МГУ им. М. В. Ломоносова

Лекции по математическому анализу

Дифференцируемые функции одной переменной

Химический факультет

2026

Оглавление

Лекция 15

Производная и дифференциал

Мы переходим к изучению функций, графики которых в некоторой окрестности заданной точки отличаются от прямых на бесконечно малую величину (см. ниже строгое определение). Про такие функции ещё можно сказать, что локально (то есть в некоторой окрестности заданной точки) они почти равны функциям, графики которых являются прямыми. Если функция локально устроена как прямая, то нам проще работать с её приближёнными значениями в некоторой окрестности точки, что удобно для приложений.

Сначала дадим определение таких функций, изучим некоторые связанные с ними понятия, а затем уже опишем, как устроены прямые, на которые в окрестностях рассматриваемых точек "похожи" эти функции.

Определение 1. *Функция f, определённая в некоторой окрестности точки a, называется* ***дифференцируемой*** *в точке a, если существуют такие число A и функция α, что при всех h из некоторой проколотой окрестности нуля выполнено равенство*

$$f(a+h)-f(a)=Ah+\alpha(h)h, \tag{1}$$

где $\lim\limits_{h\to 0}\alpha(h)=0$. При этом A и α зависят и от точки a, поэтому часто равенство (1) записывают в виде

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

Определение 2. *Функция $h\mapsto Ah$ называется* ***дифференциалом*** *функции f в точке a. Она обозначается $df(a)$ или $df|_{x=a}$, то есть $df(a)(h)=df(h)|_{x=a}=Ah$.*

Ещё раз подчеркнём, что равенство записано в фиксированной точке a, то есть оно зависит от точки a. Другими словами, число A, вообще говоря, разное при разных a.

Здесь символ $df(a)$ нужно воспринимать как обозначение функции, то есть как *цельный* символ.

Отметим два очевидных наблюдения для функции $h\mapsto Ah$: во-первых

$$\forall\lambda\in\mathbb{R}\ df(a)(\lambda h)=A\lambda h=\lambda Ah,$$

а во-вторых

$$df(a)(h_1+h_2)=A(h_1+h_2)=Ah_1+Ah_2.$$

Так как здесь A – число, то свойства очевидны. Однако позже мы увидим, что дифференциал функции многих переменных также обладает аналогичными свойствами. Выполнение этих свойств по определению означает, что *дифференциал является линейной функцией от h.*

Пример 1. ***1)*** *Пусть $f(x)=x$. Тогда $f(a+h)-f(a)=(a+h)-a=h=1\cdot h+0\cdot h$. Из этого представления имеем: $A=1,\alpha(h)=0,dx(h)|_{x=a}=h$.*

2) *Пусть $f(x)=x^2$. Тогда $f(a+h)-f(a)=(a+h)^2-a^2=2ah+h^2=2ah+h\cdot h$, то есть $A=2a,\ d(x^2)(h)|_{x=a}=2ah,\ \alpha(h)=h$.*

3) *Пусть $f(x)=x^3$. Тогда $f(a+h)-f(a)=(a+h)^3-a^3=3a^2h+3h^2a+h^3=3a^2h+(3ah+h^2)h$ то есть $A=3a^2,\ d(x^3)(h)|_{x=a}=3a^2h,\ \alpha(h)=3ah+h^2$.*

Дадим теперь определение производной функции.

Определение 3. *Если существует предел* $\lim\limits_{h\to 0}\frac{f(a+h)-f(a)}{h}$, *то он называется производной функции* f *в точке* a. *Другая форма записи предела:* $\lim\limits_{x\to a}\frac{f(x)-f(a)}{x-a}$. *Производная функции* f *в точке* a *обозначается символом* $f'(a)$ *или* $\frac{df}{dx}(a)$ *(второе обозначение позже обсудим подробно).*

В следующем предложении устанавливается связь между дифференциалом и производной.

Предложение 1. *Функция* f *дифференцируема в точке* a *тогда и только тогда, когда в точке* a *существует производная этой функции* $f'(a)$. *При этом* $df(h)|_{x=a} = f'(a)h$.

Доказательство. **Необходимость.** Если функция f дифференцируема, то справедливо равенство

$$f(a+h)-f(a) = Ah + \alpha(h)h,\ \lim_{h\to 0}\alpha(h) = 0.$$

Разделив обе части этого равенства на h, получим

$$\frac{f(a+h)-f(a)}{h} = A + \alpha(h).$$

Переходя к переделу при $h \to 0$ и учитывая, что $\lim\limits_{h\to 0}\alpha(h) = 0$, получим, что $f'(a) = A$.

Достаточность. Если функция f имеет производную, то есть существует предел

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

то по определению предела справедливо равенство

$$\frac{f(a+h)-f(a)}{h} = f'(a) + \alpha(h),$$

где $\lim\limits_{h\to 0}\alpha(h) = 0$. Домножив обе части этого равенства на h, мы получим равенство из определения дифференцируемой функции. □

Обратим внимание, что из определения дифференциала и доказанного предложения следует, что дифференциал функции f в точке a определён однозначно, так как всегда $A = f'(a)$.

Так как из примера 1, пункта 1 вытекает, что $dx(h) = h$ в любой точке a, то дифференциал можно записать в виде $df|_{x=a} = f'(a)dx$. Действительно, при любом h из проколотой окрестности нуля, о которой говорилось в определении дифференциала, справедливы равенства

$$df(h)|_{x=a} = f'(a)dx(h)|_{x=a} = f'(a)h,$$

откуда следует и равенство самих линейных функций $df(a)$ и $f'(a)dx$. При этом вторая линейная функция представляет собой произведение числа $f'(a)$ и дифференциала dx функции $f(x) = x$. Разделив в равенстве $df(a) = f'(a)dx$ обе части на dx, получим обозначение производной $f'(a)$ в виде отношения двух дифференциалов:

$$f'(a) = \frac{df(a)}{dx}.$$

Справа стоит функция, равная отношению двух линейных функций, значение которой на любом h из проколотой окрестности нуля в силу предложения выше равно $f'(a)$. Именно в таком смысле следует понимать полученное равенство, а отсюда понятно и второе обозначение для производной. Так как речь идёт о производной в точке a, то есть дифференциалы в отношении справа вычисляются в точке a, то пишут ещё $f'(a) = \frac{df}{dx}(a)$. Если речь идёт о производной, рассматриваемой сразу на множестве точек, то есть если изучают зависимость производной от точки при изменении этой точки, то возникает функция, которую принято обозначать f' или df/dx.

Мы чаще будем пользоваться обозначением f', а обозначение df/dx используется, например, в курсе дифференциальных уравнений.

Приведём несколько примеров на вычисление производных с помощью определения производной.

Пример 2. *1) Пусть $f(x) = e^x$. Найдём $\frac{df}{dx}(a)$. Имеем по определению:*

$$\frac{df}{dx}(a) = \lim_{h \to 0} \frac{e^{a+h} - e^a}{h} = e^a \lim_{h \to 0} \frac{e^h - 1}{h} = e^a.$$

Так как это верно в любой точке области определения, то при всех $x \in \mathbb{R}$ $(e^x)' = e^x$.

2) Пусть $f(x) = \cos x$. Тогда

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

Совершенно аналогично проверяется, что $\sin'(x) = \cos x$.

С помощью свойств дифференцируемых функций, которые будут выведены ниже, и теорем о производной сложной и обратной функции мы получим производные других элементарных функций и запишем *таблицу производных.* Пока в этой таблице три функции из примера выше.

Изучим *геометрическую интерпретацию производной и дифференциала.* Из равенства $h = x - a$ получим, что дифференцируемая функция может быть записана в виде $f(x) = f(a) + f'(a)(x-a) + \alpha(x)(x-a)$, где $\lim\limits_{x \to a} \alpha(x) = 0$. Это значит, что в некоторой окрестности точки a функция f приближается функцией $x \mapsto f(a) + f'(a)(x-a)$. Таким образом, локально (то есть в некоторой окрестности точки a) график функции f выглядит "почти"как прямая. Сама прямая $y = f'(a)(x-a) + f(a)$ называется *касательной* к графику функции f в точке a. Производная $f'(a)$ является тангенсом угла наклона касательной к положительному направлению оси Ox.

Мы определяем касательную для дифференцируемых функций, однако из определения дифференцируемости следует, что сама дифференцируемость равносильна тому, что значения функции в некоторой окрестности точки a отличаются от значений некоторой функции, задающей прямую, на некоторую бесконечно малую более высокого порядка по сравнению с разностью $x - a$, то есть найдётся такая окрестность $U(a)$, что $\forall x \in U(a)$ справедливо равенство

$$f(x) - (A(x-a) + f(a)) = o(x-a),\ x \to a.$$

Если выполнено это равенство, то, как мы видели, $A = f'(a)$, поэтому прямая определена единственным образом. Таким образом, *дифференцируемость в точке равносильна наличию касательной в этой точке.*

Итак, теперь можем подытожить и определить касательную так: *если найдётся такая окрестность $U(a)$, что $\forall x \in U(a)$ справедливо равенство*

$$f(x) - (A(x-a) + f(a)) = o(x-a),\ x \to a,$$

то прямая $y = A(x-a) + f(a)$ называется касательной к графику функции $y = f(x)$ в точке a.

На рисунке 1 красным цветом изображена касательная. Обратим внимание, что график

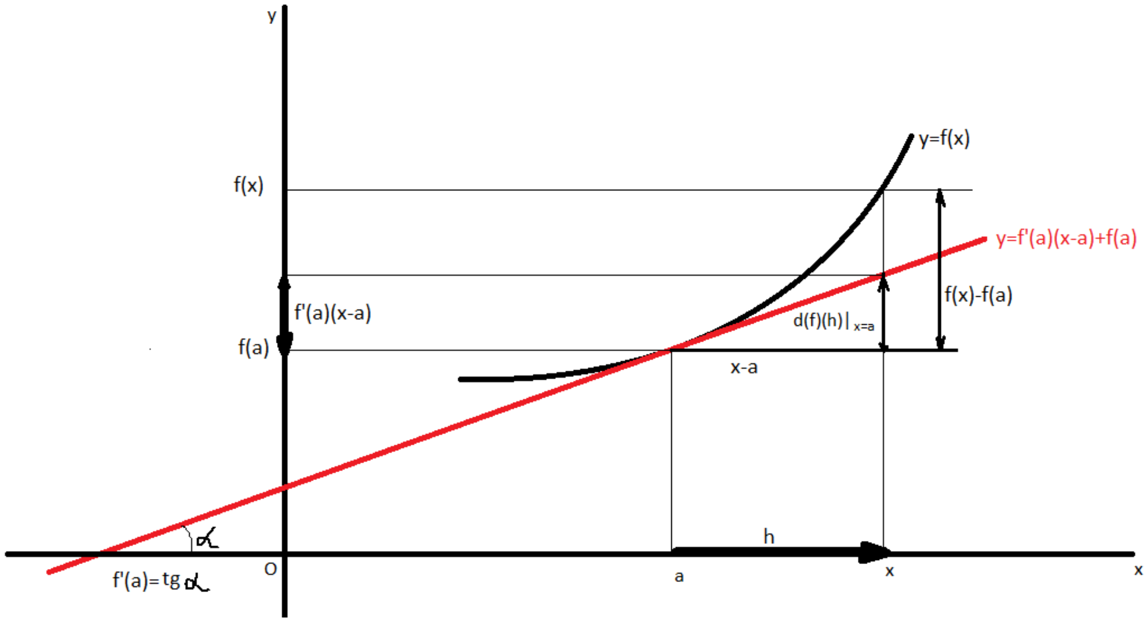

Рис. 1: Функция "сливается"с касательной.

функции и касательной неразличимы в некоторой окрестности.

При этом взаимное расположение графика функции $y = f(x)$ и графика касательной к ней в точке a могут быть весьма разнообразными. Например, в любой окрестности точки касания график функции может бесконечно много раз пересекать график касательной.

Пример 3. *Пусть*

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Имеем по определению:

$$f'(0) = \lim_{h\to 0} \frac{h^2 \sin(1/h) - 0}{h} = \lim_{h\to 0} h \sin(1/h) = 0,$$

поэтому касательная в начале координат к графику $y = f(x)$ задаётся уравнением $y = 0$, то есть совпадает с осью абсцисс. При этом при любом натуральном n в точке $x_n = \frac{1}{\pi n}$ график $y = f(x)$ пересекает ось абсцисс.

Непрерывность и дифференцируемость

Докажем простое следствие определения дифференцируемых функций.

Предложение 2. *Пусть функция f дифференцируема в точке a. Тогда f непрерывна в точке a.*

Доказательство. Из определения дифференцируемости следует, что

$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \to f(a), \ x \to a,$$

то есть предел функции в точке a равен её значению в этой точке, что по определению означает непрерывность. □

Следующий пример показывает, что обратное неверно: непрерывная в точке функция может быть недифференцируемой в этой точке. Согласно предложению 1 это равносильно тому, что такая функция в этой точке не имеет производную.

Пример 4. *Функция $y = |x|$ недифференцируема в нуле. Действительно,*

$$\lim_{h\to 0-}\frac{|h|}{h} = \lim_{h\to 0-}\frac{-h}{h} = -1,\ a\ \lim_{h\to 0+}\frac{|h|}{h} = \lim_{h\to 0-}\frac{h}{h} = 1.$$

Если бы производная в нуле существовала, то есть если бы существовал предел, то это было бы равносильно существованию и равенству односторонних пределов. Таким образом, у функции $y = |x|$ не существует производной при $x = 0$, что равносильно тому, что эта функция недифференцируема в нуле.

При этом отметим, что функция $y = |x|$ всё же является непрерывной в точке $x = 0$. Существуют функции, **непрерывные на всей прямой, но не дифференцируемые ни в одной точке.** Построение примера такой функции мы сейчас и обсудим.

Равномерная сходимость и пример Вейерштрасса

В этом разделе мы рассмотрим важное понятие равномерной сходимости, которое систематически будет изучаться при обсуждении функциональных последовательностей и рядов. При изучении этого понятия полезно повторить материал, касающийся равномерной непрерывности. Мы увидим, что в этом разделе в некоторых случаях также возникает зависимость от конкретных значений аргумента при изучении сходимости последовательности функций, а нас будут интересовать те случаи, когда такой зависимости нет.

Пусть $\{f_n(x)\}_{n=1}^{\infty}$ – последовательность функций. При каждом фиксированном x получится числовая последовательность, и можно ставить вопрос о её сходимости. Множество X всех тех x, при которых $\{f_n(x)\}_{n=1}^{\infty}$ сходится, называется *множеством сходимости* или *областью сходимости*, этой функциональной последовательности. Бывают последовательности функций, сходящиеся при всех вещественных значениях аргумента, а бывают такие, которые расходятся при любом значении аргумента (приведите примеры).

Таким образом, каждому $x \in X$ ставится в соответствие единственное число, являющееся пределом соответствующей числовой последовательности, то есть на множестве X определена функция $f : X \to \mathbb{R}$, которая задаётся равенством $f(x) = \lim\limits_{n\to\infty} f_n(x),\ x \in X$.

Оказывается, что многие свойства предельной функции f зависят от свойств функций f_n и характера сходимости последовательности функций $\{f_n\}$ к предельной.

Пример 5. *1) Рассмотрим последовательность $\{x^n\}_{n=1}^{\infty}$. При всех таких x, что $|x| < 1$, и при $x = 1$ эта последовательность сходится, а при остальных x предела нет, поэтому область сходимости X – это множество $(-1, 1]$. При этом*

$$\forall x : |x| < 1\ \lim_{n\to\infty} x^n = 0,$$

а при $x = 1$ $\lim\limits_{n\to\infty} x^n = 1$. Таким образом, предельная функция f задаётся следующим образом: $f(x) = \begin{cases} 0, & -1 < x < 1 \\ 1, & x = 1 \end{cases}$.

2) Теперь рассмотрим последовательность функций $\{\frac{x^n}{n}\}_{n=1}^{\infty}$. Здесь область сходимости $X = [-1, 1]$, так как $\forall x :\ |x| \leq 1\ |\frac{x^n}{n}| \leq \frac{1}{n} \to 0,\ n \to +\infty$, а при $|x| > 1$ последовательность расходится, так как числитель растёт экспоненциально, а знаменатель растёт, как степенная функция (см. пример $\lim\limits_{n\to+\infty} \frac{n^k}{a^n}$ $(a > 1)$ из семинарского листка). Таким образом, предельная функция f здесь – это константа 0 на отрезке $[-1, 1]$.

Обратим внимание, что предельная функция f получилась разрывной в точке 1 в первом пункте, а в пункте 2 f оказалась непрерывной на всей области сходимости. Для того, чтобы лучше понять причину этого, более внимательно рассмотрим характер сходимости последовательности из пункта 1. Зададимся числом $\varepsilon > 0$ и попробуем по этому ε подобрать такое $N \in \mathbb{N}$, чтобы при всех $n > N$ было выполнено неравенство $|f(x) - x^n| < \varepsilon$. Это N зависит не только от ε, но и от точки x. Например, в точке 0 любое N подойдет, так как при $x = 0$ у нас будет последовательность нулей, а чем ближе точка x к 1, тем больше натуральных n, для которых значения $f_n(x)$ превысят ε (при любом достаточно малом фиксированном $\varepsilon > 0$), поэтому выбор N зависит от того, в какой точке рассматривается последовательность. Для наглядного представления о характере сходимости см. рис. 2.

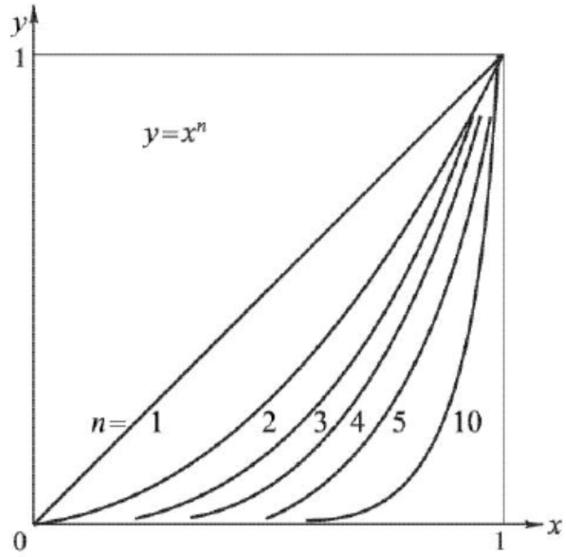

Рис. 2: Чем x ближе к 1, тем больше значения элементов последовательности

При этом в пункте 2 выбор N зависит только от $\varepsilon > 0$ так как сразу для всех $x \in [-1, 1]$ справедливо неравенство $|\frac{x^n}{n}| \leq \frac{1}{n}$ и как только $N = [1/\varepsilon]$, то при всех $n > N$ и сразу при всех $x \in [-1, 1]$ выполнено неравенство $|\frac{x^n}{n}| < \varepsilon$.

Как видим, от характера сходимости зависит непрерывность предельной функции f. Ниже мы сформулируем (без доказательства) теорему, в которой будут даны достаточные условия непрерывности предельной функции. Для этого нам потребуется следующее определение.

Определение 4. *Функциональная последовательность $\{f_n(x)\}_{n=1}^{\infty}$* ***сходится равномерно*** *к функции $f(x)$ на множестве X, если*

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} : \ \forall n > N \ \forall x \in X \ |f_n(x) - f(x)| < \varepsilon.$$

Обозначение: $f_n(x) \underset{X}{\rightrightarrows} f(x)$.

Теорема 1. *(**Непрерывность предела равномерно сходящейся последовательности**). Пусть X – область сходимости функциональной последовательности $\{f_n\}$. Пусть $\forall n \in \mathbb{N}$ $f_n(x) \in C(x_0)$, $x_0 \in X$. Пусть $f_n(x) \underset{X}{\rightrightarrows} f(x)$. Тогда $f(x) \in C(x_0)$.*

Отметим, что это условие достаточное, но не необходимое, то есть существуют неравномерно сходящиеся последовательности функций, пределы которых – непрерывные функции.

Упражнение. Доказать, что последовательность функций $f_n(x) = \frac{nx}{1+n^2x^2}$ неравномерно сходится к функции $f(x) = 0\ \forall x \in \mathbb{R}$.

С помощью теоремы о непрерывности предела равномерно сходящейся последовательности можно построить пример функции, которая в каждой точке непрерывна, но ни в одной точке не является дифференцируемой. Для этого рассмотрим следующую последовательность функций:

$$f_0(x) = \sin x,\ \ f_1(x) = \sin x + \frac{1}{2}\sin 8x,\ \ f_2(x) = \sin x + \frac{1}{2}\sin 8x + \frac{1}{4}\sin 64x,\ \ ...,$$

$$f_n(x) = \sum_{k=0}^{n} \frac{1}{2^k}\sin 8^k x.$$

Покажем, что эта последовательность функций имеет предел при каждом $x \in \mathbb{R}$, то есть областью сходимости является вся вещественная прямая. Действительно, при фиксированном x

$$\sum_{k=0}^{n} \left|\frac{1}{2^k}\sin 8^k x\right| \le \sum_{k=0}^{n} \frac{1}{2^k} \le 2,$$

то есть эти конечные суммы (называемые частичными суммами ряда $\sum\limits_{k=0}^{+\infty} \left|\frac{1}{2^k}\sin 8^k x\right|$) ограничены сверху, что по теореме Вейерштрасса для последовательностей означает сходимость последовательности функций к функции, обозначаемой $\sum\limits_{k=0}^{+\infty} \left|\frac{1}{2^k}\sin 8^k x\right|$. Таким образом, сходится ряд $\sum\limits_{k=0}^{+\infty} \frac{1}{2^k}\sin 8^k x$, а тогда можно проверить, что и ряд без модулей сходится, то есть по определению сходится его последовательность частичных сумм $\{f_n(x)\}_{n=0}^{+\infty}$. Так как наши рассуждения справедливы для любого $x \in \mathbb{R}$, то функция

$$w(x) = \sum_{k=0}^{\infty} \frac{1}{2^k}\sin(8^k x)$$

определена для любых вещественных чисел. Кроме того,

$$|w(x) - f_n(x)| = \left|\sum_{k=n+1}^{+\infty} \frac{1}{2^k}\sin 8^k x\right| \le \sum_{k=n+1}^{+\infty} \frac{1}{2^k} = \frac{1}{2^n},$$

поэтому для каждого $\varepsilon > 0$ существует такое $N = [\log_2(1/\varepsilon)]$, что при всех $n > N$ и при всех $x \in \mathbb{R}$ $|w(x) - f_n(x)| < \varepsilon$, то есть $f_n(x) \underset{X}{\rightrightarrows} w(x)$. Все функции $f_n(x)$ непрерывны, так как представляют собой суммы непрерывных синусов, поэтому по теореме о непрерывности предела равномерно сходящейся последовательности функция $w(x)$ непрерывна на всей оси.

Можно доказать, что при этом функция $w(x)$ ни в одной точке вещественной оси не является дифференцируемой.

Функция $w(x) = \sum\limits_{n=0}^{\infty} 2^{-n}\sin(8^n x)$ называется **функцией Вейерштрасса.** Отметим, что вопросы, связанные с построением непрерывной во всех точках, но нигде не дифференцируемой функции, подробно изучались в XIX веке, причём изначально предпринимались попытки доказать, что всякая непрерывная функция дифференцируема практически во всех точках, за исключением некоторых изолированных. В 1861 году Бернхард Риман на одной из своих лекций привёл следующий контрпример:

$$r(x) = \sum_{n=1}^{+\infty} \frac{\sin n^2 x}{n^2}.$$

Исследование этой функции на дифференцируемость – сложная задача, но в 1970 году всё же было доказано, что в некоторых рациональных точках эта функция дифференцируема.

Пример Вейерштрасса появился в 1872 году, причём было дано строгое доказательство непрерывности этой функции и того, что она нигде не дифференцируема. Отметим, что существуют и другие примеры функций с такими свойствами.

Доказательство того, что эта функция нигде не дифференцируема, мы проводить не будем, но для наглядности приведём её график.

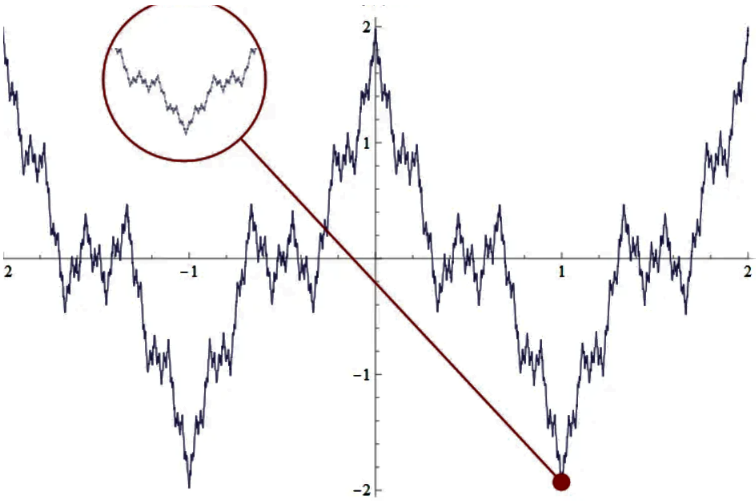

Рис. 3: функция Вейерштрасса w

Функция Вейерштрасса является самоподобной в том смысле, что увеличенные фрагменты графика в точности повторяют сам график. Ни в одной точке нельзя построить наклонную касательную, что и говорит о том, что функция Вейерштрасса недифференцируема. Конечно, это ни в коем случае не является строгим доказательством, а лишь позволяет лучше понять, как устроена такая функция.

Отметим ещё, что у этой функции в некоторых точках всё же есть бесконечная производная. В первой половине XX века А. С. Безикович построил пример непрерывной и нигде не дифференцируемой функции, у которой нет даже бесконечных производных.

Лекция 16

Правила дифференцирования и свойства дифференцируемых функций

Предложение 3. *Пусть функции f и g дифференцируемы в точке a. Тогда:*

1) *функция $\alpha f+\beta g$ дифференцируема в точке a при любых $\alpha,\ \beta\in\mathbb{R}$ и*

$$(\alpha f+\beta g)'(a)=\alpha f'(a)+\beta g'(a)$$

(свойство линейности);

2) *$f\cdot g$ дифференцируема в точке a и $(f\cdot g)'(a)=f'(a)g(a)+f(a)g'(a)$ (правило Лейбница);*

3) *если $g\neq 0$ в некоторой окрестности точки a, то $\frac{f}{g}$ дифференцируема в точке a и $\left(\frac{f}{g}\right)'(a)=\frac{f'(a)g(a)-g'(a)f(a)}{g^2(a)}$.*

Доказательство. **1)** Найдём производную функции $\alpha f+\beta g$ в точке a по определению, пользуясь арифметикой пределов:

$$\lim_{h\to 0}\frac{(\alpha f+\beta g)(a+h)-(\alpha f+\beta g)(a)}{h}=$$
$$=\lim_{h\to 0}\left(\alpha\frac{f(a+h)-f(a)}{h}+\beta\frac{g(a+h)-g(a)}{h}\right)=\alpha f'(a)+\beta g'(a).$$

2) Снова вычислим производную по определению, используя арифметику пределов:

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

3) По арифметике пределов имеем:

$$\lim_{h\to 0}\frac{f(a+h)/g(a+h)-f(a)/g(a)}{h}=\lim_{h\to 0}\frac{f(a+h)g(a)-f(a)g(a+h)}{h}\cdot\frac{1}{g(a+h)g(a)}=$$
$$=\frac{1}{(g(a))^2}\cdot\lim_{h\to 0}\left(g(a)\frac{f(a+h)-f(a)}{h}-f(a)\frac{g(a+h)-g(a)}{h}\right)=\frac{f'(a)g(a)-g'(a)f(a)}{g^2(a)}.$$

□

Приведём пример применения правил дифференцирования.

Пример 6. ***1)*** *Согласно пункту 3)*

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

Аналогично, $(\operatorname{ctg}x)'=-\frac{1}{\sin^2 x}=-1-\operatorname{ctg}^2 x$.

2) *Рассмотрим полезные функции: гиперболический синус $\operatorname{sh}x=\frac{e^x-e^{-x}}{2}$ и гиперболический косинус $\operatorname{ch}x=\frac{e^x+e^{-x}}{2}$. Согласно правилам дифференцирования*

$$(\operatorname{sh}x)'=\frac{1}{2}((e^x)'-(e^{-x})')=\frac{1}{2}\left(e^x-\left(\frac{1}{e^x}\right)'\right)=\frac{1}{2}\left(e^x-\frac{-e^x}{e^{2x}}\right)=\frac{e^x+e^{-x}}{2}=\operatorname{ch}x.$$

Аналогично, $(\operatorname{ch}x)'=\operatorname{sh}x$.

Отметим, что существует отдельный раздел математики – **гиперболическая тригонометрия**, во многом похожая на обычную тригонометрию. По аналогии с основным тригонометрическим тождеством можно записать **основное гиперболическое тождество:**

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1.$$

Упражнение. Показать, что для функции $y = \operatorname{sh} x$ существует обратная функция

$$\operatorname{arsh} x = \ln(x + \sqrt{x^2+1})$$

(обратный гиперболический синус), а для функции $y = \operatorname{ch} x$ таковой является функция

$$\operatorname{arch} x = \ln(x + \sqrt{x^2-1})\ (x \geq 1)$$

(обратный гиперболический косинус).

Следующая теорема также относится к правилам дифференцирования, но мы выделяем её в отдельную из-за большей сложности доказательства.

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

Доказательство. Так как функция g дифференцируема в точке $f(a)$, то при всех q из некоторой проколотой окрестности нуля

$$g(f(a)+q) - g(f(a)) = g'(f(a))q + \beta(q)q,\ \lim_{q\to 0} \beta(q) = 0.$$

Доопределим бесконечно малую функцию β в нуле, положив $\beta(0) := 0$, то есть доопределим функцию β в нуле по непрерывности. Теперь мы можем положить $q = f(a+h) - f(a)$. При этом при подстановке вместо q, принадлежащего проколотой окрестности нуля, приращения $f(a+h) - f(a)$ может возникнуть ситуация, при которой приращение для f равно нулю. Для того, чтобы при этом равенство для приращения g не нарушалось, мы и доопределили бесконечно малую функцию β в нуле по непрерывности. Отметим, что такое доопределение не нарушает дифференцируемость функции g, но позволяет корректно подставить $q = f(a+h) - f(a)$.

Тогда в силу дифференцируемости f (из которой вытекает, что

$$f(a+h) - f(a) = f'(a)h + \alpha(h)h$$

и $\lim\limits_{h\to 0} \alpha(h) = 0$) в точке a, получим:

$$\begin{aligned} &g(f(a+h)) - g(f(a)) = \\ &\quad = g'(f(a))(f(a+h) - f(a)) + \beta(f(a+h) - f(a))(f(a+h) - f(a)) = \\ &\quad = g'(f(a))f'(a)h + g'(f(a))\alpha(h)h + \beta(f'(a)h + \alpha(h)h)(f'(a)h + \alpha(h)h) = \\ &\quad = g'(f(a))f'(a)h + \delta(h)h. \end{aligned}$$

Мы использовали обозначение $\delta(h) = g'(f(a))\alpha(h) + \beta(f'(a)h + \alpha(h)h)(f'(a) + \alpha(h))$, а тогда из арифметики пределов и из бесконечной малости функций β и α следует, что $\delta(h) \to 0,\ h \to 0$. Таким образом, для функции $g \circ f$ получим при достаточно малых h :

$$g(f(a+h)) - g(f(a)) = g'(f(a))f'(a)h + \delta(h)h,$$

что даёт дифференцируемость функции $g \circ f$ в точке a по определению. Мы знаем, что это равносильно наличию у функции $g \circ f$ производной в точке a. Таким образом, получаем, что $(g \circ f)'(a) = g'(f(a))f'(a)$. □

В некоторой точке y, в которой дифференцируема функция g, мы можем записать дифференциал функции g в виде $dg(y) = g'(y)dy$ (см. предыдущую лекцию). Пусть f дифференцируема в точке x и $f(x) = y$. Теорема о производной сложной функции тогда в терминах дифференциалов запишется так (в точке x):

$$d(g \circ f)(x) = g'(f(x))f'(x)dx = g'(f(x))df(x) = g'(y)dy.$$

Таким образом, вне зависимости от того, является ли y независимой переменной или функцией, форма (то есть вид) первого дифференциала внешне не меняется. Это свойств дифференциала называется ***инвариантностью формы первого дифференциала.***

Упражнение. Доказать следующие формулы для дифференциалов (при тех же предположениях на функции, что и в правилах дифференцирования):

1) $d(\alpha f + \beta g) = \alpha df + \beta dg$;

2) $d(fg) = gdf + fdg$;

3) $d(f/g) = \frac{gdf - fdg}{g^2}$.

Вычислим ещё несколько производных, пользуясь теоремой о производной сложной функции.

Пример 7. *1)* $(a^x)' = (e^{x \ln a})' = e^{x \ln a} \ln a = a^x \ln a$;

2) $(\sin(\cos x))' = \sin' y \cdot y' = -\cos(\cos x) \sin x$, *где* $y = \cos x$.

Предложение 5. ***(Производная обратной функции).*** *Пусть f – непрерывная и строго монотонная функция, отображающая интервал I в интервал J. Пусть также f дифференцируема в точке $a \in I$ и $f'(a) \neq 0$. Тогда обратная функция f^{-1} дифференцируема в точке $b = f(a)$ и $(f^{-1})'(b) = \frac{1}{f'(a)}$.*

Доказательство. Существование, непрерывность и монотонность функции f^{-1} следует из теоремы об обратной функции. Воспользуемся определением производной, считая, что аргумент функции f^{-1} получает приращение q, а затем положим $q = f(a+h) - f(a)$ и воспользуемся теоремой о пределе композиции и арифметикой пределов:

$$\lim_{q \to 0} \frac{f^{-1}(f(a) + q) - f^{-1}(f(a))}{q} = \lim_{h \to 0} \frac{f^{-1}(f(a) + (f(a+h) - f(a))) - f^{-1}(f(a))}{f(a+h) - f(a)} =$$
$$= \lim_{h \to 0} \frac{a + h - a}{f(a+h) - f(a)} = \frac{1}{f'(a)}.$$

При этом $f(a+h) - f(a) \neq 0$ при $h \neq 0$, так как функция f строго монотонна. □

В дифференциалах формула запишется так: $df^{-1}(f(a)) = (df(a))^{-1}$ (что следует из формулы $df^{-1}(f(a))(h) = \frac{1}{f'(a)}h$, тогда как $df(a)(h) = f'(a)h$, то есть линейные функции $df(a)$ и $df^{-1}(f(a))$ взаимно обратны).

Прежде всего, отметим, что у этой теоремы простой геометрический смысл. Если производная функции f – это тангенс угла между касательной и положительным направлением оси Ox, то производная обратной функции – это тангенс угла наклона между касательной и положительным направлением оси Oy, который (сделайте рисунок!) равен котангенсу угле между касательной и положительным направлением оси Ox. Остаётся только вспомнить, что тангенс и котангенс одного угла взаимно обратны.

Приведём примеры применения последней теоремы.

Пример 8. *1) Пусть* $\operatorname{tg} y = x$*, тогда при* $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ $y = \operatorname{arctg} x$*. По теореме о производной обратной функции* $(\operatorname{arctg} x)' = \frac{1}{(\operatorname{tg} y)'} = \frac{1}{1+(\operatorname{tg} y)2} = \frac{1}{1+x^2}$*. Аналогично проверяется, что* $(\arcsin x)' = -(\arccos x)' = \frac{1}{\sqrt{1-x^2}}$.

2) Пусть $e^y = x$*, тогда* $\ln x = y$*.* $(\ln x)' = \frac{1}{(e^y)'} = \frac{1}{x}$.

3) Теперь, зная производную логарифма, при любом вещественном α *из теоремы о производной сложной функции имеем* $(x^\alpha)' = (e^{\alpha \ln x})' = e^{\alpha \ln x} \cdot \alpha \cdot \frac{1}{x} = \alpha x^{\alpha-1}$.

Упражнение. Найти, пользуясь теоремой о производной сложной функции, производные arsh и arch .

Теперь выпишем **таблицу производных,** все равенства в которой либо уже выведены с использованием наших теорем, либо выводятся аналогично. При этом в основании логарифма и показательной функции мы считаем, что $a > 0,\ a \neq 1$, а все функции рассматриваются на своих областях определения.

$$\mathbf{1})\ C' = 0\ \forall C = const;\ \mathbf{2})\ (x^\alpha)' = \alpha x^{\alpha-1}\ (x > 0);\ \mathbf{3})\ (e^x)' = e^x;\ \mathbf{4})\ (a^x)' = a^x \ln a (a > 0,\ a \neq 1);$$

$$\mathbf{5})\ (\ln x)' = 1/x\ (x > 0); \mathbf{6})\ (\log_a x)' = \frac{1}{x \ln a}\ (x > 0,\ a > 0,\ a \neq 1);\ \mathbf{7})\ (\sin x)' = \cos x;$$

$$\mathbf{8})\ (\cos x)' = -\sin x;\ \mathbf{9})\ (\operatorname{tg} x)' = \frac{1}{\cos^2 x} = \operatorname{tg}^2 x + 1\ (x \neq \frac{\pi}{2} + \pi k\ (k \in));$$

$$\mathbf{10})\ (\operatorname{ctg} x)' = -\frac{1}{\sin^2 x} = -\operatorname{ctg}^2 x - 1\ (x \neq \pi k\ (k \in));\ \mathbf{11})\ (\arcsin x)' = \frac{1}{\sqrt{1-x^2}}\ (-1 < x < 1);$$

$$\mathbf{12})\ (\arccos x)' = -\frac{1}{\sqrt{1-x^2}}\ (-1 < x < 1); \mathbf{13})\ (\operatorname{arctg} x)' = \frac{1}{1+x^2};$$

$$\mathbf{14})\ (\operatorname{arcctg} x)' = -\frac{1}{1+x^2};\ \mathbf{15})\ (\operatorname{sh} x)' = \operatorname{ch} x;\ \mathbf{16})\ (\operatorname{ch} x)' = \operatorname{sh} x.$$

Лекция 17

Основные теоремы дифференциального исчисления

Определение 5. *Пусть $\delta > 0$, $U_\delta(a)$ – δ-окрестность точки a Если функция f определена на множестве E, точка $a \in E$ и при всех $x \in U_\delta(a) \cap E$, выполнено неравенство $f(x) \le f(a)$, то a называется точкой локального максимума функции f. Если при тех же $x \in U_\delta(a) \cap E$ $f(x) \ge f(a)$, то f называется точкой локального минимума функции f. Если неравенство строгое, то a называется точкой строгого локального максимума или строгого локального минимума соответственно. Точки максимума и минимума называются точками локального экстремума (если неравенства нестрогие), и точками строгого локального экстремума, если неравенства строгие.*

У любой постоянной на числовом множестве функции любая точка её области определения является точкой локального экстремума (можно воспринимать их и как максимумы, и как минимумы). Если область определения функции содержит изолированные точки, то они являются точками локальных экстремумов. Точка $x = 0$ является строгим локальным минимумом функции $y = x^2$ и строгим локальным максимумом функции $y = -|x|$.

В изучаемых ниже теоремах важно иметь наглядное представление об их утверждениях, так что читателям полезно сделать чертежи при изучении формулировок.

Теорема 2. ***(Теорема Ферма)***. *Пусть функция f определена на интервале (a, b), дифференцируема в точке $c \in (a, b)$ и имеет в точке c локальный экстремум. Тогда $f'(c)$=0.*

Доказательство. Пусть в точке c функция f имеет максимум (для минимума доказательство аналогично). Тогда при всех $x > c$ справедливо неравенство $\frac{f(x)-f(c)}{x-c} \le 0$, а при всех $x < c$ – неравенство $\frac{f(x)-f(c)}{x-c} \ge 0$. Так как существует производная функции f в точке c, то есть существует предел $\lim\limits_{x\to c} \frac{f(x)-f(c)}{x-c} = f'(c)$, то существуют и равны друг другу оба односторонних предела: $\lim\limits_{x\to c-} \frac{f(x)-f(c)}{x-c} = f'(c) = \lim\limits_{x\to c+} \frac{f(x)-f(c)}{x-c}$. Таким образом, из предельного перехода в неравенствах имеем

$$(f'(c))^2 = \lim_{x\to c-} \frac{f(x)-f(c)}{x-c} \cdot \lim_{x\to c+} \frac{f(x)-f(c)}{x-c} \le 0,$$

что равносильно тому, что $f'(c) = 0$. □

Таким образом, в теореме Ферма утверждается, что если функция дифференцируема в точке экстремума, то её график в этой точке имеет касательную, параллельную оси абсцисс. Обратим внимание, что у непрерывной на отрезке функции экстремум может быть также и в точке, в которой производная не существует (например, $y = |x|$ на отрезке $[-1, 1]$.) При поиске экстремумов функции обычно изучают поведение функций в окрестности точек, где производная либо равна нулю, либо не существует. Точки, в которых производная равна нулю, не всегда являются точками экстремумов. Примером может служить функция $y = x^3$ при $x = 0$ (проверьте это).

Теорема 3. ***(Теорема Ролля)***. *Пусть:*

1) *Функция f определена и непрерывна на отрезке $[a, b]$.*

2) *Функция f дифференцируема на интервале (a, b).*

3) *$f(a) = f(b)$.*

Тогда существует такая точка $c \in (a, b)$, что $f'(c) = 0$.

Доказательство. Если функция непрерывна на отрезке $[a,b]$, то по второй теореме Вейерштрасса существуют такие точки $c_1 \in [a,b]$ и $c_2 \in [a,b]$, что $f(c_1) = m := \inf\limits_{a \le x \le b} f(x)$ и $f(c_2) = M := \sup\limits_{a \le x \le b} f(x)$. Если при этом $m = M = f(a)$, то функция f на отрезке $[a,b]$ является постоянной, поэтому в любой точке интервала (a,b) её производная равна нулю. Если $m < M$, то хотя бы одна из точек c_1 и c_2 лежит в интервале (a,b) и согласно теореме Ферма производная функции f в этой точке равна нулю. Эту точку и можно взять в качестве c. □

Все условия теоремы Ролля важны. Если отказаться от условия 1 (при выполнении условий 2 и 3), то можно рассмотреть функцию

$$f(x) = \begin{cases} x, \text{ если } x \in [0,1), \\ 0, \text{ если } x = 1 \end{cases}$$

(см. рисунок 4). Производная этой функции равна 1 в каждой точке интервала $(0,1)$.

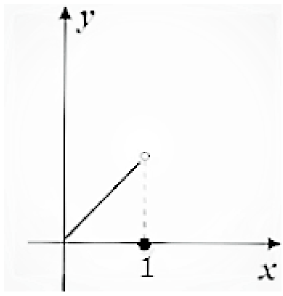

Рис. 4: Функция разрывна на отрезке, но дифференцируема на интервале

Если взять непрерывную на отрезке $[-1,1]$ функцию, но отказаться от дифференцируемости на интервале, то подойдёт функция $y = |x|$, но утверждение теоремы Ролля для неё не выполнено. Наконец, если отказаться от условия 3, то можно на любом отрезке просто взять функцию $y = x$, производная которой ни в одной точке не обращается в нуль.

Геометрический смысл теоремы Ролля состоит в том, что при условиях теоремы обязательно найдётся точка, в которой касательная к графику функции f параллельна оси абсцисс (см. рисунок 5).

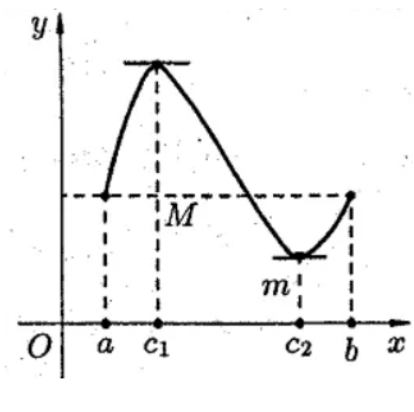

Рис. 5: Касательные параллельны оси Ox

Физический смысл теоремы Ролля состоит в том, что если x – время, а $f(x)$ – координата точки, движущейся по оси Oy, то при условии, что точка в момент времени b

возвращается в то же положение, откуда стартовала в момент a, должен наступить момент c, в который точка остановится, то есть её скорость будет равна нулю.

Теорема 4. ***(Теорема Лагранжа)***. *Пусть:*

1) *Функция f определена и непрерывна на отрезке $[a, b]$.*

2) *Функция f дифференцируема на интервале (a, b).*

Тогда существует такая точка $c \in (a, b)$, что $f(b) - f(a) = f'(c)(b - a)$.

Доказательство. Рассмотрим функцию g, определяемую равенством

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

Эта функция по свойствам непрерывных и дифференцируемых функций непрерывна на отрезке $[a, b]$, дифференцируема на интервале (a, b). Кроме того, $g(a) = 0 = g(b)$. Таким образом, функция g удовлетворяет всем условиям теоремы Ролля на $[a, b]$, поэтому существует такая точка $c \in (a, b)$, что $g'(c) = 0$, а это равносильно тому, что $f'(c) = \frac{f(b)-f(a)}{b-a}$. □

Отметим, что в доказательстве мы фактически вывели теорему Лагранжа из теоремы Ролля. Если функция принимает на концах отрезка равные значения, то легко видеть, что из теоремы Лагранжа следует теорема Ролля. Таким образом, теоремы Ролля и Лагранжа равносильны.

Геометрический смысл теоремы Лагранжа будет лучше понятен, если саму формулу из теоремы (называемую также *формулой конечных приращений*) переписать в виде $f'(c) = \frac{f(b)-f(a)}{b-a}$. Справа стоит тангенс угла наклона прямой, которая называется *секущей* (см. рисунок 6), а слева – тангенс угла наклона касательной. Таким образом, в теореме утверждается, что для дифференцируемой на интервале и непрерывной на отрезке функции найдётся точка, в которой касательная к графику функции f будет параллельна секущей, проходящей через точки $(a, f(a))$ и $(b, f(b))$.

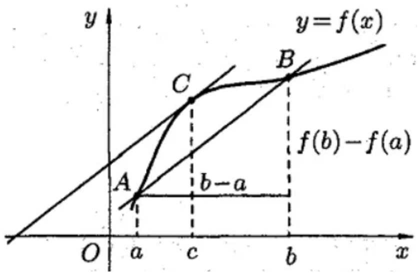

Рис. 6: Секущая AB параллельна касательной, проходящей через точку C

Физический смысл теоремы Лагранжа состоит в том, что если переменную x воспринимать как время, то величина $\frac{f(b)-f(a)}{b-a}$ – это средняя скорость материальной точки, движущейся по оси Oy, за промежуток времени $[a, b]$, а $f'(c)$ – мгновенная скорость точки в момент времени c. Таким образом, существует такой момент времени, что средняя скорость точки на промежутке времени $[a, b]$ равна её мгновенной скорости в некоторый момент времени c. Докажем два полезных следствия теоремы Лагранжа.

Предложение 6. ***(1-е следствие)*** *Пусть функция f определена и дифференцируема на интервале (a, b) и $f'(x) = 0\ \forall x \in (a, b)$. Тогда функция f постоянна на (a, b).*

Доказательство. Пусть $x_1,\ x_2 \in (a, b),\ x_1 < x_2$. Тогда на отрезке $[x_1,\ x_2]$ к функции f применима теорема Лагранжа, поэтому $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$, где $c \in (x_1,\ x_2)$. По условию $f'(c) = 0$, поэтому $f(x_1) = f(x_2)$. Так как рассуждение верно для любых точек на интервале (a, b), то, зафиксировав x_1 и беря произвольные значения $x_2 \in (a, b)$, видим, что во всех точках интервала функция f принимает равные значения. □

Предложение 7. ***(2-е следствие)*** *Пусть функция f определена и дифференцируема на интервале (a, b).*

1) *Функция f не убывает на этом интервале тогда и только тогда, когда*

$$f'(x) \geq 0\ \forall x \in (a, b).$$

2) *Функция f не возрастает на этом интервале тогда и только тогда, когда*

$$f'(x) \leq 0\ \forall x \in (a, b).$$

Доказательство. Проведём доказательство для случая неубывающей функции (для невозрастающей всё аналогично).
Необходимость. Если функция не убывает на интервале (a, b), то для любых точек

$$x_1,\ x_2 \in (a, b),$$

имеем:

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq 0.$$

Зафиксировав точку x_1, получим в силу предельного перехода в неравенствах, что

$$\lim_{x_2 \to x_1} \frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x_1) \geq 0.$$

В силу произвольности точек x_1 и x_2 получаем неотрицательность f' в каждой точке интервала (a, b).

Достаточность. Если $f'(x) \geq 0\ \forall x \in (a, b)$, то для любых точек $x_1,\ x_2 \in (a, b),\ x_1 < x_2$ по теореме Лагранжа найдётся такая $c \in (x_1,\ x_2)$, что

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) \geq 0$$

то есть f не убывает. □

Следующая теорема Коши является обобщением теоремы Лагранжа.

Теорема 5. ***(Теорема Коши)****. Пусть функции f и g:*

1) *определены и непрерывны на отрезке $[a, b]$;*

2) *дифференцируемы на интервале (a, b);*

3) *$g'(x) \neq 0\ \forall\, x \in\ (a, b)$.*

Тогда существует такая точка $c \in (a, b)$, что $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$.

Доказательство. Отметим прежде всего, что $g(a) \neq g(b)$, так как в противном случае функция g удовлетворяет всем условиям теоремы Ролля, поэтому условие 3) выполняться не может.

Снова сведём доказательство к теореме Ролля. Рассмотрим вспомогательную функцию G, определяемую равенством

$$G(x) = (g(b) - g(a))f(x) - (f(b) - f(a))g(x).$$

Эта функция непрерывна на отрезке $[a, b]$, дифференцируема на интервале (a, b) и

$$G(a) = g(b)f(a) - g(a)f(b) = G(b),$$

то есть все условия теоремы Ролля выполнены. Тогда найдётся такая точка $c \in (a, b)$, что

$$G'(c) = (g(b) - g(a))f'(c) - (f(b) - f(a))g'(c) = 0,$$

откуда и получается доказываемое равенство. □

Отметим, что если убрать условие 3 теоремы Коши, то доказано, что будет справедливо равенство $f'(c)(g(b) - g(a)) = g'(c)(f(b) - f(a))$. Это равенство равносильно тому, что

$$\begin{vmatrix} f(b) - f(a) & f'(c) \\ g(b) - g(a) & g'(c) \end{vmatrix} = 0.$$

Равенство нулю определителя означает, что столбцы матрицы образуют коллинеарные векторы, а отсюда можно получить один из возможных *физических смыслов теоремы Коши*: если вектор $(f(x), g(x))$ – это закон движения тела на плоскости, а x играет роль времени, то вектор $(f(b) - f(a), g(b) - g(a))$ будет вектором перемещения за время $b - a$, а вектор $(f'(c), g'(c))$ – это вектор мгновенной скорости в момент времени c. Таким образом, в теореме Коши утверждается, что вектор перемещения коллинеарен вектору мгновенной скорости, взятому в какой-то момент времени между моментами a и b. В трехмерном пространстве в общем случае это не верно (например, при движении по винтовой линии на поверхности цилиндра).

Упражнения.

1. Вывести из теоремы Коши теорему Лагранжа.

2. Рассмотрим следующее рассуждение при доказательстве теоремы Коши: применяя к числителю и знаменателю теорему Лагранжа, получим цепочку равенств

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

чем всё и доказано. Почему оно неверно?

3. Рассмотрим следующее рассуждение применительно к доказательству необходимости в предложении 2. *Если функция не убывает на интервале* (a, b), *то для любых точек*

$$x_1,\ x_2 \in (a, b),\ x_1 < x_2$$

имеем: $0 \le f(x_2) - f(x_1) = f'(c)(x_2 - x_1) \Leftrightarrow f'(c) \ge 0$, *где* $c \in (x_1, x_2)$. *Так как это верно для любых точек интервала, а* c *всегда лежит между* x_1 *и* x_2, *то в производная неотрицательна в каждой точке* (a, b).

Почему это рассуждение не может являться доказательством необходимости предложения 2?

4. Доказать, что если $f'(x) > 0\ \forall x \in (a, b)$, то f строго возрастает на интервале (a, b), но обратное неверно.

Теоремы Ролля, Лагранжа и Коши иногда называют теоремами о среднем (в дифференциальном исчислении), так как в каждой из них речь идёт о точке c, лежащей между концами отрезка $[a, b]$.

Лекция 18

Раскрытие неопределённостей

Сейчас мы изучим часто используемые при подсчёте пределов утверждения. Отметим, что хотя их и принято называть правилами Лопиталя, но считается, что открыл их Иоганн Бернулли, сообщивший в письме Гийому Лопиталю об этих правилах.

Так как в теоремах ниже рассматривается соотношение $\frac{f(x)}{g(x)}$, то считаем, что оно определено, то есть $g(x) \neq 0$ на рассматриваемых промежутках.

Теорема 6. ***(1-е правило Лопиталя, неопределённость вида $\frac{0}{0}$)***
Пусть:

1) *функции f и g определены и дифференцируемы на интервале (a, b);*

2) $\lim\limits_{x\to b-} f(x) = \lim\limits_{x\to b-} g(x) = 0$;

3) $g'(\xi) \neq 0 \; \forall\, \xi \in \; (a, b)$;

4) *существует конечный или бесконечный предел* $\lim\limits_{x\to b-} \frac{f'(x)}{g'(x)}$.

Тогда справедливо равенство $\lim\limits_{x\to b-} \frac{f(x)}{g(x)} = \lim\limits_{x\to b-} \frac{f'(x)}{g'(x)}$.

Доказательство. Сначала считаем, что

$$\lim_{x\to b-} \frac{f'(x)}{g'(x)} = A,$$

где A – конечное число.

Рассмотрим точки $x, y \in (a, b),\; x < y$. Так как $g'(\xi) \neq 0 \; \forall \xi \in (a, b)$, то $g(y) - g(x) \neq 0$. Таким образом, мы можем применить к функциям f и g теорему Коши:

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(c)}{g'(c)},\; c \in (x, y).$$

Домножим обе части этого равенства на $g(x)-g(y)$, а затем разделим на $g(x)$. В результате получим

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

Так как отношение производных имеет предел, то

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \; \forall x, y : b > y > x > b - \delta \;\; \left|\frac{f'(c)}{g'(c)} - A\right| < \varepsilon,$$

так как c лежит между x и y. Таким образом, имеем

$$\left|\frac{f(x)}{g(x)} - A\right| = \left|\frac{f'(c)}{g'(c)} - A - \frac{f'(c)}{g'(c)}\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}\right| \leq \left|\frac{f'(c)}{g'(c)} - A\right| + \left|\frac{f'(c)}{g'(c)}\right| \cdot \left|\frac{g(y)}{g(x)}\right| + \left|\frac{f(y)}{g(x)}\right|.$$

Зафиксировав такое x, чтобы выполнялись выписанные выше неравенства, и учитывая условие 2 теоремы, получим, что $\lim\limits_{y\to b-} |\frac{f(y)}{g(x)}|$ и $\lim\limits_{y\to b-} |\frac{g(y)}{g(x)}| = 0$. Возьмем такое $\delta > 0$, что $\left|\frac{g(y)}{g(x)}\right| < \varepsilon$ и $\left|\frac{f(y)}{g(x)}\right| < \varepsilon$. Тогда, используя то, что имеющая предел функция ограничена, и выбирая такую константу $C > 0$, что $\left|\frac{f'(c)}{g'(c)}\right| \leq C$, получим

$$\left|\frac{f'(c)}{g'(c)} - A\right| + \left|\frac{f'(c)}{g'(c)}\right| \cdot \left|\frac{g(y)}{g(x)}\right| + \left|\frac{f(y)}{g(x)}\right| < \varepsilon + C \cdot \varepsilon + \varepsilon = (2 + C)\varepsilon,$$

откуда $\lim\limits_{x\to b-}\frac{f(x)}{g(x)}=A$.

Если теперь предположить, что предел отношения производных равен ∞, то мы можем зафиксировать такое x, что выражение $\frac{f'(c)}{g'(c)}$ будет достаточно большим при всех $c\in(x,y)$, а y мы можем выбрать таким, что выражение $1-\frac{g(y)}{g(x)}$ достаточно близко к 1, а $\frac{f(y)}{g(x)}$ – к нулю. Тогда, снова используя равенство

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

получим, что и $\left|\frac{f(x)}{g(x)}\right|$ достаточно велико (рекомендуется расписать через ε и δ, как выше при конечном A). Таким образом, теорема доказана. □

Замечания. 1) В первом правиле Лопиталя можно заменить условие $x\to b-$ на $x\to a+$. При этом для доказательства мы просто сделаем замену переменной $t=a+b-x$ и сведём всё к только что доказанной теореме.

2) С учётом этого факта можно рассмотреть двустороннюю окрестность точки a и сформулировать правило Лопиталя уже не для одностороннего, а для обычного предела. При этом доказательство будет следовать из первого замечания и теоремы 1.

Сформулируем второе правило Лопиталя.

Теорема 7. ***(2-е правило Лопиталя, неопределённость вида*** $\frac{\infty}{\infty}$***)***
Пусть:

1) *функции f и g определены и дифференцируемы на интервале (a,b);*
2) $\lim\limits_{x\to b-}g(x)=\infty$;
3) $g'(\xi)\neq 0\ \forall\,\xi\in\ (a,b)$;
4) *существует конечный или бесконечный предел* $\lim\limits_{x\to b-}\frac{f'(x)}{g'(x)}$.

Тогда справедливо равенство $\lim\limits_{x\to b-}\frac{f(x)}{g(x)}=\lim\limits_{x\to b-}\frac{f'(x)}{g'(x)}$.

Доказательство. Снова сначала считаем, что

$$\lim_{x\to b-}\frac{f'(x)}{g'(x)}=A,$$

где A – конечное число.

Рассмотрим точки $x,y\in(a,b)$, но в этот раз пусть $x>y$. Так как $g'(\xi)\neq 0\ \forall\xi\in(a,b)$, то $g(x)-g(y)\neq 0$. Таким образом, мы можем применить к функциям f и g теорему Коши:

$$\frac{f(x)-f(y)}{g(x)-g(y)}=\frac{f'(c)}{g'(c)},\ c\in(y,x).$$

Снова преобразуем равенство к виду

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

Так как отношение производных имеет предел, то снова

$$\forall\varepsilon>0\ \exists\delta>0:\ \forall x,y: b>x>y>b-\delta\ \left|\frac{f'(c)}{g'(c)}-A\right|<\varepsilon,$$

так как c лежит между x и y. Таким образом, имеем

$$\left|\frac{f(x)}{g(x)}-A\right|=\left|\frac{f'(c)}{g'(c)}-A-\frac{f'(c)}{g'(c)}\frac{g(y)}{g(x)}+\frac{f(y)}{g(x)}\right|\leq\left|\frac{f'(c)}{g'(c)}-A\right|+\left|\frac{f'(c)}{g'(c)}\right|\cdot\left|\frac{g(y)}{g(x)}\right|+\left|\frac{f(y)}{g(x)}\right|.$$

На этот раз фиксируем такое y (то есть $f(y)$ и $g(y)$ – это фиксированные числа), чтобы выполнялись выписанные выше неравенства, и учитывая условие 2 теоремы, получим, что $\lim\limits_{x\to b-} |\frac{f(y)}{g(x)}| = 0$ и $\lim\limits_{x\to b-} |\frac{g(y)}{g(x)}| = 0$. Возьмем достаточно близкие к b значения x, чтобы $\left|\frac{g(y)}{g(x)}\right| < \varepsilon$ и $\left|\frac{f(y)}{g(x)}\right| < \varepsilon$. Тогда, используя то, что имеющая предел функция ограничена и выбирая такую константу $C > 0$, что $\left|\frac{f'(c)}{g'(c)}\right| \leq C$, получим

$$\left|\frac{f'(c)}{g'(c)} - A\right| + \left|\frac{f'(c)}{g'(c)}\right| \cdot \left|\frac{g(y)}{g(x)}\right| + \left|\frac{f(y)}{g(x)}\right| < \varepsilon + C\cdot\varepsilon + \varepsilon = (2+C)\varepsilon,$$

откуда $\lim\limits_{x\to b-} \frac{f(x)}{g(x)} = A$.

Рассуждения для бесконечного предела аналогичны тем, что выписаны в первом правиле Лопиталя. Таким образом, теорема доказана. □

Замечания, аналогичные сделанным после первого правила Лопиталя, справедливы и здесь.

Приведём важные примеры применения правила Лопиталя.

Пример 9. *1) Согласно второму правилу Лопиталя при всех натуральных n*

$$\lim_{x\to+\infty} \frac{x^n}{e^x} = \lim_{x\to+\infty} \frac{nx^{n-1}}{e^x} = \ldots = \lim_{x\to+\infty} \frac{n!}{e^x} = 0.$$

Это означает следующий важный факт: ***при $x \to +\infty$ экспонента растёт быстрее степени.***

2) Аналогично, $\lim\limits_{x\to+\infty} \frac{x^n}{\ln x} = \lim\limits_{x\to+\infty} \frac{nx^{n-1}}{1/x} = +\infty$. Это значит, что ***при $x \to +\infty$ степень растёт быстрее логарифма.***

3) Некоторые неопределённости можно сводить к таким, для которых применимо правило Лопиталя:

$$\lim_{x\to0+} x^x = \{0^0\} = \lim_{x\to0+} e^{x\ln x} = e^{\lim\limits_{x\to0+} x\ln x} = e^{\{0\cdot(-\infty)\}} = e^{\lim\limits_{x\to0+} \frac{\ln x}{\left(\frac{1}{x}\right)}} =$$

$$= e^{\left\{\frac{-\infty}{+\infty}\right\}} = e^{\lim\limits_{x\to0+} \frac{\frac{1}{x}}{\frac{-1}{x^2}}} = e^{\lim\limits_{x\to0+} -\frac{x^2}{x}} = e^0 = 1.$$

4) Покажем, что ***из существования предела отношения функций не следует существование предела отношения производных этих функций.*** *Действительно,*

$$\lim_{x\to+\infty} \frac{x-\sin x}{2x+\sin x} = \lim_{x\to+\infty} \frac{1-\frac{\sin x}{x}}{2+\frac{\sin x}{x}} = \frac{1}{2},$$

то есть у отношения функций $f(x) = x - \sin x$ и $g(x) = 2x + \sin x$ есть предел. При этом отношение производных $\frac{1-\cos x}{2+\cos x}$ предела не имеет, так как на последовательности

$$a_n = 2\pi n,\ n \in \mathbb{N}$$

предел этого отношения равен 0, а на последовательности

$$b_n = \pi + 2\pi n,\ n \in \mathbb{N}$$

этот предел равен 2, то есть предела у отношения производных нет в силу определения по Гейне.

Производные высших порядков

Пусть функция f дифференцируема в каждой точке множества E. Тогда её производная f' определена на E и можно ставить вопрос о её дифференцируемости. Отметим, что производной функции f нулевого порядка называется сама функция f. Производная порядка n определяется индуктивно.

Определение 6. *Если у функции f в некоторой окрестности точки a существует производная $n-1$-го порядка $f^{(n-1)}$ (то есть эта производная представляет собой определённую в некоторой окрестности точки a функцию) и эта производная дифференцируема в точке a, то производная функции f порядка n в точке a – это производная функции $f^{(n-1)}$ в точке a : $f^{(n)}(a) := (f^{(n-1)})'(a)$. В терминах пределов:*

$$f^{(n)}(a) = \lim_{h\to 0} \frac{f^{(n-1)}(a+h) - f^{(n-1)}(a)}{h}.$$

Ещё раз отметим, что, производная порядка $n-1$ определена в окрестности точки a, иначе символ $f^{(n-1)}(a+h)$ не имеет смысла.

Пример 10. *Рассмотрим функцию $f(x) = \begin{cases} x^2, x \in \mathbb{Q}, \\ 0, x \in \mathbb{I} \end{cases}$. Она непрерывна только в точке 0 и $f'(0) = \lim\limits_{h\to 0} \frac{f(h)-f(0)}{h} = \begin{cases} \lim\limits_{h\to 0} \frac{h^2}{h} = 0, h \in \mathbb{Q}, \\ \lim\limits_{h\to 0} \frac{0}{h} = 0, h \in \mathbb{I}. \end{cases}$ Таким образом, $f'(0) = 0$. При этом $f''(0)$ не существует, так как f' не определена ни в какой окрестности точки 0.*

В этом примере было рассмотрено произведение функции Дирихле и квадратичной функции, то есть $f(x) = x^2 \cdot D(x)$. Важно, что здесь берётся именно x^2, так как, например, при умножении функции Дирихле на x мы бы получили непрерывную в нуле функцию, которая не является дифференцируемой.

Контрольный вопрос. Сколько раз дифференцируема в нуле функция

$$g(x) = \begin{cases} x^{10}, x \in \mathbb{Q} \\ 0, x \in \mathbb{I} \end{cases},$$

то есть $g(x) = x^{10} \cdot D(x)$?

Следующая теорема позволяет находить производные высших порядков произведения функций. Она является обобщением формулы для вычисления первой производной произведения двух функций.

Предложение 8. ***(обобщенное правило Лейбница).*** *Пусть функции u и v n раз дифференцируемы в точке a. Тогда их произведение также n раз дифференцируемо в точке a и*

$$(u \cdot v)^{(n)}(a) = \sum_{k=0}^{n} C_n^k u^{(k)}(a) v^{(n-k)}(a). \tag{1}$$

Доказательство. Используем индукцию. При $n = 1$ $(uv)'(a) = u'(a)v(a) + v'(a)u(a)$ по правилам дифференцирования, так что база индукции доказана. Пусть формула верна при $n = m-1$, то есть

$$(uv)^{(m-1)}(a) = \sum_{k=0}^{m-1} C_{m-1}^k u^{(k)}(a) v^{(m-1-k)}(a). \tag{2}$$

Нам нужно доказать, что эта формула верна при $n = m$. Для этого мы должны продифференцировать равенство (2):

$$(uv)^{(m)}(a) = ((uv)^{(m-1)}(a))' = \left(\sum_{k=0}^{m-1} C_{m-1}^k u^{(k)}(a) v^{(m-1-k)}(a)\right)' =$$
$$= \sum_{k=0}^{m-1} C_{m-1}^k u^{(k+1)}(a) v^{(m-1-k)}(a) + \sum_{k=0}^{m-1} C_{m-1}^k u^{(k)}(a) v^{(m-k)}(a) =$$
$$= \sum_{k=1}^{m} C_{m-1}^{k-1} u^{(k)}(a) v^{(m-k)}(a) + \sum_{k=0}^{m-1} C_{m-1}^k u^{(k)}(a) v^{(m-k)}(a) =$$
$$= u(a)v^{(m)}(a) + \sum_{k=1}^{m-1} (C_{m-1}^{k-1} + C_{m-1}^k) u^{(k)}(a) v^{(m-k)}(a) + u^{(m)}(a)v(a) =$$
$$= \sum_{k=0}^{m} C_m^k u^{(k)}(a) v^{(m-k)}(a).$$

В третьей строке этой цепочки равенств мы поменяли индексы суммирования, а в конце воспользовались равенством $C_{m-1}^{k-1} + C_{m-1}^k = C_m^k$. Таким образом, теорема доказана по индукции. □

Приведём примеры применения этой формулы.

Пример 11. *Пусть $f(x) = (x^2 - x)e^{2x}$. Найдём $f^{(2025)}(0)$. Пусть $u(x) = x^2 - x$, а $v(x) = e^{2x}$. Так как $u'''(x) = u^{(4)}(x) = ... = u^{(2025)}(x) = 0$, то при применении формулы (1) справа останется лишь три слагаемых. Таким образом, получим:*

$$f^{(2025)}(0) = (2^{2025}(x^2 - x)e^{2x} + 2^{2024} \cdot 2025 \cdot (2x-1)e^{2x} + 2^{2023} \cdot 2025 \cdot 2024 \cdot e^{2x})|_{x=0} =$$
$$= 2^{2023} \cdot 4094550.$$

Формула Лейбница может быть обобщена на случай произведения большего числа дифференцируемых функций.

При вычислении производных больших порядков можно также применять индукцию и преобразования выражений, задающих функции. Соответствующие примеры будут разобраны на семинарах.

Лекция 19

Локальная формула Тейлора

Определение 7. *Многочленом Тейлора n раз дифференцируемой в точке a функции называется многочлен*

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

Отметим, что производные в точке a многочлена Тейлора совпадают с производными самой функции f в точке a (проверьте это).

Теорема 8. *Пусть функция f дифференцируема n раз в точке a. Тогда*

$$f(x) - T_n(x) = o((x-a)^n),\ x \to a.$$

Доказательство. Ещё раз важно прочитать материал про o-малое. Нам нужно доказать, что

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

Применяя к этому пределу правило Лопиталя, получим после $(n-1)$-кратного дифференцирования числителя и знаменателя:

$$\lim_{x\to a} \frac{f(x) - T_n(x)}{(x-a)^n} = \lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a) - f^{(n)}(a)(x-a)}{n!(x-a)} =$$
$$= \frac{1}{n!}\left(\lim_{x\to a} \frac{f^{(n-1)}(x) - f^{(n-1)}(a)}{(x-a)} - f^{(n)}(a)\right) = \frac{1}{n!}(f^{(n)}(a) - f^{(n)}(a)) = 0.$$

□

В доказательстве нельзя было ещё раз продифференцировать, так как производная порядка n в точке $x \neq a$ может быть не определена, ибо в условии сказано, что функция n раз дифференцируема в самой точке a, а это значит (см. определение производной n-го порядка), что в окрестности точки a определены только первые $n-1$ производные.

Таким образом, n раз дифференцируемая в точке функция хорошо приближается своим многочленом Тейлора в достаточно малой окрестности точки a. Можно записать равенство из условия теоремы в таком виде: $f(x) = T_n(x) + o((x-a)^n),\ x \to a$. Так как это равенство справедливо при достаточно близких к a значениях x, то само равенство называется *локальной формулой Тейлора.* Ещё одно название: *формула Тейлора с остаточным членом в форме Пеано.* Под остаточным членом подразумевается $o((x-a)^n)$, то есть разница между функцией и её многочленом Тейлора.

Отметим также, что если $n = 1$, то локальная формула Тейлора превращается в равенство из определения дифференцируемой функции (см. предыдущие лекции). Используя это наблюдение, доказательство теоремы 1 можно провести, используя индукцию вместо правила Лопиталя.

Упражнение. Докажите теорему 1 по индукции.

Теперь с помощью этой теоремы легко обосновываются формулы (9) – (13) из предыдущих лекций.

1) Докажем, что если $a = 0$, а $f(x) = e^x$, то

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n),\ x \to 0. \tag{9}$$

Действительно, при $x = 0$ справедливы равенства $e^x = (e^x)' = (e^x)'' = ... = (e^x)^{(n)} = 1$. Беря в качестве $f^{(k)}(a)$ значения k-й производной экспоненты в нуле при $k = 1,\ ...,\ n$ получим формулу (9).

2) Отметим, что для чётных производных синуса в нуле верны равенства

$$\sin 0 = \sin''(0) = ... = \sin^{(2k)}(0) = 0,$$

а для нечётных в нуле – равенства

$$\sin'(0) = \cos 0 = 1,\ \sin'''(0) = -\cos 0 = -1,\ ...,\ \sin^{(2k-1)}(0) = (-1)^{k-1}\cos 0 = (-1)^{k-1}$$

при всех натуральных k, то есть чётные производные синуса в точке $a = 0$ равны нулю, а нечётные в той же точке чередуются, то есть первая равна 1, вторая равна -1, третья снова равна 1 и так далее. Поэтому для синуса получаем формулу

$$\sin x = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + o(x^{2n+1}),\ x \to 0. \tag{10}$$

Отметим, что следующее слагаемое в локальной формуле Тейлора имеет степень $2n+3$, поэтому остаточный член можно записать и в виде $o(x^{2n+2}),\ x \to 0$.

3) Для косинуса формула Тейлора выводится также, как для синуса (проделайте это в качестве упражнения). В итоге получаем

$$\cos x = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + o(x^{2n}),\ x \to 0; \tag{11}$$

Здесь, аналогично случаю с синусом, допустимо писать $o(x^{2n+1}),\ x \to 0$.

4) Найдём последовательные производные логарифма:

$$(\ln(1+x))' = \frac{1}{1+x},\ (\ln(1+x))'' = \frac{-1}{(1+x)^2},$$
$$(\ln(1+x))''' = \frac{(-1)^2 \cdot 1 \cdot 2}{(1+x)^3},\ (\ln(1+x))^{(4)} = \frac{(-1)^3 \cdot 1 \cdot 2 \cdot 3}{(1+x)^4}.$$

Отсюда с помощью метода математической индукции устанавливается, что

$$(\ln(1+x))^{(n)} = \frac{(-1)^{n-1} \cdot (n-1)!}{(1+x)^n},$$

а в нуле тогда имеем

$$(\ln(1+x))^{(n)}|_{x=0} = \frac{(-1)^{n-1} \cdot (n-1)!}{(1+0)^n} = (-1)^{n-1} \cdot (n-1)!.$$

Тогда слагаемое многочлена Тейлора $\frac{f^{(k)}(a)}{k!}(x-a)^k$ при $f(x) = \ln(1+x),\ a = 0$ примет вид $\frac{(-1)^{k-1}x^k}{k}$, а формула Тейлора для логарифма будет такова:

$$\ln(1+x) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k + o(x^n),\ x \to 0. \tag{12}$$

5) Аналогично тому, как это сделано для логарифма, обоснуйте формулу

$$(1+x)^\alpha = \sum_{k=0}^{n} \frac{\alpha(\alpha-1)(\alpha-2)\cdot\ldots\cdot(\alpha-k+1)}{k!}x^k + o(x^n),\ x\to 0\ (\alpha\in\mathbb{R}). \tag{13}$$

Итак, теперь равенства, выписанные нами при изучении O-символики, обоснованы. Это пять основных равенств, а с их помощью можно получать локальные формулы Тейлора для других функций.

Формула Тейлора с с остаточным членом в общей форме. Остаточные члены в форме Лагранжа и Коши.

Если разницу между функцией и её многочленом Тейлора записать в виде формы Пеано, то никакой дополнительной информации, кроме той, что эта разница есть бесконечно малая величина при $x\to a$, мы не имеем. Иногда, например, при оценке погрешности, важно иметь более подробную информацию о поведении выражения

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

Эту информацию можно получить, имея более подробные сведения о самой функции f.

Теорема 9. ***(Формула Тейлора с остаточным членом в общей форме)****. Пусть в каждой точке отрезка с концами a и x функция f имеет n непрерывных производных, а в каждой точке интервала с концами a и x f дифференцируема $n+1$ раз. Пусть также функция g непрерывна на отрезке с концами a и x и дифференцируема в каждой точке интервала с концами a и x, причём $g'(\xi)\neq 0$ в любой точке ξ этого интервала. Тогда*

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

где c – какая-то точка из интервала с концами a и x.

Доказательство. Наша цель – применить теорему Коши. Для этого рассмотрим функцию

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

которая (проверьте!) удовлетворяет условиям теоремы Коши на отрезке с концами a и x. Вычислим производную этой функции:

$$G'(t) = f'(t) + \sum_{k=1}^{n}\left(\frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1}\right) = \frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Применяя к функциям $G(t)$ и $g(t)$ теорему Коши на отрезке с концами a и x, получим:

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

где c – точка на интервале с концами a и x. При этом $G(x)=f(x)$, а $G(a)=T_n(x;f,a)$, поэтому

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

чем всё и доказано. □

Положим теперь $g(t) = (x-t)^p$, где t лежит между точками a и x. Эта функция удовлетворяет всем условиям последней теоремы, поэтому можем подставить эту функцию в выражение для остаточного члена:

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

Точка c лежит между точками a и x, поэтому её можно представить в виде $c = a+\theta(x-a)$, где $0 < \theta < 1$. Тогда

$$R_n(x; f, a) = \frac{f^{(n+1)}(a+\theta(x-a))}{n!p}(1-\theta)^{n+1-p}(x-a)^{n+1},\ 0 < \theta < 1. \tag{3}$$

Правая часть (3) – *остаточный член в форме Шлёмильха – Роша.*

Если положить в (3) $p = 1$, то получим *остаточный член в форме Коши:*

$$R_n(x; f, a) = \frac{f^{(n+1)}(a+\theta(x-a))}{n!}(1-\theta)^{n}(x-a)^{n+1},\ 0 < \theta < 1.$$

При $p = n+1$ получим из формулы (3) *остаточный член в форме Лагранжа:*

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

Форма Лагранжа примечательна тем, что выглядит, как очередное слагаемое в многочлене Тейлора, с той лишь разницей, что производная функции вычисляется в точке c, лежащей между a и x.

Можно записать формулу Тейлора с остаточным членом в форме Лагранжа в виде

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

Эта формула является обобщением формулы конечных приращений (то есть теоремы Лагранжа).

Отметим, что хотя форма Лагранжа наиболее употребляема, но иногда её недостаточно для подходящей точности оценки остаточного члена. Тогда прибегают к форме Коши или иным формам.

Пример 12. *Зафиксируем $x > 0$ рассмотрим на отрезке $[0, x]$ функцию $f(x) = e^x$. К ней применима теорема 2. Так как производная e^x любого порядка равна e^x, то*

$$\left| e^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right| = \left| \frac{e^c}{(n+1)!}x^{n+1} \right| \le \frac{x^n \cdot e^x}{(n+1)!} \to 0, n \to \infty,$$

при некотором $c \in (0, x)$. Так как $c < x$, то $e^c < e^x$, откуда выше и появился знак неравенства. Стремление к нулю возникает, так как факториал стремится к бесконечности быстрее любой натуральной степени (проверьте это!). Таким образом, чем больше n, тем модуль

$$\left| e^x - \sum_{k=0}^{n} \frac{x^k}{k!} \right|$$

меньше. По определению сходимости бесконечного ряда это значит, что бесконечная сумма $\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ *в точности равна* e^x*. Так как точка* x *выбиралась произвольно, то равенство*

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

справедливо при всех $x \in \mathbb{R}$*, и правая часть называется* ***рядом Тейлора*** *функции* e^x*. Ряды Тейлора есть и у других элементарных функций из равенств (9)-(13). Подробнее ряды Тейлора будут изучаться на втором курсе.*

При этом такие же оценки с помощью остатка в форме Лагранжа уже нельзя провести, если вместо экспоненты рассмотреть логарифм. Для логарифма также можно получить ряд, но используя остаток в форме Коши. Действительно, остаток в форме Лагранжа для логарифма будет выглядеть так:

$$R_n(x; \ln, 0) = \frac{(-1)^n \cdot n! \cdot x^{n+1}}{(1+c)^{n+1} \cdot (n+1)!} = \frac{(-1)^n \cdot x^{n+1}}{(1+c)^{n+1} \cdot (n+1)},$$

где c *лежит между точками* 0 *и* x*. Если* $x > 1$*, то ряд* $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{k} x^n$ *расходится, так как не выполнен необходимый признак. Если* $0 < x \le 1$*, то* $0 < c < x$*, поэтому* $x < 1 + c$ *и*

$$\left| \frac{(-1)^n \cdot x^{n+1}}{(1+c)^{n+1} \cdot (n+1)} \right| \le \frac{1}{n+1} \to 0, \ n \to \infty,$$

поэтому ряд $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{k} x^n$ *сходится к функции* $f(x) = \ln(1+x)$ *при* $0 < x \le 1$*. Область определения логарифма состоит из всех* $x > -1$*, а ряд* $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{k} x^n$ *расходится при* $x = -1$*, так как является гармоническим, поэтому нужно ещё рассмотреть случай* $-1 < x < 0$*. В этом случае остаток во форме Лагранжа не позволяет сделать какие-то выводы, так как дробь* $\left| \frac{x^{n+1}}{(1+c)^{n+1}} \right|$*, представляющая собой* $n+1$*-ую степень отношения расстояния от* x *до* 0 *и от* c *до* 1*, а потому может быть как меньше* 1 *(тогда остаток в форме Лагранжа стремится к нулю), так и больше одного (тогда этот остаток бесконечно большой). Таким образом, оценку провести нельзя. Запишем остаточный член в форме Коши:*

$$R_n(x; \ln, 0) = \frac{(-1)^n \cdot n!}{(1+\theta x)^{n+1} \cdot n!}(1-\theta)^n x^{n+1} = \frac{(-1)^n (1-\theta)^n x^{n+1}}{(1+\theta x)^{n+1}}, \ 0 < \theta < 1.$$

Для того, чтобы понять, куда стремится остаточный член при $n \to \infty$*, рассмотрим дробь* $\left| \frac{x - \theta x}{1 + \theta x} \right|$*. В числителе расстояние между точками* x *и* θx*, а в знаменателе – расстояние между* -1 *и* θx*, поэтому*

$$\left| \frac{x - \theta x}{1 + \theta x} \right| = \frac{|x| - |\theta x|}{1 - |\theta x|} = 1 - \frac{1 - |x|}{1 - |\theta x|} \le 1 - (1 - |x|) = |x| < 1.$$

Таким образом, учитывая, что $-1 < x < 0$ *и* $0 < \theta < 1$*, получаем*

$$\left| \frac{(-1)^n (1-\theta)^n x^{n+1}}{(1+\theta x)^{n+1}} \right| < \left| \frac{x}{1+\theta x} \right| \cdot |x|^n < \left| \frac{x}{1+x} \right| \cdot |x|^n \to 0, \ n \to \infty.$$

Итак, равенство $\ln(1+x) = \sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{k} x^n$ *справедливо при всех* $x \in (-1, 1]$*.*

Лекция 20

Достаточное условие экстремума

Согласно теореме Ферма (смотрим предыдущие лекции), если функция f, определенная на интервале (a,b), имеет в некоторой точке c этого интервала экстремум и дифференцируема в этой точке, то $f'(c)=0$. Эта теорема даёт *необходимое условие существования экстремума*, то есть, в случае дифференцируемости функции, её точки экстремумов содержатся среди тех точек, в которых производная этой функции равна нулю. Однако не все те точки, производная в которых равна нулю, являются точками экстремума: например, у функции $y=x^3,\ x\in\mathbb{R}$, производная в нуле равна нулю, но экстремума в этой точке нет, так как эта функция возрастает на всей области определения. Таким образом, теорема Ферма не является *достаточным условием существования экстремума.*

Сейчас для тех функций, которые в точке имеют производные высших порядков, мы докажем теорему, в которой из некоторых условий на производные в точке будет следовать существование экстремума в этой точке, то есть будут доказаны условия, выполнения которых достаточно для наличия в точке экстремума.

Теорема 10. ***(Достаточное условие экстремума)*** *Пусть функция f определена в некоторой окрестности точки a и дифференцируема n раз в точке a, причём*

$$f'(a)=...=f^{(n-1)}(a)=0,$$

а $f^{(n)}(a)\neq 0$. Тогда если $n=2k$ и $f^{(n)}(a)<0$, то a – точка максимума, если $n=2k$ и $f^{(n)}(a)>0$, то a – точка минимума, а если $n=2k-1$, то экстремума нет.

Доказательство. В силу условия для функции f мы можем записать локальную формулу Тейлора в точке a, которая, с учётом равенства нулю производных, запишется в виде:

$$f(x)-f(a)=\frac{f^{(n)}(a)}{n!}(x-a)^n+o((x-1)^n),\ x\to a.$$

Последнее равенство перепишем с учетом определения o-малого:

$$f(x)-f(a)=(x-a)^n\left(\frac{f^{(n)}(a)}{n!}+h(x)\right),\ \text{где } h(x)\to 0 \text{ при } x\to a.$$

В достаточно малой окрестности точки a бесконечно малая величина $h(x)$ пренебрежимо мала по сравнению с $\frac{f^{(n)}(a)}{n!}$ поэтому знак разности $f(x)-f(a)$ совпадает со знаком выражения $\frac{f^{(n)}(a)}{n!}(x-a)^n$. Если n чётное, то $(x-a)^n\geq 0$, поэтому при $f^{(n)}(a)<0$ в достаточно малой окрестности точки a $f(x)\leq f(a)$, то есть в точке a максимум, а при $f^{(n)}(a)>0$ $f(x)\geq f(a)$ в достаточно малой окрестности точки a, то есть в точке a минимум. Если n нечётное, то выражение $\frac{f^{(n)}(a)}{n!}(x-a)^n$ меняет знак при переходе через точку a, а тогда (в достаточно малой окрестности точки a) знак меняет и разность $f(x)-f(a)$, поэтому экстремума в точке a нет. $\square$

В частности, если первая производная функции в точке равна нулю, а вторая отлична от нуля, то в этой точке функция имеет экстремум.

С помощью второго следствия теоремы Лагранжа можно сформулировать достаточные условия существования экстремума в терминах первой производной. Речь идёт о том, что *если функция f дифференцируема на интервале (a,b), $f'(c)=0$ $(c\in(a,b))$ и $f'(x)\leq 0$*

при $x \in (a, c)$, $f'(x) \geq 0$ при $x \in (c, b)$, то в точке c у функции f локальный минимум. Для максимума желающие могут сформулировать аналогичное утверждение и доказать эти утверждения в качестве упражнения.

Отметим при этом, что, конечно, бывают локальные экстремумы и в точках, в которых производная не существует, как, например, у функции $y = |x|$.

Кроме того, бывает так, что *функция в точке имеет экстремум, но её производная меняет знак в любой окрестности этой точки.* В качестве примера приведём функцию

$$f(x) = \begin{cases} x^4(2 + \sin(1/x)), \ x \neq 0, \\ 0, \ x = 0. \end{cases}$$

У f в нуле минимум, но ни в какой окрестности нуля f' не сохраняет знак, то есть сформулированное выше достаточное условие, основанное на следствии из теоремы Лагранжа, не выполнено, но экстремум есть. Действительно, при $x \neq 0$ имеем

$$f'(x) = 4x^3(2 + \sin(1/x)) - x^2 \cos(1/x)$$

и в точках вида $x_n = 1/2\pi n$ при достаточно больших n выполнено $f'(x_n) = \frac{1}{\pi^3 n^3} - \frac{1}{4\pi^2 n^2} < 0$, а точках вида $x_n = 1/(\pi + 2\pi n)$ имеем $f'(y_n) > 0$, то есть знак производной меняется в любой окрестности нуля.

Для большей наглядности приведём эскиз графика.

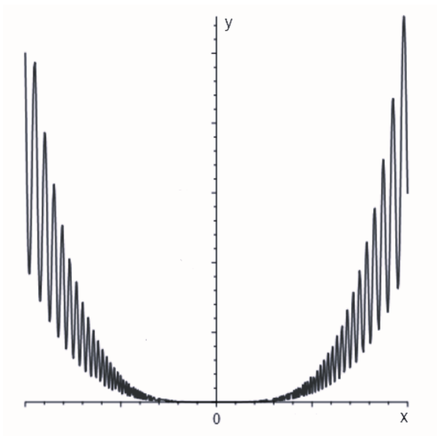

Рис. 7: В любой окрестности нуля бесконечно много осцилляций

Выпуклость

Определение 8. *Функция f, определённая на интервале (a, b), называется выпуклой (выпуклой вниз) на этом интервале, если для любых точек $x_1, x_2 \in (a, b)$ и любого числа $\lambda \in [0, 1]$ выполнено неравенство*

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2). \tag{1}$$

Если при тех же условиях выполнено противоположное неравенство, то функцию f называют вогнутой (выпуклой вверх) на интервале (a, b).

Точка $\lambda x_1 + (1 - \lambda)x_2$ при $\lambda \in [0, 1]$ лежит на отрезке с концами x_1 и x_2, так как

$$\lambda x_1 + (1 - \lambda)x_2 = x_2 - \lambda(x_2 - x_1),$$

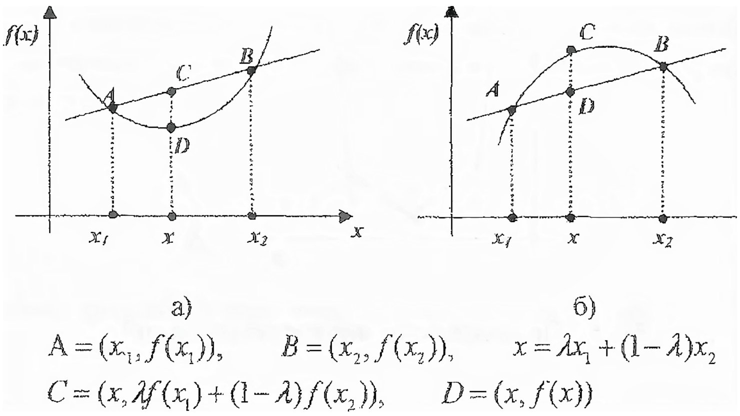

Рис. 8: а) Выпуклая функция; б) Вогнутая функция

то есть при любом $\lambda \in [0,1]$ получается точка из отрезке. Геометрически определение выпуклости означает, что всякая хорда, соединяющая две точки графика функции, *лежит не ниже графика функции*, а определение вогнутости означает, что любая хорда *лежит не выше графика функции* (см. рисунок 8).

Отметим **(интересующиеся могут доказать)**, что выпуклые и вогнутые функции на интервале (a,b), являются непрерывными на этом интервале и имеют в каждой точке интервала левые и правые производные (но они необязательно равны).

Предложение 9. *Функция f выпукла на интервале (a,b) тогда и только тогда, когда для любых таких чисел $x_1,\ x_2,\ y \in (a,b)$, что $x_1 < y < x_2$, выполнено неравенство*

$$\frac{f(y)-f(x_1)}{y-x_1} \le \frac{f(x_2)-f(y)}{x_2-y}. \tag{2}$$

Прежде чем доказывать это предложение, отметим, что его геометрический смысл состоит в том, что тангенсы углов наклонов секущих не убывают с увеличением абсцисс точек пересечения с графиком функции f. Отсюда следует, что любая прямая пересекает график выпуклой функции в двух точках, либо (например, в случае функций $y = x$ или $y = |x|$) совпадает с частью графика или всем графиком.

Доказательство. **Необходимость.** Если функция выпукла, то для неё выполнено неравенство (1). Любая точка $y \in (x_1, x_2)$ может быть записана в виде $y = \lambda x_1 + (1-\lambda)x_2$ при некотором $\lambda \in (0,1)$. Выразим λ из последнего равенства: $\lambda = \frac{y-x_2}{x_1-x_2}$ и запишем неравенство (1), подставив такое λ:

$$f(y) \le \frac{y-x_2}{x_1-x_2}f(x_1) + \left(1 - \frac{y-x_2}{x_1-x_2}\right)f(x_2) = \frac{y-x_2}{x_1-x_2}f(x_1) + \frac{x_1-y}{x_1-x_2}f(x_2).$$

Используя равенство $x_1 - x_2 = x_1 - y + y - x_2$ и домножая обе части на $x_1 - y + y - x_2$ (это выражение меньше нуля!), получим:

$$(x_1-y)f(y) + (y-x_2)f(y) \ge (y-x_2)f(x_1) + (x_1-y)f(x_2)$$
$$\Leftrightarrow (f(y)-f(x_1))(x_2-y) \le (f(x_2)-f(y))(x_2-y).$$

Последнее неравенство равносильно неравенству (2).

Достаточность. Нам нужно вывести из неравенства (2) неравенство (1). Отметим, что при $\lambda = 1$ и $\lambda = 0$ неравенство (1) превращается просто в равенство. При $\lambda \in (0,1)$ любое число $y \in (x_1, x_2)$ представимо в виде $y = \lambda x_1 + (1-\lambda)x_2$. Подставив это равенство в

(2) и проведя преобразования из доказательства необходимости в обратном порядке (что возможно в силу равносильности этих преобразований), получим неравенство (1). □

Упражнение. Сформулируйте и докажите аналогичное утверждение для вогнутых функций.

Все замечания, сделанные перед доказательством предложения 1, верны и для вогнутых функций, только теперь тангенсы углов наклонов хорд не возрастают.

Теперь выясним, как связаны дифференцируемость функции и выпуклость. Выпуклые функции необязательно дифференцируемы, но если рассматривать только те выпуклые функции, которые дифференцируемы, то можно получить ряд новых полезных понятий и утверждений.

Предложение 10. ***(Следствие предложения 1)*** *Пусть функция f дифференцируема на интервале (a, b). Тогда выпуклость f на этом интервале равносильна тому, что функция f' не убывает на (a, b).*

Доказательство. Пусть условия на числа x_1, x_2, y те же, что в предложении 1.

Необходимость. Если функция f выпукла, то для неё выполнено неравенство (2). Устремляя y к x_1, в неравенстве (2), получим в силу предельного перехода в неравенствах

$$f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Устремляя в неравенстве (2) y к x_2, получаем

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_2).$$

Таким образом,

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

то есть f' не убывает на отрезке (a, b).

Достаточность. Пусть производная f' не убывает на (a, b). На отрезках $[x_1, y]$ и $[y, x_2]$ к функции f применима теорема Лагранжа, то есть найдутся такие точки $c_1 \in (x_1, y)$ и $c_2 \in (y, x_2)$, что $f(y) - f(x_1) = f'(c_1)(y - x_1)$ и $f(x_2) - f(y) = f'(c_2)(x_2 - y)$. Тогда в силу неубывания производной

$$\frac{f(y) - f(x_1)}{y - x_1} = f'(c_1) \leq f'(c_2) = \frac{f(x_2) - f(y)}{x_2 - y},$$

то есть выполнено неравенство (2), которое равносильно выпуклости функции f на отрезке (a, b). □

Функция f' также может быть дифференцируемой на интервале (a, b). Так как неубывание функции на интервале равносильно неотрицательности её производной (2-е следствие теоремы Лагранжа), то можем сформулировать критерий выпуклости для дважды дифференцируемых на интервале функций.

Предложение 11. ***(Следствие предложения 2)*** *Пусть функция f дважды дифференцируема на интервале (a, b). Тогда выпуклость f на этом интервале равносильна тому, что функция f'' неотрицательна на (a, b).*

Доказательство. Это утверждение прямо следует из 2-го следствия теоремы Лагранжа и предложения 2. □

Предложение 12. ***(Следствие предложения 2)*** *Если функция f дифференцируема на интервале (a, b), то она выпукла тогда и только тогда, когда при всех $x, x_0 \in (a, b)$ выполнено неравенство $f(x) \geq f'(x_0)(x - x_0) + f(x_0)$.*

Доказательство. Необходимость. На любом отрезке, содержащемся в интервале (a, b) функция f удовлетворяет условиям теоремы Лагранжа, поэтому найдётся такая точка c, принадлежащая интервалу с концами x и x_0, что $f(x) - f(x_0) = f'(c)(x - x_0)$. В силу предложения 2 выражение $(f'(c) - f'(x_0))(x - x_0)$ всегда неотрицательно, поэтому

$$f(x) - f'(x_0)(x - x_0) - f(x_0) = (f'(c) - f'(x_0))(x - x_0) \geq 0,$$

чем всё и доказано.

Достаточность. Фиксируем $y = x_0$. Тогда неравенство из условия перепишется в виде

$$f(x) \geq f'(y)(x - y) + f(y),$$

что равносильно

$$\frac{f(x) - f(y)}{x - y} \geq f'(y) \text{ при } x > y \text{ и } \frac{f(x) - f(y)}{x - y} \leq f'(y) \text{ при } x < y.$$

Пусть $x_1 < y < x_2$, где $x_1,\ x_2 \in (a, b)$. Положим в неравенстве из условия $x = x_1$, а затем в нём же положим $x = x_2$. Тогда

$$\frac{f(y) - f(x_1)}{y - x_1} \leq f'(y) \leq \frac{f(x_2) - f(y)}{x_2 - y}$$

Таким образом, выполнен критерий выпуклости из предложения выше, так как y и x_1, x_2 выбираются произвольно и $x_1 < y < x_2$. □

Функция $y = f'(x_0)(x - x_0) + f(x_0)$ задаёт уравнение касательной к графику функции f в точке x_0. Таким образом, геометрически это предложение означает, что график выпуклой функции лежит над любой касательной к этому графику или совпадают с частью касательной (см. рис. 9).

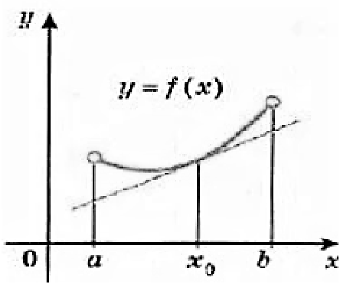

Рис. 9: График функции над касательной

Упражнение. Сформулируйте и докажите аналогичные предложениям 2 – 4 утверждения для вогнутых функций.

Лекция 21

Точки перегиба и построение графиков

Предположим, что функция f имеет n производных в точке a, причём производные со второй по $n-1$-ю равны нулю, а производная порядка n отлична от нуля (о первой производной мы не делаем никаких дополнительных предположений.) Запишем локальную формулу Тейлора для функции f в точке a в следующем виде:

$$f(x) - (f'(a)(x-a) + f(a)) = \frac{f^{(n)}(a)}{n!}(x-a)^n + o((x-a)^n).$$

Слева стоит разность значений функции f и касательной к графику этой функции в точке a. Используя те же соображения, что в достаточном условии экстремума (см. предыдущую лекцию), мы можем показать, что в достаточно малой окрестности точки a левая часть не меняет знак при чётном n и этот знак совпадает со знаком $f^{(n)}(a)$. Таким образом, график функции f в достаточно малой окрестности лежит или над касательной, или под касательной.

Если же n нечётное, то знак разности меняется при переходе через точку a.

Если слева, в некотором интервале с правым концом a, функция выпукла (вогнута), а справа, в некотором интервале с левым концом a, функция вогнута (выпукла), то говорят, что функция f в точке a имеет *перегиб,* а саму точку a называют *точкой перегиба* (см. рис. 10). Например, график функции $y = x^3$ в нуле имеет имеет перегиб. Мы даём определение

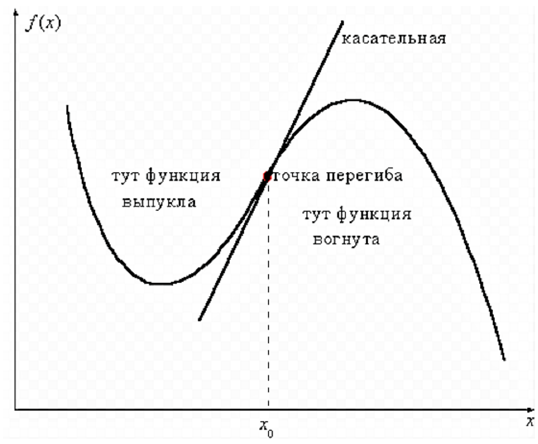

Рис. 10: В точке $a = x_0$ функция f имеет перегиб

точки перегиба только для дифференцируемых в этой точке функций.

Однако взаимное расположения касательной и графика функции, аналогичное тому, что на рисунке 3, ещё не гарантирует, что точка касания – это точка перегиба, если, например, функция осциллирует в любой окрестности точки касания. В качестве примера приведём функцию

$$f(x) = \begin{cases} x^3(2+\sin(1/x)),\ x \neq 0, \\ 0,\ x = 0. \end{cases}$$

Если в определении выпуклости неравенство строгое, то функция называется *строго выпуклой.* Например, функция $f(x) = x^2$ строго выпукла на $\mathbb{R}$, а функция $g(x) = |x|$ выпукла, но не строго выпукла на этом же множестве (почему?). Для строго выпуклых функций справедливы все те же утверждения, что для выпуклых, только в предложениях

1 и 3 неравенства строгие, в предложении 2 f' строго возрастает, а в предложении 4 касательная не может совпадать с частью графика, а обязательно имеет с графиком только одну общую точку, а все остальные точки графика лежат над касательной.

Аналогично определяются строго вогнутые функции. Подумайте самостоятельно, что меняется в предложениях 1 – 4, если речь идёт о строго вогнутых функциях.

Построение графика функции

Мы обсудили свойства монотонности и выпуклости функции, дали определение экстремумов функции, разобрали понятие точки перегиба, то есть обсудили многие основные характеристики графика функции. Сейчас будут даны определения асимптот, после чего мы запишем план исследования графика функции, позволяющий строить эскизы многих графиков с отражением практически всех основных свойств.

Определение 9. *Прямая $x = a$ называется вертикальной асимптотой графика функции f, если хотя бы один из односторонних пределов $\lim\limits_{x\to a-0} f(x)$ или $\lim\limits_{x\to a+0} f(x)$ равен $-\infty$, либо $+\infty$.*

Например, прямая $x = 0$ – вертикальная асимптота графика функции $y = 1/x$.

Определение 10. *Прямая $y = \alpha x+\beta$ называется наклонной асимптотой графика функции f при $x \to +\infty$, если $f(x) = \alpha x + \beta + o(1),\ x \to +\infty$.*

При $x \to -\infty$ наклонная асимптота определяется аналогично.

Приведём без доказательства теорему, позволяющую находить наклонные асимптоты с помощью пределов.

Теорема 11. *Наклонная асимптота функции f при $x \to +\infty$ существует тогда и только тогда, когда одновременно выполняются два условия:*
1) $\lim\limits_{x\to+\infty} \frac{f(x)}{x} = \alpha,\ \alpha \in \mathbb{R}$;
2) $\lim\limits_{x\to+\infty} (f(x) - \alpha x) = \beta,\ \beta \in \mathbb{R}$.

Упражнение. 1. Может ли график функции пересекать наклонную асимптоту?
2. Как определить, сверху или снизу приближается график функции f к наклонной асимптоте при $x \to +\infty$?

Далее выпишем общую схему построения графика функции. Все термины, используемые в ней, либо изучались в нашем курсе, либо встречались в школьном курсе.

Общая схема построения графика функции

1. Найти область определения функции.
2. Исследовать функцию на чётность, периодичность, промежутки знакопостоянства. Найти точки пересечения графика с осями координат.
3. Исследовать значения функции на границах области определения, определить характеры точек разрыва и найти вертикальные асимптоты.
4. Найти наклонные асимптоты.
5. Определить промежутки монотонности и экстремумы.
6. Найти промежутки выпуклости, вогнутости и точки перегиба.
7. Построить график функции.

Некоторые неравенства (приложения выпуклости)

Доказанные в этом разделе неравенства находят применение во многих разделах математики.

Теорема 12. ***(Неравенство Йенсена)*** *Пусть функция f выпукла на интервале (a, b). Тогда при любых $x_1,\ x_2,\ \dots,\ x_n \in (a, b)$ и всех таких $\lambda_1 \geq 0,\ \lambda_2 \geq 0,\ \dots,\ \lambda_n \geq 0$, что $\lambda_1 + \dots + \lambda_n = 1$, выполняется неравенство $f(\lambda_1 x_1 + \dots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \dots + \lambda_n f(x_n)$.*

Доказательство. Воспользуемся индукцией по n.

При $n = 2$ неравенство Йенсена превращается в неравенство (1), которое справедливо, так как функция f выпукла, поэтому база индукция очевидна.

Предположим, что при $n = k$ неравенство верно, то есть

$$f(\lambda_1 x_1 + \dots + \lambda_k x_k) \leq \lambda_1 f(x_1) + \dots + \lambda_k f(x_k).$$

Из предположения индукции выведем справедливость неравенства Йенсена при $n = k+1$. Пусть $\lambda := \sum\limits_{i=1}^{k} \lambda_i$. Точка $\frac{\lambda_1}{\lambda}x_1 + \dots + \frac{\lambda_k}{\lambda}x_k$ принадлежит интервалу (a, b) (это легко проверяется по индукции – проверьте!). Кроме того, $\frac{\lambda_i}{\lambda} \geq 0,\ i = 1, \dots, k$, и $\frac{\lambda_1}{\lambda} + \dots + \frac{\lambda_k}{\lambda} = 1$, поэтому в силу выпуклости f и предположения индукции получим:

$$\begin{aligned} f(\lambda_1 x_1 + \dots + \lambda_k x_k + \lambda_{k+1} x_{k+1}) &= f(\lambda \left(\frac{\lambda_1}{\lambda} x_1 + \dots + \frac{\lambda_k}{\lambda} x_k \right) + \lambda_{k+1} x_{k+1}) \leq \\ \leq \lambda f \left(\frac{\lambda_1}{\lambda} x_1 + \dots + \frac{\lambda_k}{\lambda} x_k \right) + \lambda_{k+1} f(x_{k+1}) &\leq \lambda \left(\frac{\lambda_1}{\lambda} f(x_1) + \dots + \frac{\lambda_k}{\lambda} f(x_k) \right) + \lambda_{k+1} f(x_{k+1}) = \\ &= \lambda_1 f(x_1) + \lambda_2 f(x_2) + \dots + \lambda_{k+1} f(x_{k+1}). \end{aligned}$$

Теорема доказана. □

Приведём пример применения неравенства Йенсена. Неравенство, которое мы докажем, уже встречалось нам на семинарах.

Пример 13. ***(Неравенство между средним арифметическим и средним геометрическим).*** *Для положительных чисел $x_1,\ x_2,\ ,\dots,\ x_n$ справедливо неравенство*

$$\sqrt[n]{x_1 \cdot \dots \cdot x_n} \leq \frac{x_1 + \dots + x_n}{n}.$$

Для доказательства рассмотрим функцию $y = -\ln x$. Так как $y'' = 1/x^2 > 0$ при $x > 0$, то функция выпукла на промежутке $(0, +\infty)$, поэтому к ней применимо неравенство Йенсена (полагаем, что $\lambda_1 = \dots = \lambda_n = 1/n$):

$$-\ln \left(\frac{1}{n} x_1 + \dots + \frac{1}{n} x_n \right) \leq \frac{1}{n}(-\ln x_1) + \dots + \frac{1}{n}(-\ln x_n) = -\ln(x_1 \cdot \dots \cdot x_n)^{1/n},$$

что равносильно $\ln\left(\frac{1}{n}x_1 + \dots + \frac{1}{n}x_n\right) \geq \ln(x_1 \cdot \dots \cdot x_n)^{1/n}$. Беря экспоненту от обеих частей последнего неравенства, получаем требуемое. Отметим, что, беря экспоненту от обеих частей, мы используем монотонное возрастание функции $y = e^x$, которое следует из того, что $(e^x)' = e^x > 0\ \forall x \in \mathbb{R}$ (хотя мы доказали возрастание, когда строили экспоненту).

Метод касательных (метод Ньютона)

При изучении свойств непрерывных функций мы доказывали, что непрерывная на отрезке $[a, b]$ функция f, принимающая на его концах значения разных знаков, имеет хотя бы один корень на этом отрезке. Метод доказательства этой теоремы (в котором мы делим отрезок пополам и выбираем ту половину, на концах которой функция принимает значения разных знаков) можно трактовать также как способ приближенного вычисления корня функции. Такой метод приближенного вычисления корня называется методом бисекции. Так как на n-ом шаге метода бисекции мы получаем отрезок длины $\frac{b-a}{2^n}$, то последовательность точек сходится к решению со скоростью геометрической прогрессии.

Сейчас мы обсудим метод, который при достаточно хороших условиях даёт ещё более впечатляющую скорость сходимости.

Итак, пусть на интервале (a, b) задана функция f, которая дважды дифференцируема на этом интервале, причём выполнены следующие условия:

$$\forall\, x \in (a, b)\ 0 < A \leq f'(x) \text{ и } B \geq f''(x) > 0$$

для некоторых чисел A и B. Пусть, кроме того,

$$\exists\, c \in (a, b)\ :\ f(c) = 0.$$

Тогда в силу ограничений на f этот корень на интервале (a, b) единственный, так как функция f строго возрастает. Отметим также, что в силу положительности второй производной f строго выпукла на (a, b).

Возьмём точку $x_1 \in (a, b),\ x_1 > c$ и построим касательную к графику функции f в точке x_1:

$$g_1(x) = f'(x_1)(x - x_1) + f(x_1).$$

Пусть x_2 – точка пересечения этой касательной с осью абсцисс. Тогда (в чём мы убедимся ниже) $x_2 \in (a, b)$. В точке x_2 построим касательную к графику функции f, через x_3 обозначим точку её пересечения с осью абсцисс и так далее (на рисунке 11 см. графическую иллюстрацию).

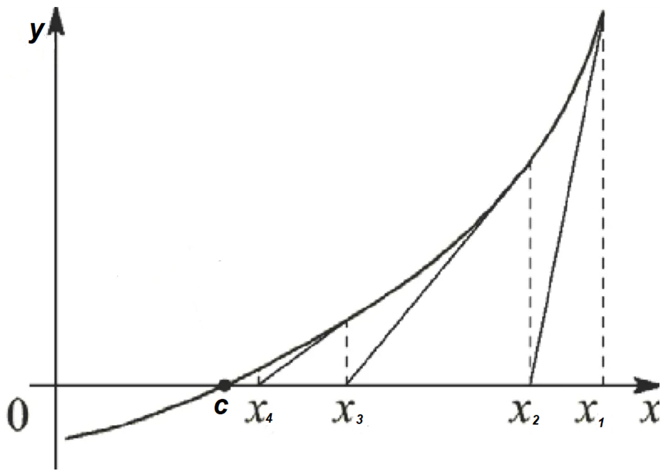

Рис. 11: Построение первых четырёх приближений корня функции f

Получив точку x_n, построим в ней касательную. Выразим явно точку x_{n+1} :

$$g_n(x) = f'(x_n)(x - x_n) + f(x_n) = 0\ \Leftrightarrow x = x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Отметим, что так как f' всюду отлична от нуля, то x_{n+1} определен корректно.

С помощью теоремы Вейерштрасса докажем, что рекуррентная последовательность $\{x_n\}$ имеет предел. Прежде всего отметим, что так как функция f строго выпукла, то, как следует из критерия выпуклости через касательные, она всегда принимает значения, не меньшие, чем функция, задающая касательную к ней, причём равенство достигается только в точке касания. Тогда

$$f(x_{n+1}) > g_n(x_{n+1}) = 0,$$

откуда в силу возрастания f имеем $x_{n+1} > c$ при всех натуральных n. Итак, доказано, что последовательность $\{x_n\}$ ограничена снизу. Тогда $\forall\ n \in \mathbb{N}\ f(x_n) > 0$, поэтому

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

то есть последовательность $\{x_n\}$ убывающая. Тогда по теореме Вейерштрасса у неё есть предел, который обозначим A. В рекуррентном соотношении $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ перейдём к пределу:

$$A = A - \frac{f(A)}{f'(A)} \ \Leftrightarrow\ f(A) = 0 \ \Leftrightarrow\ A = c.$$

Итак, доказано, что $\lim\limits_{n\to\infty} x_n = c$.

Оценим скорость сходимости последовательности $\{x_n\}$ к c. Прежде всего заметим, что при $t \in (a,b)$ к f можно применить теорему о разложении по формуле Тейлора с остаточным членом в форме Лагранжа:

$$f(t) = f(x_n) + f'(x_n)(t - x_n) + \frac{f''(d)}{2}(t - x_n)^2,$$

где d лежит на интервале с концами x_n и t. Отсюда получим

$$f(x_n) = f(t) - f'(x_n)(t - x_n) - \frac{f''(d)}{2}(t - x_n)^2.$$

Положив $t = c$, имеем:

$$\begin{aligned}
|c - x_{n+1}| = \left|c - x_n + \frac{f(x_n)}{f'(x_n)}\right| = \\
&= \left|c - x_n + \frac{f(c) - f'(x_n)(c - x_n) - \frac{f''(d)}{2}(c - x_n)^2}{f'(x_n)}\right| = \\
&= \left|-\frac{f''(d)}{2f'(x_n)}(c - x_n)^2\right| \le \\
&\le \frac{B}{2A}(c - x_n)^2 = M(c - x_n)^2,
\end{aligned}$$

где $M = B/2A$. Применяя в точности те же рассуждения к разности $|c - x_n|$, получим оценку

$$|c - x_{n+1}| \le M \cdot (M(c - x_{n-1})^2)^2 = M^3(c - x_{n-1})^4.$$

После n применений этой оценки получим

$$|c - x_{n+1}| \le M^{2^n - 1}(c - x_1)^{2^n}.$$

Если теперь предположить, что изначально $|c - x_1| < \frac{1}{2M}$, то последовательность $\{x_n\}$ будет стремиться к c также, как последовательность $\{\frac{1}{2^{2^n}}\}$ стремится к нулю, то есть этот метод гораздо быстрее даёт хорошие приближения к корню по сравнению с методом бисекций.

Описанный метод нахождения приближенного значения корня заданной функции называется *методом касательных* или *методом Ньютона.* Отметим, что если начальное приближение (то есть точка x_1) недостаточно близко к решению, то метод может не сойтись. По крайней мере для того, чтобы метод работал, надо выбирать такой интервал (a, b), на котором для заданной функции f верны все указанные ограничения.

В предыдущих лекциях рассматривалась итерационная формула Герона. Теперь читатели могут убедиться, что она построена с помощью применения метода Ньютона к приближенному вычислению положительного корня функции $f(x) = x^2 - a$ $(a > 0)$.

Если функция выпукла вверх и возрастает, то точки, в которых проводятся касательные, нужно брать слева от корня. Полезно разобрать все возможные комбинации направления выпуклости и характера монотонности.

Существуют функции, для которых метод Ньютона даёт расходящуюся последовательность, какую бы окрестность корня мы ни брали. В качестве упражнения попробуйте применить метод Ньютона к функции

$$f(x) = \begin{cases} x + x^2 \sin \frac{2}{x}, \ x \neq 0, \\ 0, \ x = 0. \end{cases}$$

В силу того, что производная в любой окрестности нуля меняет знак, найти корень с помощью метода Ньютона не удастся.

Полезно также рассмотреть функции, для которых скорость сходимости меньше, чем для рассмотренных нами функций и понять, почему так бывает. В качестве примера можно рассмотреть функцию $f(x) = x + \sqrt[3]{x^4}$.

Метод хорд и комбинированный метод

Отметим, не вдаваясь в подробности, что при разумных ограничениях на заданную функцию существует, например, ещё метод хорд для нахождения корня этой функции. Мы не будем его подробно описывать, а лишь приведём иллюстрацию на рисунке 12.

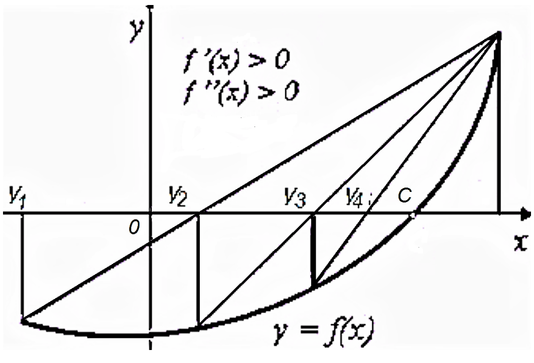

Рис. 12: Первые четыре приближения корня функции f с помощью метода хорд

Иногда для нахождения корня применяется комбинированный метод: на первом шаге находят y_1 по методу хорд и x_1 – по методу касательных, на втором находят y_2 и x_2. Корень c всегда лежит между точкой, найденной по методу хорд и точкой, найденной на том же шаге по методу Ньютона. Если на k-ом шаге получается отрезок, длина которого $|x_k - y_k|$ не превосходит точность, с которой нужно найти приближенный корень, то вычисления на этом шаге заканчивают.