В любой области помимо профессиональной основы
первостепенную роль играет хорошее настроение
и резонанс с отдалёнными сферами бытия.

МА-2. *ПРОИЗВОДНАЯ И ДИФФЕРЕНЦИАЛ*

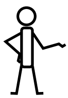
§1. Производная

Производная $f'(x)$ — это ***мгновенная скорость изменения*** функции $f(x)$ в точке x. Если x — время[1], а $f(x)$ — пройденный путь, производная $f'(x)$ — это обычная мгновенная скорость в момент времени x. В ситуации «x — время», производную обозначают точкой сверху, $\dot{f}$.

Подсознание имеет смысл приучать к интерпретации $f'(x)$ как скорости изменения $f(x)$ при изменении аргумента x, *независимо от природы* x и $f(x)$. Скажем, вполне естественно говорить о скорости $f'(x)$ изменения плотности $f(x)$ атмосферы по мере увеличения высоты x.

Теперь о формальном **определении**.

• ***Производной*** *функции $f(x)$ в точке x называется предел*



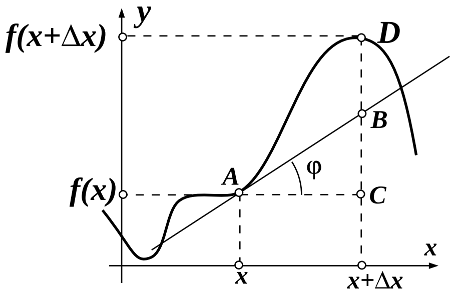
$$\boxed{f'(x) = \lim_{\triangle x \to 0} \frac{f(x+\triangle x) - f(x)}{\triangle x},} \qquad (1)$$

если таковой существует. Операцию взятия производной называют ***дифференцированием****. При существовании предела (1) функцию называют дифференцируемой в точке*[2] *x.*

[1]Коль скоро внутри возникает дискомфорт, когда время обозначается не буквой t, — над этим стоит поработать. Подсознание не должно приклеивать суть к случайным обстоятельствам.

[2]И говорят, что $f(x)$ *дифференцируема на множестве* $X \subset \mathbb{R}$, если она дифференцируема в любой точке $x \in X$.

На картинке ситуация выглядит так



Касательная к графику $y = f(x)$ в точке A тем точнее приближает дугу AD, чем меньше Δx. Разумеется, это само по себе сказано не вполне точно. Во-вторых, — ниоткуда не следует. Так будет, когда всё хорошо, т. е. существует предел (1). Тогда при уменьшении Δx вертикальная прямая CD движется влево, и $BD/DC \to 0$, что и является эквивалентом существования предела (1), который в терминах геометрического рисунка оказывается равным $\operatorname{tg}\varphi$, т. е. *производная численно равна тангенсу угла наклона касательной к графику $f(x)$ в точке x*,

$$\boxed{f'(x) = \operatorname{tg}\varphi}$$

Имеет смысл сразу осознать, что существование предела (1) — штука достаточно естественная и широко распространёная. Пусть $f(x) = x^2$, тогда в соответствии с (1)

$$f'(x) = \lim_{\triangle x \to 0} \frac{(x + \triangle x)^2 - x^2}{\triangle x} = \lim_{\triangle x \to 0} (2x + \triangle x) = 2x.$$

Совсем просто получается $x' = 1$. **(?)** Почти так же легко вычисляются производные других элементарных функций. Но прежде чем этим заниматься, мы рассмотрим некоторые вспомогательные инструменты, облегчающие взятие производных конкретных функций.

§2. Правила дифференцирования

1. Если известна производная $f'(x)$, то для $y = cf(x)$, c — константа, можно сразу написать $y' = cf'(x)$, ибо

$$\frac{cf(x+\triangle x) - cf(x)}{\triangle x} = c\frac{f(x+\triangle x) - f(x)}{\triangle x} \to cf'(x).$$

- $(\pi x^2)' = 2\pi x, \qquad (3x)' = 3.$

2. *Производная суммы равна сумме производных*

$$\boxed{(f+g)' = f' + g'}$$

Доказательство — в качестве упражнения.

- $(x^2 - 7x)' = 2x - 7.$

3. *Производная произведения* равна

$$\boxed{(fg)' = f'g + fg'} \tag{2}$$

◀ Пусть $y = f \cdot g$, а $\triangle f$, $\triangle g$ — обозначают приращения функций, например, $\triangle f = f(x+\triangle x) - f(x)$. Тогда

$$y' = \lim_{\triangle x \to 0} \frac{(f+\triangle f)(g+\triangle g) - fg}{\triangle x} = \lim_{\triangle x \to 0}\left(\frac{\triangle f}{\triangle x}g + f\frac{\triangle g}{\triangle x} + \frac{\triangle f}{\triangle x}\triangle g\right),$$

что и даёт (2), поскольку

$$\frac{\triangle f}{\triangle x} \to f', \quad \frac{\triangle g}{\triangle x} \to g', \quad \frac{\triangle f}{\triangle x}\triangle g \to f'\triangle g \to 0. \quad ▶$$

- $(x^3)' = (x^2 \cdot x)' = 2x \cdot x + x^2 \cdot 1 = 3x^2.$

4. *Правило дифференцирования сложной функции (композиции функций)*,

$$\boxed{y = f(g(x)) \quad \to \quad y' = f'\,g'}$$

Здесь f дифференцируется «по всему своему аргументу, т. е. по $g(x)$». Допустим, $y = (3x)^2$. Полагаем $g(x) = 3x$, $f(\cdot) = (\cdot)^2$. Дифференцирование квадрата (по $3x$) даёт $2 \cdot 3x$, $(3x)' = 3$, поэтому $y' = 18x$.

◀ Доказательство занимает одну строчку,

$$y' = \lim_{\triangle x \to 0} \frac{f(g + \triangle g) - f(g)}{\triangle g} \frac{\triangle g}{\triangle x} = f'\, g'. \quad ▶$$

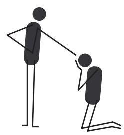
Понятно, что цепное правило дифференцирования сложной функции можно индуктивно продолжить

$$y = f(g(h(x))) \quad \Rightarrow \quad y' = f'_g\, g'_h\, h'_x, \tag{3}$$

и так далее. Нижний индекс в (3) подчёркивает (напоминает), по какому аргументу производится дифференцирование.

5. *Производная частного*

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



легко устанавливается с помощью предыдущих правил. Частное можно считать произведением f и $1/g$, а производную $(1/g)'$ вычислять как производную сложной функции[3].

6. *Производная обратной функции*. Если для $y = f(x)$ определена обратная функция $x = f^{-1}(y)$, то[4]

$$\boxed{x'_y = \frac{1}{y'_x}}$$

что легко получается[5] предельным переходом в очевидном равенстве $\triangle x/\triangle y = 1/(\triangle y/\triangle x)$.

---

[3]Здесь мы, немного забегая вперёд, имеем в виду $\left(\dfrac{1}{x}\right)' = -\dfrac{1}{x^2}$.

[4]Ещё раз обратите внимание на использование нижнего индекса, который указывает, по какому аргументу производится дифференцирование. Это удобный приём, широко используемый в анализе.

[5]Мы избегаем «разжевывания», которое иногда создаёт иллюзию исчерпывающего объяснения, но всегда топит суть в трясине деталей.

Человек, увы, мыслит в «обратном направлении» всегда хуже. Трижды набивший оскомину пример тюбика, из которого выдавить пасту легче, чем загнать её обратно, образно отражает положение дел. Именно в связи с трудностями манипулирования обратными функциями — использование правила $x'_y = 1/y'_x$ часто вызывает затруднения. Осваивать такие вещи лучше всего на примерах. Определим производную арксинуса, исходя из формулы для производной синуса, $(\sin x)' = \cos x$. Пусть $y = \sin x, \quad x = \arcsin y$. Тогда

$$x'_y = \frac{1}{y'_x} = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}},$$

что даёт необходимый результат. При желании буквы x, y теперь можно поменять местами, $(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$.

§3. Дифференциалы

Определение производной (1) означает

$$f'(x) = \frac{f(x+\triangle x) - f(x)}{\triangle x} + \xi(\triangle x), \tag{4}$$

где $\xi(\triangle x) \to 0$ при $\triangle x \to 0$. Равенство (4) запишем в виде[6]

$$\triangle f = f'(x) \cdot \triangle x + o(\triangle x), \tag{5}$$

где $o(\triangle x) = \xi(\triangle x)\triangle x$, и потому $o(\triangle x)$ не просто стремится к нулю при $\triangle x \to 0$, а гораздо быстрее, чем $\triangle x$. Точнее говоря,

$$\frac{o(\triangle x)}{\triangle x} \to 0 \quad \text{при} \quad \triangle x \to 0. \tag{6}$$

Условие (6) определяет ***о-малое***, $o(\cdot)$, как стенографический трюк, удобный и широко используемый. Для усвоения полезно объяснить соседу, что такое

$$o(x), \qquad o(x^3), \qquad o(x^2-1), \qquad o\left(\frac{1}{x}\right).$$

[6] Здесь $\triangle f = f(x+\triangle x) - f(x)$.

Два слова о понятии ***порядка малости***, позволяющем в одно касание уходить от неуклюжих пояснений. Пусть $\alpha(x)$ и $\beta(x)$ — бесконечно малые величины, т. е. функции, убывающие до нуля при $x \to 0$, и

$$\frac{\alpha(x)}{\beta(x)} \to 0 \quad \text{при} \quad x \to 0.$$

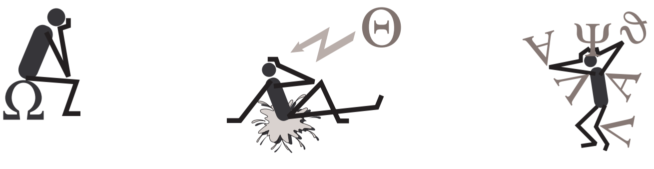
Тогда говорят, что $\alpha(x)$ имеет *более высокий порядок малости* чем $\beta(x)$. На этом языке $o(\triangle x)$ — величина более высокого порядка малости чем $\triangle x$.

Перейдём теперь к главной теме параграфа.

• *Линейная часть приращения $\triangle y$, равная $f'(x)\triangle x$ в представлении (5), называется* ***дифференциалом*** *функции $y = f(x)$ в точке x, и обозначается $dy = df$.*

Приращение $\triangle x$ в этом контексте обозначают как dx, и вместо (5) в линейном приближении имеем

$$df = f'(x) \cdot dx. \tag{7}$$



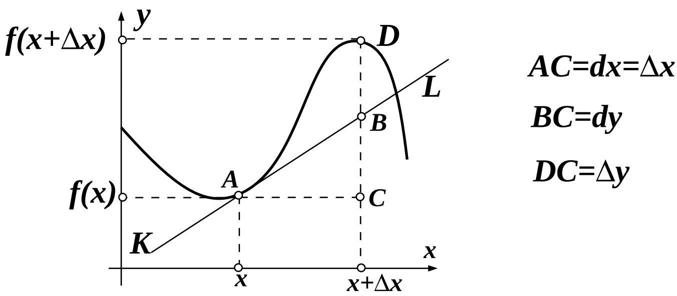
Поскольку рассмотрению дифференциалов сопутствует обычно большая неразбериха, остановимся на некоторых реверансах, хотя это и сопряжено с издержками типа «объяснения анекдота». Часто определяют df как *линейную функцию*, что у новичков уводит почву из под ног. Это функция в том же самом смысле, что и в случае $y = 2x$, переменная y — функция, но при данном значении аргумента y — это уже число. Точно также dy при заданном dx — это число. И потому с дифференциалами можно обращаться как с числами, делить одно на другое, сокращать. В частности, деление (7) на dx даёт

$$f'(x) = \frac{df}{dx}, \tag{8}$$

что служит ещё одним общепринятым обозначением производной.

Кроме того, дифференциалы — это отнюдь не «маленькие дельта», как иногда думают. Они могут принимать любые значения.

Ситуацию (7) адекватно описывает следующее представление. Функция $y = f(x)$ в окрестности точки x аппроксимируется линейным приближением $s = f'(x)r$, где r и s обозначают линейные приращения вдоль осей x, y. Обозначая r через dx, s — через dy, имеем (7). На рисунке



линейная аппроксимация $y = f(x)$ представляет собой прямую KL, касающуюся графика $f(x)$ в точке A, и описываемую линейной функцией $df = f'(x) \cdot dx$ в координатах $\{dx, dy\}$ с началом отсчёта в точке A.

Всё это не имело бы особого смысла, если бы речь шла только о ситуациях типа (5), (7). Выигрыш от рассмотрения дифференциалов невелик, а головной боли много[7]. Выгоды, конечно, есть. Например, при работе с приращениями функций приходится использовать приближённые равенства $\triangle y \approx f'(x)\triangle x$, таская из строчки в строчку хвосты

[7]Поначалу действительно не ясно, зачем вводятся дифференциалы. Тот небольшой выигрыш, который они дают, на первом курсе выглядит неубедительно, ибо пока ничего не известно о «территориях», где всё это может хорошо «выстрелить». Но если присмотреться, то это хорошо «стреляет» и в самых простых ситуациях.

нелинейных добавок $o(\cdot)$. Понятие дифференциала освобождает от этих неудобств, позволяя писать для линейных частей приращений абсолютно строгие равенства. Это особенно удобно в более сложных ситуациях. Например, для $y = f(x)g(x)$, имеем[8]

$$\boxed{dy = df \cdot g + f \cdot dg},$$

что сразу получается умножением формулы производной произведения двух функций на dx. Одна лишь возможность обращаться с df/dx, как с обыкновенной дробью, порождает массу удобств. Например, $\dfrac{df}{dx}\dfrac{dx}{dt} = \dfrac{df}{dt}$.

• **Инвариантность формы дифференциала.** Допустим, в $y = f(x)$ производится замена $x = u(\tau)$, после которой x становится зависимой переменной. И хотя теперь уже $\triangle x \neq dx$ — равенство $dy = f'(x)dx$ *сохраняется*, в чем легко убедиться используя формулу дифференцирования сложной функции:

$$dy = y'_\tau d\tau = y'_x x'_\tau d\tau = y'_x dx.$$

Это свойство называют *инвариантностью формы дифференциала.*

Подобные «пустячки» пускают анализ в другое русло. Дифференциалы шаг за шагом обретают другой облик. Становится ясно, что линейные аппроксимации функций представляют собой удобную и продуктивную категорию мышления. Оказывается, их (аппроксимации – дифференциалы) в своей нише можно полезным образом комбинировать, оставляя за бортом кандалы нелинейных добавок.

Формула для производной dy/dx, допускающая сокращения как в обыкновенных дробях (но не буквы d!), часто оказывается предпочтительнее. Например, в случае параметрического задания кривой,

$$x = \varphi(\tau), \qquad y = \psi(\tau),$$

производную y'_x с помощью дифференциалов можно вычислить, не восстанавливая зависимости $y(x)$:

[8] Не говоря уж о $d(f + g) = df + dg$ или $d\zeta f = \zeta df$.

$$y'_x = \frac{dy}{dx} = \frac{y'_\tau d\tau}{x'_\tau d\tau} = \frac{\psi'(\tau)}{\varphi'(\tau)}$$

• ***Дифференциалы высших порядков.*** *Вторым дифференциалом* d^2y *функции* $y = f(x)$ *называется (первый) дифференциал функции* $dy = f'(x)dx$ *(как функции* x*, но не* dx*), т. е.*

$$d^2y = d(f'(x))dx = (f''(x)dx)dx = f''(x)dx^2.$$

Дифференциалы более высокого порядка определяются индуктивно[9], $d^ny = f^{(n)}(x)dx^n$.

В отличие от дифференциалов первого порядка, для которых всегда $dy = f'(x)dx$ (инвариантность формы), приведённые формулы для дифференциалов d^n $(n > 1)$ справедливы лишь в предположении независимости аргумента x. Если $x = \varphi(t)$, то

$$d^2y = [f(\varphi(t))]''_{t^2}dt^2 = [f'(\cdot)\varphi'(\cdot)]'dt^2 = (f''\cdot(\varphi')^2 + f'\cdot\varphi'')dt^2 = f''dx^2 + f'd^2x.$$

[9]Вторая производная $f''(x)$ получается дифференцированием $f'(x)$. Дифференцируя $f(x)$ n раз, — будем иметь n-ю производную $f^{(n)}(x)$.