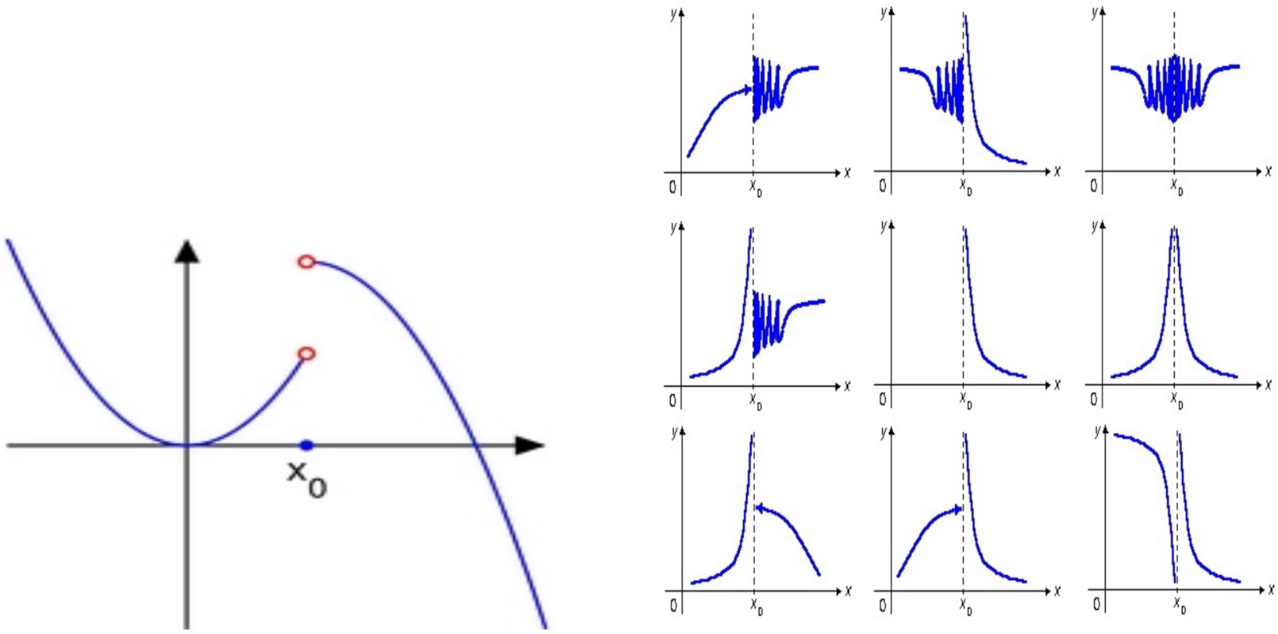


Рис. 2. На левой картинке показан разрыв 1-го рода, на правой картинке — почти все возможные варианты разрыва 2-го рода

Примеры. 1) Функция *знак* определена как:

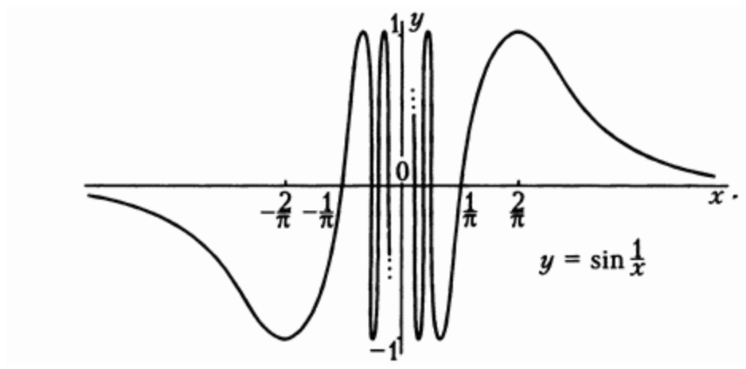
$$sgn(x) = \begin{cases} 0, & x = 0, \\ 1, & x > 0, \\ -1, & x < 0. \end{cases}$$

У неё есть единственный разрыв в точке 0, являющийся разрывом 1-го рода.

2) Функция

$$f(x) = \begin{cases} \sin(1/x), & x \neq 0, \\ 0, & x = 0 \end{cases}$$

разрывна в точке 0. В этой точке у неё разрыв 2-го рода.



3) *Функция Дирихле* определена следующим образом:

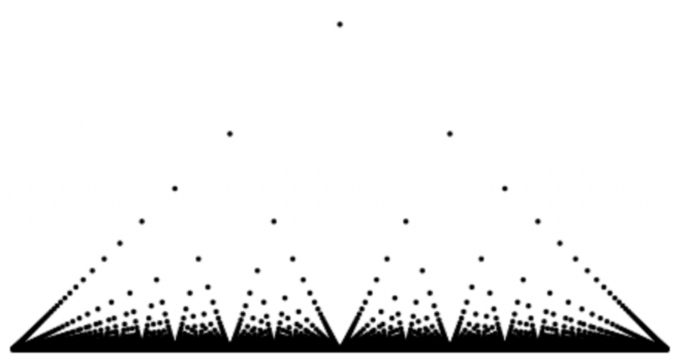
$$D(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

У функции Дирихле в каждой точке разрыв 2-го рода.

2) Родственницей функции Дирихле является *функция Римана*:

$$R(x) = \begin{cases} 1, & x = 0, \\ 1/q, & x = p/q, (p, q) = 1, \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Функция Римана разрывна в рациональных точках (все они являются разрывами 1-го рода) и непрерывна в иррациональных.



Элементарные функции

Элементарной функцией называется функция, которую можно задать одним аналитическим выражением, составленным из основных элементарных функций с помощью четырёх арифметических действий и композиций функций. Основными элементарными функциями называются функции $f(x) = const, f(x) = x, f(x) = a^x$ и $f(x) = \ln x$, степенная функция x^α, все тригонометрические функции и обратные к ним. Таким образом, основная элементарная функция — это кирпичик, из которого мы строим более сложные элементарные функции с помощью обычных операций. На семинарах мы корректно определим функции $a^x, \ln x$ и x^α и докажем, что они непрерывны в своих областях определения. Для тригонометрическич функций и обратных к ним мы не будем этого доказывать.

С помощью основной элементарной функции определяются *гиперболические функции*:

$$\operatorname{sh} x = \frac{e^x - e^{-x}}{2}, \quad \operatorname{ch} x = \frac{e^x + e^{-x}}{2}, \quad \operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x}, \quad \operatorname{cth} x = \frac{\operatorname{ch} x}{\operatorname{sh} x},$$

эти функции называются соответсвенно *гиперболическим синусом, гиперболическим косинусом, гиперболическим тангенсом и гиперболическим котангенсом.* Также по аналогии с тригонометрическими функциями определяются *гиперболический секанс и гиперболический косеканс*:

$$sch\ x = 1/\operatorname{ch} x, \quad csch\ x = 1/\operatorname{sh} x.$$

Аналогия с тригонометрическими функциями заключается в том, что для гиперболических функций выполняются например такие тождества

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1, \quad \operatorname{sh}(x \pm y) = \operatorname{sh} x \operatorname{ch} y \pm \operatorname{ch} x \operatorname{sh} y, \quad \operatorname{ch}(x \pm y) = \operatorname{ch} x \operatorname{ch} y \pm \operatorname{sh} x \operatorname{sh} y,$$

$$\operatorname{th}(x \pm y) = \frac{\operatorname{th} x \pm \operatorname{th} y}{1 \pm \operatorname{th} x \operatorname{th} y}, \quad \operatorname{ch}^2(x/2) = \frac{\operatorname{ch} x + 1}{2}, \quad \operatorname{sh}^2(x/2) = \frac{\operatorname{ch} x - 1}{2}.$$

Также верны формулы двойного, тройного, кратного “угла”, суммы и разности синусов, косинусов и тангенсов. Все эти формулы легко выводятся из определений. Также легко видеть, что гиперболические синус и тангенс — нечётные функции (т. е. $f(-x) = -f(x)$), гиперболические косинус и котангенс — чётные (т. е. $f(-x) = f(x)$).

Лекция 14
Производные

Пусть функция f определена в окрестности точки x_0.

Определение производной 1. *Если существует конечный предел*

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}=A,$$

то функция f называется дифференцируемой в точке x_0, число A называется производной функции f в точке x_0 и обозначается $f'(x_0)$.

Определение производной 2. *Функция f называется дифференцируемой в точке x_0, если при некотором A для её значений в точках x_0+h при $h\to 0$ справедливо представление $f(x_0+h)=f(x_0)+Ah+o(h)$, при этом число A называется производной функции f в точке x_0 и обозначается $f'(x_0)$.*

Оба определения эквивалентны, это просто различная формулировка одной и той же формулы:

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}=A\Leftrightarrow\lim_{h\to 0}\frac{f(x_0+h)-f(x)}{h}=A\Leftrightarrow f(x_0+h)=f(x_0)+Ah+o(h).$$

Утверждение. *Дифференцируемая в точке x_0 функция непрерывна в этой точке.*

Доказательство. Следует из определения 2: если $f(x_0+h)=f(x_0)+Ah+o(h)$, при $h\to 0$, то $\lim\limits_{h\to 0}f(x_0+h)=f(x_0)$. $\square$

Замечание. 1) Функция f дифференцируема в точке x_0 тогда и только тогда, когда

$$\lim_{x\to x_0+0}\frac{f(x)-f(x_0)}{x-x_0}=\lim_{x\to x_0-0}\frac{f(x)-f(x_0)}{x-x_0}$$

при условии, что оба предела существуют и конечны. Они называются "односторонними производными". Более точно, предел при $x\to x_0+0$ называется "правой производной", а предел при $x\to x_0-0$ — левой.

2) Не верно, что всякая непрерывная в x_0 функция дифференцируема в x_0: функция $|x|$ является непрерывной в нуле, но не дифференцируемой в нём. Убедитесь в этом!

Определение. *Функция $f\colon E\to\mathbb{R}$ называется дифференцируемой на E, если она дифференцируема в каждой точке множества E.*

Обозначение. $C^1(E)$ — множество дифференцируемых функций на E. Из предыдущего утверждения следует, что $C^1(E)\subset C^0(E)$.

Приведем некоторое количество примеров вычисления производных по определению.

$$(x^n)'=\lim_{h\to 0}\frac{(x+h)^n-x^n}{h}=\lim_{h\to 0}\frac{x^n+nx^{n-1}h+\ldots+h^n-x^n}{h}=nx^{n-1};$$

$$(a^x)'=\lim_{h\to 0}\frac{a^{x+h}-a^x}{h}=\lim_{h\to 0}\frac{a^x(a^h-1)}{h}=a^x\lim_{h\to 0}\frac{a^h-1}{h}=a^x\ln a;$$

$$(\sin x)'=\lim_{h\to 0}\frac{\sin(x+h)-\sin x}{h}=\lim_{h\to 0}\frac{2\cos(x+h/2)\sin(h/2)}{h}=\lim_{h\to 0}\frac{\sin(h/2)}{h/2}\lim_{h\to 0}\cos(x+h/2)=\cos x;$$

аналогично $(\cos x)'=-\sin x$.

Арифметические правила дифференцирования.

Теорема. *Пусть функции f и g дифференцируемы в точке x_0, внутренней точке области определения функций.*

1. Линейная комбинация $\varphi = \alpha f + \beta g (\alpha, \beta \in \mathbb{R})$ дифференцируема в точке x_0 и $\varphi' = \alpha f' + \beta g'$.

2. Произведение $f \cdot g$ дифференцируемо в точке x_0 и $(f \cdot g)' = f' \cdot g + f \cdot g'$ (правило Лейбница).

3. Пусть $g(x_0) \neq 0$, тогда частное f/g дифференцируемо в точке x_0 и

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

В частности

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

Доказательство.

$$1.\ \alpha f(x_0+h) + \beta g(x_0+h) - \alpha f(x_0) - \beta g(x_0) = \alpha(f(x_0+h) - f(x_0)) + \beta(g(x_0+h) - g(x_0)) =$$
$$= \alpha(f'(x_0)h + o(h)) + \beta(g'(x_0)h + o(h)) = (\alpha f'(x_0) + \beta g'(x_0))h + o(h),\ h \to 0.$$

$$2.\ f(x_0+h) \cdot g(x_0+h) - f(x_0) \cdot g(x_0) = f(x_0+h) \cdot g(x_0+h) - f(x_0) \cdot g(x_0+h) +$$
$$+ f(x_0) \cdot g(x_0+h) - f(x_0) \cdot g(x_0) = (f'(x_0)h + o(h)) \cdot (g(x_0) + o(1)) + f(x_0) \cdot (g'(x_0)h + o(h)) =$$
$$= (f'(x_0)g(x_0) + f(x_0)g'(x_0))h + o(h)\, h \to 0.$$

$$3.\ \frac{f(x_0+h)}{g(x_0+h)} - \frac{f(x_0)}{g(x_0)} = \frac{f(x_0+h)g(x_0) - f(x_0)g(x_0+h)}{g(x_0+h)g(x_0)} =$$
$$= \frac{f(x_0+h)g(x_0) - f(x_0)g(x_0) + f(x_0)g(x_0) - f(x_0)g(x_0+h)}{g(x_0+h)g(x_0)} =$$
$$= \frac{(f(x_0+h) - f(x_0))g(x_0) + f(x_0)(g(x_0) - g(x_0+h))}{g(x_0+h)g(x_0)} = \frac{f'(x_0)g(x_0)h - f(x_0)g'(x_0)h + o(h)}{g^2(x_0) + o(1)} =$$
$$= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}h + o(h),\ h \to 0.\ \square$$

Пример. $(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{(\sin x)' \cos x - \sin x (\cos x)'}{\cos^2 x} = \frac{1}{\cos^2 x}$.

Теорема о дифференцировании сложной функции. *Пусть функция f определена в окрестности точки x_0 и дифференцируема в точке x_0. Пусть функция g определена в окрестности точки $f(x_0)$ и дифференцируема в точке $f(x_0)$. Тогда функция $g \circ f$ определена в окрестности точки x_0, дифференцируема в этой точке и*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Доказательство. По предположению $f(x_0+h) - f(x_0) = f'(x_0)h + o(h)$, $g(f(x_0)+t) - g(f(x_0)) = g'(f(x_0))t + o(t)$ при $h, t \to 0$. Положим во второй формуле $t = f(x_0+h) - f(x_0)$. В силу первой формулы имеем $t = O(h)$, значит $o(t) = o(h)$ при $h, t \to 0$. Поэтому

$$g(f(x_0)+h) - g(f(x_0)) = g'(f(x_0))(f(x_0+h) - f(x_0)) + o(h) = g'(f(x_0))(f'(x_0)h + o(h)) + o(h) =$$
$$= g'(f(x_0))f'(x_0)h + o(h).$$

При этом дифференцируемая в точке x_0 функция f непрерывна в этой точке, поэтому при достаточно малом h величина $f(x_0+h)$ попадает в ту окрестность точки $f(x_0)$, в которой определена функция g. $\square$

Пример. $(\sin x^2)' = (\sin)'(x^2)(x^2)' = (\cos x)2x = 2x \cos x$.

Дифференцируемость обратной функции

Пусть функция f непрерывна и строго монотонна в окрестности точки x_0. На семинарах мы показали, что в окрестности точки $f(x_0)$ определена непрерывная обратная функция f^{-1} со значениями в окрестности точки x_0.

Теорема о производной обратной функции. *Пусть функция f дифференцируема в точке x_0 и пусть $f'(x_0) \neq 0$. Тогда функция f^{-1} дифференцируема в точке $f(x_0)$ и*

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}$$

Доказательство. Имеем $\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0) \neq 0$. Хотим показать, что

$$\lim_{t\to 0} \frac{f^{-1}(f(x_0)+t) - f^{-1}(f(x_0))}{t} = \frac{1}{f'(x_0)}.$$

Перепишем искомое равенство в эквивалентном виде

$$\lim_{t\to 0} \frac{t}{f^{-1}(f(x_0)+t) - f^{-1}(f(x_0))} = f'(x_0).$$

Обозначим $u = f^{-1}(f(x_0)+t)$, тогда $f(u) = f(x_0)+t$ и $t = f(u) - f(x_0)$, очевидно, $u \to x_0$ при $t \to 0$. Поэтому

$$\lim_{t\to 0} \frac{t}{f^{-1}(f(x_0)+t) - f^{-1}(f(x_0))} = \lim_{u\to x_0} \frac{f(u)-f(x_0)}{u-x_0} = f'(x_0).$$

Пример 1. $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$, $(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$, $x \in (-1,1)$. Действительно, пусть $x = \sin y$, $y = \arcsin x$. Эти функции взаимно обратны, если $x \in [-1,1]$ и $y \in [-\pi/2, \pi/2]$. Теперь, если $\cos y \neq 0$ или, что тоже, $|x| \neq 1$, то

$$y' = \frac{1}{(\sin y)'} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}}.$$

Знак перед корнем выбран с учётом того, что $\cos y > 0$ при $y \in [-\pi/2, \pi/2]$. Далее, так как $\arcsin x + \arccos x = \pi/2$ при $x \in (-1,1)$, то $(\arccos x)' + (\arcsin x)' = 0$, откуда находим формулу для $(\arccos x)'$.

Пример 2. $(\operatorname{arctg} x)' = \frac{1}{1+x^2}$. Действительно, пусть $x = \operatorname{tg} y, y = \operatorname{arctg} x$. Эти функции взаимно обратны, если $x \in (-\infty, \infty)$ и $y \in [-\pi/2, \pi/2]$. Теперь

$$y' = \frac{1}{(\operatorname{tg} y)'} = \frac{1}{\frac{1}{\cos^2 y}} = \cos^2 y = \frac{1}{1+\operatorname{tg}^2 y} = \frac{1}{1+x^2}$$

Таблица производных

Эта таблица содержит все основные формулы производных, зная которые можно вычислять производные других функций. По крайней мере некоторые из этих формул надо знать наизусть: 1-9, 11. Формулы 13-15 желательно помнить, как похожие на формулы 5-7. Формулы 17-20 приведены в ознакомительных целях.

	Функция f	Производная f'	О.Д.З.
1.	$const$	0	
2.	x^α	$\alpha x^{\alpha-1}$	$x > 0$ при $\alpha \in \mathbb{R}$, $x \in \mathbb{R} \setminus \{0\}$ при $\alpha \in \mathbb{Q}$
3.	a^x	$a^x \ln a$	$x \in \mathbb{R},\ a > 0,\ a \neq 1$
4.	$\log_a \lvert x \rvert$	$\frac{1}{x \ln a}$	$x \neq 0,\ a > 0,\ a \neq 1$
5.	$\sin x$	$\cos x$	
6.	$\cos x$	$-\sin x$	
7.	$\operatorname{tg} x$	$\frac{1}{\cos^2 x}$	$x \neq \frac{\pi}{2} + \pi k$
8.	$\operatorname{ctg} x$	$-\frac{1}{\sin^2 x}$	$x \neq \pi k$
9.	$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$\lvert x \rvert < 1$
10.	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$	$\lvert x \rvert < 1$
11.	$\operatorname{arctg} x$	$\frac{1}{1+x^2}$	
12.	$\operatorname{arcctg} x$	$-\frac{1}{1+x^2}$	
13.	$\operatorname{sh} x$	$\operatorname{ch} x$	
14.	$\operatorname{ch} x$	sh	
15.	$\operatorname{th} x$	$\frac{1}{\operatorname{ch}^2 x}$	
16.	$\operatorname{cth} x$	$-\frac{1}{\operatorname{sh}^2 x}$	$x \neq 0$
17.	$\operatorname{arsh} x = \ln\left(x + \sqrt{1+x^2}\right)$	$\frac{1}{\sqrt{1+x^2}}$	
18.	$\operatorname{arch} x = \ln\left(x \pm \sqrt{x^2-1}\right)$	$\pm\frac{1}{\sqrt{x^2-1}}$	$\lvert x \rvert > 1$
19.	$\operatorname{arth} x = \frac{1}{2} \ln \frac{1+x}{1-x}$	$\frac{1}{1-x^2}$	$\lvert x \rvert < 1$
20.	$\operatorname{arcth} x = \frac{1}{2} \ln \frac{x+1}{x-1}$	$\frac{1}{1-x^2}$	$\lvert x \rvert > 1$

Рис. 3. Таблица производных.

Лекция 15

Геометрический смысл производной. Касательная как предел секущих

Будем предполагать, что кривая задана графиком функции f. Секущей, проходящей через точку $(x_0, f(x_0))$, называется прямая, проходящая через эту точки и ещё точку $(x_0+h, f(x_0+h))$ графика функции.

Уравнение секущей:

$$y - f(x_0) = \frac{1}{h}\big(f(x_0+h) - f(x_0)\big)(x - x_0).$$

При $x = x_0$ имеем $y = f(x_0)$, при $x = x_0 + h$ имеем $y = f(x_0 + h)$, то есть это в самом деле секущая; при каждом $h \neq 0$ — своя секущая.

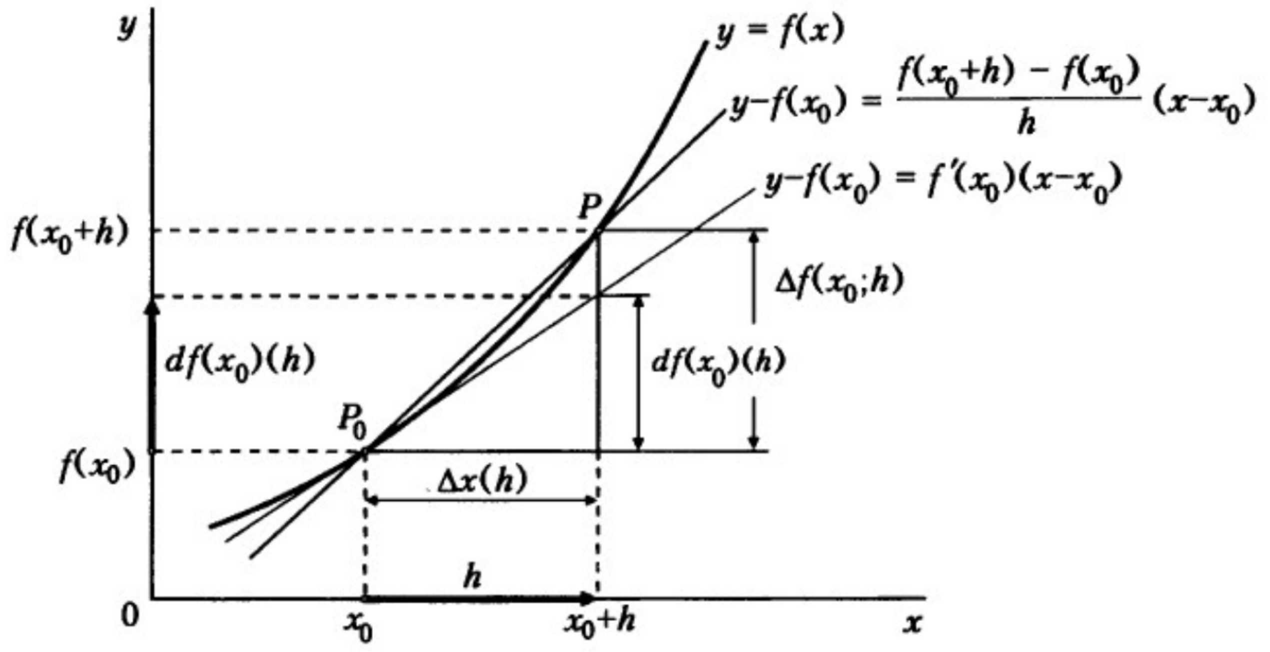


Рис. 4. Касательная и секущая.

Теперь устремим h к нулю. Если функция f дифференцируема в точке x_0, то уравнение "предельной секущей" будет иметь вид $y - f(x_0) = f'(x_0)(x - x_0)$. Эта прямая проходит через точку $(x_0, f(x_0))$, тангенс угла её наклона равен производной, эта прямая называется касательной (или касательной к графику f в точке x_0).

Здесь есть одна тонкость: мы перешли к пределу в уравнении прямой. Что это значит формально — прямая стремится к прямой — не слишком понятно. Пересекающиеся в некоторой точке прямые, даже с очень близкими углами наклона, "на бесконечности" далеки друг от друга. Однако, пусть мы будем писать уравнение каждой прямой в виде $y - f(x_0) = k(x - x_0)$. При разных k это будут прямые, проходящие через точку $(x_0, f(x_0))$. Пусть число k зависит от параметра h. И пусть существует предел $k(h)$ при $h \to 0$ такой, что $k(h) \to k(0)$. Тогда вполне естественно считать, что прямая $y - f(x_0) = k(0)(x - x_0)$ есть предельное положение прямых $y - f(x_0) = k(h)(x - x_0)$ при $h \to 0$. По такому определению у дифференцируемой функции касательная $y - f(x_0) = f'(x_0)(x - x_0)$ есть предельное положение секущей: $y - f(x_0) = h^{-1}(f(x_0+h) - f(x_0))(x - x_0))$, при $h \to 0$. Секущая проходит через точки $(x_0, f(x_0)$ и $(x_0 + h, f(x_0 + h))$.

Дифференциал

Определение. *Линейная функция $h \mapsto f'(x_0)h$ называется дифференциалом функции f в точке x_0.*

Обозначения. $df(x_0), d_{x_0}f, Df(x_0), D_{x_0}f$.

Часто говорят, что дифференциал — *главная линейная часть приращения дифференцируемой функции.* В самом деле, *приращение функции* $f(x_0+h)-f(x_0)$ имеет вид суммы значения дифференциала Ah и малых более высокого порядка $o(h)$.

Замечание 1. Можно задать вопрос: раз дифференциал — линейное отображение, то из какого векторного пространства в какое он бьёт? Дифференциал — это линейная функция, для вещественных функций линейная функция — это умножения на число. Это функция определённая на приращениях h к x_0 и принимающая значения в множестве приращений к значениям функции. То есть, дифференциал, как и всякое линейное отображение, должен переводить начало координат в начало координат. Иными словами, если параллельно переместить начало координат на плоскости в точку $(x_0, f(x_0))$, то в новых координатах дифференциал будет настоящей линейной функцией. При этом у нас были переменные x и y, появились новые переменные $x-x_0$ и $y-f(x_0)$. Была числовая прямая "иксов", стала числовая прямая приращений x. Была числовая прямая "игреков", стала числовая прямая приращений y. Эти сдвинутые числовые прямые называют "касательные пространства" и обозначают $T_{x_0}\mathbb{R}$ и $T_{f(x_0)}\mathbb{R}$. Дифференциал — это линейная функция из $T_{x_0}\mathbb{R}$ в $T_{f(x_0)}\mathbb{R}$:

$$d_{x_0}f\colon T_{x_0}\mathbb{R} \to T_{f(x_0)}\mathbb{R}.$$

Эта терминология для функций из $\mathbb{R}$ в $\mathbb{R}$ не обязательна и используется редко. Однако, когда мы в будущем дойдем до дифференцирования отображений из $\mathbb{R}^m$ в $\mathbb{R}^n$, вот там терминология станет уместной.

Замечание 2. Лейбниц использовал вместо обозначения $f'(x)$ для производной обозначение $\frac{df}{dx}$. Это обозначение очень удобно. Вообще-то, это единый символ: производная f в точке x. Однако, это одновременно и дробь. И в некоторых формулах именно так её и удобно использовать, по крайней мере для запоминания. Мы будем использовать это обозначение иногда по-разному:

$$f'(x_0) = \frac{df}{dx}(x_0) = \frac{df(x_0)}{dx} = \frac{d}{dx}f(x)|_{x=x_0}.$$

Следующее утверждение — это всего лишь "наведение порядка" в дифференциалах функций и касательных пространствах, в которых действуют эти дифференциалы, из теоремы о производной сложной функции.

Утверждение о дифференциале сложной функции. *Дифференциал композиции $g \circ f$ в точке x_0 является композицией дифференциалов $dg(f(x_0)) \circ df(x_0)$.*

Замечание. В условиях теоремы об обратной функции дифференциал обратного отображения — обратное отображение к дифференциалу функции f. Дифференциал $df(x_0)$ действует из касательного пространства $T_{x_0}\mathbb{R}$ в касательное пространство $T_{f(x_0)}\mathbb{R}$, дифференциал $df^{-1}(f(x_0))$ действует из касательного пространства $T_{x_0}\mathbb{R}$ в касательное пространство $T_{f(x_0)}\mathbb{R}$. Геометрически теорема об производной обратной функции означает следующее. График обратной функции симметричен графику функции относительно биссектрисы $y=x$. Поэтому касательная в какой-то точке также симметрична соответствующей касательной. Соответственно, произведение тангенсов углов наклона симметричных касательных равно 1. Это и утверждает теорема о проbзводной обратной функции.

Основные теоремы дифференциального исчисления

Пусть на некотором интервале U задана непрерывная функция f.

Определение. *Внутренняя точка x_0 называется точкой локального минимума, если в некоторой окрестности $U(x_0) \subset U$ $f(x_0) \leqslant f(x)$ для всякой точки $x \in U(x_0)$. Внутренняя точка x_0 называется точкой строгого локального минимума, если в некоторой проколотой окрестности $\overset{\circ}{U}(x_0) \subset U$ $f(x_0) < f(x)$ для всякой точки $x \in \overset{\circ}{U}(x_0)$. Величина $f(x_0)$ называется локальным минимумом функции f. Аналогично определяется точка (строгого) локального максимума и локальный максимум функции f. Точки локального максимума и минимума называются точками локального экстремума.*

Теорема Ферма. *Пусть в точке x_0 локального экстремума функция f дифференцируема. Тогда $f'(x_0) = 0$.*

Доказательство. Пусть, например, x_0 — точка локального максимума. Тогда

$$f'(x_0) = \lim_{h\to 0+0} \frac{f(x_0+h) - f(x_0)}{h} \leqslant 0, \quad f'(x_0) = \lim_{h\to 0-0} \frac{f(x_0+h) - f(x_0)}{h} \geqslant 0,$$

поэтому $f'(x_0) = 0$. $\square$

Замечание 1. Теорема Ферма справедлива исключительно для внутренних точек области определения.

Теорема Ферма утверждает, что равенство $f'(x_0) = 0$ является необходимым условием локального экстремума. Для того, чтобы найти максимум дифференцируемой функции можно сначала найти все нули производной. Если их окажется конечное количество, то переберем их по очереди и найдем ту, в которой достигается максимум. Если требуется найти максимум функции на отрезке, то надо не забыть, что он может достигаться в крайних точках отрезка. Поэтому к списку “подозрительных на максимум” точек, в которых $f' = 0$ надо добавить ещё границы отрезка.

Теорема Ролля. *Пусть функция f непрерывна на отрезке $[a, b]$, дифференцируема на интервале (a, b), и пусть $f(a) = f(b)$. Тогда для некоторой точки $c \in (a, b)$ справедливо соотношение $f'(c) = 0$.*

Доказательство. Так как функция f непрерывна на компакте, то найдутся точки x_m и x_M, в которых функция f достигает своего минимума и максимума на $[a, b]$. Если хотя бы одна из этих точек внутренняя, то утверждение теоремы следует из теоремы Ферма. Если обе эти точки граничные, то получается, что минимум функции равен максимуму, то есть функция f постоянная и её производная равна 0 во всех точках интервала. $\square$

Теорема Лагранжа. *Пусть функция f непрерывна на отрезке $[a, b]$, дифференцируема на интервале (a, b). Тогда для некоторой точки $c \in (a, b)$ справедливо соотношение*

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

Теорему Лагранжа называют также “формула Лагранжа” и “формула конечных приращений”.

Доказательство. Рассмотрим функцию

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a), \quad F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

Эта функция удовлетворяет всем условиям теоремы Ролля: $F(a) = F(b) = f(a)$. Значит, в некоторой точке $c \in (a, b)$ справедливо равенство $F'(c) = 0$. $\square$

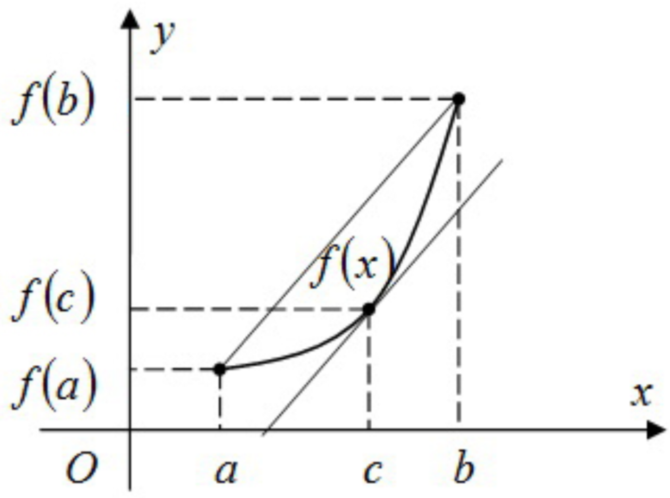


Рис. 5. К теореме Лагранжа.

Следствие 1 (критерий постоянства). *Если $f' = 0$ на интервале, то $f = const$. Более того, если $f' = g'$ на интервале, то $f - g = const$.*

Следствие 2 (признак монотонности). *Если $f' \geqslant 0$, то функция f не убывает, если $f' > 0$, то функция строго возрастает. Если $f' \leqslant 0$, то функция f не возрастает, если $f' < 0$, то функция строго убывает.*

Следствие 3 (условие Липшица). $|f(b) - f(a)| \leqslant \sup\limits_{(a,b)} |f'| \cdot |b - a|$ (это, конечно имеет смысл, когда супремум $\sup\limits_{(a,b)}$ конечен).

Определение. *Говорят, что функция f удовлетворяет условию Липшица с константой Липшица L, если существует константа $L > 0$ така, что при всех x, y справедливо неравенство $|f(x) - f(y)| \leqslant L|x - y|$. Такую функцию называют "липшицева функция".*

Следствие 3 означает, что если у функции на промежутке есть *ограниченная производная*, то функция удовлетворяет условию Липшица.

Пример. Приведём любопытный пример. Оказывается, существует непрерывная монотонная функция из отрезка в отрезок, которая не является постоянной, но при этом имеет производную, равную нулю в почти всех точках. Эта функция называется *канторовой лестницей*. Её определение таково. В точках 0 и 1 значение функции принимается равным соответственно 0 и 1. Далее интервал $(0, 1)$ разбивается на три равные части $\left(0, \frac{1}{3}\right)$, $\left(\frac{1}{3}, \frac{2}{3}\right)$ и $\left(\frac{2}{3}, 1\right)$. На среднем сегменте полагаем $F(x) = \frac{1}{2}$. Оставшиеся два сегмента снова разбиваются на три равные части каждый, и на средних сегментах $F(x)$ полагается равной $\frac{1}{4}$ и $\frac{3}{4}$. Каждый из оставшихся сегментов снова делится на три части, и на внутренних сегментах $F(x)$ определяется как постоянная, равная среднему арифметическому между соседними, уже определенными значениями $F(x)$. На остальных точках единичного отрезка определяется по непрерывности. Нетрудно видеть, что производная канторовой лестницы определена и равна нулю во всех точках, кроме канторова множества.

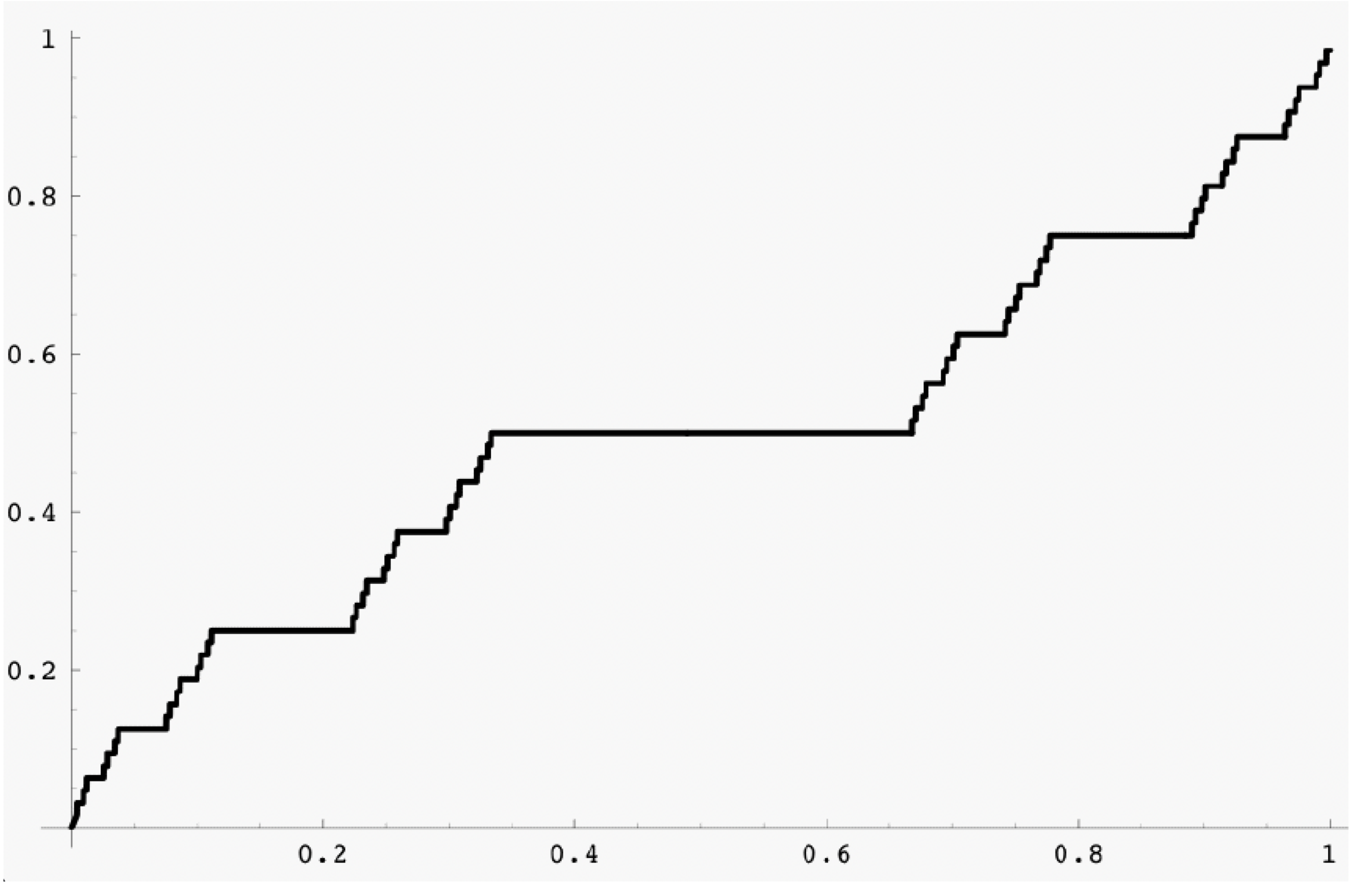


Рис. 6. Канторова лестница.

Лекция 16

Теорема Коши. *Пусть даны две функции f и g, непрерывные на отрезке $[a,b]$ и дифференцируемые на интервале (a,b). Тогда найдется такая точка $c \in (a,b)$, что*

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

Доказательство. Функция $F(x) = f(x)\big(g(b)-g(a)\big) - g(x)\big(f(b)-f(a)\big)$ удовлетворяет всем условиям теоремы Ролля: $F(a) = F(b) = f(a)g(b) - g(a)f(b)$. Поэтому в некоторой точке $F'(c) = 0$. □

Следствие. *Пусть, дополнительно, $g'(x) \neq 0$ на (a,b). Тогда $g(a) \neq g(b)$ и в условиях теоремы Коши*

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

Доказательство. Соотношение $g(a) \neq g(b)$ следует из теоремы Ролля: если бы это было не так, в некоторой точке производная обратилась бы в 0. Требуемое равенство следует из теоремы Коши. □

Раскрытие неопределённостей. Правило Лопиталя

В этом разделе мы вернёмся к вычислению пределов функции $\lim\limits_{x\to x_0} f(x)$. Самый простой случай — это когда функция f непрерывная в точке x_0. По определению непрерывности, этот предел равен $f(x_0)$ и всё. Этот случай настолько прост, что его можно не рассматривать.

Трудности возникают, когда простых теорем о пределах оказывается недостаточно. Например, $f(x) = g(x)/h(x)$ и $\lim\limits_{x\to x_0} g(x) = \lim\limits_{x\to x_0} h(x) = 0$ или $\lim\limits_{x\to x_0} g(x) = \lim\limits_{x\to x_0} h(x) = \infty$. Ещё варианты сложностей: $f(x) = g(x) - h(x)$ и $\lim\limits_{x\to x_0} g(x) = \lim\limits_{x\to x_0} h(x) = \infty$; $f(x) = g(x) \cdot h(x)$ и $\lim\limits_{x\to x_0} g(x) = 0$, $\lim\limits_{x\to x_0} h(x) = \infty$. Рассмотрим именно такие случаи, все они называются «неопределённостями». Найти предел с неопределённостью называется «раскрыть неопределённость». Уже перечисленные неопределённости называются неопределённость $0/0, \infty/\infty, \infty-\infty$ и $0\cdot\infty$. Эквивалентными преобразованиями и упрощениями иногда часто можно сводить одни неопределённости к другим, самый просто пример: $0\cdot\infty = 0/(1/\infty)$, последняя неопределённость уже имеет вид $0/0$.

Правило Лопиталя $[\frac{0}{0}]$**.** *Пусть функции f и g определены и дифференцируемы в окрестности точки x_0 и пусть $f(x_0) = g(x_0) = 0$. Пусть $g'(x) \neq 0$ на этой окрестности. Пусть существует предел $\lim\limits_{x\to x_0} \dfrac{f'(x)}{g'(x)} = A$. Тогда существует предел $\lim\limits_{x\to x_0} \dfrac{f(x)}{g(x)} = A$.*

Доказательство. Заметим, что

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{g(x) - g(x_0)}$$

Функции f и g удовлетворяют всем условиям теоремы Коши, поэтому при некотором $\xi \in (x, x_0)$ справедливо равенство

$$\frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(\xi)}{g'(\xi)}.$$

Осталось заметить, что при $x \to x_0$ также $\xi \to x_0$. $\square$

Есть и другие утверждения о ситуациях, в которых из существования предела отношения производных следует существование предела отношения самих функций. Их тоже называют правилом Лопиталя. Например, приведённое утверждение можно переформулировать для односторонних пределов. Сформулируем без доказательства ещё одно утверждение.

Правило Лопиталя $[\frac{\infty}{\infty}]$**.** *Пусть при $x \in (x_0, x_0 + h)$ функции f и g определены и дифференцируемы и пусть $f(x) \to \infty$, $g(x) \to \infty$ при $x \to x_0 + 0$. Пусть $g'(x) \neq 0$ в $(x_0, x_0 + h)$ и пусть существует предел $\lim\limits_{x\to x_0+0} \dfrac{f'(x)}{g'(x)} = A$. Тогда существует предел $\lim\limits_{x\to x_0+0} \dfrac{f(x)}{g(x)} = A$.*

Доказательство можно найти в книге Фихтенгольца, стр. 320-321.

Замечание. Правило Лопиталя применимо только в случае, когда исходная дробь является неопределённостью! Только в этом случае!

Пример.

$$\lim_{x\to 0} \frac{e^x - \ln(1+x) - \cos x}{x^2} = \left[\frac{0}{0}\right] = \lim_{x\to 0} \frac{(e^x - \ln(1+x) - \cos x)'}{(x^2)'} =$$
$$= \lim_{x\to 0} \frac{e^x - 1/(1+x) + \sin x}{2x} = \left[\frac{0}{0}\right] = \lim_{x\to 0} \frac{e^x + 1/(1+x)^2 + \cos x}{2} = \frac{3}{2}.$$

Односторонняя производная. Теорема Дарбу

В этом разделе мы частично ответим на вопрос: «Для каких функций f найдется функция F такая, что $F' = f$?».

Бывает так, что у графика функции в некоторой точке есть «излом». Например, как у графика функции $y = |x|$ в нуле. На математическом языке это означает, что существуют односторонние пределы

$$f'_+(x_0) = \lim_{h\to 0+0} \frac{f(x_0+h) - f(x_0)}{h}, \quad f'_-(x_0) = \lim_{h\to 0-0} \frac{f(x_0+h) - f(x_0)}{h},$$

не равные между собой. Аналогично, к исследованию таких пределов приводит определение производной в конце отрезка.

Величина $f'_+(x_0)$ называется *правой производной в точке* x_0, величина $f'_-(x_0)$ называется *левой производной в точке* x_0, обе эти величины называются *односторонними производными (аналогично пределам, правому, левому, одностороннему).*

Опеределение. *Про функцию f говорят, что она дифференцируема на отрезке, если она дифференцируема на интервале и в концах отрезка существуют односторонние производные.*

Теорема Дарбу. *Производная функции на отрезке принимает все промежуточные значения.*

Лемма. *Пусть функция f дифференцируемая на $[a, b]$ и пусть односторонние производные $f'_+(a)$ и $f'_-(b)$ имеют разный знак. Тогда $f'(c) = 0$ в некоторой точке $c \in (a, b)$.*

Доказательство. Пусть, для определённости, $f'_+(a) > 0$ и $f'_-(b) < 0$. Тогда максимум функции f (который достигается в некоторой точке отрезка) не может быть ни в одном из концов отрезка. Пусть $f(c) = \max f$, тогда по теореме Ферма $f'(c) = 0$. Лемма доказана. $\square$

Доказательство. Перейдём к доказательству теоремы Дарбу. Пусть $f'(x_1) \neq f'(x_2)$ и $\xi \in [f'(x_1), f'(x_2)]$. Покажем, что $f'(c) = \xi$ в некоторой точке $c \in (x_1, x_2)$.

Рассмотрим вспомогательную функцию $F(x) = f(x) - \xi x$. Так как $\xi \in [f'(x_1), f'(x_2)]$, то в точках x_1, x_2 производная этой функции принимает значения разных знаков. В силу доказанной леммы в некоторой точке с интервала (x_1, x_2) производная F' обращается в нуль, $f'(c) = \xi$. $\square$

Подчеркнём, что никаких предположений о непрерывности производной не было. Отсюда следует, что не все функции могут быть производными других функций: они должны удовлетворять свойству Дарбу — принимать все промежуточные значения, как это делали непрерывные функции.

Формула Тейлора

Производные высших порядков. Если функция f дифференцируема на интервале, то возникает другая функция f' — её производная. Если она дифференцируема, то эту функцию тоже можно дифференцировать. Производная от производной называется второй производной и обозначается f'' или $\frac{d^2f}{dx^2}$. Иногда вторую производную обозначают f''_{xx}. Точно также определяют следующие производные: $f''', f^{IV} = f^{(4)}, ..., f^{(n)}$. Обозначают либо римскими «цифрами», либо в круглых скобках.

Упражнение. Докажите формулу Лейбница:

$$(uv)^{(n)} = \sum_{k=0}^{n} C_n^k u^{(n-k)} v^{(k)}.$$

Формула Тейлора для многочлена. Рассмотрим многочлен P_n степени n. Ясно, что у многочлена степени n производная — также многочлен, только степени на 1 ниже. Таким образом, $P_n^{(n)}$ — это константа, $P_n^{(n+1)} = 0$. Выпишем формулы для производных многочлена, в частности их значений в нуле:

$$P_n(x) = c_0 + c_1x + \ldots + c_nx^n, \quad P_n(0) = c_0;$$
$$P_n'(x) = c_1 + 2c_2x + 3x_3^2 + \ldots + nc_nx^{n-1}, \quad P_n'(0) = c_1;$$
$$P_n''(x) = 2c_2 + 2 \cdot 3 \cdot c_3x + \ldots + n(n-1)c_nx^{n-2}, \quad P_n''(0) = 2c_2;$$
$$\ldots$$
$$P_n^{(n)}(x) = n!c_n, \quad P_n^{(n)}(0) = n!c_n.$$

Выразим теперь (из формул правого столбца) коэффициенты c_k через значения производных в нуле и перепишем многочлен в виде:

$$P_n(x) = P_n(0) + \frac{1}{1!}P_n'(0)x + \frac{1}{2!}P''(0)x^2 + \frac{1}{3!}P'''(0)x^3 + \ldots + \frac{1}{n!}P_n^{(n)}(0)x^n.$$

Справедлива аналогичная, формально более общая, формула:

$$P_n(x) = P_n(x_0) + \frac{1}{1!}P_n'(x_0)(x-x_0) + \frac{1}{2!}P''(x_0)(x-x_0)^2 + \frac{1}{3!}P'''(x_0)(x-x_0)^3 + \ldots + \frac{1}{n!}P_n^{(n)}(x_0)(x-x_0)^n.$$

Первая формула называется *разложением многочлена Тейлора в нуле, вторая — в точке x_0*. Это точные формулы. Есть аналогичные формулы, справедливые не только для многочленов, но для любых функций, которые имеют достаточное количество производных в окрестности точки, в которой выписывается разложение.

Формула Тейлора.

Определение. *Пусть в окрестности точки x_0 задана функция $f(x)$, у которой есть n производных. Тогда по f можно выписать многочлен Тейлора*

$$T_n(x_0; x) = f(x_0) + \frac{1}{1!}f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \ldots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n.$$

Определение. *Функцию $r_n(x_0; x) = f(x) - T_n(x_0; x)$ называют n−м остаточным членом формулы Тейлора $f(x) = T_n(x_0; x) + r_n(x_0; x)$.*

Эта формула имеет практический смысл, если мы увидим, что остаточный член маленький, например, $r_n(x_0; x) = o((x - x_0)^n)$ при малых $x - x_0$.

Замечание. Многочлен Тейлора подобран так, чтобы у остаточного члена $r(x_0; x)$ его значение в точке x_0 и значения всех его производных до порядка n включительно были равны 0

$$r(x_0; x) = r'(x_0; x) = r''(x_0; x) = \ldots = r^{(n)}(x_0; x) = 0, \text{ при } x = x_0.$$

Теорема. *Пусть на некотором интервале U определена и $n+1$ раз дифференцируема функция f. Пусть $x_0 \in U$. Пусть функция $\varphi(x)$ непрерывна и дифференцируема на U, причём $\varphi'(x) \neq 0$ на U (кроме, может быть, точки x_0). Тогда для любой точки $x \in U$ найдётся точка ξ, лежащая между x и x_0 такая, что*

$$r_n(x_0; x) = \frac{\varphi(x) - \varphi(x_0)}{\varphi'(\xi)n!}f^{(n+1)}(\xi)(x - \xi)^n.$$

Доказательство. Рассмотрим случай $x_0 < x$. Случай $x_0 > x$ аналогичен.

Рассмотрим вспомогательную функцию

$$F(t) = f(x) - T_n(t; x) = f(x) - f(t) - \frac{1}{1!}f'(t)(x - t) - \frac{1}{2!}f''(t)(x - t)^2 - \ldots - \frac{1}{n!}f^{(n)}(t)(x - t)^n$$

переменной t. Эта функция непрерывна и дифференцируема по крайней мере один раз во всех точках отрезка $[x_0, x]$, причём

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

Теперь применим к функциям $F(t)$ и $\varphi(t)$ на отрезке $[x_0, x]$ теорему Коши и получим, что в некоторой точке $\xi \in (x_0, x)$ справедливо равенство

$$\frac{F(x) - F(x_0)}{\varphi(x) - \varphi(x_0)} = \frac{F'(\xi)}{\varphi'(\xi)}.$$

Теперь $F(x) = 0, F(x_0) = r_n(x_0; x)$ и теорема доказана. $\square$

Остаточный член в форме Коши. Положим $\varphi(t) = x - t$, тогда

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

Остаточный член в форме Лагранжа. Положим $\varphi(t) = (x-t)^{n+1}$, тогда

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

Замечание. Обратите внимание на то, что функции φ' при $\varphi(t) = (x-t)^{n+1}$ обращаются в 0 при $t = x$. В условиях теоремы Коши, на которую мы ссылаемся, функция φ не должна обращаться в 0 на интервалах (x, x_0) или (x_0, x), про концы ничего не требуется.

Остаточный член в форме Пеано. Пусть на некотором интервале U определена и n раз дифференцируема функция f. Пусть $x_0 \in U$. Тогда

$$r_n(x_0; x) = o((x - x_0)^n).$$

Доказательство. Доказательство следует из замечания на странице 55 и следующей леммы, которая имеет самостоятельный интерес.

Лемма. *Пусть функция φ определена в окрестности точки x_0 и пусть*

$$\varphi(x_0) = \varphi'(x_0) = \varphi''(x_0) = \ldots = \varphi^{(n)}(x_0) = 0.$$

Тогда $\varphi(x) = o((x - x_0)^n)$.

Доказательство. По индукции. При $n = 1$ утверждение леммы следует из формулы $\varphi(x) = \varphi(x_0) + \varphi'(x_0)(x - x_0) + o(x - x_0)$, которая является определением дифференцируемости φ в точке x_0. Пусть утверждение леммы справедливо при $n = k$: из $\varphi(x_0) = \varphi'(x_0) = \varphi''(x_0) = \ldots = \varphi^{(k)}(x_0) = 0$ следует $\varphi(x) = o((x - x_0)^k)$.

Теперь пусть $\varphi(x_0) = \varphi'(x_0) = \varphi''(x_0) = \ldots = \varphi^{(k+1)}(x_0) = 0$. Функция $\psi(x) = \varphi'(x)$ имеет k производных и $\psi(x_0) = \psi'(x_0) = \psi''(x_0) = \ldots = \psi^{(k)}(x_0) = 0$. По предположению индукции $\psi(x) = o((x - x_0)^k)$. Теперь по теореме Лагранжа

$$\varphi(x) - \varphi(x_0) = \psi(\xi)(x - x_0) = o((\xi - x_0)^k)(x - x_0) = o((x - x_0)^{k+1}), \quad \varphi(x_0) = 0. \ \square$$

Связь формы Пеано остаточного члена с формами Лагранжа и Коши лучше всего видна при дополнительном предположении о непрерывности функции $f^{(n+1)}$. В этом случае

$$|r_n(x_0; x)| \leqslant |f^{(n+1)}| \cdot |x - x_0|^{n+1}$$

и получается даже более сильное (чем форма Пеано) утверждение $r_n(x_0; x) = O((x - x_0)^{n+1})$.

Замечание. Многочлен Тейлора $T_n(x_0;x)$ — единственный многочлен степени n, удовлетворяющий соотношению $f(x)-P_n(x)=o((x-x_0)^n)$. В самом деле, из соотношений $f(x)-P_n(x)=o((x-x_0)^n)$ и $f(x)-T_n(x_0;x)=o((x-x_0)^n)$ следует совпадение многочленов $P_n(x)$ и $T_n(x_0;x)$ переменной x. Это верно, так как если многочлен $P_n(x)-T_n(x_0;x)$ есть $o((x-x_0)^n)$, то он есть тождественный нуль.

Резюмируем содержание последних конструкций.

Формула Тейлора с остаточным членом в форме Пеано. Пусть функция f определена и n раз дифференцируема в некоторой окрестности точки x_0. Тогда

$$f(x)=f(x_0)+\frac{1}{1!}f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\ldots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n+o((x-x_0)^n).$$

Формула Тейлора с остаточным членом в форме Лагранжа. Пусть функция f определена и $n+1$ раз дифференцируема в некоторой окрестности точки x_0. Тогда

$$f(x)=f(x_0)+\frac{1}{1!}f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\ldots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n+\frac{1}{(n+1)!}f^{(n+1)}(\xi)(x-x_0)^{n+1}.$$

Формулу Тейлора в нуле (при $x_0=0$) называют *формулой Маклорена.*

Ряд Тейлора. Пусть функция f имеет в окрестности точки x_0 все производные. Тогда можно выписать ряд

$$\sum_{k=0}^{\infty}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k=f(x_0)+\frac{1}{1!}f'(x_0)(x-x_0)+\frac{1}{2!}f''(x_0)(x-x_0)^2+\ldots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n+\ldots$$

Этот ряд называется *рядом Тейлора функции f в точке x_0*. Его можно выписать для любой бесконечно дифференцируемой функции f . Не надо думать, что ряд Тейлора всегда “сходится” к функции, по которой он построен.

Контрпример. Функция $f(x)=e^{-x^{-2}},(f(0)=0)$ бесконечно дифференцируема, в том числе в нуле, где все её производные равны нулю. То есть ряд Тейлора — нулевой ряд. А функция нулю равна только в нуле.

Замечание. Формулы и ряд Тейлора чётной функции содержат только слагаемые с чётными степенями, формулы и ряд Тейлора нечётной функции содержат только слагаемые с нечётными степенями.

Формулы Тейлора, которые надо знать на память.

$$
\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + O(x^{n+1}),\\
e^{-x} &= 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \ldots + \frac{(-1)^n}{n!}x^n + O(x^{n+1}),\\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + O(x^{2n+3}),\\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + \frac{(-1)^n}{(2n)!}x^{2n} + O(x^{2n+2}),\\
\operatorname{tg} x &= x + \frac{x^3}{3} + O(x^5),\\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + \frac{(-1)^{n+1}}{n}x^n + O(x^{n+1}),\\
(1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \ldots + \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}x^n + O(x^{n+1}),\\
\operatorname{sh} x &= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + \frac{1}{(2n+1)!}x^{2n+1} + O(x^{2n+3}),\\
\operatorname{ch} x &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + \frac{1}{(2n)!}x^{2n} + O(x^{2n+2}).
\end{aligned}
$$

Лекция 17

Выпуклые функции

Определение. *Функция $f\colon (a,b) \to \mathbb{R}$ называется выпуклой вниз или просто выпуклой, если для любых точек $x, y \in (a,b)$ и любых $\alpha, \beta \geqslant 0, \alpha + \beta = 1$ справедливо неравенство*

$$f(\alpha x + \beta y) \leqslant \alpha f(x) + \beta f(y).$$

Геометрически условие выпуклости означает, что график функции лежит под любой хордой.

Определение. *Функция $f\colon (a,b) \to \mathbb{R}$ называется выпуклой вверх или просто вогнутой, если для любых точек $x, y \in (a,b)$ и любых $\alpha, \beta \geqslant 0, \alpha + \beta = 1$ справедливо неравенство*

$$f(\alpha x + \beta y) \geqslant \alpha f(x) + \beta f(y).$$

Геометрически условие выпуклости означает, что график функции лежит над любой хордой.

Примеры: функции x^2, e^x выпуклые, $\ln x$ — вогнутая, $1/x$ выпуклая при $x > 0$ и вогнутая при $x < 0$.

Замечание. Говорят также, что функция выпуклая, если её *надграфик — выпуклое множество*, т.е. любые две точки можно этого множества можно соединить отрезком, целиком лежащим в этом множестве . Функция вогнутая, если её *подграфик* — выпуклое множество. Надграфик — это часть плоскости, расположенная над графиком функции f. Аналитически он задаётся так: $epi\, f := \{(x,y) \mid y \geqslant f(x)\}$. Подграфик — это часть плоскости, расположенная под графиком функции. Аналитически он задаётся так: $hyp\, f := \{(x,y) \mid y \leqslant f(x)\}$.

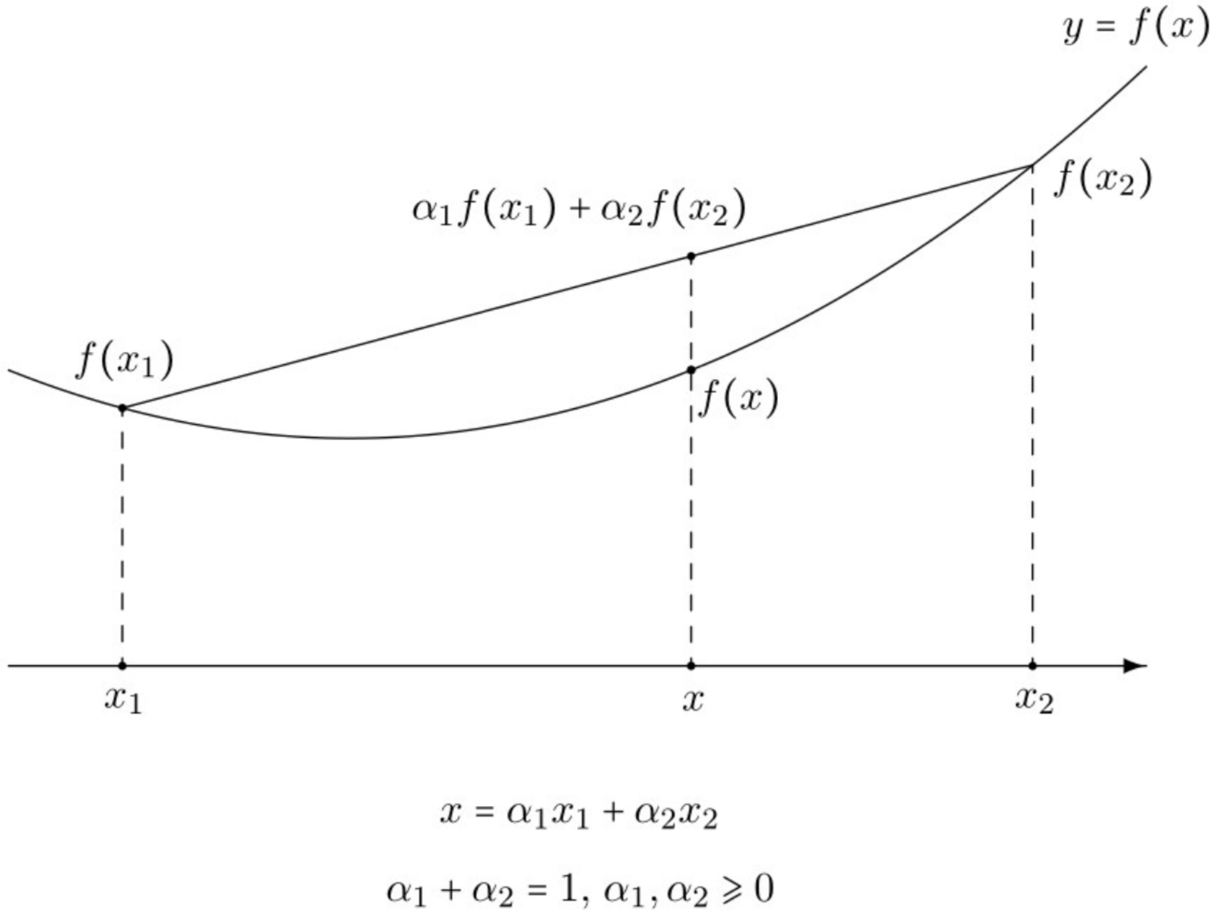


Рис. 7. Выпуклая функция

Теорема. *Выпуклая (вогнутая) функция непрерывна.*

Доказательство. Выберем точку x_0, в которой будем доказывать непрерывность. Выберем ещё 2 точки, $x_1 < x_0$ и $x_2 > x_0$. Проведем хорду через точки $(x_1, f(x_1))$ и $(x_0, f(x_0))$, а также хорду через точки $(x_2, f(x_2))$ и $(x_0, f(x_0))$ и продлим эти хорды на весь промежуток $[x_1, x_2]$. По определению выпуклости график функции находится между пересекающимися в точке $(x_0, f(x_0))$ продолженными отрезками. В самом деле, на $[x_0, x_2]$ график f ниже хорды непосредственно по определению. А если бы этот график хоть в одной точке x_* был бы ниже продолжения другой хорды, то хорда, соединяющая точки $(x_1, f(x_1))$ и $(x_*, f(x_*))$ в точке x_0 лежала бы ниже графика функции f (см. рис.8), что противоречило бы определению выпуклости. Непрерывность в точке x_0 теперь следует из теоремы о двух милиционерах. $\square$

Лемма. *Пусть f выпуклая на (a, b) функция и $a < x_1 < x < x_2 < b$. Тогда*

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

Доказательство. Из соотношений $x = \alpha_1 x_1 + \alpha_2 x_2$ и $\alpha_1 + \alpha_2 = 1$ следуют равенства

$$\alpha_1 = \frac{x_2 - x}{x_2 - x_1}, \quad \alpha_2 = \frac{x - x_1}{x_2 - x_1}.$$

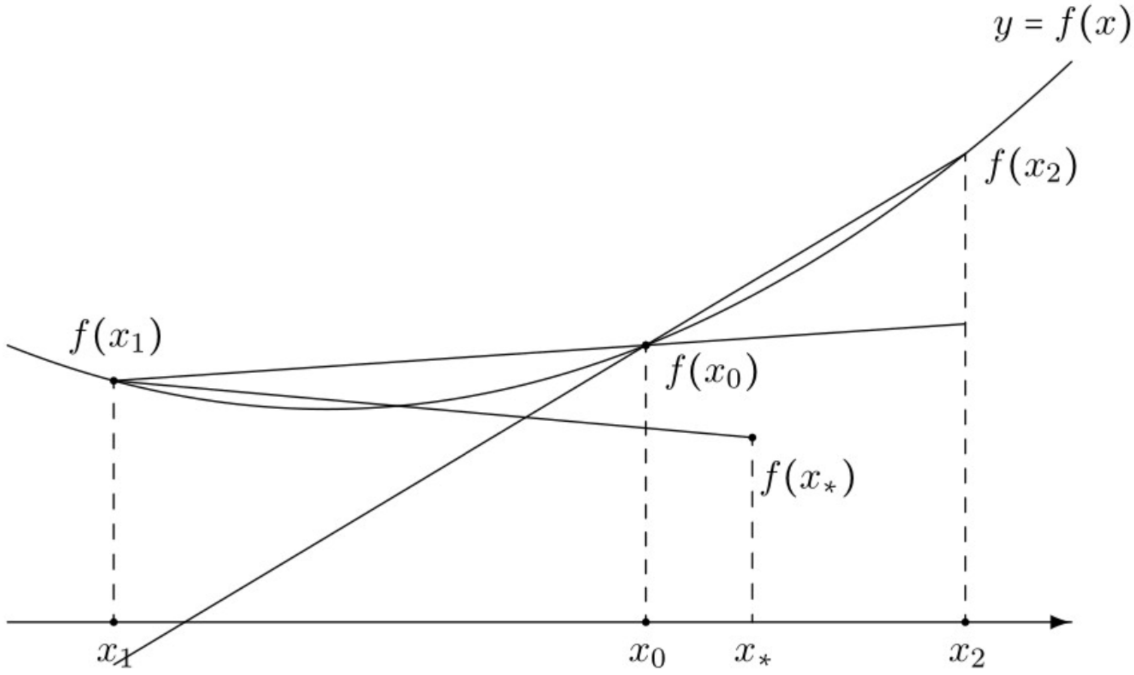


Рис. 8. Непрерывность выпуклой функции

Тогда основное условие выпуклости $f(\alpha_1 x_1 + \alpha_2 x_2) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2)$ можно переписать в виде

$$f(x) \leqslant \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2).$$

Учитывая $x_1 \leqslant x \leqslant x_2$ и $x_1 < x_2$ получаем

$$(x_2 - x) f(x_1) + (x_1 - x_2) f(x) + (x - x_1) f(x_2) \geqslant 0.$$

Каждое из неравенств искомого двойного неравенства ему эквивалентно! В частности,

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

Следствие (критерий выпуклости 1). *Дифференцируемая функция выпукла, если и только если её производная возрастает (не убывает).*

Доказательство. Перейдём сначала к пределу $x \to x_1$, а затем к пределу $x \to x_2$ (по отдельности) в двойном неравенстве

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

Получим

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

т.е. $f'(x_1) \leqslant f'(x_2)$. Обратно. Пусть f' не убывает. По теореме Лагранжа для всех точек $x < z < y$ найдутся $\xi_1 \in (x, z)$ и $\xi_2 \in (z, y)$ такие, что

$$\frac{f(z) - f(x)}{z - x} = f'(\xi_1), \quad \frac{f(y) - f(z)}{y - z} = f'(\xi_2).$$

Тогда т.к. f' не убывает, то

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

Раскрывая пропорцию и упрощая выражение, получим

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

Заметим, что

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

Тогда вводя обозначение $\alpha = \frac{y-z}{y-x}$, получаем $\frac{z-x}{y-x} = 1-\alpha$ и $z = \alpha x + (1-\alpha)y$ и предыдущее неравенство даёт

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y),$$

т.е. f выпукла по определению. $\square$

Замечание Производная дифференцируемой вогнутой функции f убывает: берем выпуклую функцию $-f$, её производная $-f'$ возрастает, значит, производная f' убывает.

Пусть функция f дважды дифференцируема на интервале (a,b). Всякая дифференцируемая функция не убывает, если и только если её производная неотрицательна (следствие из теоремы Лагранжа). Поэтому справедлив

Критерий выпуклости 2. *Дважды дифференцируемая функция выпукла, если и только если её вторая производная неотрицательная.*

Именно этот критерий является основным при практическом построении графиков. Есть незначительное различие между словами "возрастает" и "не убывает". Можно говорить о строго выпуклых (вогнутых) функциях, для которых все неравенства строгие, в наших построениях в этом нет нужны.

Пример. Функция $f(x) = x^2 + \sin x$ выпуклая на $\mathbb{R}$.

На семинарах мы докажем следующие утверждения:

Утверждение. *Пусть функция f выпукла на некотором интервале (a,b). Тогда она удовлетворяет на любом меньшем отрезке $\Delta = [c,d] \subset (a,b)$ условию Липшица.*

Неравенство Йенсена. *Пусть функция $f\colon (a,b) \to \mathbb{R}$ выпуклая, $\alpha_1, \ldots, \alpha_n \geqslant 0, \sum\limits_{k=1}^{n} \alpha_k = 1$, и $x_k \in (a,b)$. Тогда*

$$f(\alpha_1 x_1 + \ldots + \alpha_n x_n) \leqslant \alpha_1 f(x_1) + \ldots + \alpha_n f(x_n).$$

Достаточное условие локального экстремума

Из теоремы Ферма следует необходимое условие $f'(x_0) = 0$ локального экстремума дифференцируемой функции. Точки, в которых выполняется достаточное условие экстремума, называются *критическими точками.* Перейдём к достаточным условиям.

Пусть функция f дифференцируема на интервале (a,b) и пусть x_0 — критическая точка функции f, то есть $f'(x_0) = 0$.

Теорема. *Пусть функция f имеет в окрестности x_0 производные до порядка $n > 1$ включительно. Пусть $f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0$ и $f^{(n)}(x_0) \neq 0$. Тогда при нечётном n точка x_0 не является точкой локального экстремума. При чётном n точка x_0 является точкой* *локального экстремума, если* $f^{(n)}(x_0) > 0$ *— локального минимума, если* $f^{(n)}(x_0) < 0$ *— локального максимума.*

Доказательство. Используем формулу Тейлора с остаточным членом в форме Пеано:

$$f(x) = f(x_0) + \frac{1}{1!}f'(x_0)(x-x_0) + \frac{1}{2!}f''(x_0)(x-x_0)^2 + \ldots + \frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n + o((x-x_0)^n).$$

По предположению$f'(x_0) = \ldots = f^{(n-1)}(x_0) = 0$, таким образом $f(x) = f(x_0) + \frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n + o((x-x_0)^n)$. Пусть n нечётное, тогда при малых $|x-x_0|$ знак величины $f(x)-f(x_0)$ совпадает со знаком величины $f^{(n)}(x_0)(x-x_0)^n$, совпадающим со знаком $f^{(n)}(x_0)(x-x_0)$. Таким образом, сколь угодно близко к x_0 есть как точки, в которых $f(x) > f(x_0)$, так и точки, в которых $f(x) < f(x_0)$. Если n чётное, то знак величины $f(x)-f(x_0)$ совпадает со знаком величины $f^{(n)}(x_0)$. Поэтому в проколотой окрестности точки x_0 справедливы соотношения $f(x) > f(x_0)$ при $f^{(n)}(x_0) > 0$ и $f(x) < f(x_0)$ при $f^{(n)}(x_0) < 0$. $\square$

Замечание. Основной случай при исследовании критической точки — случай $n = 2$. Критическая точка x_0 является точкой локального минимума, если $f''(x_0) > 0$ и точкой локального максимума, если $f''(x_0) < 0$.

Лекция 18

Непрерывность функции многих переменных

В этом разделе мы рассмотрим некоторые необходимые сведения о функциях многих переменных. Предполагается известным определение функции многих переменных: формально, имеем множество E в $\mathbb{R}^n$ и отображение $f\colon E \to \mathbb{R}$. Будем писать $f(x)$, полагая, что в $\mathbb{R}^n$ выбран базис и координаты вектора x — числа $x_i, i = \overline{1,n}$, т.е. $x = (x_1, \ldots, x_n)$. Также предполагается известным, как мерить обычное расстояние в $\mathbb{R}^n$ *стандартным образом*: будем обозначать расстояние на плоскости между точками x и y через $|x-y|$, как и на прямой. Однако, это просто обозначение. Формула же такова

$$|x-y| = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2},\text{ где } x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n).$$

Эта функция расстояния называется *евклидовой метрикой* на $\mathbb{R}^n$.

Определение. *Множество $U \subset \mathbb{R}^n$ с евклидовой метрикой называется открытым, если для любой точки $a = (a_1, \ldots, a_n) \in U$ найдется шар*

$$B_\varepsilon(a) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n\colon\ |(x_1, \ldots, x_n) - (a_1, \ldots, a_n)| < \varepsilon\} \subset U$$

достаточно малого радиуса $\varepsilon > 0$ с центром в точке a.

Замечание. $B_\varepsilon(a)$ — это *открытый* шар, так как неравенство строгое.

Пример. Всякий открытый шар в $\mathbb{R}^n$ с евклидовой метрикой — действительно открытое множество.

Определение. *Окрестностью $U(a)$ точки $a = (a_1, \ldots, a_n) \in E$ называется любое открытое множество в $\mathbb{R}^n$ с евклидовой метрикой, содержащее точку a. Проколотая окрестность точки a — окрестность без a.*

Определение. *Рассмотрим последовательность точек $(x_k)_k$, $x_k(x_1^k, \ldots, x_n^k) \in E$ (верхний индекс k — это не степень, а номер последовательности!). Такая последовательность называется сходящейся к точке $a = (a_1, \ldots, a_n) \in E$, если вне любой окрестности точки a лежит лишь конечное количество точек последовательности. Формально говоря:*

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N} : \forall k > N\ |x_k - a| < \varepsilon.$$

Обозначение. $x_k \to a$ при $k \to \infty$.

Теорема. $(x_1^k, \ldots, x_n^k) \to (x_1, \ldots, x_n)$ при $k \to \infty$, если и только если для любого $i = 1, \ldots, n$ верно, что $x_i^k \to x_i$ при $k \to \infty$.

Доказательство. Пусть $(x_1^k, \ldots, x_n^k) \to (x_1, \ldots, x_n)$. Сходимости $x_i^k \to x_i$ при $k \to \infty$ следуют из соотношений

$$|x_i - x_i^k| \leqslant |(x_1, \ldots, x_n) - (x_1^k, \ldots, x_n^k)|.$$

Обратно: пусть для любого $i = 1, \ldots, n$ верно, что $x_i^k \to x_i$ при $k \to \infty$. Зафиксируем $\varepsilon > 0$. По определению, существует $N(\varepsilon)$ такое, что при $k > N(\varepsilon)$ справедливо $|x_i - x_i^k| < \varepsilon/\sqrt{n}$. Следовательно,

$$|(x_1, \ldots, x_n) - (x_1^k, \ldots, x_n^k)| = \sqrt{\sum_{i=1}^{n}(x_i - x_i^k)^2} < \varepsilon.$$

По определению $(x_1^k, \ldots, x_n^k) \to (x_1, \ldots, x_n)$. $\square$

Определение. *Точка $x_* \in \mathbb{R}^n$ называется предельной точкой множества E, если в любой окрестности точки x_* найдется $x \in E : x \neq x_*$.*

Эквивалентно можно сказать, что точка $x_* \in \mathbb{R}^n$ называется предельной точкой множества E, если в любой окрестности точки x_* найдется бесконечное множество точек из E.

Определение. *Множество называется замкнутым, если оно содержит все свои предельные точки.*

Пример. Замкнутый шар $\overline{B}_\varepsilon(a) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n\colon\ |(x_1, \ldots, x_n) - (a_1, \ldots, a_n)| \leqslant r\}$ с центром в a и радиуса r — замкнутое множество.

Определение. *Точка называется граничной для множества E, если в любой её окрестности есть и точки из E, и точки не из E. Множество всех граничных точек называется границей и обозначается ∂E.*

Точно так же, как и на прямой, справедливы следующие свойства открытых и замкнутых множеств: открытое минус замкнутое — открытое, замкнутое минус открытое — замкнутое, объединение любого набора открытых множеств открыто; пересечение конечного набора открытых множеств открыто; объединение конечного набора замкнутых множеств замкнуто; пересечение любого набора замкнутых множеств замкнуто. Единственное, что на прямой устроено по-другому — это открытые множества. На прямой каждое открытое множество — это объединение не более чем счётного набора непересекающихся промежутков. Уже на плоскости открытые множества могут быть устроены очень сложно. То есть, представляем мы себе открытое множество как нечто с кусочно-гладкой границей, без этой самой границы. Но это просто потому, что нам сложно представить себе что-то более сложное.

Упражнение. Придумайте пример трёх открытых множеств на плоскости, имеющих общую границу. Это сложная конструкция!

Определение. *Множество E называется компактным, если из любого его покрытия открытыми множествами можно выбрать конечное открытое подпокрытие. Множество E называется секвенциально компактным, если из любой последовательности $(x_k)_k \subset E$ можно извлечь сходящуюся к элементу E подпоследовательность.*

Остаётся верной следующая теорема о компактных множества.

Теорема. *Множество на плоскости секвенциально компактно $\Leftrightarrow$ оно компактно $\Leftrightarrow$ оно ограничено и замкнуто.*

По существу доказательство этой теоремы не отличается от доказательства аналогичной теоремы про компактные подмножества $\mathbb{R}$.

Предел и непрерывность функции многих переменных

Определение. *Пусть a — предельная точка множества $E \subset \mathbb{R}^n$. Число A называется пределом функции $f\colon E \to \mathbb{R}$ при $x \to a$, обозначается*

$$A = \lim_{x \to a} f(x),$$

если

$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall x \in E,\ 0 < |x - a| < \delta \Rightarrow |f(x) - A| < \varepsilon.$$

Отличие в этом определении от определения предела на $\mathbb{R}$ в том, что теперь точки x и a принадлежат $\mathbb{R}$, т.е. у них есть n координат. Так же как и на прямой, можно использовать **определение предела по Гейне**: *для любой последовательности $(x_k)_k$, $x_k \to a$, $x_k \neq a$ справедливо $f(x_k) \to A$, если и только если $A = \lim\limits_{x \to a} f(x)$.*

Определив понятие предела функции многих переменных и предельной точки множества, перейдём к следующему определению, которое полностью аналогично определению непрерывной функции одной переменной.

Определение. *Функция $f\colon E \to \mathbb{R}$ называется непрерывной в точке $x_0 \in E$, если*

$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall x \in E \cap U_\delta(x_0) \Rightarrow f(x) \in U_\varepsilon(f(x_0)).$$

Иначе можно сказать, что $f\colon E \to \mathbb{R}$ непрерывна в точке $x_0 \in E$, если для любой окрестности $U(f(x_0))$ найдётся окрестность $U(x_0)$ такая, что $f(E \cap U(x_0)) \subset U(f(x_0))$.

Упражнение. Почему в предыдущем определении действительно можно так сказать?

Заметим, что всякая функция f непрерывна в изолированной точке области определения: существует окрестность $U(x_0)$ такая, что $f(E \cap U(x_0)) = f(x_0) \subset U(f(x_0))$. Так же как и в случае функции одной переменной, непрерывность функции f в точке $x_0 \in E$, где x_0 — предельная точка E, эквивалента $\lim\limits_{x \to x_0} f(x) = f(x_0)$ или более эффектно $\lim\limits_{x \to x_0} f(x) = f(\lim\limits_{x \to x_0} x)$.

Теорема. *Функция $f\colon E \to \mathbb{R}$ непрерывна в точке $x_0 \in E$, предельной для E, тогда и только тогда, когда $\lim\limits_{x \to x_0} f(x) = f(x_0)$.*

Доказательство этой теоремы ничем не отличается от доказательства аналогичной теоремы для функции одной переменной.

Основные свойства непрерывных функции двух переменных

1. Непрерывная функция на компакте ограничена.
2. Непрерывная функция на компакте достигает максимума и минимума.
3. Если $f(x_0) = A > 0$, то в некоторой окрестности точки (x_0) функция f принимает значения, больше $A/2$.

Мы здесь привели необходимые нам свойства непрерывных функций. Все формулировки и доказательства абсолютно аналогичны таким же утверждениям для функций из $\mathbb{R}$ в $\mathbb{R}$.

Норма

Определение. *Нормой на векторном пространстве V над $\mathbb{R}$ мы будем называть любую функцию $x \mapsto ||x||$, удовлетворяющую следующим свойствам:*
1. $||x|| \geqslant 0, ||x|| = 0 \Leftrightarrow x = 0_V$ (0_V — ноль пространства V);
2. $\forall \alpha \in \mathbb{R}$ справедливо $||\alpha x|| = |\alpha| \cdot ||x||$;
3. $||x + y|| \leqslant ||x|| + ||y||$.

Выражение $||x - y||$ задаёт в пространстве V *метрику*. В частности, стандартная *евклидова* метрика на $\mathbb{R}^n$ задаётся нормой $|(x_1, \ldots, x_n)| = \sqrt{\sum_{i=1}^n x_i^2}$.

Определение. *Пусть есть вещественное векторное пространство V, в котором заданы две нормы $|| \cdot ||_1$ и $|| \cdot ||_2$. Эти нормы называются эквивалентными, если существуют числа $C > c > 0$, такие что для любого $x \in V$ справедливы неравенства*

$$c||x||_1 \leqslant ||x||_2 \leqslant C||x||_1.$$

Соотношение эквивалентности норм транзитивно: если норма $||\cdot||_1$ эквивалентна норме $||\cdot||_2$ и норма $|| \cdot ||_1$ эквивалентна норме $|| \cdot ||_3$, то норма$|| \cdot ||_2$ эквивалентна норме $|| \cdot ||_3$.

На семинарах мы покажем, что если последовательность $(x_n)_n \subset E$ сходится к $x \in E$ по какой-то одной норме, то она сходится и по эквивалентной норме. Более того, выясним, что верна следующая теорема

Теорема. *Все нормы в конечномерном пространстве эквивалентны.*

Дифференцируемость функций многих переменных

Пусть дана функция многих переменных $f\colon \mathbb{R}^n \to \mathbb{R}$.

Определение. *Частной производной* $f'_{x_k} = \partial f/\partial x_k$ *функции* f *по координате* x_k *называется предел*

$$\lim_{h\to 0} \frac{f(x_1,\ldots,x_k+h,\ldots,x_n) - f(x_1,\ldots,x_k,\ldots,x_n)}{h},$$

если он существует и конечен.

Пример. Рассмотрим $f(x,y) = \sin(xy^2)$. Тогда $\frac{\partial f}{\partial x} = y^2\cos(xy^2)$, $\frac{\partial f}{\partial y} = 2yx\cos(xy^2)$.

Будем писать, что $\alpha(h) = o(h)$ или $\alpha(h) = o(||h||)$, где $h \in \mathbb{R}^n$ если $\lim\limits_{h\to 0}\alpha(h)\cdot||h||^{-1} = 0$ (*предел по норме* пространства $\mathbb{R}^n$). Как мы знаем, в конечномерном пространстве все нормы эквивалентны, поэтому o-малое по разным нормам одно и то же.

Пусть дана функция многих переменных $f\colon \mathbb{R}^n \to \mathbb{R}$.

Определение дифференцируемости функции многих переменных. *Функция* f *называется дифференцируемой в точке* x*, если существует линейное отображение* $A\colon \mathbb{R}^n \to \mathbb{R}$*, такое что* $f(x+h) - f(x) - Ah = o(||h||)$*, то есть*

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

где $0 = (0,\ldots,0) \in \mathbb{R}^n$.

Замечание. Так как в конечномерном пространстве все нормы эквивалентны, поэтому определение дифференцируемости не зависит от нормы.

Линейное отображение A называют *производной функции* f. Обозначают $f'(x)$.

Так же как и в случае одной переменной определяется понятие дифференциала функции.

Определение. *Линейная функция* $h \mapsto f'(x_0)h$ *называется дифференциалом функции* f *в точке* x_0.

Обозначения. $df(x_0), d_{x_0}f, Df(x_0), D_{x_0}f$.

Где определен дифференциал $d_{x_0}f$? Линейное отображение — дифференциал $d_{x_0}f$ в точке x_0 — действует в пространстве приращений к точке x_0. И значения это линейное отображение принимает в множестве приращений к функции f, причём начало координат переходит в начало координат. Эти приращения образуют важную структуру — *касательное пространство* $T_{x_0}\mathbb{R}^n$. Мы применяем линейный оператор дифференциал всегда к приращениям аргумента. *Никогда не к* x*, всегда к* h. Геометрически это означает, что вместо исходного пространства векторов x (“радиус-векторов торчащих из начала координат”) мы рассматриваем множество векторов h, торчащих из x_0. Его тоже можно рассматривать, как векторное пространство той же размерности. В $\mathbb{R}^n$ и в $T_{x_0}\mathbb{R}^n$ разные базисы.

Итак, линейное отображение $d_{x_0}f$ действует из касательного пространства $T_{x_0}\mathbb{R}^n$. А куда оно действует? В такое же касательное пространство приращений к $f(x)$, обозначаемое $T_{f(x_0)}\mathbb{R}$. Буква T от слова *tangent (касательная)*. Значение дифференциала $d_{x_0}f$ на векторе $T_{x_0}\mathbb{R}^n$ есть *число* $d_{x_0}fh \in T_{f(x_0)}\mathbb{R}$ — *главная линейная часть приращения отображения при малых* h.

Замечание. Не путать касательное пространство с касательной прямой, плоскостью, гиперплоскостью и т.п.! Касательная плоскость — это чисто геометрический объект.

Пусть в $\mathbb{R}^n$ есть структура *евклидова пространства*, то есть задано скалярное произведение $\langle\cdot,\cdot\rangle$ такое, что $\langle x,y\rangle = \sum_{i=1}^n x_i y_i$ в стандартном базисе в $\mathbb{R}^n$. Вспомним простой факт из линейной алгебры: всякое линейное отображение $L\colon \mathbb{R}^n \to \mathbb{R}$ (линейный функционал) может быть записано при некотором $l \in \mathbb{R}^n$ в виде $L(h) = \langle l,h\rangle$.

Теорема. *Если функция $f(x_1,\ldots,x_n)$ дифференцируема в точке x, то существуют все частные производные в точке x и $l_k = f'_{x_k}$.*

Доказательство. Пусть функция f дифференцируема в точке $x = (x_1,\ldots,x_m)$. Тогда найдется вектор $l = (l_1,\ldots,l_n)$ такой, что при $h = (h_1,\ldots,h_n)$ справедливо соотношение

$$\lim_{h\to 0}\frac{|f(x_1+h_1,\ldots,x_n+h_n) - f(x_1,\ldots,x_n) - \sum_{k=1}^n l_k h_k|}{||h||} = 0.$$

Положим в этом равенстве $h_k = 0$ при всех $k \neq j$ и $h_j = t$, получим

$$\lim_{t\to 0}\frac{|f(x_1,\ldots,x_j+t,\ldots,x_m) - f(x_1,\ldots,x_m) - l_j t|}{|t|} = 0.$$

Это означает, что у f есть частная производная по x_j, причём она равна l_j $\square$

Таким образом, мы определили вид вектора l, задающего дифференциал функции многих переменных:

$$l = (l_1,\cdots,l_l) = \left(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}\right) = \nabla f = grad\ f.$$

Этот вектор называется *градиентом функции f (в точке x).*

Иными словами, для всякой дифференцируемой в точке x функции f ее дифференциал имеет вид $df(x)h = \langle l,h\rangle$, равенство справедливо при всех $h \in T_x\mathbb{R}^n$. Вектор $grad\ f$ имеет компоненты $\partial f/\partial x_k$, он может быть символически записан в виде $\nabla f(x)$, где ∇ — это символ, обозначающий “вектор”:

$$\nabla = \left(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_m}\right).$$

Это не вектор в $\mathbb{R}^n$ — пока нужно это воспринимать просто как обозначение, символ. Более глубокий смысл этого обозначения вам откроется на старших курсах. Сейчас же можно подразумевать, что ∇ — это такой *дифференциальный оператор*, который определен на дифференцируемых функциях f (нужной размерности) и переводящей функцию в её градиент $grad\ f$.

Лекция 19

На прошлой лекции мы показали, функция f дифференцируемая в точке x, если $f(x+h)-f(x)-\langle\nabla f, h\rangle = o(||h||)$. Дифференциал функции — линейное отображение (функционал), градиент — вектор.

Производная функции по вектору и по направлению

Определение. *Пусть $f\colon \mathbb{R}^m \to \mathbb{R}$. Производная по направлению вектора $v \in \mathbb{R}^m$:*

$$D_v f(x) := \lim_{t\to 0}\frac{f(x+vt)-f(x)}{t} = \langle v, grad\ f\rangle = \langle v, \nabla f\rangle.$$

Эта формула следует из формулы–определения дифференцируемости: $f(x+h)-f(x) = \langle h, \nabla f\rangle + o(h)$, при $h = vt, t \in \mathbb{R}$ получается $f(x+vt)-f(x) = \langle v, \nabla f\rangle t + o(t)$.

Производная функции $f(x+th)$ переменной t — это будет как раз производная по вектору h.

Производная по направлению — производная по вектору $e = (\cos\alpha_1, \ldots, \cos\alpha_m)$ единичной длины:

$$D_e f(x) = \langle e, grad\ f\rangle = \langle e, \nabla f\rangle = \sum_{k=1}^{m}\frac{\partial f}{\partial x_k}\cos\alpha_k;$$

$\cos\alpha_k$ — “направляющие косинусы”, углы α_k — это углы между e и базисными векторами e_k.

Частная производная — производная по направлению координатного вектора

Характеристическое свойство градиента. *Направление градиента — направление наибольшей скорости возрастания функции.*

Замечание. Предполагая, что градиент имеет направление, мы сразу допускаем, что он ненулевой. Это означает, что из всех производных $D_e f$ максимальное значение (при разных e) достигается на направлении e, совпадающем с направлением градиента.

Доказательство. Скалярное произведение равно произведению длин векторов на косинус угла между ними:

$$D_e f = \langle e, \nabla f\rangle \leqslant |\nabla f|,$$

по неравенству Коши-Буняковского-Шварца. Как известно равенство в нём достигается тогда и только тогда, когда e коллинеарно ∇f. Если e и ∇f противонаправлены, то $\langle e, \nabla f\rangle$ отрицательно и функци f убывает в этом направлении.Если же e и ∇f сонаправлены, то $\langle e, \nabla f\rangle$ положительно и функци f возратсает в этом направлении. $\square$

Теорема (достаточное условие дифференцируемости). *Если функция $f(x_1, \ldots, x_n)$ имеет частные производные в окрестности точки x, и они непрерывны в x, то f дифференцируема в точке x.*

Доказательство. Пусть у функции $f(x_1, \ldots, x_n)$ в окрестности точки $(x_1, \ldots, x_n)$ есть все частные производные f'_x , причём они непрерывны в этой окрестности. Докажем, что тогда справедливо равенство

$$\lim_{h\to 0}\frac{|f(x_1+h_1, \ldots, x_n+h_n) - f(x_1, \ldots, x_n) - \sum_{k=1}^{n} f'_{x_k}(x_1, \ldots, x_n)h_k|}{||h||} = 0,$$

которое в точности означает дифференцируемость функции f. Для сокращения обозначений положим, что $n = 3$, желающие могут посмотреть полный вариант доказательства в Зориче. Используем теорему Лагранжа по разным переменным: $f(a,b,c) - f(a+\Delta, b, c) = f'_{x_1}(a+$

$\theta_1\Delta, b, c)\Delta$ некотором $\theta_1 \in (0,1)$, также $f(a,b,c) - f(a,b+\Delta,c) = f'_{x_2}(a, b+\theta_2\Delta, c)\Delta$ некотором $\theta_2 \in (0,1)$ и $f(a,b,c) - f(a,b,c+\Delta) = f'_{x_3}(a,b,c+\theta_3\Delta)\Delta$ некотором $\theta_3 \in (0,1)$. Тогда

$$f(x_1+h_1, x_2+h_2, x_3+h_3) - f(x_1,x_2,x_3) = f(x_1+h_1, x_2+h_2, x_3+h_3) - f(x_1+h_1, x_2+h_2, x_3)+$$
$$+f(x_1+h_1,x_2+h_2,x_3) - f(x_1+h_1,x_2,x_3) + f(x_1+h_1,x_2,x_3) - f(x_1,x_2,x_3) =$$
$$= \frac{\partial f}{\partial x_1}(x_1+h_1, x_2+h_2, x_3+\theta_1 h_3)h_3 + \frac{\partial f}{\partial x_2}(x_1+h_1, x_2+\theta_2 h_2, x_3)h_2 + \frac{\partial f}{\partial x_3}(x_1+\theta_3 h_1, x_2, x_3)h_1$$

Из непрерывности всех частных производных следует искомое утверждение: $f'_{x_i}(x_1+h_1, x_2+h_2, x_3+\theta_1 h_3) = f'_{x_i}(x_1,x_2,x_3) + o(1)$, $i = 1,2,3$ и $o(1)h_1 + o(1)h_2 + o(1)h_3 = o(h)$. $\square$

Линейность операции дифференцирования. *Пусть заданы две функции $f, g\colon \mathbb{R}^n \to \mathbb{R}$, пусть они определены в окрестности точки $x \in \mathbb{R}^n$ и пусть они обе дифференцируемы в точке x. Тогда функция $\xi = \alpha f + \beta g$ дифференцируема в точке x и $d\xi = \alpha df + \beta dg$.*

Доказательство. $\xi(x+h) - \xi(x) = \alpha f(x+h) + \beta g(x+h) - (\alpha f(x) + \beta g(x)) = \ldots = (\alpha d_x f + \beta d_x g)h + o(h)$. $\square$

Градиент произведения и частного

Верны формулы

$$\nabla(fg) = g\nabla f + f\nabla g, \quad \nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f\nabla g}{g^2}.$$

Единственное условие — дифференцируемость функций f и g и ненулевое значение g для частного. Для доказательства надо сказать, что формулы

$$(fg)'_{x_k} = f'_{x_k}g + fg'_{x_k}, \quad \left(\frac{f}{g}\right)'_{x_k} = \frac{f'_{x_k}g - fg'_{x_k}}{g^2},$$

для частных производных — это обычные формулы дифференцирования функции одной переменной. А из этих формул и определения градиента следуют формулы для градиента.

Упражнение. Докажите это.

Дифференцируемость отображений $\mathbb{R}^m \to \mathbb{R}^n$

Определение дифференцируемости отображения $\mathbb{R}^m \to \mathbb{R}^n$. *Отображение $F : \mathbb{R}^m \to \mathbb{R}^n$ называется дифференцируемым в точке x, если существует линейное отображение $A\colon \mathbb{R}^m \to \mathbb{R}^n$, такое что $F(x+h) - F(x) - Ah = o(||h||)$, то есть*

$$\lim_{h\to 0} \frac{||F(x+h) - F(x) - Ah||}{||h||} = 0$$

или, что то же,

$$\forall \varepsilon > \ \exists \delta > 0 : \forall h \in \mathbb{R}^m, ||h|| \leqslant \delta \Rightarrow ||h||^{-1}||F(x+h) - F(x) - Ah|| \leqslant \varepsilon.$$

Здесь в числителе и знаменателе стоят нормы пространств $\mathbb{R}^n$ и $\mathbb{R}^m$ соответственно. В конечномерном пространстве все нормы эквивалентны, поэтому можно не говорить, какая именно норма использована, производная не зависит от нормы.

Линейное отображение $L(h) = Ah$ называют по разному (дифференциал, касательно отображение) и обозначают по-разному: $dF(x), DF(x), d_xF$. Значения производной — это векторы $dF(x)h, DF(x)h, h$ — это приращение аргумента, $F(x+h) - F(x)$ — приращение функции.

Мы применяем линейный оператор дифференциал всегда к приращениям аргумента. Никогда не к x, всегда к h. Геометрически это означает, что вместо исходного пространства векторов x ("торчащих из начала координат") мы рассматриваем множество векторов h, торчащих из x. Его тоже можно рассматривать, как линейное пространство той же размерности. В $\mathbb{R}^m$ и в $T_x\mathbb{R}^m$ разные базисы. Итак, линейное отображение $dF(x)$ действует из касательного пространства $T_x\mathbb{R}^m$. А куда оно действует? В такое же касательное пространство приращений к $F(x)$, обозначаемое $T_{F(x)}\mathbb{R}^n$. Буква T от слова *tangent* (касательная). Значение дифференциала $dF(x)$ на векторе $h \in T_x\mathbb{R}^m$ есть вектор $dF(x)h \in T_{F(x)}\mathbb{R}^n$ — главная линейная часть приращения отображения при малых h.

Не путаем касательное пространство с касательной прямой, плоскостью, гиперплоскостью! Касательная плоскость — это чисто геометрический объект.

Матрица Якоби и якобиан

Пусть в $\mathbb{R}^m$ и в $\mathbb{R}^n$ выбраны базисы. Пусть отображение $F\colon \mathbb{R}^m \to \mathbb{R}^n$ дифференцируемо в точке x. Это значит, что дифференцируемы n функций $F_k, k = 1, \dots, n$. Как мы знаем,

$$F_k(x+h) - F_k(x) = \langle \nabla F_k, h\rangle + o(h) = \sum_{j=1}^m \frac{\partial F_k}{\partial x_j} h_j + o(h).$$

Здесь k — номер функции, $h = (h_1, \dots, h_m)$. Таким образом,

$$\begin{pmatrix} F_1(x+h) - F_1(x) \\ \dots \\ F_n(x+h) - F_n(x) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m \frac{\partial F_1}{\partial x_j} h_j \\ \dots \\ \sum_{j=1}^m \frac{\partial F_n}{\partial x_j} h_j \end{pmatrix} + o(h) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_m} \\ \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_m} \end{pmatrix} \begin{pmatrix} h_1 \\ \dots \\ h_m \end{pmatrix} + o(h)$$

Матрицу

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_m} \\ \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_m} \end{pmatrix}$$

называют *матрицей Якоби. Якобианом* называется *определитель* матрицы Якоби, если матрица квадратная ($m = n$). **Матрица может быть и не квадратная, якобиан бывает только у квадратных матриц.**

Итак, производная — это линейный оператор из $\mathbb{R}^m$ в $\mathbb{R}^n$. Если в обоих этих пространствах выбраны базис, то каждый линейный оператор — это оператор умножения на $m \times n$ матрицу Якоби.

Из доказанной теоремы про достаточное условие дифференцируемости скалярной функции многих переменных следует, что если все $n \times m$ частных производных непрерывны в окрестности точки x, то отображение дифференцируемо в этой точке x.

Утверждение. *Отображение F дифференцируемо в точке x, если и только если дифференцируемы все n скалярнозначных функций F_k.*

Доказательство: Утверждение следует из определений:

$$\lim_{h\to 0} \frac{||F(x+h) - F(x) - Ah||}{||h||} = 0 \Leftrightarrow \lim_{h\to 0} \frac{|F_k(x+h) - F_k(x) - (Ah)_k|}{||h||} = 0 \ \forall k = \overline{1, n},$$

здесь $(Ah)_k$ — $k-$ая координата вектора Ah. Для доказательства эквивалентности нужно использовать специальную норму в числителе $||(x_1, \dots, x_n)|| = \max_k |x_k|$ и теорему о двух милиционерах. Проделайте это самостоятельно. $\square$

Лекция 20
Неявная функция

Одним из важных способов задания функций является уравнение. Рассматривается функция $F(x, y)$ двух вещественных переменных и уравнение $F(x, y) = 0$. Если из этого уравнения можно выразить y через x, то говорят, что это уравнение задаёт *неявную функцию*.

Мы будем изучать уравнение $F(x, y) = 0$ в окрестности точки (x_0, y_0), которая удовлетворяет этому уравнению, то есть $F(x_0, y_0) = 0$.

Примеры. 1) $F(x, y) = x^2+y^2-1 = 0,\ x_0 = 0, y_0 = -1$. Легко видеть, что при $x \in (-a, a), a \in (0, 1)$ уравнение $F(x, y) = 0$ задает функцию $y = -\sqrt{1 - x^2}$.

2) $F(x, y) = x^2 + y^2 = 0,\ x_0 = y_0 = 0$. Легко видеть, что уравнение $F(x, y) = 0$ не задает функцию $y = f(x)$.

Пусть даны $\delta, \sigma > 0$ и $F(x, y), x \in [x_0 - \delta, x_0 + \delta], y \in [y_0 - \sigma, y_0 + \sigma]$ — непрерывная функция, причем при каждом $x \in [x_0 - \delta, x_0 + \delta]$ уравнение $F(x, y) = 0$ имеет единственное решение $y = f(x) \in [y_0 - \sigma, y_0 + \sigma]$.

Теорема. *f — непрерывная функция.*

Доказательство. На отрезке $[x_0-\delta, x_0+\delta]$ нет изолированных точек. Тогда для доказательства непрерывности f в точке x_* выберем последовательность $(x_n)_n, x_n \to x_*$. Покажем, что существует предел $f(x_n)$ при $n \to \infty$ (он обязательно конечен так как функция f ограничена).

Если это так при любой последовательности $x_n \to x_*$, то, в частности, и для последовательности $x_1, x_*, x_2, x_*, x_3, \ldots$, а уж если последовательность $f(x_1), f(x_*), f(x_2), f(x_*), f(x_3), \ldots$ сходится, то и $f(x_n) \to f(x_*)$, отсюда следует непрерывность f в точке x_*.

Если это не так (не существует предела $f(x_n)$), то в $(x_n)_n$ найдутся подпоследовательности $x'_k = x_{n_k}$ и $x''_k = x_{m_k}$ такие, что пределы $y' = \lim\limits_{k\to\infty} f(x'_k)$ и $y'' = \lim\limits_{k\to\infty} f(x''_k)$ существуют (ввиду непрерывности F) и различны: $y' \neq y''$. Теперь из равенств $F(x'_k, f(x'_k)) = F(x''_k, f(x''_k)) = 0$ из непрерывности F следуют равенства $F(x_*, y') = F(x_*, y'') = 0$, а это противоречит единственности решения уравнения $F(x, y) = 0$ при $x = x_*$. $\square$

Утверждение 2. *Пусть $F(x, y)$ строго монотонна по переменной y при каждом $x \in [x_0 - \delta, x_0 + \delta]$. Пусть $F(x, y_0 + \sigma) > 0, F(x, y_0 - \sigma) < 0$ при каждом $x \in [x_0 - \delta, x_0 + \delta]$. Тогда уравнение $F(x, y) = 0$ имеет единственное решение $y = f(x)$ при каждом $x \in [x_0 - \delta, x_0 + \delta]$ и f — непрерывная функция.*

Это утверждение следует из теоремы о промежуточном значении и Утверждения 1.

Теорема о неявной функции. *Пусть $F(x_0, y_0) = 0$, функция F непрерывна в окрестности U точки (x_0, y_0), пусть определена частная производная $F'_y(x, y)$, которая также непрерывна в этой окрестности. Пусть $F'_y(x_0, y_0) \neq 0$. Тогда найдутся δ и σ такие, что на прямоугольнике $x \in [x_0 - \delta, x_0 + \delta], y \in [y_0 - \sigma, y_0 + \sigma]$ определена непрерывная неявная функция $y = f(x)$.*

В рассмотренных чуть выше примерах условие $F'_y = 2y$. В первом примере $F'_y \neq 0$ и неявная функция определена, во втором примере основное условие не выполнено и нет никакой неявной функции.

Доказательство. Доказательство сводится к тому, что в условиях теоремы при некоторых δ и σ выполняются условия Утверждения 2. Докажем это для случая, когда $F'_y(x_0, y_0) > 0$. В этом случае в некоторой окрестности $U' \subset U$ точки (x_0, y_0) частная производная F'_y (при фиксированном x_0) также положительна. Это следует из непрерывности функции F'_y и из свойства непрерывных функций: если непрерывная функция в некоторой точке положительна, то она положительна и в некоторой окрестности этой точки. Поэтому при каждом x, достаточно близком к x_0, функция $y \mapsto F(x, y)$ строго монотонно возрастает. В частности, $F'_y(x_0, y) > 0$ и функция $y \mapsto F(x_0, y)$ строго монотонно возрастает. Следовательно, при малом $\sigma > 0$ выполнены неравенства $F(x_0, y_0 + \sigma) > 0, F(x_0, y_0 - \sigma) < 0$. Поэтому найдется такое $\delta > 0$, что при $x \in [x_0 - \delta, x_0 + \delta]$ также $F(x, y_0 + \sigma) > 0, F(x, y_0 - \sigma) < 0$ (ввиду непрерывности F). □

Эту теорему легко обобщить для случая, когда уравнение имеет вид $F(x_1, \ldots, x_n, y) = 0$: *Если F и F'_y — непрерывные функции, $F(x_1^0, ..., x_n^0, y^0) = 0$ и $F'_y(x_1^0, ..., x_n^0, y^0) \neq 0$, то в окрестности точки $(x_1^0, ..., x_n^0, y^0)$ решения уравнения $F = 0$ описываются неявной функцией $y = f(x_1, \ldots, x_n)$.* Доказательство совершенно аналогично.

Условия дифференцируемости неявной функции

Пусть в некотором *выпуклом* множестве D значений x, y существует функция $F'_x(x, y)$. Рассмотрим разность $F(x + h, y) - F(x, y)$. По теореме Лагранжа при при всех x, y, h найдется точка $\theta \in (0, 1)$, при которой верно равенство $F(x + h, y) - F(x, y) = F'_x(x + \theta h, y)h$. Чтобы это доказать достаточно рассмотреть функцию $F(x + th, y)$ переменной t, ее производная по t равна $F'_x(x + th, y)h$, применяем теорему Лагранжа на промежутке $t \in [0, 1]$, получаем требуемое. Выпуклость множества D требовалась для того, чтобы отрезок, соединяющий точки (x, y) и $(x + h, y)$ целиком лежал в D.

Аналогично $F(x, a) - F(x, b) = F'_y(x, \theta a + (1 - \theta)b)(a - b)$.

Пусть выполнены условия теоремы о неявной функции, неявная функция $y = f(x)$ определена в некоторой окрестности точки (x_0, y_0), причём, дополнительно, пусть определена и непрерывна функция $F'_x(x, y)$.

Тогда из тождества $F(x, f(x)) \equiv 0$ вытекает, что

$$0 = F(x+h, f(x+h)) - F(x, f(x)) = F(x+h, f(x+h)) - F(x+h, f(x)) + F(x+h, f(x)) - F(x, f(x)).$$

По теореме Лагранжа при некоторых $\theta_1, \theta_2 \in (0, 1)$ верны равенства

$$F(x + h, f(x + h)) - F(x + h, f(x)) = F'_y(x + h, \theta_1 f(x + h) + (1 - \theta_1)f(x))(f(x + h) - f(x)),$$

$$F(x + h, f(x)) - F(x, f(x)) = F'_x(x + \theta_2 h, f(x))h,$$

поэтому

$$\frac{f(x + h) - f(x)}{h} = -\frac{F'_x(x + \theta_2 h, f(x))}{F'_y(x + h, \theta_1(x + h) + (1 - \theta_1)f(x))}.$$

Делить можно так как $F'_y(x_0, y_0) \neq 0$ по предположению, поэтому (в силу непрерывности F'_y) при (x, y) близких к (x_0, y_0) также $F'_y(x, y) \neq 0$.

Теперь осталось перейти в полученном равенстве двух дробей к пределу при $h \to 0$. Справа предел существует по непрерывности функций F'_x и F'_y, значит, он существует слева. Таким образом,

$$f'(x) = -\frac{F'_x(x, f(x))}{F'_y(x, f(x))}.$$

Обычно все это формулируют как **вторую половину теоремы о неявной функции:** *если в окрестности точки (x_0, y_0) определена и непрерывна также частная производная F'_x, то неявная функция $f(x)$ дифференцируема и*

$$f'(x) = -\frac{F'_x(x, f(x))}{F'_y(x, f(x))}.$$

Производная композиции функции двух переменных и двух функций

Пусть задана “очень сложная” функция: $f(t) = F(x(t), y(t))$. Она является композицией функции F и функций $x(t)$ и $y(t)$. Или композицией функции F и *вектор-функции* $t \mapsto (x, y)$. Пусть существуют производные x' и y', пусть существуют частные производные F'_x, F'_y, причём последние производные непрерывные.

Утверждение. *Тогда функция f дифференцируема и*

$$f'(t) = F'_x(x(t), y(t))x'(t) + F'_y(x(t), y(t))y'(t).$$

Доказательство. В самом деле

$$F(x(t+h), y(t+h)) - F(x(t), y(t)) = F(x(t+h), y(t+h)) - F(x(t), y(t+h)) + F(x(t), y(t+h)) - F(x(t), y(t)).$$

К тому же

$$F(x(t+h), y(t+h)) - F(x(t), y(t+h)) = F'_x(\theta x(t+h) + (1-\theta)x(t), y(t+h))(x(t+h) - x(t))$$

и

$$F(x(t), y(t+h)) - F(x(t), y(t)) = F'_y(x(t), \theta_1 y(t+h) + (1-\theta_1)y(t))(y(t+h) - y(t)).$$

Тогда

$$\frac{F(x(t+h), y(t+h)) - F(x(t), y(t))}{h} = F'_x(\ldots, y(t+h))\frac{x(t+h) - x(t)}{h} + F'_y(x(t), \ldots)\frac{y(t+h) - y(t)}{h}.$$

При $h \to 0$, очевидно, $\theta x(t+h) + (1-\theta)x(t) \to x(t), \theta_1 y(t+h) + (1-\theta_1)y(t) \to y(t)$. Тогда переходя к пределу, получаем нужную формулу. $\square$

Пример. Как дифференцировать $x(t)^{y(t)}$. Сначала преобразуем это выражение в $e^{y(t)\ln(x(t))}$ и дальше используем теорему о производной сложной функции. Теперь мы видим, что нужно положить $F(x, y) = x^y$, тогда $F'_x = yx^{y-1}, F'_y = x^y \ln x$, поэтому

$$(x(t)^{y(t)})' = y(t)x(t)^{y(t)-1}x'(t) + x(t)^{y(t)}\ln(x(t))y(t).$$

АППЕНДИКС

Существование и единственность вещественных чисел (набросок)

Модель вещественных чисел, которую мы сейчас рассмотрим, использует упорядоченное поле рациональных чисел. Мы уже видели, что рациональные числа образуют упорядоченное поле. Также они удовлетворяют Аксиоме Архимеда.

Последовательность $(x_n)_n \in \mathbb{Q}$ называется фундаментальной, если

$$\forall \varepsilon \in \mathbb{Q}\, \varepsilon > 0\ \exists N \in \mathbb{N} \colon \forall k, n > N \Rightarrow |x_k - x_m| < \varepsilon$$

Это совершенно такое же определение, как и было раньше, для вещественных последовательностей. Только теперь ε — рациональное, а не любое вещественное.

Мы будет говорит, что $x_n \to 0$ при $n \to \infty$, если

$$\forall \varepsilon \in \mathbb{Q}\, \varepsilon > 0\ \exists N \in \mathbb{N} \colon \forall k > N \Rightarrow |x_k| < \varepsilon$$

Фундаментальные последовательности $(x_n)_n$ и $(y_n)_n$ назовём эквивалентными, если $(x_n - y_n)_n \to 0$, обозначаем $(x_n)_n \sim (y_n)_n$. Это действительно отношение эквивалентности.

Разобьём множество фундаментальных последовательностей на классы эквивалентности, в каждый класс отнесём все эквивалентные между собой фундаментальные последовательности. Обозначать классы будем так: $[(x_n)_n]$ — класс, содержащий фундаментальную последовательность $(x_n)_n$ и все другие эквивалентные $(x_n)_n$ последовательности. Покажем, что эти сложные объекты — классы эквивалентных фундаментальных последовательностей рациональных чисел — можно дополнить операциями и отношением порядка так, чтобы получилось $\mathbb{R}$. Любая сходящаяся в $\mathbb{Q}$ фундаментальная последовательность $(x_n)_n$ окажется в одном классе с постоянной последовательностью $y_n = \lim\limits_{n\to\infty} x_n$. Такие классы естественно соответствуют рациональным числам.

Пусть $[(x_n)_n]$ — класс эквивалентности, содержащий фундаментальную последовательность $(x_n)_n$, $[(y_n)_n]$ — класс эквивалентности, содержащий фундаментальную последовательность $(y_n)_n$. Тогда по определению положим $[(x_n)_n] + [(y_n)_n]$ — класс эквивалентности, содержащий фундаментальную последовательность $(x_n + y_n)_n$, положим $[(x_n)_n] \cdot [(y_n)_n]$ — класс эквивалентности, содержащий фундаментальную последовательность $(x_n \cdot y_n)_n$. Ясно, что последовательности $(x_n \cdot y_n)_n$ и $(x_n + y_n)_n$ также фундаментальны (проверьте корректность такого определения!).

Это мы определили операции в множестве классов эквивалентности. Нейтральные элементы — это $(e_n)_n, e_n = 0$ для сложения и $(e_n)_n, e_n = 1$ для умножения. Обратный элемент по сложению $-[(x_n)_n]$ — это, очевидно, $[(-x_n)_n]$. Обратный элемент по умножению существует для всех классов, кроме $[(0)_n]$. Каждая фундаментальная последовательность $(x_n)_n$, не лежащая в классе $[(0)_n]$, начиная с некоторого N не обращается в 0, поэтому можно корректно определить для $[(x_n)_n] \neq [(0)_n]$ обратный по умножению элемент: $[(x_n)_n]^{-1} = [(x_{n+N}^{-1})_{n+N}]$. Справедливость всех остальных свойств поля (коммутативность и ассоциативность сложения и умножения и дистрибутивность этих операций) легко следует из определений и этих же свойств для $\mathbb{Q}$.

Порядок в множестве классов эквивалентности также просто вводится: $[(x_n)_n] \geqslant [(y_n)_n]$, если

$$\forall \varepsilon \in \mathbb{Q}\, \varepsilon > 0\ \exists N \in \mathbb{N} \colon \forall n > N \Rightarrow x_n + \varepsilon \geqslant y_n$$

Упражнение. Проверьте, что
1) из $[(x_n)_n] \geqslant [(y_n)_n]$ следует $[(x_n)_n] + [(z_n)_n] \geqslant [(y_n)_n] + [(z_n)_n]$;
2) из $[(x_n)_n], [(y_n)_n] \geqslant [(0)_n]$ следует $[(x_n)_n] \cdot [(y_n)_n] \geqslant [(0)_n]$.

Таким образом, множество классов фундаментальных последовательностей уже стало упорядоченным полем. Осталось проверить, что на нём выполнена аксиома непрерывности и конструкция будет завершена.

Пусть даны два непустых множества A и B классов эквивалентности. Предполагается, что все классы из A меньше всех классов из множества B. Нужно доказать, что есть класс, который меньше или равен всех классов из B и больше или равен всех классов из A. Множество A не пусто, в нём есть хотя бы один класс $[(a_n)_n]$, где $(a_n)_n$ — фундаментальная последовательность рациональных чисел. Множество B не пусто, в нём есть хотя бы один класс $[(b_n)_n]$, справедливо отношение $[(a_n)_n] < [(b_n)_n]$. Теперь возьмём середину промежутка между классами $[(a_n)_n], [(b_n)_n]$, класс $[((a_n+b_n)/2)_n]$. То, что полусумма фундаментальных последовательностей — фундаментальная последовательность, очевидно. Справедлив один из трёх вариантов:

1) Этот класс — искомый, он больше всех классов из A и меньше всех классов из B.

2) На промежутке $[[(a_n)_n], [((a_n + b_n)/2)_n]]$ есть классы и из A, и из B.

3) На промежутке $[((a_n + b_n)/2)_n], [(b_n)_n]]$ есть классы и из A, и из B.

В первом случае утверждение доказано. Во втором и третьем случае это непростая проверка. Проделав эту проверку, получаем следующую теорему.

Теорема. *Множество классов эквивалентности фундаментальных последовательностей (с введёнными операциями и отношением эквивалентности) является упорядоченным полем с аксиомой непрерывности, то есть множеством вещественных чисел.*

Единственность множества вещественных чисел

Утверждение. *Пусть $\mathbb{R}_1$ и $\mathbb{R}_2$ — два упорядоченных поля, элементы которых удовлетворяют аксиоме непрерывности. Тогда существует биекция $F\colon \mathbb{R}_1 \to \mathbb{R}_2$, сохраняющая порядок и операции поля:*

$$\forall x, y \in \mathbb{R}_1\colon F(x + y) = F(x) + F(y),\ F(x \cdot y) = F(x) \cdot F(y), x \leqslant y \Rightarrow F(x) \leqslant F(y);$$

и такая, что F и F^{-1} — непрерывные отображения.

Доказательство. Для доказательства этого утверждения построим биекцию F в явном виде. Обозначим через 0_i и $1_i, (i = 1, 2)$ нейтральные элементы обоих полей. Положим $F(0_1) = 0_2$ и $F(1_1) = 1_2$.

Операции и соотношение порядка в полях $\mathbb{R}_1$ и $\mathbb{R}_2$ будем обозначать одинаково, хотя на самом деле это операции и соотношение в разных множествах.

Далее, положим $F(1_1 + 1_1) = 1_2 + 1_2$, и также со всеми натуральными числами. Аналогично определим биекцию на всех целых числах и на всех рациональных числах. Очевидно, что таким образом определённая биекция сохраняет все операции и порядок.

На остальных элементах множества $\mathbb{R}_1$, не представимых в виде m_1/n_1, где m_1, n_1 — целые числа (иррациональных числах), действуем следующим образом. Для каждого иррационального $\alpha \in \mathbb{R}_1$ рассмотрим множество $A \subset \mathbb{R}_1$ рациональных чисел $r \in R_1, r > \alpha$ и рассмотрим в $\mathbb{R}_2$ множество $F(A)$ чисел $F(r), r \in A$. Положим по определению $F(\alpha) = \inf F(A)$.

Теперь нужно проверить, что определённое таким образом отображение есть биекция, и что она удовлетворяет условиям утверждения.

То, что F — биекция на рациональных числах, мы уже отмечали. Про единственность точной нижней границы подмножества $\mathbb{R}$ нам известно. То, что у разных лучей, не содержащих вершину, разные точные нижние грани — очевидно. Поэтому F — биекция. Теперь надо аккуратно проверить, что F сохраняет операции и отношение порядка.

Упражнение. Проделайте это!

То, что биекция F — монотонная, вроде очевидно по построению. Непрерывность построенной биекции F и обратной биекции F^{-1} доказывается аналогично ранее доказанным теоремам о монотонных отображениях. $\square$

ОСНОВНАЯ ЛИТЕРАТУРА:

[1] В.А. Зорич. Математический анализ, том 1. МЦНМО, 2002.

[2] Г.М. Фихтенгольц. Курс дифференциального и интегрального исчисления, том 1. Издание 6, Наука, 1966.

[3] А.М. Красносельский. Конспект лектора. Математический анализ. 1 курс, 1 модуль, 2020.

[4] В.И. Богачёв. Математический анализ: конспект лекций. Мат. фак. ВШЭ, 1 курс, осень 2021.

[5] С.М. Львовский. Лекции по математическому анализу. МЦНМО, 2008.

[6] У. Рудин. Основы математического анализа. Мир, 1976.