15. Символы o и O, теорема о среднем, формула Тейлора

Начнем эту лекцию с того, что введем два часто используемых в анализе обозначения. Именно: пусть f и g — две функции переменной x, обе стремящиеся к нулю при $x \to a$,

Обозначение 15.1. Если $\lim_{x\to a}(f(x)/g(x)) = 0$, будем писать: что $f(x) = o(g(x))$; если существуют такие $C > 0$, $\delta > 0$, что $|f(x)| \leqslant C \cdot |g(x)|$, если $0 < |x-a| < \delta$, то будем писать, что $f(x) = O(g(x))$.

(Здесь подразумевается, конечно, что $g(x) \neq 0$ при всех x, достаточно близких к a).

Неформально говоря, соотношение $f(x) = o(g(x))$ означает, что функция f стремится к нулю быстрее, чем функция g (говорят еще: f — бесконечно малая более высокого порядка, чем g)); соотношение $f(x) = O(g(x))$ означает: что f стремится к нулю не быстрее, чем g.

У определения 15.1 имеются различные вариации. Например, вместо $x \to a$ можно предполагать, что $x \to \infty$. Можно также считать, что переменная не непрерывна, а дискретна, заменив функции f и g на последовательности $\{f_n\}$ и $\{g_n\}$. Можно, наконец, предполагать, что функции f и g (или последовательности $\{f_n\}$ и $\{g_n\}$) обе стремятся к бесконечности, а не к нулю.

Пусть, например, $0 < m < n$. Тогда $x^n = o(x^m)$ при $x \to 0$ и $x^m \to x^n$ при $x \to \infty$.

В силу самого определения ясно, что слова «$f(x) = o(g(x))$» имеют смысл только вместе с уточнением «при $x \to a$». Тем не менее на практике это уточнение всегда опускают, полагая, что из контекста ясно, что куда стремится.

Для начинающих символы o и O могут представлять известную трудность, поскольку их использование представляют собой довольно серьезную вольность речи (например, одним и тем же символом $o(x)$ могут обозначаться совершенно разные функции). Тем не менее такая символика и удобна, и общепринята, так что давайте к ней привыкать.

Приведем примеры использования символа o. Начнем с очень простой, но важной переформулировки.

Предложение 15.2. *Пусть $f\colon (p,q) \to \mathbb{C}$ — функция и $a \in (p;q)$. Тогда следующие два условия эквивалентны:*

(1) $f'(a) = c$;

(2) $f(x) = f(a) + c(x-a) + o(x-a)$.

(В условии (2), конечно, подразумевается уточнение «при $x \to a$».)

Доказательство. (1) $\Rightarrow$ (2). Нам надо доказать, что

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

Это действительно так, поскольку

$$\lim_{x\to a} \frac{f(x)-f(a)-c(x-a)}{x-a} = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} - c, \tag{15.1}$$

и правая часть равна нулю ввиду определения производной.

(2) $\Rightarrow$ (1). На сей раз нам дано, что

$$\lim_{x\to a} \frac{f(x)-f(a)-c(x-a)}{x-a} = 0;$$

формула (15.1) показывает, что это соотношение равносильно тому, что $\lim_{x\to a} \frac{f(x)-f(a)}{x-a} - c$. □

А теперь покажем, как o-символика может работать.

Предложение 15.3. *Пусть* $f\colon (p,q) \to \mathbb{R}$ *— дифференцируемая функция, и пусть для точки* $a \in (p;q)$ *имеем* $f'(a) > 0$. *Тогда существует такое* $\varepsilon > 0$, *что* $f(x) > f(a)$ *при* $a < x < a+\varepsilon$ *и* $f(x) < f(a)$ *при* $a-\varepsilon < x < a$.

Можно сказать, что это предложение — локальный аналог достаточного признака возрастания функции. Ясно, что аналогичное утверждение верно и для случая $f'(a) < 0$.

Доказательство. Положим $f'(a) = c > 0$. Идея доказательства состоит в следующем. Ввиду предложения 15.2 имеем

$$f(x) = f(a) + c(x-a) + o(x-a);$$

если теперь, скажем, $x > a$, то $f(x)-f(a) = c(x-a)+o(x-a)$, где первое слагаемое положительно, так как $c > 0$, а второе слагаемое, коль скоро оно стремится к нулю быстрее, чем $x-a$, будет (при достаточно малом $x-a$) много меньше по модулю, чем $c(x-a)$, так что после прибавления этого слагаемого к $c(x-a)$ результат останется положительным. Аккуратно эта идея проводится следующим образом. Обозначим «остаточный член» $f(x)-f(a)-c(x-a)$ через $r(x)$. Поскольку $r(x) = o(x-a)$, существует такое $\varepsilon > 0$, что при $0 < |x-a| < \varepsilon$ имеем $|r(x)| < \frac{c}{2}|x-a|$, и тем самым знак выражения $c(x-a) + r(x) = f(x) - f(a)$ совпадает со знаком выражения $c(x-a)$ (если $|v| < |u|$, то знаки чисел u и $u+v$ совпадают). Стало быть, если $a < x < a+\varepsilon$, то знак $f(x) - f(a)$ совпадает со знаком $c(x-a)$ и тем самым положителен: а если $a - \varepsilon < x < a$, то знак $f(x) - f(a)$ опять же совпадает со знаком $c(x-a)$ и тем самым отрицателен. □

Из доказанного предложения немедленно вытекает«теорема Ферма» (необходимое условие экстремума). Более того, из этого же предложения нетрудно получить доказательство достаточного условия возрастания (предложение 13.14), не опирающееся на теорему Лагранжа: если $f' > 0$ всюду, то из предложения видно, что $f(v) > f(u)$, если $v > u$ и v близко к u, а переход к случаю, когда v далеко от u, делается с помощью той же техники, что в нашем доказательств теоремы о промежуточном значении.

Наша следующая тема — теорема о среднем, являющаяся естественным обобщением теорем Ролля и Лагранжа из предыдущей лекции. Естественнее всего формулировать и понимать ее геометрически.

Напомним (см. предыдущую лекцию), что *параметризованной кривой* (мы будем иногда говорить просто «кривой») на плоскости называется непрерывное отображение

$$\gamma\colon [a;b] \to \mathbb{R}^2, \qquad t \mapsto (\gamma_1(t), \gamma_2(t))$$

(непрерывность означает, что функции γ_1 и γ_2 непрерывны). Если функции γ_1 и γ_2 к тому же дифференцируемы на интервале $(a;b)$, будем говорить, что γ — *гладкая* параметризованная кривая. Если в этой ситуации $c \in (a;b)$, то вектор $\gamma'(c) = (\gamma_1'(c), \gamma_2'(c))$ будем называть *вектором скорости* в точке c (или «в момент c»: параметр c полезно представлять себе как время, а отображение γ — как описание движения материальной точки по плоскости). Стоит иметь в виду, что если $\gamma\colon [a;b] \to \mathbb{R}^2$ — гладкая кривая, то образ $\gamma([a;b])$ может не во всех точках выглядеть как гладкая кривая: на нем могут быть «точки самопересечения», «точки заострения», «изломы» и прочее — см. рис. 1.

Теорема 15.4 (Теорема о среднем). *Пусть $[a;b] \subset \mathbb{R}$ — отрезок и $\gamma = (\gamma_1, \gamma_2)\colon [a;b] \to \mathbb{R}$ — отображение, непрерывное на $[a;b]$ и дифференцируемое на $(a;b)$. Предположим также, что $\gamma(a) \neq \gamma(b)$. Тогда существует такая точка $c \in (a;b)$, что вектор скорости $\gamma'(c) = (\gamma_1'(c), \gamma_2'(c))$ коллинеарен вектору $\overline{\gamma(a), \gamma(b)}$.*

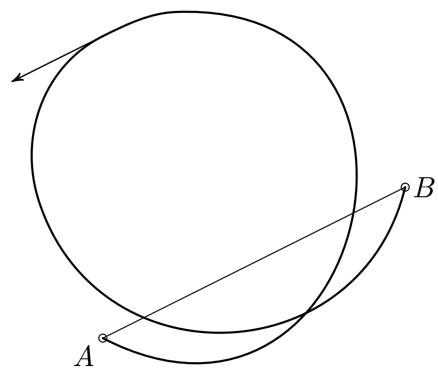

Рис. 1. Кривая и касательный вектор

(Напомним, что нулевой вектор коллинеарен любому.)

Доказательство. Для всякого $\varphi \in \mathbb{R}$ обозначим через $R_\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2$ поворот плоскости на угол φ:

$$R_\varphi \colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\varphi & -\sin\varphi \\ sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Поскольку R_φ — линейное отображение, оно коммутирует с взятием производной: вектор скорости кривой $R_\varphi \circ \gamma$ в момент t получается из вектора скорости кривой γ в момент t поворотом на угол φ. Вот формальная выкладка:

$$((\cos\varphi\gamma_1(t) - \sin\varphi\gamma_2(t))', (\sin\varphi\gamma_1(t) + \cos\varphi\gamma_2(t))') =$$
$$= (\cos\varphi\gamma_1'(t) - \sin\varphi\gamma_2'(t), \sin\varphi\gamma_1'(t) + \cos\varphi\gamma_2'(t)) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} \gamma_1'(t) \\ \gamma_2'(t) \end{pmatrix}.$$

Поскольку при этом, разумеется, вектор $\overline{R_\varphi\gamma(a), R_\varphi\gamma(b)}$ также получается из вектора $\overline{\gamma(a), \gamma(b)}$ поворотом на угол φ, а коллинеарность векторов при повороте сохраняется, утверждение теоремы верно для кривой γ тогда и только тогда, когда оно верно для кривой $R_\varphi \circ \gamma$. Выберем теперь φ, для которого отрезок $[R_\varphi(\gamma(a); R_\varphi(\gamma(b)))]$ будет горизонтален. Это означает, что если $R_\varphi \circ \gamma = (\delta_1, \delta_2)$, то $\delta_2(a) = \delta_2(b)$, то есть функция $\delta_2 \colon [a; b] \to \mathbb{R}$ удовлетворяет условиям теоремы Ролля. Значит, для некоторого $c \in (a; b)$ имеем $\delta_2'(c) = 0$; это означает, что касательный вектор к кривой $R_\varphi \circ \gamma$ в момент c также горизонтален, то есть коллинеарен вектору, соединяющего концы кривой. Ввиду сказанного выше из этого следует, что и касательный вектор к исходной кривой в момент c будет коллинеарен вектору $\overline{\gamma(a), \gamma(b)}$. $\square$

Следствие 15.5 (теорема Коши)**.** *Пусть функции $\varphi, \psi\colon [a;b] \to \mathbb{R}$ непрерывны на $[a;b]$ и дифференцируемы на $(a;b)$, причем $\psi(a) \neq \psi(b)$ и $\psi'(t) \neq 0$ для всех $t \in (a;b)$. Тогда существует такое $c \in (a;b)$, что*

$$\frac{\varphi'(c)}{\psi'(c)} = \frac{\varphi(b) - \varphi(a)}{\psi(b) - \psi(a)}.$$

Доказательство. Примените теорему 15.4 к кривой $t \mapsto (\psi(t), \varphi(t))$. $\square$

Наша следующая тема — формула Тейлора; для формулировки и доказательства основных результатов нам понадобится терминология и обозначения, относящиеся к высшим производным. С этого и начнем.

Пусть $f\colon (p;q) \to \mathbb{R}$ (или в $\mathbb{C}$ — это все равно) — дифференцируемая функция; тогда можно рассмотреть функцию $f'\colon (p;q) \to \mathbb{R}$; если эта функция также дифференцируема, то говорят, что функция f *дважды дифференцируема* на интервале $(p;q)$; производная функции f' в точке $c \in (a;b)$ обозначается $f''(c)$ и называется *второй производной* функции f в точке c. Если f дважды дифференцируема, то можно рассмотреть ее вторую производную как функцию $f''\colon (p;q) \to \mathbb{R}$ на $(a;b)$; если функция f'' также дифференцируема, то говорят, что f трижды дифференцируема, и т.д. При этом если вторую и третью производную обычно обозначают с помощью двух и трех штрихов, то в общем случае k-ю производную функции f обозначают $f^{(k)}$.

Определение 15.6. Если функция $f\colon (a;b) \to \mathbb{C}$ является k раз дифференцируемой на интервале $(a;b)$ и при этом функция $f^{(k)}\colon (a;b) \to \mathbb{C}$ непрерывна, то говорят, что функция f *принадлежит классу C^k* на интервале $(a;b)$. Множество всех функций класса C^k на $(a;b)$ обозначается $C^k(a;b)$. Если функция принадлежит классу C^k при всех $k \in \mathbb{N}$, то говорят, что функция f *принадлежит классу C^∞*; это равносильно тому, что она бесконечно дифференцируема.

Понятие «функция класса C^k» является более ограничительным, чем «бесконечно дифференцируемая функция», но оно является более «правильным» — в первую очередь потому, что с ним удобнее работать.

Теперь, переходя непосредственно к формуле Тейлора, рассмотрим такую задачу. Пусть f — функция (скажем, бесконечно дифференцируемая); предположим, что мы хотим представить ее в виде суммы степенного ряда

$$f(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \ldots + c_n(x-a)^n + \ldots, \tag{15.2}$$

сходящегося в некоторой окрестности точки a. Давайте прикинем, какими при этом должны быть коэффициенты c_i.

Проще всего с коэффициентом c_0: подставим в обе части равенства (15.2) значение $x = a$, тогда все слагаемые в правой части, кроме первого, обратятся в нуль, и получим, что $c_0 = f(a)$. Чтобы найти c_1, продифференцируем обе части (15.2):

$$f'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \ldots + nc_n(x-a)^n + \ldots$$

(вообще-то степенной ряд — это не конечная сумма, и законность такого дифференцирования надо обосновывать; поскольку наше рассуждение не строгое, а прикидочное, мы эту трудность проигнорируем). Если теперь подставить в обе части $x = a$, то в правой части обратятся в нуль все слагаемые кроме первого, откуда $c_1 = f'(a)$.

Теперь ясно, как найти c_k для произвольного k. Именно, давайте k раз продифференцируем обе части равенства (15.2) (правую часть опять будем дифференцировать почленно, не задумываясь о законности этой операции). Легко видеть, что k-я производная от $(x-a)^n$ равна

$$n(n-1)\ldots(n-k+1)(x-a)^{n-k}$$

(в частности, она тождественно равна нулю при $n < k$). Поэтому в результате k-кратного дифференцирования равенства (15.2) получится вот что:

$$f^{(k)}(x) = k!c_k + (k+1)k\ldots 2c_{k+1}(x-a) + \ldots + n(n-1)\ldots(n-k+1)(x-a)^{n-k} + \ldots$$

Подставляя теперь $x = a$, получаем, что $f^{(k)}(a) = k!c_k$, откуда

$$c_k = f^{(k)}(a)/k!; \tag{15.3}$$

эта формула годится и при $k = 0$, если считать, что нулевая производная от функции f — это сама f, а $0! = 1$.

Ряд $\sum_{n=0}^{\infty} c_n(x-a)^n$,где коэффициенты c_k задаются формулой (15.2), называется *рядом Тейлора* функции f (в точке a).

Теперь обсудим, как все эти неформальные прикидки связаны с реальностью. Оказывается, что почленное дифференцирование степенного ряда (правой части равенства (15.2)) действительно законно внутри его круга сходимости (мы это докажем в дальнейшем); беда в том, что даже для бесконечно дифференцируемой функции f ее ряд Тейлора не обязан к ней сходиться нигде (кроме, конечно, самой точки a). Максимум, чего можно гарантировать для произвольной бесконечно дифференцируемой функции — не разложение (15.2), а некоторый его суррогат, называемый формулой Тейлора. Подчеркнем, что этот «суррогат» на самом деле чрезвычайно важен.

Теорема 15.7. *Пусть $f\colon (p;q) \to \mathbb{R}$ — функция класса C^n, и пусть $a \in (p;q)$. Тогда*

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

Доказательство. Положим

$$r(x) = f(x) - f(a) - f'(a)(x-a) - \frac{f''(a)}{2!}(x-a)^2 - \ldots - \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Ясно, что r — также функция класса C^k на интервале $(p;q)$, а также что

$$r(a) = r'(a) = \ldots = r^{(n)}(a) = 0.$$

Пусть теперь $x \in (p;q)$; применим теперь «теорему Коши» (следствие 15.5) к отрезку, соединяющему a и x, положив $\psi(x) = (x-a)^r$ и $\varphi(x) = r(x)$; согласно этой теореме, для некоторого ξ_1, лежащего между a и x, имеем

$$\frac{r'(\xi_1)}{n(\xi_1-a)^{n-1}} = \frac{r(x)-r(a)}{(x-a)^n-(a-a)^n} = \frac{r(x)}{(x-a)^n}.$$

Применяя теорему Коши к отрезку, соединяющему a и ξ_1 и функциям $x \mapsto n(x-a)^{n-1}$ и $x \mapsto r'(x)$, получаем, что для некоторого ξ_2, лежащего между a и ξ_1, имеем

$$\frac{r''(\xi_2)}{n(n-1)(\xi_2-a)^{n-2}} = \frac{r'(\xi_1)}{n(\xi_1-a)^{n-1}} = \frac{r(x)}{(x-a)^n}.$$

Продолжая в том же духе, получим, что существует такое число ξ_n, лежащее между x и a, что $\frac{r(x)}{(x-a)^n} = r^{(n)}(\xi_n)/n!$. Мы доказали следующую лемму.

Лемма 15.8. *В условиях теоремы, если $x \in (p;q)$, то существует число η, лежащее между a и x, для которого*

$$\frac{r(x)}{(x-a)^n} = \frac{r^{(n)}(\eta)}{n!}.$$

Теперь можно завершить доказательство следующим образом. Нам необходимо установить, что $\lim_{x\to a} r(x)/(x-a)^n = 0$. Пусть $\{x_k\}$ — последовательность (отличных от a) чисел, сходящаяся к a; достаточно установить, что $\lim_{n\to\infty} r(x_k)/(x_k-a)^n = 0$. Ввиду леммы для всякого n существует η_n, лежащее между a и x_n, для которого $r(x_k)/(x_k-a)^n = r^{(n)}(\eta_k)n!$. По теореме о двух милиционерах имеем $\lim \eta_k = a$; поскольку функция $r^{(k)}$ непрерывна (так как $r \in C^k(p;q)$), имеем

$$\lim_{n\to\infty} r^{(n)}(\eta_k) = r^{(n)}(a) = 0.$$

Все доказано. $\square$