3. Начала анализа

3.1. Последовательности и ряды.

Напомним: числовая последовательность — это функция на натуральном ряде: $\{x_1, x_2, \cdots, x_n, \cdots\}$ (короче: $\{x_n\}_{n\in\mathbb{N}}$, $x_n \in \mathbb{R}$ или $\mathbb{C}$). Иногда всю последовательность мы будем обозначать одной буквой и писать так: $x = (x_1, x_2, \cdots, x_n, \cdots)$ или так: $x = \{x_n\}_{n\in\mathbb{N}}$.

Совокупность всех вещественных последовательностей будем иногда обозначать $\mathbb{R}^{\mathbb{N}}$ (хотя логичнее было бы обозначать её $(\mathbb{R}^{\mathbb{N}})'$).

Элементы из $\mathbb{R}^{\mathbb{N}}$ можно складывать (покоординатно) :

$$x = \{x_n\}_{n\in\mathbb{N}},\ y = \{y_n\}_{n\in\mathbb{N}} \Rightarrow x + y \overset{\text{def}}{=} \{x_n + y_n\}_{n\in\mathbb{N}}$$

и умножать на вещественные числа:

$$x = \{x_n\}_{n\in\mathbb{N}}, a \in \mathbb{R} \Rightarrow ax \overset{\text{def}}{=} \{ax_n\}_{n\in\mathbb{N}}.$$

Определение 3.1. Множество элементов X любой природы, в котором определены две операции: сложения $+$ и умножения на (вещественное) число, такие что X является абелевой группой по сложению и удовлетворяет таким аксиомам, касающимся умножения на число: a) $1 \cdot x = x$, b) $\alpha(\beta x) = (\alpha\beta)x$, c) $(\alpha+\beta)x = \alpha x + \beta x$, d)$\alpha(x + y) = \alpha x + \alpha y$, называется *векторным пространством.* Отображение $\Lambda : X \to Y$ (где Y — другое векторное пространство) называется *линейным*, если $\Lambda(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 \Lambda x_1 + \alpha_2 \Lambda x_2$

Пример 3.1. Примером векторного пространства является $\mathbb{R}$ и его обобщения $\mathbb{R}^n$ и $(\mathbb{R}^n)'$. Векторным пространством является и пространство последовательностей $\mathbb{R}^{\mathbb{N}}$.

Задача 3.1 (на доказательство). *Совокупность сходящихся последовательностей образует векторное пространство. Сходящиеся последовательности можно перемножать (покоординатно).*

Иначе говоря, если каждая из двух последовательностей сходится, то сходится и их сумма, и их произведение; если последовательность сходится, то сходится и её произведение на любое вещественное число.

Это легко следует из определений, но можно посоветовать читателю освоиться (решив эту и некоторые сходные задачи) с “$\varepsilon - \delta$-хозяйством”.

Совокупность сходящихся последовательностей обозначают буквой c.

Векторное пространство, являющееся одновременно кольцом, называется *алгеброй.* Из задачи 2.4 следует что c является алгеброй (аксиомы векторного пространства и кольца проверяются легко).

Существует несколько замечательных *функционалов* на пространствах последовательностей, т. е. $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$. Приведем два из них. Это

—*верхняя (и нижняя) грани* :

$$x = \{x_n\}_{n\in\mathbb{N}} \mapsto \sup_{n\in\mathbb{N}} x_n \quad (x = \{x_n\}_{n\in\mathbb{N}} \mapsto \inf_{n\in\mathbb{N}} x_n).$$

Эти функционалы определены всюду на $\mathbb{R}^{\mathbb{N}}$ и принимают значения в $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ — *расширенной вещественной прямой.*

Третий функционал определён лишь на *сходящихся* последовательностях из $\mathbb{R}^{\mathbb{N}}$. Это — *предел*:

$$x = \{x_n\}_{n\in\mathbb{N}} \mapsto \lim_{n\to\infty} x_n.$$

Рассмотрим, например, последовательность $\{x_n\}_{n\in\mathbb{N}}, x_1 = 2, x_n = (-1)^n(1 - 1/n), n \geq 2$. Для нее

$$\sup x_n = 2, \quad \inf x_n = -1,$$

а предела не существует.

Задача 3.2 (на доказательство). *Докажите, что если* $\{x_n\}_{n\in\mathbb{N}}$ *и* $\{y_n\}_{n\in\mathbb{N}}$ *— две сходящиеся последовательности, а* α *и* β *— числа (из* $\mathbb{R}$ *или* $\mathbb{C}$*), то*
a) $\lim_{n\to\infty}(\alpha x_n + \beta y_n) = \alpha\lim_{n\to\infty} x_n + \beta\lim_{n\to\infty} y_n$*;*
b) $\lim_{n\to\infty} x_n y_n = \lim_{n\to\infty} x_n \lim_{n\to\infty} y_n$*;*
c) если в дополнение к условию сходимости $\lim_{n\to\infty} y_n \neq 0$*, то* $\lim_{n\to\infty}\frac{x_n}{y_n} = \frac{\lim_{n\to\infty} x_n}{\lim_{n\to\infty} y_n}$.

Таким образом, lim является *линейным* функционалом на пространстве c.

Функционалы sup (inf) — *сублинейны* (*суперлинейны*) (по определению функционал $p : X \to \mathbb{R}$, заданный на векторном пространстве X сублинеен (суперлинеен), если $p(x_1+x_2) \le p(x_1)+p(x_2)$ $(\ge p(x_1)+p(x_2))$ $\forall x_1,\ x_2 \in X$ и $p(ax) = ap(x)$ $\forall x \in X, a \in \mathbb{R}_+$). (Докажите все это.)

Совокупность всех этих простых утверждений (о пределе суммы, произведения, частного последовательностей, произведения последовательности на число), доказательство которых мы предоставили читателю, далее именуется *исчислением пределов*. Мы не сочли обязательным уделять этому исчислению большое внимание: понятие предела играет в анализе вспомогательную роль. (Тему "Пределы" разумно отрабатывать на семинарских занятиях).

Займёмся вычислением лишь особо важных пределов.

Предел для числа e. Докажем здесь существование предела последовательности, которая у нас уже встречалась: $x_n = \lim_{n\to\infty}(1+\frac{1}{n})^n$. Применив формулу бинома Ньютона, представим x_n в виде суммы $\sum_{k=0}^n C_n^k n^{-k}$. При этом легко понять, что $C_n^k n^{-k} \le 2^{-k+1}$, откуда следует, что $x_n \le 3$, а из неравенств $C_n^k n^{-k} \le C_{n+1}^k (n+1)^{-k}$ выводится, что $x_n < x_{n+1}$, и значит, x_n имеет предел, который обозначают буквой e. Совершенно аналогично доказывается, что $\lim_{n\in\mathbb{N}}(1+\frac{x}{n})^n$ также существует при любом x. Это приводит к одному из определений функции: $x \mapsto e^x$.

Чуть далее будут найдены ещё два важных предела.

Определение 3.2. С каждой последовательностью $\{x_n\}_{n\in\mathbb{N}}$ можно связать другую последовательность $\{s_n\}_{n\in\mathbb{N}}$, $s_n := \sum\limits_{k=1}^{n} x_k$. Последовательность $\{s_n\}_{n\in\mathbb{N}}$ называют последовательностью *частичных сумм ряда* $x_1+x_2+\cdots+x_n+\cdots$, а x_n — *общим членом ряда*. Ряд коротко записывают так: $\sum_{k\in\mathbb{N}} x_k$. Если последовательность s_n сходится к числу S, то S называют *суммой ряда* $\sum_{k\in\mathbb{N}} x_k$. При этом ряд $\sum_{k\in\mathbb{N}} x_k$ называют *сходящимся*. Если же последовательность $\{s_n\}_{n\in\mathbb{N}}$ не имеет предела (или имеет предел, равный $+\infty$ или $-\infty$), то говорят, что ряд $\sum_{k\in\mathbb{N}} x_k$ *расходится*. Если сходится ряд неотрицательных чисел $|x_1|+|x_2|+\cdots+|x_n|+\cdots$, то ряд $\sum\limits_{n\in\mathbb{N}} x_k$ называют *абсолютно сходящимся*. Если ряд сходится неабсолютно, говорят, что ряд сходится *условно*.

Пример 3.2. *Сумма геометрической прогрессии:* $x_n = q^{n-1}$, $n \in \mathbb{N}$. Частичная сумма ряда $1+q+q^2+\cdots+q^n+\cdots$ равна $s_n = \frac{q^n-1}{q-1}$, если $q \neq 1$ и n, если $q = 1$, так что если $|q| < 1$, то ряд абсолютно сходится и его сумма $S = S(q) = \dfrac{1}{1-q}$. Если $q \ge 1$, ряд расходится.

Лемма 3.1. *Необходимым условием сходимости ряда* $x_1+\cdots+x_n+\cdots$ *является стремление* x_n *к нулю.*

Доказательство. Действительно, если ряд сходится и имеет сумму S, то из определения сходимости ряда

$$\forall\varepsilon > 0\ \ \exists N \in \mathbb{N} : |s_n - S| < \varepsilon\ \ \forall n \ge N. \qquad (i)$$

Взяв $n \ge N+1$, получаем, что $|x_n| \overset{\mathrm{Id}}{=} |s_n - s_{n-1}| \overset{\mathrm{Id}}{=} |s_n - S + S - s_{n-1}| \overset{(i)}{<} 2\varepsilon \overset{\mathrm{def}}{\Longrightarrow} \lim\limits_{n\to\infty} x_n = 0$. □

Это условие не является достаточным.

Пример 3.3. (*гармонический ряд*): $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$.

Имеем:

$$s_{2^n} = 1 + (\frac{1}{2} + \frac{1}{4}) + (\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}) + (\frac{1}{9} + \cdots + \frac{1}{16}) + \cdots$$
$$+ (\frac{1}{2^{n-1}+1} + \cdots + \frac{1}{2^n}) > \frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2} = \frac{n}{2},$$

следовательно, гармонический ряд расходится (точнее, сходится к бесконечности), хотя его общий член стремится к нулю.

Задача 3.3 (на доказательство). *Пусть $\sum_{n\in\mathbb{N}} x_n$ и $\sum_{n\in\mathbb{N}} y_n$ — два абсолютно сходящихся ряда, ξ и η — их суммы. Тогда ряд $(\sum_{n\in\mathbb{N}} x_n)(\sum_{n\in\mathbb{N}} y_n) = \sum_{\mathrm{j}\in\mathbb{N}} \sum_{k=1}^{j-1}(x_k y_{j-k})$ —абсолютно сходится и его сумма равна $\xi\eta$.*

3.2. Признаки сходимости

Признак Коши.*Для того, чтобы ряд $\sum_{n\in\mathbb{N}} x_n =$ сходился, необходимо и достаточно, чтобы для любого $\varepsilon > 0$ нашелся бы номер $N \in \mathbb{N}$ такой, что для любых n и m, превосходящих N, было выполнено неравенство $|s_n - s_m| < \varepsilon$, т. е., чтобы последовательность $s_n = x_1 + \cdots + x_n,\ n \in \mathbb{N}$ была фундаментальна.*

Доказательство. Если $\{s_n\}_{n\in\mathbb{N}}$ фундаментальна, она сходится в силу полноты $\mathbb{R}$, что (по определению) означает сходимость ряда $\sum_{n\in\mathbb{N}} x_n$.

Если ряд $\sum_{n\in\mathbb{N}} x_n$ сходится, то (по определению) сходится $\{s_n\}_{n\in\mathbb{N}}$ и значит, последовательность s_n фундаментальна. □

Признак Даламбера. *Пусть (для некоторого положительного C) общий член ряда x_n допускает оценку $|x_n| \le Cq^n$. Тогда, если $0 < q < 1$, то ряд $\sum_{n\in\mathbb{N}} x_n = x_1 + x_2 + \cdots + x_n + \cdots$ сходится абсолютно.*

(Другой (не эквивалентный) вариант: *если существует предел $\lim_{n\to\infty}\left|\frac{x_{n+1}}{x_n}\right| = \alpha$, то при $\alpha < 1$ ряд сходится абсолютно, при $\alpha > 1$ — расходится*).

Доказательство. Выберем по $\varepsilon > 0$ число N так, чтобы $\frac{Cq^{N+1}}{1-q} < \varepsilon$. Тогда для $n > m > N$ получим (НТ — нер-во тр-ка):

$$|s_n - s_m| \overset{\mathrm{Id}}{=} |x_{m+1} + \cdots + x_n| \overset{HT}{\le} |x_{m+1}| + \cdots + |x_n| \overset{Cond}{\le} C(q^{m+1} + \cdots + q^n)$$
$$< \frac{Cq^{N+1}}{1-q} < \varepsilon \Longrightarrow$$

□

Применим признак Даламбера для ещё одного описания числа e и функции $x \mapsto e^x$. Было доказано, что последовательность $x_n = (1 + \frac{1}{n})^n$ сходится к e. Выбрав $N \in \mathbb{N}$ из бинома Ньютона получим (при $n > N$), что $x_n = \xi_{nN} + \eta_{nN}$, где $\xi_{nN} = \sum_{k=0}^{N} C_n^k n^{-k}$, $\eta_{nN} \le 2^{-N+1}$ (мы воспользовались неравенством $C_n^k n^{-k} \le 2^{-k+1}$). При этом $\lim_{n\to\infty} \xi_{nN} = \sum_{k=0}^{N} \frac{1}{N!}$. Отсюда следует, что $e = \sum_{n=0}^{\infty} \frac{1}{n!}$. Аналогично доказывается, что $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Наконец, из теоремы о произведении рядов получаем, что $e^{a+b} = e^a e^b$ (см. задачу 3.3).

Признак Лейбница. *Пусть в ряде $x_1 - x_2 + x_3 - x_4 + \cdots$, числа x_n монотонно убывают и стремятся к нулю. Тогда этот ряд сходится.*

Доказательство. Из условия заключаем, что отрезки $\Delta_1 = [x_1 - x_2, x_1], \Delta_2 = [x_1 - x_2 + x_3 - x_4,\ x_1 - x_2 + x_3], \ldots$ вложены друг в друга и их длины стремятся к нулю. В силу аксиомы Кантора существует единственное число S общее для всех отрезков. Из определений следует, что $\lim\limits_{n\to\infty} s_n = S$. □

Примеры.

1. Ряд $1 - \dfrac{1}{2} + \dfrac{1}{3} - \dfrac{1}{4} + \cdots$ сходится условно (ибо гармонический ряд расходится). (Доказывается, его сумма равна $\ln 2$.)

2. Ряд $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$ (ряд Грегори – Лейбница) также сходится условно (читатель, надеюсь, легко докажет, что ряд чисел обратных к нечетным, расходится). (Доказывается, что сумма ряда равна $\pi/4$.)

В заключение найдём два важных предела функций.

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

Доказательство основывается на следующем утверждении, которое студенческий фольклор именует *леммой о двух милиционерах: пусть три функции f_1, f_0 и f_2 определены на интервале $]x_0 - h, x_0 + h[$, кроме, быть может, точки x_0, и при этом а) $f_1(x) \le f_0(x) \le f_2(x)$ и кроме того $\lim_{x\to x_0} f_1(x) = \lim_{x\to x_0} f_2(x) = A$. Тогда $\lim_{x\to x_0} f_0(x) = A$.*

Это утверждение без труда выводится из определения предела. А далее надо применить лемму о двух милиционерах к $f_1(x) = \frac{\sin x \cos x}{2}$, $f_0(x) = \frac{x}{2}$ и $f_2(x) = \frac{\sin x}{2\cos x}$ и далее — поделить на $\sin x$ и перейти к пределам.

Но в этом рассуждении есть одна деталь, которую мы здесь как бы пропускаем, что площадь сектора единичной окружности с углом x равна $\sin\frac{x}{2}$. Это верно, но пока не было объяснено, что такое площадь. Примем здесь это утверждение на веру (в школе оно как-то объясняется); теория меры, в частности, определение площади, излагается в шестой главе.

Второй предел: $\lim_{x\to 0} \frac{e^x - 1}{x} = 1$ легко следует из обоих определений функции $x \to e^x$: $e^x = \lim_{n\to\infty}(1 + \frac{x}{n})^n$ и $e^x = \sum_{k\in\mathbb{N}} \frac{x^n}{n!}$.

3.3. Начала анализа (дифференцирование и интегрирование)

Приведём замечательные слова И. Ньютона, в которых он вкратце объясняет суть математического анализа.

“Для пояснения искусства анализа, — пишет Ньютон, — надо привести некоторые образцы задач [...]. Все заключающиеся в них затруднения можно свести к двум следующим проблемам относительно пути [...]:

1) Пусть дана длина проходимого пути; требуется найти скорость движения в данное время.

2) Пусть дана скорость движения; требуется найти длину пройденного пути.

3.4. Что такое производная?

Последуем за Ньютоном. В его “первом образце задач” фигурирует понятие скорости. Что же такое скорость? Какую величину показывает спидометр, когда мы едем на автомобиле?

Если движение равномерно и прямолинейно, т. е. если в равные промежутки времени проходятся равные расстояния, то скорость — это отношение длины пройденного пути к затраченному времени для любого участка пути. При этом длина пройденного пути $s(t)$ в момент времени t с начала движения из точки с координатой x_0 равна $s(t) = vt + x_0$. А если движение неравномерно, то как разумным образом определить *скорость в данный момент времени?* Как определить $v(t_0)$ — скорость в момент времени t_0, если “дана длина пройденного пути”, — как пишет Ньютон, — т. е. функция $t \to s(t)$?

За время $t_0 + \Delta t$ пройден путь $s(t_0 + \Delta t) - s(t_0)$. Если бы скорость в течение времени Δt была постоянной, она была бы равна $\frac{s(t_0+\Delta t) - s(t_0)}{\Delta t}$. При этом чем меньше отрезок времени Δt, тем более мы приближаемся к числу, которое и следует считать “скоростью в данный момент времени". Иначе говоря, $v(t_0)=\lim_{\Delta t\to 0}\frac{s(t_0+\Delta t)-s(t_0)}{\Delta t}$. Мы пришли к общепринятому ныне определению производной, которое было дано Коши (хотя, конечно, Ньютон именно так и понимал суть дела. Ньютон писал: "Количества, которые в продолжении любого времени постоянно стремятся к равенству и ранее конца этого времени приближаются друг к другу ближе, нежели на любую заданную величину, будут в пределе равны". Если применить это высказывание к числу $v(t_0)$ и "количеству" $\frac{s(t_0+\Delta t)-s(t_0)}{\Delta t}$, мы приходим к определению Коши скорости, как производной пути).

Приведём теперь определение Коши производной для произвольной функции f одного переменного.

Пусть функция f определена в окрестности точки x_0.

Определение 3.3. Говорят, что функция f *дифференцируема в этой точке*, если существует конечный предел

$$\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h}=:f'(x_0). \tag{1}$$

Иначе определение Коши можно переформулировать так: функция f дифференцируема в точке x_0, если имеет место такое представление:

$$f(x_0+h)=f(x_0)+f'(x_0)h+r(h),\quad \lim_{h\to 0}|r(h)|/|h|=0. \tag{2}$$

Предел (1) называют *производной* f в точке x_0.

Лейбниц исходил из геометрического определения. На нашем современном языке производная это тангенс угла наклона касательной к графику функции в данной точке.

Исторический комментарий. Термин "производная" и обозначение f' ввел Лагранж (в семидесятые годы 18 столетия), обозначение $\dfrac{dy}{dx}$ принадлежит Лейбницу – 1684 г. Определение (2) допускает далёкое (бесконечномерное) обобщение (о котором будет рассказано в пятом разделе). Это определение было дано Фреше (1913)

Если функция дифференцируема в каждой точке, то после дифференцирования возникает новая функция, обозначаемая $f'(\cdot)$ (или $x\to f'(x)$), называемая *производной* f.

Вычислим производные некоторых элементарных функций.

Пример 3.4 Производная степенной функции. Из формулы бинома Ньютона следует, что $(x_0+h)^n=x_0+nx_0^{n-1}h+r(h),\quad |r(h)/|h|\to 0$, откуда

$$(x^n)'=nx^{n-1}.$$

Пример 3.5 Производные синуса и косинуса. Из предельного соотношения $\lim\limits_{h\to 0}\dfrac{\sin h}{h}=1$ следует, что

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

$$\lim_{h\to 0}\cos(x+\frac{h}{2})\lim_{h\to 0}\frac{\sin\frac{h}{2}}{\frac{h}{2}}=\cos x,$$

т.е. $(\sin x)'=\cos x$. Аналогично $(\cos x)'=-\sin x$.

Пример 3.6 Производная экспоненты. Из соотношения $\lim\limits_{h\to 0}\dfrac{e^h-1}{h}=1$ следует $\lim\limits_{h\to 0}\dfrac{e^{x+h}-e^x}{h}=e^x\lim\limits_{h\to 0}\dfrac{e^h-1}{h}=e^x$
$\Longrightarrow (e^x)'=e^x$.

Замечание 3.1. Функция, имеющая производную в точке, непрерывна в этой точке. Функции $x\to|x|$ и $x\to x\ln x$ не дифференцируемы в нуле (последняя функция вообще не определена при $x<0$, а предел, определяющий производную, справа $(h\to+0)$ равен бесконечности). Функция $f(x)\Longrightarrow x\sin 1/x$ при $x\neq 0$ и нулю при $x=0$ не только не дифференцируема в точке $x=0$, но не имеет производной в этой точке ни конечной, ни бесконечной, ни справа, ни слева.

Существует непрерывные функции, не имеющие производной ни в одной точке. Первые примеры были найдены Больцано (в 1830 г, но опубликованы в конце века) и Вейерштрассом (1860 г).

3.5. Основные теоремы дифференциального исчисления функций одного переменного.

Из теорем об исчислении пределов следует

Теорема 3.2 (об исчислении производных). *Если f и g дифференцируемы в точке $\widehat{x}$, то*

a) функция $af+bg$ (при любых $a,b\in\mathbb{R}$) также дифференцируема в $\widehat{x}$ и при этом $(af+bg)'(\widehat{x})=af'(\widehat{x})+bg'(\widehat{x})$;

b)функция fg также дифференцируема в $\widehat{x}$ и при этом $(fg)'(\widehat{x})=f'(\widehat{x})g(\widehat{x})+f(\widehat{x})g'(\widehat{x})$;

c) если $g'(\widehat{x})\neq 0$, то функция $\frac{f}{g}$ также дифференцируема в $\widehat{x}$ и при этом $(\frac{f}{g})'(\widehat{x})=\frac{f'(\widehat{x})g(\widehat{x})-f(\widehat{x})g'(\widehat{x})}{g^2(\widehat{x})}$;

d) Если функция f дифференцируема в точке $\widehat{x}$, а функция g — в точке $\widehat{y}=f(\widehat{x})$, то суперпозиция $h=g\circ f$ этих функций дифференцируема в точке $\widehat{x}$ и при этом

$$h'(\widehat{x})=g'(\widehat{y})f'(\widehat{x}).$$

(теорема Лейбница о суперпозиции).

Доказательство. Утверждения a) — c) предоставляется доказать самостоятельно, утверждение d) будет доказано в пятом разделе в многомерном случае.

Пусть $f(\cdot)$ — вещественная функция на отрезке $[a,b]$. Точка $\widehat{x}\in]a,b[$ называется точкой локального минимума (максимума) f, если существует такая окрестность U точки $\widehat{x}$, что $f(x)\geq f(\widehat{x})$ $(f(x)\leq f(\widehat{x}))$ $\forall\ x\in U$ (или можно сказать так: существует число $\delta>0$ такое, что $]x-\delta,\widehat{x}+\delta[\subset[a,b]$ и $f(x)\geq f(\widehat{x})(f(x)\leq f(\widehat{x}))$ $\forall\ x\in]\widehat{x}-\delta,\widehat{x}+\delta[$.

Теорема 3.3 (Ферма; необходимое условие экстремума). *Пусть функция f_0 определена на некотором отрезке $[a,b]$, точка $\widehat{x}$ лежит внутри этого отрезка и f_0 дифференцируема в $\widehat{x}$. Тогда если $\widehat{x}$ доставляет локальный минимум или максимум функции f_0, то $f_0'(\widehat{x})=0$.*

Далее при доказательствах мы выделяем "пункт ноль", где кратко обозначаем те определения или утверждения (которые обычно встречались ранее), на которые опираемся при доказательстве.

Доказательство. 0.1. Определение производной. 0.2. Определение предела.

1. Допустим, что $f'(\widehat{x})=k>0$. Из 0.2 найдем $\delta>0$ такое, что $|h|<\delta\Longrightarrow\left|\frac{r(h)}{h}\right|<\left(\frac{k}{2}\right)|x|$ (где $r(h)$ из 0.1 есть $f(\widehat{x}+h)-f(\widehat{x})-kh$). Тогда при $h>0$

$$f(\widehat{x}+h)\overset{0.1}{=}f(\widehat{x})+kh+r(h)\overset{(i)}{\geq}f(\widehat{x})+kh-\frac{k}{2}h\overset{\mathrm{Id}}{=}f_0(\widehat{x})+\frac{k}{2}h>f(\widehat{x}),$$

т.е. $\widehat{x}$ — не максимум. Аналогично (при $h<0$) показываем, что $\widehat{x}$ не минимум. Случай $k<0$ аналогичен. □

Исторический комментарий. Простейший вариант этой теоремы (для полиномов) был сформулирован Ферма в 1638 г.

Теорема 3.4 (Ролля). *Пусть a) f непрерывна на $[a,b]$ (b), дифференцируема во всех внутренних точках отрезка (b) и при этом $f(a)=f(b)$ (c). Тогда существует внутренняя точка отрезка ξ, для которой $f'(\xi)=0$.*

Доказательство. 0.1. Следствие из принципа компактности Вейерштрасса(см. §2): *непрерывная функция на отрезке достигает своего максимума и минимума.* 0.2. Теорема Ферма.

1. Возможно одно из двух: либо максимальное и минимальное значения функции f (а они в силу 0.1 достигаются) совпадают, либо различаются. В первом случае f — константа и f' — тождественный нуль, т. е. в этом случае теорема верна. Во втором случае либо точка минимума, либо точка максимума лежит внутри отрезка. В силу 0.2 в этой точке производная должна быть нулем. □

Исторический комментарий. Ролль также доказал свою теорему для полиномов (чисто алгебраическими средствами) в 1690 г.

Теорема 3.5 (Лагранжа о среднем; формула конечных приращений). *В условиях (a) и (b) теоремы Ролля найдется точка ξ, $a < \xi < b$ такая, что $f(b) - f(a) = (b-a)f'(\xi)$.*

Теорема Лагранжа является простым следствием теоремы Ролля: достаточно положить $g(x) = f(x) - \dfrac{f(b)-f(a)}{b-a}(x-a)$ и применить теорему Ролля. □

Исторический комментарий. Эта теорема, доказана Лагранжем в 1797 г.

Следствие 3.6 (о монотонности). *Если f непрерывна на отрезке $[a,b]$ и дифференцируема на интервале $]a,b[$, и в любой точке интервала производная неотрицательна (положительна), то функция не убывает (возрастает).*

Доказательство. $x_1 < x_2 \Longrightarrow f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) \geq 0 \ (> 0)$. □

Следствие 3.7 (о функции с нулевой производной). *Непрерывная на отрезке $[a,b]$ и дифференцируемая в любой точке интервала $]a,b[$ функция постоянна на $[a,b]$ тогда и только тогда, когда ее производная равна нулю в любой точке $]a,b[$.*

Доказательство сразу следует из теоремы Лагранжа.

Теорема 3.8 (о дифференцируемости обратной функции). *Пусть функции $f : [a,b] \to [c,d]$ и*

$$f^{-1} : [c,d] \to [a,b]$$

— взаимно обратны и непрерывны в точках $x_0 \in]a,b[$ и $f(x_0) = y_0 \in]c,d[$. Если f дифференцируема в точке x_0 и $f'(x_0) \neq 0$, то f^{-1} дифференцируема в y_0 и

$$f^{-1\prime}(y_0) = (f'(x_0)^{-1}).$$

Доказательство. Используя простейшие свойства пределов, получаем:

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} \stackrel{def}{=} \lim_{x \to x_0} \frac{x - x_0}{f(x) - f(x_0)} \stackrel{def}{=} \frac{1}{f'(x_0)}$$

□

Из последних двух теорем и доказанных формул $(e^x)' = e^x$, $(\sin x)' = \cos x$, $(\cos x)' = -\sin x$ легко выводятся формулы для дифференцирования произвольных элементарных функций: $(a^x)' = a^x \ln a$ $(x \in \mathbb{R},\ a > 0,\ a \neq 1)$, $(\log_a)'|x| = \frac{1}{x \ln a}$ $(x \in \mathbb{R} \setminus 0)$, $a > 0,\ a \neq 1$, $x^\alpha = e^{\alpha \ln x} \Rightarrow (x^\alpha)' = e^{\alpha \ln x} \frac{\alpha}{x} = \alpha x^{\alpha - 1}$, $(\operatorname{tg} x)' = \cos^{-2} x$ $(x \neq= \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z})$, $(\operatorname{ctg} x)' = -\sin^{-2} x$ $(x \neq k\pi,\ k \in \mathbb{Z})$, $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$, $(|x| \neq 1)$, $(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$, $(|x| < 1)$, $(\arctan x)' = \frac{1}{1+x^2}$, $(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}$ и т.д.

Теорема 3.9 (Коши). *Пусть $t \mapsto x(t)$ и $t \mapsto y(t)$ — функции, непрерывные на отрезке $[\alpha,\beta]$ и дифференцируемые на интервале $]\alpha,\beta[$, причем $x'(t) \neq 0$ $\forall t \in]\alpha,\beta[$. Тогда найдется точка $\tau \in]\alpha,\beta[$ такая, что*

$$\frac{y(\beta) - y(\alpha)}{x(\beta) - x(\alpha)} = \frac{y'(\tau)}{x'(\tau)}.$$

Доказательство. 0.1. Суперпозиция непрерывных функций непрерывна. 0.2. Суперпозиция дифференцируемых функций дифференцируема. 0.3. Теорема об обратной функции. 0.4. Теорема Лагранжа.

1. 0.3. $\Leftarrow x(\cdot)$ строго монотонна и имеет обратную функцию $x \to t(x)$, определенную и непрерывную на отрезке $[a,b]$, $a = x(\alpha)$, $b = x(\beta)$ и дифференцируемую в $]a,b[$. Тогда $y(t) = y(t(x)) = f(x)$, т.е. $f(\cdot)$ есть суперпозиция $y(\cdot)$ и $t(\cdot)$. Из 0.1 и 0.2 следует, что можно применить теорему Лагранжа, т.е.

$$y(\beta) - y(\alpha) \stackrel{def}{=} y(t(b)) - y(t(a)) \stackrel{0.4}{=} y'_t(t(\xi))t'_x(\xi)(b-a) \stackrel{0.3}{=} \frac{y'_t(t(\xi))}{x'_t(t(\xi))}(x(\beta) - x(\alpha))$$

□

3.6. Старшие производные и формула Тейлора

Если функция, заданная на некотором подмножестве прямой, дифференцируема в любой точке этого подмножества, то возникает, как говорилось, новая функция $x \mapsto f'(x)$. Функция f' может также быть дифференцируемой. Ее производная $(f')'$ называется второй производной от f и обозначается одним из символов f'' или $\frac{d^2 f(x)}{dx^2}$. По индукции определяется производная произвольного n-го порядка: $f^{(n)}(x) = (f^{(n-1)})'(x)$. Кроме $f^{(n)}(x)$ принято также обозначение Лейбница $\frac{d^n f(x)}{dx^n}$. Имеет место

Формула Лейбница. Если функции f и g имеют на общем множестве $X \subset \mathbb{R}$ производные до порядка n включительно, то для n-й производной их произведения справедлива следующая формула

$$(fg)^{(n)} = \sum_{k=0}^{n} \mathrm{C}_n^k f^{(n-k)} g^{(k)}$$

Эта формула доказывается методом математической индукции.

Формула Тейлора. Если на отрезке $[x_0, x]$ функция f непрерывна вместе со своими производными до порядка n включительно, а во внутренних точках отрезка она имеет производную порядка $n+1$, то имеет место формула

$$f(x_0 + h) = f(x_0) + \frac{f'(x_0)}{1!}h + \cdots + \frac{f^{(n)}(x_0)}{n!}h^n + r_n(x_0, h),$$

где $r_n(x_0, h) = \frac{1}{(n+1)!} f^{(n+1)}(\xi) h^{n+1}$, $\xi \in]x_0, x_0 + h[$ (остаточный член в форме Лагранжа).

Доказательство. Положим

$$g(t) = f(x_0 + h) - (f(t) + \frac{f'(t)}{1!}(x_0 + h - t) + \cdots + \frac{f^{(n)}(t)}{n!}(x_0 + h - t)^n)$$

Из определения $g(\cdot)$ и условий теоремы видно, что g непрерывна на отрезке $[0,h]$ и дифференцируема в его внутренних точках, причём

$$g'(t) = -(f'(t) - \frac{f'(t)}{1!} - \frac{f''(t)}{2!}(x_0 + h - t) + \frac{f''(t)}{2!}(x_0 + h - t) + \cdots) = \\ -\frac{f^{(n+1)}(t)}{n!}(x_0 + h - t)^n$$

□

3.7. Что такое интеграл?

Снова последуем за Ньютоном. Пусть нам известна скорость $t \to v(t)$ движения некоей частицы. Как узнать путь, пройденный ею за отрезок времени $[t_0, t_1]$?

Снова рассмотрим маленький отрезок времени от момента τ до $\tau + \Delta t$. За малый отрезок времени скорость изменится несущественно, и потому длина пути примерно будет равна $v(\tau)\Delta t$. А весь путь за время движения от t_0 до t_1 будет примерно равен сумме $\sum_{k=0}^{n-1} v(\tau_k)\Delta\tau_k$, где $t_0 = \tau_0 < \tau_1 < \ldots < \tau_{n-1} < \tau_n = t_1$ — разбиение отрезка на мелкие отрезочки $[\tau_k, \tau_{k+1}]$, $0 \le k \le n-1$ длины $\Delta\tau_k$. Для определённости можно положить $\tau_{kn} = t_0 + k\frac{t_1-t_0}{n}$, $0 \le k \le n-1$ и рассмотреть последовательность чисел $I_n = \sum_{k=0}^{n-1} v(\tau_{kn})\Delta\tau_{kn}$. Предел этой последовательности и будет длиной пути, пройденного за время от t_0 до t_1. Этот предел обозначают символом $\int_{t_0}^{t_1} v(t)dt$ и называют *определённым интегралом от функции* $v(\cdot)$ *на отрезке интегрирования* $[t_0, t_1]$. Мы получили, что $s(t_1) - s(t_0) = \int_{t_0}^{t_1} v(t)dt$. Но с другой стороны мы знаем уже, что $v(t) = s'(t)$. Таким образом, мы пришли к основной формуле интегрального и дифференциального исчисления — формуле Ньютона – Лейбница:

$$s(t_1) - s(t_0) = \int_{t_0}^{t_1} s'(t)dt.$$

В лейбницевых обозначениях, когда переменная x, а функция f, формула Ньютона – Лейбница имеет вид:

$$f(x_1) - f(x_0) = \int_{x_0}^{x_1} f'(x)dx.$$

Функция $F(\cdot)$ такая, что $F' = f$ называется *первообразной* F. В силу того, что константы имеют производную тождественно равную нулю, получаем, что первообразная определяется неоднозначно. Совокупность первообразных обозначают $\int f(x)dx$ и называют *неопределённым интегралом.* Таким образом, первообразная F функции f описывается так: $F = \int f(x) + C$. Из формул дифференцирования элементарных функций следуют такие формулы для первообразных:

$\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C, \quad \alpha \ne -1;$
$\int \frac{dx}{x} = \ln|x| + C;$
$\int a^x dx = \frac{1}{\ln a}a^x + C, \quad 0 < a \ne 1, \quad \int e^x dx = e^x + C;$
$\int \sin x dx = -\cos x + C;$ $\int \cos x dx = \sin x + C;$
$\int \frac{1}{\cos^2 x}dx = \operatorname{tg} x + C;$
$\int \frac{1}{\sin^2 x}dx = -\operatorname{ctg} x + C;$
$\int \frac{1}{\sqrt{1-x^2}}dx = \arcsin x + C;$
$\int \frac{1}{1-x^2}dx = \arctan x + C;$
$\int \operatorname{sh} x dx = \operatorname{ch} x + C;$
$\int \operatorname{ch} x dx = \operatorname{sh} x + C;$
$\int \frac{1}{\operatorname{ch}^2 x}dx = \operatorname{th} x + C;$
$\int \frac{1}{\operatorname{sh}^2 x}dx = -\coth x + C;$
$\int \frac{1}{\sqrt{x^2 \pm 1}}dx = \ln|x + \sqrt{1 \pm x^2}| + C;$
$\int \frac{1}{1-x^2}dx = \frac{1}{2}\ln|\frac{1+x}{1-x}| + C.$

Вычисление неопределённых и определённых интегралов также замечательная тренировка в области математического анализа и математики вообще.

(Но при этом каждый теперь уже знает, что любой предел, любая производная, график любой элементарной функции, любой определённый и неопределённый интеграл из любых учебников и задачников (за ничтожными исключениями) могут быть вычислены нажатием нескольких кнопок компьютера, оборудованного достаточно богатой стандартной программой.)