Г. В. Дятлов

ОСНОВЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА

для студентов-физиков

ЛЕКЦИИ

2. Дифференциальное исчисление функций одной переменной

Новосибирск ○ 2014

УДК 517.1
ББК 22.16
Д998

Дятлов Г. В.

Основы математического анализа для студентов-физиков. Лекции. 2. Дифференциальное исчисление функций одной переменной / Г. В. Дятлов. — Новосибирск: Издательство Института математики, 2014. — 62 с.

ISBN 978–5–86134–149–3

Второй раздел курса лекций, прочитанных автором студентам физического факультета Новосибирского государственного университета в 2014/15 учебном году. Включает в себя элементы дифференциальноего исчисления функций одной переменной.

Для студентов специальностей с углубленным изучением математического анализа.

УДК 517.1
ББК 22.16
Д998

Д$\frac{1602070000–02}{\text{Я}82(03)–14}$ Без объявл.

ISBN 978–5–86134–149–3

§ 1. Определение и основные свойства производной

1.1. Производная — определение, геометрический и физический смысл.

Определение производной. Пусть $f : \langle a, b\rangle \to \mathbb{R}$ — функция, заданная на промежутке $\langle a, b\rangle$, и $x_0 \in \langle a, b\rangle$ — фиксированная точка этого промежутка. Если существует конечный предел

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

то его называют *производной функции f в точке x_0* и обозначают через $f'(x_0)$ или $\dfrac{df}{dx}(x_0)$.

Если $f(x)$ имеет производную $f'(x)$ в каждой точке $x \in \langle a, b\rangle$, то возникает функция

$$x \mapsto f'(x),$$

которую также называют *производной функции* $f(x)$ (*на промежутке*).

Таким образом, когда говорят о производной, то либо имеют в виду число, если это производная в фиксированной точке, либо функцию, если рассматривают производную в каждой точке промежутка. В дальнейшем для нас производная будет ассоциироваться преимущественно с функцией.

Существование конечного предела в определении производной с учетом равенства нулю предела знаменателя гарантирует, что предел числителя также равен нулю, а это означает, что $\lim\limits_{x \to x_0} f(x) = f(x_0)$, следовательно, если функция имеет производную в некоторой точке, то она в такой точке непрерывна. Обратное утверждение неверно — непрерывность не гарантирует наличия производной.

Физическая интерпретация производной. Если x — время, а $f(x)$ — координата прямолинейно движущегося тела, то отношение $\dfrac{f(x) - f(x_0)}{x - x_0}$ — это средняя скорость на промежутке $[x_0, x]$, а $f'(x_0)$ — мгновенная скорость в момент времени x_0.

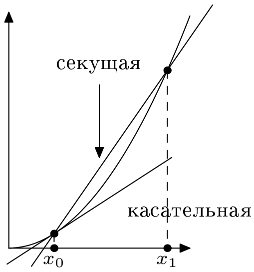

Рис. 1.1.

Геометрическая интерпретация производной. Рассмотрим функцию $f : (a, b) \to \mathbb{R}$ и зафиксируем точку $x_0 \in (a, b)$. Выберем произвольно еще одну точку $x_1 \in (a, b)$ и проведем прямую, проходящую через точки $(x_0, f(x_0))$ и $(x_1, f(x_1))$ графика $f(x)$. Эта прямая называется *секущей*, проходящей через данные точки. Ее уравнение имеет вид

$$y = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) + f(x_0).$$

Если функция $f(x)$ имеет в точке x_0 производную $f'(x_0)$, то при $x_1 \to x_0$ угловой коэффициент секущей стремится к производной $f'(x_0)$, а уравнение секущей переходит в уравнение

$$y = f'(x_0)(x - x_0) + f(x_0) = f'(x_0)x + (f(x_0) - f'(x_0)x_0),$$

которое задает некоторую прямую. Ее называют *касательной к графику функции* $f(x)$ *в точке* x_0. Имея в виду эти построения, говорят, что касательная является предельным положением секущей, а производная — угловым коэффициентом касательной к графику функции.

Определение дифференциала. Пусть $f : \langle a, b\rangle \to \mathbb{R}$ — функция, заданная на промежутке $\langle a, b\rangle$, и $x_0 \in \langle a, b\rangle$ — фиксированная точка данного промежутка. Если существует такое линейное отображение $L : \mathbb{R} \to \mathbb{R}$, что

$$f(x_0 + h) = f(x_0) + L(h) + o(h), \quad h \to 0,$$

то говорят, что f *дифференцируема в точке* x_0, а отображение L называют *дифференциалом функции* $f(x)$ *в точке* x_0 и обозначают символом $df(x_0)(h)$.

Некоторая необычность ситуации в определении дифференциала в том, что здесь функция зависит от двух переменных, каждый из которых заключен в отдельные скобки. Первый аргумент — та точка, в которой все происходит, в контексте определения она фиксирована. Второй аргумент — это величина приращения. В этой записи набор символов $df(x_0)$ надо воспринимать как единое обозначение для дифференциала в данной точке.

Замечание о линейных и аффинных функциях. Отображение $L : \mathbb{R} \to \mathbb{R}$, обладающее следующим свойством:

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y), \quad \alpha, \beta \in \mathbb{R}, \quad x, y \in \mathbb{R},$$

называют *линейным*. Иначе говоря, отображение линейно, если оно переводит линейную комбинацию элементов из области определение в соответствующую линейную комбинацию образов этих элементов. Линейное отображение можно рассматривать не только на $\mathbb{R}$, но и на любом векторном пространстве — важна возможность составлять линейные комбинации элементов со скалярами, т. е. числами.

В случае, когда L рассматривается на $\mathbb{R}$, разнообразие линейных отображений невелико — все они имеют вид $L(h) = a \cdot h$, $h \in \mathbb{R}$. Заметим, что график любого линейного отображения проходит через начало координат. В школе линейными называли функции вида $f(x) = ax + b$ в честь того, что график такого отображения — прямая на плоскости, не обязательно проходящая через начало координат. У нас такое отображение будет называться *аффинным*.

Оперируя имеющимися в нашем распоряжении понятиями, можно сказать, что дифференциал — это *главная линейная часть приращения функции*. Действительно, запишем определяющее дифференциал равенство в виде

$$f(x_0 + h) - f(x_0) = L(h) + o(h), \quad h \to 0.$$

В левой части последнего равенства стоит приращение функции на приращении аргумента h. В правой части первое слагаемое — это дифференциал, т. е. линейная функция вида

$$L(h) = a \cdot h.$$

Наконец, завершает сумму слагаемое, бесконечно малое по сравнению с h, что указывает на то, что $L(h)$ — не что иное как главная часть того, что расположено слева, т. е. приращения исходной функции.

В равенстве, характеризующем дифференциал, приращение аргумента записано с использованием величины h отклонения точки от фиксированной точки x_0. Можно записывать это приращение, используя не величину отклонения, а положение точки x, т. е. используя величину $h = x - x_0$.

Конечно, можно было ставить задачу выделить главную часть приращения функции какого-то другого вида, например, постоянную или квадратичную. Однако постоянные функции недостаточно квалифицированно показывают скорость изменения функции вблизи данной точки, а квадратичные уже не столь просты, как линейные. Иначе говоря, линейные функции образуют класс достаточно просто устроенных функций, показывающих свойства функции вблизи данной точки.

Замечание. Сопоставляя равенство, характеризующее дифференциал, т. е.

$$f(x) = f(x_0) + L(x - x_0) + o(x - x_0), \quad x \to x_0,$$

и уравнение касательной, т. е.

$$y = f(x_0) + f'(x_0)(x - x_0),$$

можно обнаружить, что они похожи между собой, и это намек на связь дифференциала с производной. С другой стороны, видно, насколько различаются касательная и сама функция вблизи данной точки, а именно они различаются на добавку $o(h)$. Ясно, что если x приближается к x_0, то разность между функцией и касательной стремится к нулю, причем довольно быстро: скорость ее исчезновения выше скорости, с которой происходит приближение x к x_0.

Теорема 1 (о производной и дифференциале). *Пусть $f : \langle a, b\rangle \to \mathbb{R}$, $x_0 \in \langle a, b\rangle$. Тогда функция $f(x)$ имеет производную $f'(x_0)$ в точке x_0 тогда и только тогда, когда f имеет дифференциал $df(x_0)$, т. е. дифференцируема в точке x_0, при этом*

$$df(x_0)(h) = f'(x_0)h, \quad h \in \mathbb{R}.$$

Доказательство. Необходимость. Предположим, что существует $f'(x_0)$. Это означает существование предела

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0). \qquad (*)$$

Введем обозначение для разности:

$$\alpha(h) = \frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0).$$

Из $(*)$ следует, что

$$\alpha(h) = \frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0) \to 0 \quad \text{при } h \to 0.$$

Преобразования последнего равенства приводят к такому:

$$f(x_0+h) = f(x_0) + f'(x_0)h + \alpha(h)h.$$

Здесь ввиду того, что $\alpha(h) \to 0$ при $h \to 0$, последнее слагаемое представляет собой $o(h)$ при $h \to 0$. Тем самым существование производной гарантирует наличие равенства, характеризующего дифференцируемость, и из него видно, что дифференциал есть линейное отображение

$$df(x_0)(h) = f'(x_0)h.$$

Достаточность. Предположим, что $f(x)$ имеет дифференциал $df(x_0)$, т. е. по определению

$$f(x) = f(x_0) + df(x_0)(x-x_0) + o(x-x_0).$$

Преобразуем это равенство с целью выделения требуемого разностного отношения:

$$\frac{f(x)-f(x_0)}{x-x_0} = \frac{df(x_0)(x-x_0)}{x-x_0} + \frac{o(x-x_0)}{x-x_0}. \qquad (\dagger)$$

Последнее слагаемое стремится к нулю при $x \to x_0$, т. е. это $o(1)$. Заметим, что $o(1)$ — просто обозначение в терминах асимптотических сравнений для бесконечно малой функции. Далее, ввиду линейности отображения $df(x_0)$ множитель перед ним вида $\dfrac{1}{x-x_0}$ можно внести в аргумент и записать, что

$$\frac{1}{x-x_0}\cdot df(x_0)(x-x_0) = df(x_0)\left(\frac{x-x_0}{x-x_0}\right) = df(x_0)(1).$$

Тем самым первое слагаемое в правой части (†) представляет собой постоянную функцию, и она имеет предел. Следовательно, левая часть имеет предел, который, с одной стороны, равен $f'(x_0)$, а с другой — $df(x_0)(1)$. Стало быть, производная функции в точке x_0 существует и равна $f'(x_0) = df(x_0)(1)$.

Теорема доказана.

Замечание о форме записи дифференциала.

Сначала разберемся с тем, как воспринимать символ dx. Есть две точки зрения на этот символ — как у математиков и как у физиков. Сначала посмотрим, как на это смотрят математики.

Рассмотрим тождественное отображение $y = x$ на некотором промежутке. Его производная в каждой точке x_0 равна 1, т. е. $x'(x_0) = 1$ (возможно, не самое удачное обозначение для производной функции $y = x$, но возможное — так обозначается сама функция). По предыдущей теореме у этого отображения есть дифференциал и в каждой точке он действует по формуле

$$dx(x_0)(h) = 1 \cdot h = h.$$

Поскольку действие не зависит от точки x_0, первый аргумент x_0 опускают и пишут $dx(h) = h$. Таким образом, dx — это просто дифференциал тождественного отображения, который также действует тождественно. Однако между тождественным отображением и его дифференциалом есть существенная разница: первое действует на точки $x \in \langle a, b\rangle$, а дифференциал, будучи линейным отображением, определен на $\mathbb{R}$ и действует на приращения h. Позже, при изучении функций многих переменных, будем говорить, что дифференциал при фиксированном x_0 определен на касательном пространстве и действует на касательные векторы.

Имея отображение dx, можно записать без указания аргументов связь между производной и дифференциалом

$$df(x_0)(h) = f'(x_0)h = f'(x_0)dx(h),$$

установленную в теореме 1 и представляющую собой равенство линейных функций или, можно сказать, разные записи одного и того же линейного отображения. Опуская последний аргумент h, приходим к такому выражению:

$$df(x_0) = f'(x_0)dx,$$

где в каждой из частей равенства находятся линейные функции. Эта формула означает, что дифференциал $df(x_0)$ получается из тождественного линейного отображения dx при помощи умножения на $f'(x_0)$ или что $f'(x_0)$ — коэффициент пропорциональности между функциями $df(x_0)$ и dx. Это объясняет обозначение $\dfrac{df}{dx}(x_0)$ для производной $f'(x_0)$.

Теперь обсудим, как dx воспринимают физики. Если приращение аргумента Δx большое, то дифференциал приближает приращение функции с некоторой ошибкой:

$$\Delta f(x_0) = f(x) - f(x_0) = f'(x_0)\underbrace{(x - x_0)}_{\Delta x} + \underbrace{o(x - x_0)}_{\text{ошибка}},$$

где ошибка исчезает быстрее, чем приращение аргумента. А если приращение настолько малое, что ошибкой можно пренебречь, то такое приращение называют бесконечно малым приращением и обозначают через dx, при этом имеет место равенство

$$df(x_0) = f'(x_0)\,dx,$$

в котором dx — бесконечно малое приращение аргумента, а $df(x_0)$ — бесконечно малое приращение функции, $f'(x_0)$ — коэффициент пропорциональности между ними. Тем самым у физиков dx — это просто приращение аргумента, $df(x_0)$ — приращение функции, только бесконечно малое. И когда физики говорят «записать дифференциал функции», то имеют в виду намерение выразить бесконечно малое приращение функции через бесконечно малое приращение аргумента. У математиков это линейная функции, т. е. другой объект.

Пример. Возьмем какую-либо простую функцию, например, $f(x) = ax^2 + bx + c$, и найдем ее дифференциал в точке x_0. Если руководствоваться теоремой 1, то для нахождения дифференциала надо взять производную $f'(x) = 2ax + b$, рассмотреть ее в точке x_0 и затем записать дифференциал $df(x_0)(h) = (2ax_0 + b) \cdot h$. Если учесть замечание, то можно записать дифференциал в виде

$$df(x_0) = (2ax_0 + b)\,dx.$$

Мы можем говорить о скорости изменения функции на двух языках — на языке производных и на языке дифференциалов. Для функций одной переменной эти языки очевидно эквивалентны. Понятнее язык производной, поэтому дальнейшие результаты будем формулировать на языке производных, хотя их можно переформулировать на языке дифференциалов.

1.3. Общие правила дифференцирования.

Теорема 2 (о производной и алгебраических операциях). *Пусть $f, g : \langle a, b\rangle \to \mathbb{R}$ имеют производные в точке $x_0 \in \langle a, b\rangle$. Тогда их* *сумма, произведение и частное тоже имеют производные, причем*

$$\begin{aligned}(f+g)'(x_0) &= f'(x_0)+g'(x_0),\\ (fg)'(x_0) &= f'(x_0)g(x_0)+f(x_0)g'(x_0),\\ \left(\frac{f}{g}\right)'(x_0) &= \frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g^2(x_0)}\end{aligned}$$

(последнее при условии, что $g(x_0)\neq 0$*).*

Доказательство. Первое утверждение совсем простое, третье предполагает дополнительный объем технических деталей, связанных с исключением нуля в знаменателе, поэтому докажем только второе.

Используя непрерывность f и g в точке x_0, имеем

$$\begin{aligned}\frac{(fg)(x)-(fg)(x_0)}{x-x_0} &= \frac{f(x)g(x)-f(x_0)g(x)+f(x_0)g(x)-f(x_0)g(x_0)}{h}\\ &= \frac{f(x)-f(x_0)}{h}g(x)+\frac{g(x)-g(x_0)}{x-x_0}f(x_0)\\ &\to f'(x_0)g(x_0)+f(x_0)g'(x_0), \quad x\to x_0.\end{aligned}$$

Теорема доказана.

Теорема 3 (о производной композиции). *Пусть* $f:\langle a,b\rangle\to\langle c,d\rangle$, $g:\langle c,d\rangle\to\mathbb{R}$ *и* $x_0\in\langle a,b\rangle$*. Если*

(1) f *имеет производную* $f'(x_0)$ *в точке* $x_0\in\langle a,b\rangle$,

(2) g *имеет производную* $g'(y_0)$ *в точке* $y_0=f(x_0)$,

то композиция $(g\circ f)(x)=g(f(x))$ *имеет производную в точке* x_0 *и*

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

Доказательство. По условию существует $g'(y_0)$, стало быть,

$$g(y)=g(y_0)+g'(y_0)(y-y_0)+\alpha(y)(y-y_0), \tag{$*$}$$

где $\alpha(y)$ определена формулой

$$\alpha(y)=\begin{cases}\dfrac{g(y)-g(y_0)}{y-y_0}-g'(y_0) & \text{при } y\neq y_0,\\ 0 & \text{при } y=y_0\end{cases}$$

и $\alpha(y)\to 0$, $y\to y_0$. Заметим, что при таком определении функция $\alpha(y)$ непрерывна в точке y_0. Подставляя $y=f(x)$ и $y_0=f(x_0)$ в равенство $(*)$ и пользуясь непрерывностью $\alpha(y)$ и $f(x)$ в нужных точках, получаем

$$g(f(x))=g(f(x_0))+g'(f(x_0))(f(x)-f(x_0))+\alpha(f(x))(f(x)-f(x_0)),$$

откуда

$$\frac{g(f(x))-g(f(x_0))}{x-x_0} = g'(f(x_0))\underbrace{\frac{f(x)-f(x_0)}{x-x_0}}_{\to f'(x_0)} + \underbrace{\frac{f(x)-f(x_0)}{x-x_0}}_{\to f'(x_0)}\underbrace{\alpha(f(x))}_{\to 0}$$

$$\to g'(f(x_0))f'(x_0), \quad x\to x_0.$$

При переходе к пределу в последнем слагаемом учтено то обстоятельство, что, $\alpha(f(x)) \to 0$ по теореме о непрерывности композиции, а дробь, на которую это выражение умножается, ограничена, ибо имеет предел.

Теорема доказана.

Теорема 4 (о производной обратной функции). *Пусть* $f : \langle a,b\rangle \to \mathbb{R}$ *обратима и* $x_0 \in \langle a,b\rangle$. *Если*

(1) $f(x)$ *имеет ненулевую производную* $f'(x_0) \neq 0$,

(2) $f^{-1}(y)$ *непрерывна в точке* $y_0 = f(x_0)$,

то $f^{-1}(y)$ *имеет производную в точке* y_0, *причем*

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

Доказательство. По условию теоремы

$$\frac{f(x)-f(x_0)}{x-x_0} \to f'(x_0), \quad x\to x_0. \tag{$*$}$$

Поскольку $f(x)$ обратима, $f(x)-f(x_0)\neq 0$ при $x\neq x_0$. Значит, по теореме $10'$ из гл. 1 дробь в $(*)$ можно «перевернуть»:

$$\frac{x-x_0}{f(x)-f(x_0)} \to \frac{1}{f'(x_0)}, \quad x\to x_0.$$

Воспользуемся теоремой 16 из гл. 1 о пределе композиции и сделаем замену $x = f^{-1}(y)$ для y, близких к y_0, при этом $x_0 = f^{-1}(y_0)$. Отметим, что

- $f^{-1}(y)\neq x_0$ при $y\neq y_0$;
- $f^{-1}(y)\to f^{-1}(y_0) = x_0$, так как f^{-1} непрерывна в y_0 по условию.

В итоге

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

Теорема доказана.

1.3. Производные элементарных функций.

Теорема 5 (о производных элементарных функций). *Имеют место равенства:*

$$(a^x)' = a^x \ln a, \quad x \in \mathbb{R}, \tag{1}$$

$$(\log_a x)' = \frac{1}{x \ln a}, \quad x > 0, \tag{2}$$

$$(x^\alpha)' = \alpha x^{\alpha-1}, \quad x > 0, \tag{3}$$

$$(\sin x)' = \cos x, \quad x \in \mathbb{R}, \tag{4}$$

$$(\cos x)' = -\sin x, \quad x \in \mathbb{R}, \tag{5}$$

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x}, \quad x \in \mathbb{R},\ x \neq \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}, \tag{6}$$

$$(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}, \quad x \in \mathbb{R},\ x \neq k\pi,\ k \in \mathbb{Z}, \tag{7}$$

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}, \quad -1 < x < 1, \tag{8}$$

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}, \quad -1 < x < 1, \tag{9}$$

$$(\operatorname{arctg} x)' = \frac{1}{1+x^2}, \quad x \in \mathbb{R}, \tag{10}$$

$$(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}, \quad x \in \mathbb{R}. \tag{11}$$

Доказательство. Равенства (1)–(5) доказывается на основе замечательных пределов с использованием теоремы 16 гл. 1 о пределе композиции по одной схеме:

$$\begin{aligned}\frac{a^x - a^{x_0}}{x - x_0} &= a^{x_0}\frac{a^{x-x_0}-1}{x-x_0}\\ &= a^{x_0}\ln a\frac{\exp((x-x_0)\ln a)-1}{(x-x_0)\ln a} \to a^{x_0}\ln a, \quad x \to x_0;\end{aligned}$$

$$\begin{aligned}\frac{\log_a x - \log_a x_0}{x - x_0} &= \frac{\log_a\left(\frac{x}{x_0}\right)}{\frac{x-x_0}{x_0}}\frac{1}{x_0}\\ &= \frac{\ln\left(\frac{x-x_0}{x_0}+1\right)}{\frac{x-x_0}{x_0}}\frac{1}{x_0\ln a} \to \frac{1}{x_0\ln a}, \quad x \to x_0;\end{aligned}$$

$$\frac{x^\alpha - x_0^\alpha}{x - x_0} = x_0^\alpha \frac{\left(\frac{x}{x_0}\right)^\alpha - 1}{x - x_0} = x_0^{\alpha-1} \frac{\left(\frac{x}{x_0}\right)^\alpha - 1}{\frac{x}{x_0} - 1} \to \alpha x_0^{\alpha-1}, \quad x \to x_0.$$

$$\frac{\sin x - \sin x_0}{x - x_0} = \frac{2 \sin \frac{x-x_0}{2} \cos \frac{x+x_0}{2}}{x - x_0} = \underbrace{\frac{\sin \frac{x-x_0}{2}}{\frac{x-x_0}{2}}}_{\to 1} \cos \frac{x + x_0}{2} \to \cos x_0.$$

Производные тангенса и котангенса вычисляются как производные отношения по теореме 2. Проделаем выкладки, например, для тангенса:

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

Формулы (8)–(11) связаны с производными обратных тригонометрических функций и находятся по теореме о производной обратной функции. Найдем, к примеру, производную арксинуса, остальные находятся аналогично. Пусть $f(x) = \sin x$, $f^{-1}(y) = \arcsin y$, $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $y \in [-1, 1]$. По предыдущему шагу $\sin x$ имеет *ненулевую* производную на $(-\frac{\pi}{2}, \frac{\pi}{2})$. Можно также показать, что обратная функция $\arcsin y$ непрерывна (однако это выходит за рамки нашего курса). По теореме 4

$$(\arcsin y_0)' = (f^{-1}(y_0))' = \frac{1}{f'(x_0)} = \frac{1}{\cos x_0} = \frac{1}{\sqrt{1 - \sin^2 x_0}} = \frac{1}{\sqrt{1 - y_0^2}}.$$

Теорема доказана.

§ 2. Приращения дифференцируемых функций

2.1. Производная в точке экстремума.

Определение локального экстремума. Пусть $f : \langle a, b\rangle \to \mathbb{R}$ — функция на промежутке $\langle a, b\rangle$. Точку $x_0 \in \langle a, b\rangle$ называют *точкой локального максимума* f, если существует такая окрестность U точки x_0, что для любого $x \in U \cap \langle a, b\rangle$ выполнено неравенство $f(x) \leqslant f(x_0)$, при этом говорят, что f имеет локальный максимум в точке x_0. Точку $x_0 \in \langle a, b\rangle$ называют *точкой локального минимума* f, если существует такая окрестность U точки x_0, что для любого $x \in U \cap \langle a, b\rangle$ выполнено неравенство $f(x) \geqslant f(x_0)$, при этом говорят, что f имеет локальный минимум в точке x_0.

Точки локального максимума и минимума называют *точками локального экстремума.*

Определение внутренней точки. Точку $x_0 \in \langle a, b\rangle$ называют *внутренней*, если существует такая ее окрестность $U = (x_0-\delta, x_0+\delta)$, $\delta > 0$, которая содержится в $\langle a, b\rangle$.

В полной мере свойство, характеризующее внутренние точки, скажется в многомерном случае, в нашем, одномерном, достаточно понимать, что внутренние точки промежутка $\langle a, b\rangle$ — это точки интервала (a, b).

Теорема 6 (Ферма о необходимом условии локального экстремума). *Если функция* $f : \langle a, b\rangle \to \mathbb{R}$

(1) *имеет локальный экстремум во внутренней точке* $x_0 \in (a, b)$,

(2) *имеет производную* $f'(x_0)$ *в точке* x_0,

то $f'(x_0) = 0$.

Доказательство. Не теряя общности, предположим, что x_0 — точка локального максимума. Поскольку x_0 одновременно точка локального максимума и внутренняя точка, существует окрестность $U = (x_0-\delta, x_0+\delta) \subset \langle a, b\rangle$ такая, что

$$\forall x \in U \quad f(x) \leqslant f(x_0).$$

Для точек $x \in (x_0-\delta, x_0)$ имеем

$$\frac{f(x)-f(x_0)}{x-x_0} \geqslant 0,$$

откуда, переходя к пределу, по теореме $6'$ из гл. 1 получаем $f'(x_0) \geqslant 0$. Аналогично, для точек $x \in (x_0, x_0+\delta)$ имеем

$$\frac{f(x)-f(x_0)}{x-x_0} \leqslant 0,$$

откуда $f'(x_0) \leqslant 0$. В итоге $f'(x_0) = 0$.

Теорема доказана.

2.2. Теоремы Ролля, Лагранжа и Коши.

Если точка, плавно двигаясь по прямой, возвращается в начальную точку, значит в какой-то момент она имеет нулевую скорость.

Теорема 7 (Ролля). *Если функция* $f : [a, b] \to \mathbb{R}$

(1) *непрерывна на* $[a, b]$,

(2) *имеет производную (по крайней мере) на* (a, b),

(3) $f(a) = f(b)$,

то существует такая точка $\xi \in (a, b)$, *в которой* $f'(\xi) = 0$.

Доказательство. По теореме Вейерштрасса функция f, будучи непрерывной функцией на отрезке, принимает на $[a, b]$ наибольшее и наименьшее значения, т. е.

$$\exists x_{\min}, x_{\max} \in [a, b]\ \forall x \in [a, b] \quad f(x_{\min}) \leqslant f(x) \leqslant f(x_{\max}).$$

Если $f(x_{\min}) = f(x_{\max})$, то f постоянна и $f'(x) = 0$ во всех точках $x \in (a, b)$. Если же $f(x_{\min}) < f(x_{\max})$, то в силу условия (3) хотя бы одна из точек $x_{\min}$, $x_{\max}$ внутренняя, и по теореме Ферма в этой точке f имеет нулевую производную.

Теорема доказана.

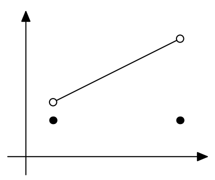

Рис. 2.1.

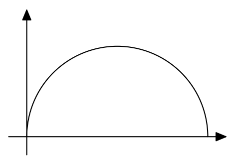

Рис. 2.2.

Замечание. Все условия теоремы 7 важны. Взяв функцию типа той, которая изображена на рис. 2.1, т. е. функцию, не обязательно непрерывную на всем отрезке, видим, что у нее производная может оказаться отличной от нуля во всех внутренних точках.

Тот факт, что достаточно ожидать наличие производной только во внутренних точках промежутка, показывает пример функции, изображенной на рис. 2.2. У этой функции нет производной в концевых точках, но тем не менее внутри есть точка, в которой производная обращается в нуль.

Теорема 8 (Лагранжа о конечном приращении). *Если функция* $f : [a, b] \to \mathbb{R}$

(1) *непрерывна на* $[a, b]$,

(2) *имеет производную (по крайней мере) на* (a, b),

то существует такая точка $\xi \in (a, b)$, *что*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi). \qquad (*)$$

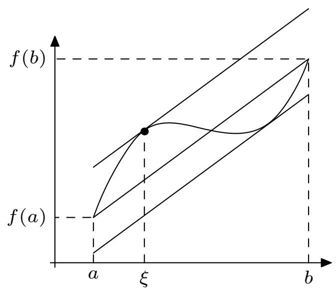

Рис. 2.3.

Содержание теоремы имеет простую геометрическую интерпретацию (рис. 2.3). В равенстве $(*)$ слева записан угловой коэффициент секущей, проходящей через точки $(a, f(a))$ и $(b, f(b))$, а справа — угловой коэффициент касательной в некоторой точке ξ. Теорема утверждает, что внутри промежутка есть хотя бы одна точка, в которой эти величины окажутся равными, т. е. касательная будет параллельна секущей.

Доказательство. Рассмотрим функцию

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

которая получается из функции f вычитанием линейной функции, проходящей через точки $(a, f(a))$ и $(b, f(b))$. Она вдобавок к первым двум условия теоремы удовлетворяет условию $F(a) = F(b)$. По теореме Ролля существует $\xi \in (a, b)$ такая, что

$$0 = F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

Теорема доказана.

Следствие (о конечном приращении). *Если функция f удовлетворяет условиям теоремы Лагранжа и существует такое C, что $|f'(x)| \leqslant C$ для любого $x \in (a, b)$, то*

$$|f(b) - f(a)| \leqslant C|b - a|.$$

Образно говоря, если мы едем со скоростью, не превышающей C км/ч, то за один час более чем C км не проедем.

Теорема 9 (Коши о конечном приращении). *Если функции* $f, g : [a, b] \to \mathbb{R}$

(1) *непрерывны на* $[a, b]$,

(2) *имеют производные* (*по крайней мере*) *на* (a, b),

то существует такая точка $\xi \in (a, b)$, *что*

$$\det \begin{pmatrix} f'(\xi) & f(b) - f(a) \\ g'(\xi) & g(b) - g(a) \end{pmatrix} = 0.$$

Если $g(b) \neq g(a)$, *то результат теоремы можно записать в виде отношения:*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}. \tag{†}$$

Ясно, что теорема Коши является обобщением теоремы Лагранжа.

Доказательство. Рассмотрим функцию

$$F(x) = \det \begin{pmatrix} f(x) & f(b) - f(a) \\ g(x) & g(b) - g(a) \end{pmatrix} = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)),$$

которая, как легко проверить, удовлетворяет всем условиям теоремы Ролля, откуда заключаем, что существует такая точка $\xi \in (a, b)$, что

$$0 = F'(\xi) = (g(b) - g(a))f'(\xi) - (f(b) - f(a))g'(\xi).$$

Теорема доказана.

Посмотрим, что означает теорема Коши с геометрической точки зрения. Если значения функций f и g суть координаты точки на плоскости, т. е. это точка $(g(x), f(x))$, то при изменении x на плоскости появляется кривая с начальной точкой $(g(a), f(a))$ и конечной $(g(b), f(b))$. В левой части формулы (†) записан тангенс угла наклона секущей, проходящей через точки $(g(a), f(a))$ и $(g(b), f(b))$, а правой — тангенс угла наклона касательной к кривой в некоторой точке. Согласно теореме существует такая точка на кривой, в которой касательная будет наклонена под тем же углом к оси абсцисс, что и секущая.

§ 3. Формула Тейлора

3.1. Производные высших порядков.

Определение старших производных. Пусть функция $f : \langle a, b\rangle \to \mathbb{R}$ имеет производную в окрестности $(x_0 - \delta, x_0 + \delta)$ точки x_0. Если функция $f'(x)$ имеет производную $(f')'(x_0)$, то говорят, что f *имеет вторую производную* или *производную второго порядка* в точке x_0 и обозначают ее через $f''(x_0)$.

Старшие производные, называемые также *производными высших порядков*, определяют по индукции: если f имеет $(k-1)$-ю производную $f^{(k-1)}(x)$ в некоторой окрестности точки x_0 и эта производная дифференцируема в точке x_0, то говорят, что f *имеет k-ю производную в точке* x_0 и обозначают ее через $f^{(k)}(x_0) = (f^{(k-1)})'(x_0)$.

Важно понимать, что для определения производной порядка k в некоторой точке x_0 надо иметь производные порядка $k-1$ в точках x из некоторой окрестности точки x_0.

3.2. Глобальная формула Тейлора.

Теорема 10 (формула Тейлора с остаточным членом в форме Лагранжа). *Пусть $f : (a, b) \to \mathbb{R}$ имеет $n+1$ производную на (a, b) и $x_0 \in (a, b)$. Тогда для любого $x \in (a, b)$ имеет место равенство, называемое формулой Тейлора с остаточным членом в форме Лагранжа:*

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n}_{\substack{= p_n(x) \\ \text{полином Тейлора функции } f \text{ степени } n}} +$$

$$+ \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}}_{\substack{\text{остаточный член} \\ \text{в форме Лагранжа}}}, \quad x \in (a, b),$$

где ξ — некоторая точка, лежащая строго между x и x_0.

В формулировке теоремы имеется в виду, что сначала берется точка x и затем гарантируется существование точки ξ.

Доказательство. Шаг 1 (идея). Условия теоремы напоминают условия теоремы Лагранжа, поэтому и обоснование, скорее всего, будет связано с теоремой Лагранжа. Более того, формула Тейлора при $n = 0$ есть не что иное как равенство в теореме Лагранжа: существует такая точка ξ, расположенная между x_0 и x, что

$$f(x) = f(x_0) + f'(x_0)(x - x_0).$$

В данной случае полином Тейлора состоит из одного члена — константы $f(x_0)$.

Идея доказательства в общем случае такова: надо взять некоторую удачную функцию и применить к ней подходящую теорему о среднем. Теорема Лагранжа для этой цели слабовата, мы применим теорему Коши. Займемся выбором функции.

Шаг 2 (общая формула). Пусть x_0, x фиксированы и, не теряя общности, предположим, что $x_0 < x$. Остаточный член в формуле Тейлора представляет собой разность между функцией и полиномом Тейлора, т. е.

$$r_n(x, x_0) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k,$$

и надо показать существование такой точки $\xi \in (x_0, x)$, что

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

Рассмотрим «хитрую» функцию

$$F(t) = f(x) - \left(f(t) + \frac{f'(t)}{1!}(x - t) + \cdots + \frac{f^{(n)}(t)}{n!}(x - t)^n \right), \quad t \in [x_0, x],$$

представляющую собой остаточный член, где точка x, в которой берется значение $f(x)$, фиксирована, а точка, около которой записывается полином Тейлора, меняется. В случае, когда $n = 1$, функция $F(t)$ имеет следующий смысл. Между точками x_0 и x берем точку t и записываем полином Тейлора для этой точки. Если $n = 1$, то полином Тейлора, состоящий из двух членов, первой степени.

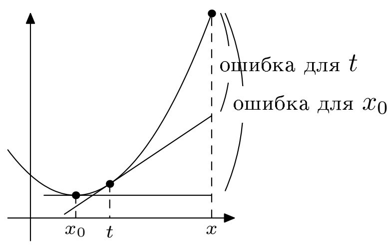

Рис. 3.1.

Изобразим на рисунке график некоторой функции f, точки x_0, x, t и график полинома Тейлора для $n = 1$ с начальной точкой t (рис. 3.1). Это прямая, наилучшим образом приближающая график функции f вблизи t. Будем эту точку перемещать, вследствие чего касательная также будет перемещаться. При каком положении точки ошибка станет самой маленькой? Когда $t = x$. А при удалении от точки x ошибка нарастает. В том случае, когда $t = x_0$, значение $F(x_0)$ станет равным величине остаточного члена, которую мы и хотим оценить. Из рисунка видно, что остаточный член — это ошибка, возникающая в результате приближения.

Отметим свойства функции $F(t)$:

(1) $F(t)$ непрерывна на $[x_0, x]$ и имеет производную на (x_0, x) (и даже на $[x_0, x]$);

(2) $F(x_0) = r_n(x)$;

(3) $F(x) = 0$;

(4) для производной имеет место равенство

$$F'(t) = -\underbrace{f'(t) + \frac{f'(t)}{1!}}_{=0} - \underbrace{\frac{f''(t)}{1!}(x-t) + \frac{f''(t)}{1!}(x-t)}_{=0} - \dots - \frac{f^{(n+1)}(t)}{n!}(x-t)^n = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Имея намерение применить теорему Коши, возьмем произвольную функцию $\varphi : (a, b) \to \mathbb{R}$, непрерывную на $[x_0, x]$ и имеющую ненулевую производную $\varphi'(x) \neq 0$ на (x_0, x). Применим к паре функций $F(t)$, $\varphi(t)$ теорему Коши, согласно которой существует такое $\xi \in (x_0, x)$, что

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

или, с учетом перечисленных свойств $F(t)$,

$$\frac{-\frac{f^{(n+1)}(\xi)(x-\xi)^n}{n!}}{\varphi'(\xi)} = \frac{0 - r_n(x)}{\varphi(x) - \varphi(x_0)}$$

или, окончательно,

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

Полученная форма остаточного члена называется *общей формой*, из которой конкретные формы получаются в результате выбора конкретных функций в качестве φ.

Шаг 3 (форма Лагранжа). Взяв $\varphi(t) = (x-t)^{n+1}$, получаем

$$\frac{\varphi(x)-\varphi(x_0)}{\varphi'(\xi)} = \frac{0-(x-x_0)^{n+1}}{-(n+1)(x-\xi)^n} = \frac{(x-x_0)^{n+1}}{(n+1)(x-\xi)^n},$$

и (†) принимает вид

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

Теорема доказана.

3.3. Локальная формула Тейлора.

Теорема 11 (формула Тейлора с остаточным членом в форме Пеáно). *Пусть $f:(a,b)\to\mathbb{R}$ имеет n производных в точке $x_0\in(a,b)$. Тогда для $x\in(a,b)$ имеет место формула*

$$f(x) = \underbrace{\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}_{\text{полином Тейлора}} + \underbrace{o((x-x_0)^n)}_{\substack{\text{остаточный член}\\ \text{в форме Пеано}}}, \quad x\to x_0.$$

Кроме того, если $q_n(x)$ — многочлен степени n такой, что

$$f(x) = q_n(x) + o((x-x_0)^n), \quad x\to x_0,$$

то $q_n(x)\equiv p_n(x)$.

Доказательство проведем при более жестких, чем в условии, ограничениях на функцию f, а именно для случая, когда f имеет производную порядка $n+1$ на (a,b), непрерывную на нем. Выведем результат теоремы из теоремы 10.

По теореме 10 существует такая точка ξ между x_0 и x, что остаточный член имеет вид

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

Рассмотрим функцию

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

По теореме Вейерштрасса функция $f^{(n+1)}(\xi)$ как функция от ξ, будучи непрерывной, ограничена на $[x_0, x]$, т. е. существует такое $C > 0$, что $|f^{(n+1)}(\xi)| \leqslant C$ для всех $\xi \in [x_0, x]$. Значит,

$$0 \leqslant |\alpha(x)| = \left(\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0) \right) \leqslant C|x - x_0| \to 0,$$

и

$$\alpha(x) \to 0 \quad \text{при } x \to x_0$$

по теореме о зажатой функции.

По определению o-малого для остаточного члена имеем

$$r_n(x) = \alpha(x)(x - x_0)^n = o((x - x_0)^n), \quad x \to x_0.$$

Теорема доказана.

Замечание 1. Формула Тейлора — это представление функции $f(x)$ в виде

$$f(x) = p_n(x; x_0) + r_n(x; x_0),$$

где $p_n(x; x_0)$ — полином Тейлора, а $r_n(x; x_0)$ — остаточный член. Вид полинома Тейлора для всех формул один и тот же, а остаточный член, представляющий собой ошибку приближения, записывается в различным формах — Лагранжа, Пеано, Коши, интегральной. Выбор вида остаточного члена зависит от задачи, которую вы решаете с помощью этого разложения. Если считаете пределы, то надо использовать формулу Тейлора с остаточным членом в форме Пеано, а если приближаете функцию, то — в форме Лагранжа.

Замечание 2 (о формуле Тейлора и ряде Тейлора). Предположим, что $f(x)$ имеет все производные на (a, b). Зафиксируем $x_0 \in (a, b)$ и запишем формулу Тейлора

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + r_n(x; x_0).$$

Пусть x тоже фиксировано, а $n \in \mathbb{N}$ меняется. Что может произойти при $n \to \infty$? В нормальном случае будет $r_n(x, x_0) \to 0$ при $n \to \infty$ и функция $f(x)$ становится равной пределу:

$$f(x) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Этот предел называют *суммой ряда Тейлора* и обозначают через

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Итак, в хорошем случае

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Надо понимать разницу между формулой Тейлора и рядом Тейлора. Формула Тейлора дает **приближение** функции полиномом Тейлора с некоторой точностью, которую можно оценивать. Она выражена остаточным членом. Ряд Тейлора — это **точное** представление функции в виде предела последовательности полиномов. В формуле Тейлора

$$f(x) = f(x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \dots$$

последнее многоточие означает остаточный член, который можно записать и оценить, и математики его здесь записывают. В ряде Тейлора многоточие означает, что нужно продолжить суммирование и понимать бесконечную сумму как предел.

Физики на практике пользуются формулой Тейлора, называя ее рядом. Указывать физикам на неточность в употреблении терминологии не стоит по двум причинам: во-первых, это ни к чему хорошему не приводит, а во-вторых, это никак не влияет на качество получаемых при этом научных результатов.

Пример (конечно-разностное приближение производной). Пусть $f : [a,b] \to \mathbb{R}$ имеет достаточное число производных. Кстати, словосочетание «достаточное число производных» употребляют тогда, когда собираются проводить какие-то преобразования и в результате придется дифференцировать может быть раз, а может быть два, и сколько раз надо дифференцировать, до такого порядка и предполагается наличие производной. Пусть все производные непрерывны. Предположим, что значения функции известны только в точках $x_i = a + (i-1)h$, где $h = \dfrac{b-a}{n-1}$, $i = 1, \dots, n$. Такие точки называют *узлами*. Эта ситуация достаточно типична. Предположим, что есть какая-то зависимость, например, электрический сигнал. Он определен в каждый момент времени, а записываем мы его с некоторой частотой и фиксируем его значения в некоторых точках (рис. 3.2). Введем обозначение для значения в i-м узле: $f_i = f(x_i)$. Задача: выразить (приближенно) производную $f'(x_i)$ через значения в каких-то соседних узлах, возможно, с участием данного узла.

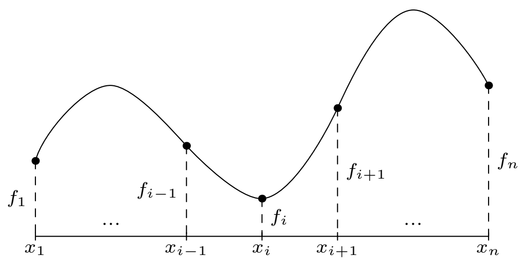

Рис. 3.2.

Будем считать, что узел не граничный. Какие есть для этого возможности? Можно посмотреть на определение производной или на теорему Лагранжа и увидеть, что там производная приближенно равна отношению приращения функции к приращению аргумента, которое нередко называют *конечно-разностным отношением.*

Способ 1:

$$f'(x_i) \approx \frac{f_{i+1} - f_i}{h} \tag{$*$}$$

Какова ошибка этого приближения? Воспользуемся для ее оценки формулой Тейлора в окрестности точки x_i для значения в точке x_{i+1} с остаточным членом в форме Лагранжа:

$$f(x_{i+1}) = f(x_i) + f'(x_i)(x_{i+1} - x_i) + \frac{f''(\xi)}{2!}(x_{i+1} - x_i)^2,$$

где ξ — некоторая точка из (x_i, x_{i+1}). Преобразуем последнюю формулу для получения разностного отношения:

$$\left| f'(x_i) - \frac{f_{i+1} - f_i}{h} \right| = \left| \frac{f''(\xi)}{2!} \right| h. \tag{$\dagger$}$$

В правой части равенства (†) записана ошибка. Если вторая производная ограничена, т. е. существует такое C, что $|f''(x)| \leqslant C$ для $x \in [a, b]$, то ошибка ограничена сверху величиной $\dfrac{C}{2!}h$. Ее можно посчитать и гарантировать, что ошибка на каждом шаге не больше указанной величины. Кроме того, ошибка есть величина $O(h)$ при стремлении h к нулю. Это означает, что формула (†) обеспечивает первый порядок аппроксимации, т. е. ошибка убывает как первая степень h.

Способ 2 (центральная разность):

$$f'(x_i) \approx \frac{f_{i+1} - f_{i-1}}{2h}. \qquad (**)$$

Посчитаем ошибку. По формуле Тейлора при разложении в окрестности точки x_i для точек x_{i-1}, x_{i+1} имеем

$$f(x_{i+1}) = f(x_i) + f'(x_i)h + \frac{f''(x_i)}{2!}h^2 + \frac{f'''(\xi_1)}{3!}h^3,$$

$$f(x_{i-1}) = f(x_i) - f'(x_i)h + \frac{f''(x_i)}{2!}h^2 - \frac{f'''(\xi_2)}{3!}h^3,$$

где ξ_1 — некоторая точка между x_{i+1} и x, а ξ_2 — некоторая точка между x_{i-1} и x. Подставим полученные разложения в формулу $(**)$:

$$\left| f'(x_i) - \frac{f_{i+1} - f_{i-1}}{2h} \right| = \left| \frac{f'''(\xi_1)}{3!}\frac{h^2}{2} + \frac{f'''(\xi_2)}{3!}\frac{h^2}{2} \right| \qquad (‡)$$

В предположении, что третья производная ограничена на всем промежутке константой C, получаем оценку для ошибки:

$$\left| f'(x_i) - \frac{f_{i+1} - f_{i-1}}{2h} \right| \leqslant \frac{C}{3!}h^2 = O(h^2).$$

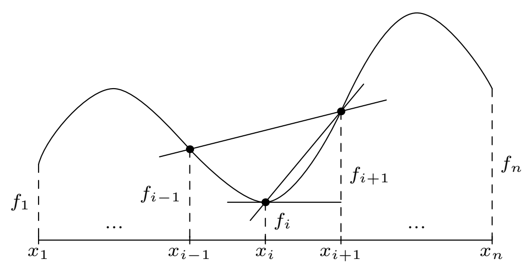

Рис. 3.3.

Такая формула дает второй порядок аппроксимации. Действительно, это хорошо видно из рисунка (рис. 3.3). В случае приближения по способу 1 секущая проходит справа и только она участвует при первом способе приближения, ошибка заметная, а в случае центральной разности секущая проходит через соседние два узла и ее угол наклона ближе к углу наклона касательной. Это подтверждается математически: правая разность имеет первый порядок аппроксимации, центральная — второй.

Аналогично можно получить формулы с более высоким порядком аппроксимации, но для этого, во-первых, придется задействовать большее количество узлов, а во-вторых, потребовать от функции больший порядок гладкости, т. е. наличие производных более высоких порядков.

Можно записать приближение и для второй производной, но для этого потребуется по крайней мере три узла, можно третью, и т. д.

Теорема 12 (формула Тейлора для основных элементарных функций). *Имеют место равенства*

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + r_n(x) \\ &= \sum_{k=0}^{n} \frac{x^k}{k!} + r_k(x), \quad x \in \mathbb{R}, \end{aligned} \tag{1}$$

$$\begin{aligned} \ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{n+1}\frac{x^n}{n} + r_n(x) \\ &= \sum_{k=1}^{n} \frac{(-1)^{k-1}x^k}{k} + r_n(x), \quad -1 < x < 1, \end{aligned} \tag{2}$$

$$\begin{aligned} \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + r_{2n+1} \\ &= \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + r_{2n+1}, \quad x \in \mathbb{R}, \end{aligned} \tag{3}$$

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + r_{2n}(x) \\ &= \sum_{k=0}^{n} \frac{(-1)^k x^{2k}}{(2k)!} + r_{2n}(x), \quad x \in \mathbb{R}, \end{aligned} \tag{4}$$

$$\begin{aligned} (1+x)^\mu &= 1 + \mu x + \frac{\mu(\mu-1)}{2!}x^2 + \dots + \frac{\mu(\mu-1)\dots(\mu-n+1)}{n!}x^n + r_n(x) \\ &= 1 + \sum_{k=1}^{n} \frac{\mu(\mu-1)\cdots(\mu-k+1)}{k!}x^k, \quad -1 < x < 1. \end{aligned} \tag{5}$$

Доказательство. Запишем формулу Тейлора и посчитаем производные:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \dots + \frac{f^{(n)}(0)}{n!}x^n + r_n(x).$$

(1) Для функции $f(x) = e^x$ имеем

$$f'(x) = f''(x) = \dots = f^{(n)}(x) = e^x.$$

Значит, $f^{(k)}(0) = 1$, и формула (1) получена.

(2) Дифференцируем логарифм:

$$(\ln(1+x))' = \frac{1}{1+x} = (1+x)^{-1}, \quad (\ln(1+x))'' = (-1)(1+x)^{-2},$$

$$(\ln(1+x))''' = (-1)(-2)(1+x)^{-3},$$
$$\dots,$$
$$(\ln(1+x))^{(k)} = (-1)(-2)\dots(-k+1)(1+x)^{-k} = (-1)^{k-1}(k-1)!(1+x)^{-k}.$$

В итоге

$$\ln(1+x) = 0 + \frac{1}{1!}x - \frac{x^2}{2!} + \frac{2!x^3}{3!} - \dots + (-1)^{n-1}(n-1)!(1+x)^{-n} + r_n(x)$$
$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{n-1}\frac{x^n}{n} + r_n(x).$$

(3) Продифференцируем синус:

$$(\sin x)' = \cos x,\ (\sin x)'' = -\sin x,\ (\sin x)''' = -\cos x,\ (\sin x)^{\mathrm{iv}} = \sin x.$$

Эти производные можно записать с использованием сдвига:

$$(\sin x)^{(k)} = \sin\left(x + \frac{k\pi}{2}\right).$$

Аналогично можно найти

$$(\cos x)^{(k)} = \cos\left(x + \frac{k\pi}{2}\right).$$

Следовательно, все производные в нуле четных порядков у синуса равны нулю, а нечетного равны 1 или -1 с чередованием знаков. У косинуса аналогично производные нечетных порядков равны нулю, а четных 1 или -1 с чередованием знаков. Поэтому

$$\sin x = 0 + \frac{1}{1!}x + 0 - \frac{1}{3!}x^3 + 0 + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + r_{2n+1}(x).$$

Формула для косинуса выписывается аналогично.

Формула для бинома обосновывается тем же путем, и подробно это записывать не будем.

Теорема доказана.

Пример. Приблизим $\sin x$ полиномом Тейлора и найдем точность приближения. Запишем несколько первых членов в формуле Тейлора для синуса и определимся с формой остаточного члена:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + 0 \cdot x^6 + \frac{\sin^{(7)} \xi}{7!} x^7,$$

где ξ — некоторая точка между 0 и x. Поскольку $|\sin^{(7)} \xi| \leqslant 1$ при всех $\xi \in \mathbb{R}$, ошибка по модулю ограничивается сверху величиной $\dfrac{1}{7!}x^7$. Теперь можно обратиться к возможной постановке вопроса. Можно взять некоторый промежуток и поинтересоваться, какова будет оценка ошибки приближения полиномом Тейлора на данном промежутке, а можно задать величину оценки и интересоваться, на каком промежутке эта величина гарантируется. Ясно, что если x увеличивать, то ошибка станет большой, а при малых x ошибка мала. Например, ошибка для $|x| \leqslant 1$ не больше чем $\dfrac{1}{7!}$.

В качестве поучительного упражнения можно взять какое-либо программное средство для вычисления и графического представления математических объектов и сопоставить на одной координатной плоскости графики синуса и приближающих его полиномов. Например, в Mathematica достаточно набрать

$$\mathrm{Plot}[\{\mathrm{Sin}[x], x - x\hat{\ }3/6,\ x - x\hat{\ }3/6 + x\hat{\ }5/120\}, \{x, 0, 3 * \mathrm{Pi}\}].$$

§ 4. Исследование функции при помощи производной

4.1. Монотонность и точки экстремума.

В § 2 была доказана теорема Ферма, согласно которой для дифференцируемой функции необходимым условием экстремума является обращение в нуль производной в точке экстремума. Однако необходимое условие не позволяет гарантировать наличие экстремума в данной точке, оно лишь может гарантировать, что в каких-то точках экстремума нет, а именно если производная во внутренней точке отлична от нуля, то в такой точке экстремума нет. Для исследования точки на экстремум нужны достаточные условия, и один из вариантов таких условий будет дан в следующей теореме.

Теорема 13 (достаточное условие локального экстремума.) *Пусть* $f:(a,b)\to\mathbb{R}$ *имеет вторую производную* $f''(x)$ *на* (a,b) *и* $x_0\in(a,b)$.

(1) *Если* $f'(x_0)=0$ *и* $f''(x_0)>0$, *то* x_0 *— точка локального минимума.*

(2) *Если* $f'(x_0)=0$ *и* $f''(x_0)<0$, *то* x_0 *— точка локального максимума.*

(3) *Если* $f'(x_0)=0$ *и* $f''(x_0)=0$, *то ничего об экстремуме в точке* x_0 *утверждать невозможно.*

Доказательство. Ограничимся доказательством п. (1), п. (2) доказывается аналогично, в п. (3) достаточно привести соответствующие примеры.

Предположим, что $f'(x_0)=0$ и $f''(x_0)>0$, и запишем формулу Тейлора для функции f в окрестности точки x_0:

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\alpha(x)(x-x_0)^2, \quad (*)$$

где $\alpha(x)\to 0$ при $x\to x_0$. По предположению $f'(x_0)=0$, стало быть, соответствующее слагаемое в $(*)$ обращается в нуль. Преобразуем формулу $(*)$:

$$f(x)-f(x_0)=\left(\frac{f''(x_0)}{2!}+\alpha(x)\right)(x-x_0)^2. \quad (\dagger)$$

По определению предела для любого $\varepsilon>0$ существует проколотая окрестность U точки x_0 такая, что $|\alpha(x)|<\varepsilon$ для $x\in U$. Наша цель — найти такую окрестность точки x_0, в которой сумма $\dfrac{f''(x_0)}{2!}+\alpha(x)$ окажется положительной, а тогда и разность $f(x)-f(x_0)$ будет положительной в такой окрестности и тем самым x_0 — точка минимума. Для достижения этой цели естественно в качестве ε взять половину значения $\dfrac{f''(x_0)}{2!}$. Получаем, что для $\varepsilon=\dfrac{1}{2}\dfrac{f''(x_0)}{2!}$ (>0) существует такая проколотая окрестность $U\subset(a,b)$ точки x_0, в которой $|\alpha(x)|<\varepsilon$, а тогда

$$\frac{f''(x_0)}{2!}+\alpha(x)>\frac{1}{2}\frac{f''(x_0)}{2!}>0.$$

В итоге оказалось, что есть такая проколотая окрестность точки x_0, в которой

$$f(x) - f(x_0) = \underbrace{\left(\frac{f''(x_0)}{2!} + \alpha(x)\right)}_{>0 \text{ в } U} \underbrace{(x - x_0)^2}_{>0 \text{ при } x \neq x_0} > 0.$$

Результат п. (1) доказан.

(3) Рассмотрим функции $f_1(x) = x^3$, $f_2(x) = x^4$, $f_3(x) = -x^4$. Ясно, что у всех этих функций первые две производные в нуле равны нулю. Однако f_1 не имеет локального экстремума, f_2 имеет минимум, а f_3 — максимум в точке 0.

Теорема доказана.

На самом деле есть более общая теорема, в которой задействованы производные более высоких порядков, и которая утверждает, что если производная нечетного порядка оказалась равной нулю, а следующего четного порядка отлична от нуля, то в такой точке есть экстремум, вид которого зависит от знака производной четного порядка в этой точке. Если же производная нечетного порядка в точке отлична от нуля, то экстремума в такой точке нет.

Теорема 14 (критерий монотонности). *Если $f : [a, b] \to \mathbb{R}$ непрерывна на $[a, b]$ и дифференцируема на (a, b).*

(1) *$f(x)$ неубывающая тогда и только тогда, когда $f'(x) \geqslant 0$ на (a, b),*

(2) *$f(x)$ (строго) возрастающая тогда и только тогда, когда $f'(x) \geqslant 0$ на (a, b) и не существует промежутка, на котором $f'(x) = 0$,*

(3) *$f(x)$ постоянна тогда и только тогда, когда $f'(x) = 0$ на (a, b).*

Доказательство. Шаг 1 (f неубывающая $\Rightarrow f'(x) \geqslant 0$). Поскольку $f(x)$ неубывающая, для любых $x_1, x_2 \in [a, b]$, $x_1 < x_2$, будет $f(x_1) \leqslant f(x_2)$. Значит,

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

Переходя к пределу в неравенстве при $x_2 \to x_1$, получаем

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

Шаг 2 (f неубывающая $\Leftarrow f'(x) \geqslant 0$). Пусть $x_1, x_2 \in [a, b]$, $x_1 < x_2$. Покажем, что $f(x_1) \leqslant f(x_2)$. По теореме Лагранжа найдется такая точка $\xi \in (x_1, x_2)$, что

$$f(x_2) - f(x_1) = \underbrace{f'(\xi)}_{\geqslant 0}\underbrace{(x_2 - x_1)}_{>0} \geqslant 0.$$

Шаг 3 (f возрастающая $\Rightarrow$ не существует промежутка …). Будем рассуждать методом от противного. Предположим, что существует такой промежуток $[x_1, x_2]$, на котором $f'(x) = 0$. По теореме Лагранжа найдется такая точка $\xi \in (x_1, x_2)$, что

$$f(x_2) - f(x_1) = \underbrace{f'(\xi)}_{=0}(x_2 - x_1) = 0,$$

что противоречит строгой монотонности $f(x)$.

Шаг 4 (f возрастающая $\Leftarrow f'(x) \geqslant 0$ и не существует промежутка …). Неубывание функции $f(x)$ следует из шага 2. Докажем, что $f(x)$ строго монотонна, методом от противного. Предположим, что f не строго возрастающая, и найдем промежуток, на котором $f'(x) = 0$. Ввиду отсутствия строго возрастания существуют такие точки $x_1, x_2 \in [a, b]$, $x_1 < x_2$, для которых $f(x_1) = f(x_2)$. Из этого равенства и неубывания f вытекает, что $f(x)$ постоянна на всем промежутке $[x_1, x_2]$ и тем самым производная на нем нулевая.

Теорема доказана.

4.2. Выпуклость функции и точки перегиба.

Определение выпуклой функции. Функцию $f : \langle a, b\rangle \to \mathbb{R}$, заданную **на промежутке**, называют *выпуклой*, если для любых $x_1, x_2 \in \langle a, b\rangle$ и любых чисел $\alpha_1, \alpha_2 \geqslant 0$ таких, что $\alpha_1 + \alpha_2 = 1$, выполнено неравенство

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2). \qquad (*)$$

Выпуклую функцию $f : \langle a, b\rangle \to \mathbb{R}$ называют *строго выпуклой*, если для $x_1 \neq x_2$ и $\alpha_1, \alpha_2 > 0$ неравенство $(*)$ строгое.

Выражение $\alpha_1 x_1 + \alpha_2 x_2$ с указанными в определении выпуклости требованиями к α_1, α_2 называют *выпуклой комбинацией точек x_1, x_2 с коэффициентами α_1, α_2*.

Аналогично определяются *вогнутая* и *строго вогнутая функции* — нужно только поменять знак неравенства на противоположный. Иногда о вогнутых функциях говорят как о *выпуклых вверх*. Можно также сказать, что вогнутость f равносильна выпуклости $-f$.

Геометрический смысл выпуклости функции. Изобразим координатную плоскость, на оси абсцисс возьмем какой-либо промежуток $\langle a, b\rangle$ и выберем в нем пару точек x_1, x_2. Точка $\alpha_1 x_1 + \alpha_2 x_2$ с указанными выше требованиями к α_1, α_2 расположена между x_1 и x_2 и делит отрезок $[x_1, x_2]$ в отношении $\alpha_1 : \alpha_2$, считая от x_1. Это можно пронаблюдать, беря конкретные значения α_1 и α_2. Если $\alpha_1 = 1$, а тогда $\alpha_2 = 0$, мы находимся в левом конце x_1. При $\alpha_2 = 1$, $\alpha_1 = 0$ попадаем в правый конец x_2. Если $\alpha_1 = \alpha_2 = \dfrac{1}{2}$, то точка расположена в середине промежутка.

В левой части неравенства $(*)$ стоит значение функции в некоторой точке промежутка $[x_1, x_2]$. В правой части расположена выпуклая комбинация значений $f(x_1)$, $f(x_2)$, т. е. точка из промежутка с концами $f(x_1)$, $f(x_2)$ на оси ординат, расположенная в той же пропорции, что и значения аргумента x_1, x_2 на отрезке $[x_1, x_2]$. График функции, сопоставляющей точке $x = \alpha_1 x_1 + \alpha_2 x_2$ значение $\alpha_1 f(x_1) + \alpha_2 f(x_2)$, представляет собой отрезок с концами $(x_1, f(x_1))$, $(x_2, f(x_2))$, и неравенство $(*)$ означает следующее. Взяв произвольные точки x_1, x_2 промежутка $\langle a, b\rangle$ и рассмотрев часть графика функции на промежутке с концами в x_1, x_2, можно увидеть, что она расположена не выше отрезка, соединяющего концы графика (рис. 4.1). Об этом также говорят, что дуга графика выпуклой функции на любом отрезке из данного промежутка расположена ниже (не выше) хорды, соединяющей концы этой дуги.

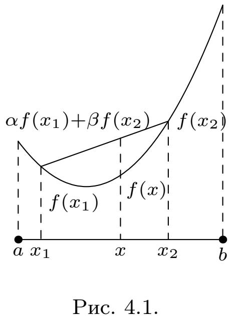

Рис. 4.1.

Определение выпуклого множества. Множество M в векторном пространстве называют *выпуклым*, если

$$\forall P_1, P_2 \in M \ \forall \alpha_1, \alpha_2 \geqslant 0, \ \alpha_1 + \alpha_2 = 1, \quad \alpha_1 P_1 + \alpha_2 P_2 \in M,$$

т. е. вместе с любыми двумя точками множество содержит весь отрезок, соединяющий эти точки.

Определение надграфика функции. Для функции $f : \langle a, b \rangle \to \mathbb{R}$ множество

$$\mathrm{epi}(f) = \{(x, y) \mid x \in \langle a, b \rangle, \ y \geqslant f(x)\}$$

называют *надграфиком* функции f.

Замечание (геометрический смысл выпуклости). Функция выпукла тогда и только тогда, когда ее надграфик — выпуклое множество. Это утверждение вытекает непосредственно из сопоставления определений выпуклого множества и выпуклой функции.

Теорема 15 (критерий выпуклости функции). *Непрерывная на $\langle a, b \rangle$ функция $f : \langle a, b \rangle \to \mathbb{R}$, имеющая вторую производную $f''(x)$ на (a, b), выпукла на $\langle a, b \rangle$ тогда и только тогда, когда $f''(x) \geqslant 0$ на (a, b).*

По ходу доказательства мы получим еще пару критериев выпуклости, однако, по-видимому, критерий в терминах второй производной наиболее наглядный и чаще всего применяемый на практике.

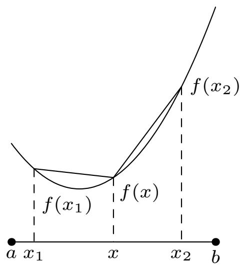

Рис. 4.2.

Доказательство. Шаг 1 (другое определение выпуклости). Докажем следующее утверждение. *Функция $f : \langle a, b \rangle \to \mathbb{R}$, заданная на промежутке $\langle a, b \rangle$, выпукла тогда и только тогда, когда для любых $x_1, x, x_2 \in \langle a, b \rangle$ таких, что $x_1 < x < x_2$, выполнено неравенство*

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

Суть сформулированного утверждения видна из рисунка (рис. 4.2). Расположив на оси абсцисс точки x_1, x, x_2 в указанном в утверждении порядке и проведя секущие через точки графика, соответствующие точкам x, x_2 и x_1, x, находим, что для выпуклой функции тангенс угла наклона правой секущей не меньше чем тангенс угла наклона левой. Запись этого обстоятельства в виде формулы и есть содержание утверждения.

Заметим, что любой точке $x \in (x_1, x_2)$ взаимно однозначно соответствуют коэффициенты $\alpha_1, \alpha_2 \in (0, 1)$ такие, что $\alpha_1 + \alpha_2 = 1$ и $x = \alpha_1 x_1 + \alpha_2 x_2$. Действительно, по заданным α_1, α_2 элемент x получается непосредственно. Обратно, пусть $x \in (x_1, x_2)$. Положим

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

Нетрудно проверить, что выпуклая комбинация с такими α_1, α_2 дает x.

В характеризующем выпуклость неравенстве

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2),$$

$$\alpha_1, \alpha_2 \geqslant 0, \quad \alpha_1 + \alpha_2 = 1, \quad x = \alpha_1 x_1 + \alpha_2 x_2,$$

добавив в левой части множитель $\alpha_1 + \alpha_2$, равный 1, сгруппируем выражения, собирая члены с α_1 в левой части, а с α_2 — в правой:

$$\alpha_1(f(x) - f(x_1)) \leqslant \alpha_2(f(x_2) - f(x)),$$

или, иначе,

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

откуда получаем требуемое неравенство.

Все проделанные переходы в доказательстве были равносильными, стало быть, доказан весь критерий, т. е. необходимость и достаточность.

Шаг 2 (критерий в терминах первой производной). *Функция f выпукла тогда и только тогда, когда ее производная $f'(x)$ неубывающая.*

Необходимость. Пусть f выпукла. Тогда по шагу 1 для любых $x_1, x_2, x \in \langle a, b \rangle$ таких, что $x_1 < x < x_2$, имеем

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

Перейдем в этом неравенстве к пределу при $x \to x_1$. Слева получится производная, а справа — значение стоящей в правой части функции в точке x_1, т. е.

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

Точно так же можно перейти к пределу при $x \to x_2$ и получить такое неравенство:

$$\frac{f(x_2) - f(x_1)}{x2 - x_1} \leqslant f'(x_2).$$

Совмещая два последние неравенства, получаем, что для любых x_1, x_2 таких, что $x_1 < x_2$, справедливо неравенство $f'(x_1) \leqslant f'(x_2)$. Таким образом, первая производная выпуклой функции неубывающая.

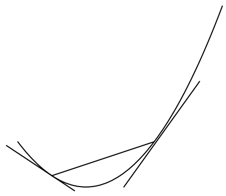

Рис. 4.3.

Геометрическое содержание этого факта также достаточно прозрачно. Изобразим выпуклую функцию. Процедура доказательства означает, что если точки расположены так, как дано в условии, то сначала идет касательная в точке x_1, затем секущая, угол наклона которой не меньше чем у касательной, а после этого — касательная в точке x_2 также с не меньшим углом наклона (рис. 4.3).

Достаточность. Пусть производная f' функции f неубывающая. Докажем, что для любых $x_1 < x < x_2$ выполнено неравенство

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

По теореме Лагранжа существуют точка $\xi_1 \in (x_1, x)$ такая, что

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1),$$

и точка $\xi_2 \in (x, x_2)$ такая, что

$$\frac{f(x_2) - f(x)}{x_2 - x} = f'(\xi_2).$$

Производная неубывающая, поэтому $f'(\xi_1) \leqslant f'(\xi_2)$, откуда и следует требуемое неравенство.

Шаг 3 (критерий в терминах второй производной). По теореме 14 о критерии монотонности $f'(x)$ неубывающая тогда и только тогда, когда $f''(x) \geqslant 0$ на (a, b).

Теорема доказана.

Полезность свойства выпуклости по крайней мере в том, что у выпуклой на отрезке функции всегда есть глобальный минимум. Кроме того, из соображений выпуклости можно получить много весьма полезных неравенств, которые другим путем не получаются. На простом примере покажем, как получаются хорошо известные неравенства, хотя бы в целях демонстрации идеи.

Пример (как из выпуклости получать неравенства). Возьмем функцию $f(x) = -\ln x$. Нетрудно проверить, что она выпуклая. Действительно, $f''(x) = \dfrac{1}{x^2} > 0$. Запишем для нее определение выпуклости: для любых $x_1, x_2 > 0$, $\alpha_1, \alpha_2 \in [0, 1]$, $\alpha_1 + \alpha_2 = 1$, имеем

$$-\ln(\alpha_1 x_1 + \alpha_2 x_2) \leqslant -\alpha_1 \ln x_1 - \alpha_2 \ln x_2. \qquad (*)$$

Преобразуем неравенство $(*)$:

$$\ln(x_1^{\alpha_1} x_2^{\alpha_2}) \leqslant \ln(\alpha_1 x_1 + \alpha_2 x_2).$$

Натуральный логарифм — функция возрастающая, следовательно,

$$x_1^{\alpha_1} x_2^{\alpha_2} \leqslant \alpha_1 x_1 + \alpha_2 x_2.$$

В частности, при $\alpha_1 = \alpha_2 = \dfrac{1}{2}$ получаем

$$\sqrt{x_1 x_2} \leqslant \frac{x_1 + x_2}{2}.$$

Получили неравенство, связывающее среднее арифметическое (в правой части) и среднее геометрическое (в левой).

Можно доказать, что для выпуклой функции верно более общее, чем в ее определении, неравенство, а именно для любых $x_1, x_2, \ldots, x_n$ и любых $\alpha_1, \alpha_2, \ldots, \alpha_n \in [0, 1]$ таких, что $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$, верно неравенство Йенсена

$$f(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2) + \cdots + \alpha_n f(x_n).$$

Оно аналогично неравенству выпуклости, но в отличие от него содержит n элементов вместо двух. Доказательство его можно провести по индукции.

Применив неравенство Йенсена к логарифму, приходим к неравенству
$$\sqrt[n]{x_1 \cdot \ldots \cdot x_n} \leqslant \frac{x_1 + \cdots + x_n}{n}.$$
Заметим, что неравенство для двух элементов легко получается и из других соображений, а последнее неравенство из других соображений не получается.

Используем выпуклость еще одной функции. Рассмотрим функцию $f(x) = \dfrac{1}{x}$, $x > 0$. Легко проверить, что она выпуклая. По определению
$$\frac{1}{\alpha_1 x_1 + \alpha_2 x_2} \leqslant \frac{\alpha_1}{x_1} + \frac{\alpha_2}{x_2}.$$
В частном случае $\alpha_1 = \alpha_2 = \dfrac{1}{2}$ получаем
$$\frac{1}{\frac{x_1+x_2}{2}} \leqslant \frac{1}{2}\left(\frac{1}{x_1} + \frac{1}{x_2}\right),$$
или, что равносильно,
$$\frac{2}{\frac{1}{x_1} + \frac{1}{x_2}} \leqslant \frac{x_1 + x_2}{2}.$$
Последнее неравенство означает, что среднее гармоническое (в левой части) не больше среднего арифметического.

Для полноты картины не хватает сопоставления среднего гармонического и среднего геометрического. Соотношение между этими средними получается, если в первом неравенстве заменить x_1 на $\dfrac{1}{x_1}$ и x_2 на $\dfrac{1}{x_2}$. В результате преобразований окажется, что среднее гармоническое будет самым маленьким, посередине будет среднее геометрическое.

4.3. Асимптоты.

Определение асимптот. Пусть даны функция $f : X \to \mathbb{R}$ и x_0 — предельная точка множества X. Говорят, что f *имеет вертикальную асимптоту в точке* x_0, если бесконечен хотя бы один из односторонних пределов, т. е. $\lim\limits_{x \to x_0 - 0} f(x) = \pm\infty$ либо $\lim\limits_{x \to x_0 + 0} f(x) = \pm\infty$. Может получиться и так, что оба эти пределы бесконечны.

Пусть $+\infty$ — предельная точка множества $\mathrm{dom}(f)$. Говорят, что f *имеет наклонную асимптоту* $y = ax + b$ при $x \to +\infty$, если

$$\lim_{x\to+\infty}(f(x) - (ax + b)) = 0.$$

Аналогичное определение можно дать и для $x \to -\infty$. Мы будем подробно рассматривать только случай $x \to +\infty$, случай $x \to -\infty$ полностью аналогичен.

Последнее равенство напоминает определение главной части функции в предельной точке. А именно, его можно записать в виде

$$f(x) = ax + b + o(1).$$

Вопрос: можно ли утверждать, что $ax + b$ является главной частью $f(x)$? Ответ оставим читателям.

Как искать асимптоты? Вопрос о поиске вертикальных асимптот есть не что иное как вопрос о существовании бесконечного предела функции в точке. Разберемся с наклонными асимптотами.

Замечание (о нахождении асимптот). Функция $f(x)$ имеет асимптоту $y = ax + b$ в том и только в том случае, если существуют соответственно равные коэффициентам пределы

$$a = \lim_{x\to+\infty}\frac{f(x)}{x}, \quad b = \lim_{x\to+\infty}(f(x) - ax). \tag{$*$}$$

Тем самым для нахождения асимптот функции f сначала изучаем наличие первого из пределов в $(*)$, и если он есть, ищем второй. Если и он есть, то имеется асимптота с найденными коэффициентами. Действительно, пусть f имеет асимптоту $y = ax + b$, т. е.

$$\lim_{x\to+\infty}(f(x) - (ax + b)) = 0.$$

Тогда $f(x) - ax \to b$, и получаем существование правого предела в $(*)$, равного b. Далее,

$$\frac{f(x) - (ax + b)}{x} = \frac{f(x)}{x} - a - \frac{b}{x},$$

откуда

$$\frac{f(x)}{x} = \frac{f(x) - (ax + b)}{x} + a + \frac{b}{x}. \tag{$\dagger$}$$

При $x \to +\infty$ отношения в правой части (†) стремятся к нулю, откуда и получаем, что $a = \dfrac{f(x)}{x}$. В обратную сторону доказывается аналогично.

При нахождении асимптот появляется необходимость вычислять предел отношения $\dfrac{f(x)}{g(x)}$, где пределы числителя и знаменателя бесконечные. О такой ситуации говорят как о *неопределенности вида* $\dfrac{\infty}{\infty}$. Наиболее содержательны примеры сопоставления показательной, степенной и логарифмической функций. Эти функции имеют разные скорости возрастания на бесконечности, и соответствующие пределы отношения бесконечны. В случае изучения асимптот конечность предела отношения говорит о том, что данная функция и линейная имеют сопоставимый рост на бесконечности. Нахождение предела в ситуации указанной выше неопределенности облегчает следующая

Теорема 16 (правило Бернулли — Лопиталя). *Пусть даны функции* $f, g : \langle a, b\rangle \to \mathbb{R}$ *и* x_0 *— предельная точка промежутка* $\langle a, b\rangle$. *Если*

(1) f, g *имеют производные* f', g', *причем* $g'(x) \neq 0$ *для* $x \in \langle a, b\rangle$ *за исключением, возможно, точки* x_0.

(2a) $f(x) \to 0$, $g(x) \to 0$ *при* $x \to x_0$

или

(2b) $g(x) \to +\infty$, *а* $f(x)$ *любая.*

(3) *существует* $\lim\limits_{x\to x_0} \dfrac{f'(x)}{g'(x)}$.

Тогда существует предел

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

Иначе говоря, если есть потребность найти предел, в котором возникает неопределенность вида «нуль на нуль» или «бесконечность на бесконечность», то можно найти отдельно производные числителя и знаменателя и озаботиться нахождением предела отношения производных. Если он есть, то существует и предел отношения функций, равный пределу отношения производных.

Если рассматривается неопределенность вида $\dfrac{0}{0}$, то обычно проще воспользоваться асимптотическими разложениями, а если неопределенность вида $\dfrac{\infty}{\infty}$, то справиться с ней при помощи асимптотики проблематично, и тогда можно применить правило Лопиталя.

Мы не будем доказывать теорему 16. Случай неопределенности вида $\frac{0}{0}$ легко обосновывается с помощью теоремы Коши и особого интереса не представляет. Доказательство в случае неопределенности вида $\frac{\infty}{\infty}$ интересное, но технически сложное и требует определенного времени.

§ 5. Первообразная

Определение первообразной. Функцию $F : \langle a, b\rangle \to \mathbb{R}$ называют *первообра́зной* или *неопределенным интегралом* функции $f : \langle a, b\rangle \to \mathbb{R}$, если

$$F'(x) = f(x) \quad \text{для любого } x \in \langle a, b\rangle.$$

Функцию $F : \langle a, b\rangle \to \mathbb{R}$ называют *обобщенной первообразной* функции $f : \langle a, b\rangle \to \mathbb{R}$, если

(1) F непрерывна на $\langle a, b\rangle$;

(2) $F'(x) = f(x)$ на $\langle a, b\rangle$ за исключением, возможно, конечного числа точек.

Для (обобщенной) первообразной функции $f(x)$ используют обозначение

$$\int f(x)\, dx.$$

Замечание (об интеграле и первообразной). Первообразная — это операция, обратная к операции дифференцирования, а интеграл, который рассматривается в следующей главе, — это площадь подграфика. Интеграл по отрезку $[a, b]$ от функции f обозначают символом $\int\limits_a^b f(x)\, dx$. Тем самым в результате нахождения первообразной находится функция, а результат нахождения интеграла — число. Их обозначения похожи друг на друга, но имеют отличия. Это является отражением тесной связи интеграла и первообразной, а именно если есть некоторая первообразная $F(x)$ функции f на $[a, b]$, то при некоторых условиях на функцию f есть равенство

$$\int\limits_a^b f(x)\, dx = F(b) - F(a),$$

называемое *формулой Ньютона — Лейбница*, связывающее интеграл с первообразной. Но это будет подробно рассмотрено в следующей главе.

Записываемый в обозначении символ dx указывает на то, что взятие первообразной есть операция, обратная к операции нахождения дифференциала. А именно, если $F(x)$ — первообразная функции $f(x)$ и мы хотим проверить, что это действительно первообразная, то надо ее продифференцировать, т. е. взять от нее дифференциал, и тогда мы должны получить дифференциал исходной функции, тем самым должно быть выполнено равенство $dF(x) = F'(x)\,dx = f(x)\,dx$, в котором появляется dx.

Пример. Покажем, зачем бывает нужна обобщенная первообразная. Если

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

то для нее обобщенной первообразной будет функция $F(x) = |x|$. Действительно, функция $|x|$ дифференцируема во всех точках, кроме нуля, и ее производная там, где существует, совпадает с функцией $\operatorname{sgn} x$. В нуле производной у модуля нет, но для обобщенной первообразной это несущественно. Важно, что модуль — функция непрерывная и ее производная всюду за исключением конечного множества точек (в данном случае одноточечного) совпадает в исходной функцией. Конечно, можно было бы по разные стороны взять функции, представляющие собой сдвиги модуля вверх или вниз на разные величины. Совпадение производных в таком случае есть, а непрерывности нет.

Если добавить к функции константу, то ее производная не изменится.

Теорема 17 (о множестве первообразных). Пусть $f : \langle a, b\rangle \to \mathbb{R}$ — некоторая функция.

(1) *Если $F(x)$ — (обобщенная) первообразная $f(x)$, то $F(x) + C$, где C — произвольная константа, тоже (обобщенная) первообразная $f(x)$.*

(2) *Если $F_1(x)$ и $F_2(x)$ — две (обобщенные) первообразные $f(x)$, то*

$$F_2(x) = F_1(x) + C$$

с некоторой константой $C \in \mathbb{R}$.

Доказательство. Первое утверждение очевидно следует из линейности операции дифференцирования и того, что производная постоянной функции равна нулю.

Пусть $x_1 < \dots < x_n$ — все точки промежутка $\langle a, b\rangle$, в которых не выполняется условие $F_j'(x) = f(x)$, $j = 1, 2$, хотя бы для одной из обобщенных первообразных $F_1(x)$, $F_2(x)$. Обозначим $x_0 = a$, $x_{n+1} = b$. На каждом из промежутков (x_k, x_{k+1}), $k = 0, \dots, n$, верно $F_2'(x) = F_1'(x) = f(x)$, следовательно, по теореме о монотонности $F_2 - F_1 = c_k$, где константы c_k свои для каждого промежутка. В силу непрерывности $F_1(x)$ и $F_2(x)$ все константы c_k равны между собой, т. е. $F_2(x) = F_1(x) + C$ с некоторой единой константой C.

Теорема доказана.

Теорема 17 указывает причину, по которой первообразную рассматривают на промежутке. Если рассматривать функция при нахождении ее первообразной например на двух промежутках, то теорема 17 была бы неверна. Имея две первообразных и пытаясь выразить одну через другую, мы на одном из промежутков имели бы одну константу, а на втором — другую, и поскольку эти промежутки не соприкасаются и никак друг с другом не связаны, появился бы разнобой, т. е. первообразные отличались бы на свою константу на каждом из промежутков, что нежелательно (как будет видно из дальнейшего).

Замечание об обозначении первообразной. Знаком $\int$ обычно обозначают множество всех первообразных или какую-либо из этих первообразных. Поэтому если найдена какая-то первообразная F функции f, то пишут

$$\int f(x)\,dx = F(x) + C.$$

Здесь отражено следующее содержание: найдя одну из первообразных, мы находим их все.

Возникает естественный вопрос: как искать первообразную? Для этого используется некоторый базис в виде таблицы первообразных, которая получается из таблицы производных, если в ней считать заданной производную и получать функцию, от которой эта производная была взята. Кроме этого есть три утверждения, которые помогают искать первообразные функций, получаемых из тех, первообразные которых указаны в таблице. Это линейность, интегрирование по частям и замена переменной.

Приведем таблицу простейших неопределенных интегралов:

$$\int x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1} + C, \quad \alpha \neq -1,\ x > 0;$$

$$\int \frac{dx}{x} = \ln|x| + C, \quad x \neq 0;$$

$$\int a^x\,dx = \frac{a^x}{\ln a} + C, \quad a > 0,\ a \neq 1,\ x \in \mathbb{R}; \quad \int e^x\,dx = e^x + C, \quad x \in \mathbb{R};$$

$$\int \sin x\,dx = -\cos x + C, \quad \int \cos x\,dx = \sin x + C, \quad x \in \mathbb{R};$$

$$\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C, \quad x \in (k\pi, (k+1)\pi),\ k \in \mathbb{Z};$$

$$\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C, \quad x \in (-\pi/2 + k\pi, \pi/2 + k\pi),\ k \in \mathbb{Z};$$

$$\int \frac{dx}{1+x^2} = \begin{cases} \operatorname{arctg} x + C, \\ -\operatorname{arcctg} x + C, \end{cases} \quad x \in \mathbb{R};$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \arcsin x + C, \\ -\arccos x + C, \end{cases} \quad x \in [-1, 1];$$

$$\int \frac{dx}{1-x^2} = \frac{1}{2}\ln\left|\frac{1+x}{1-x}\right| + C, \quad |x| \neq 1;$$

$$\int \frac{dx}{\sqrt{x^2+1}} = \ln(x + \sqrt{x^2+1}) + C,$$

$$\int \frac{dx}{\sqrt{x^2-1}} = \ln(x + \sqrt{x^2-1}) + C,\ |x| > 1.$$

Теорема 18 (о линейности первообразной). *Если $f, g : \langle a, b\rangle \to \mathbb{R}$ имеют (обобщенные) первообразные, то любая их линейная комбинация $\alpha f(x) + \beta g(x)$, $\alpha, \beta \in \mathbb{R}$, тоже имеет (обобщенную) первообразную, причем*

$$\int (\alpha f(x) + \beta g(x))\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx + C.$$

Доказательство. Очевидно следует из линейности дифференцирования.

Теорема 19 (формула интегрирования по частям для первообразной). *Если $f, g : \langle a, b\rangle \to \mathbb{R}$ дифференцируемы и $f'g$ имеет первообразную, то fg' тоже имеет первообразную и*

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx.$$

Доказательство. По теореме о производной произведения

$$(f(x)g(x))' = f'(x)g(x) + f(x)g(x)',$$

или

$$f(x)g'(x) = (f(x)g(x))' - f'(x)g(x).$$

Функция справа имеет имеет первообразную по условию. Значит, первообразная есть и у функции в левой части, при этом

$$\int f(x)g'(x)\,dx = \int \Big((fg)'(x) - f'(x)g(x)\Big)\,dx$$
$$= f(x)g(x) - \int f'(x)g(x)\,dx.$$

Теорема доказана.

На практике пишут так:

$$\int f(x)g'(x)\,dx = \int f(x)\,dg(x) = f(x)g(x) - \int g(x)\,df(x)$$
$$= f(x)g(x) - \int g(x)f'(x)\,dx.$$

Пример. Найдем первообразную $\int \ln x\,dx$. Здесь не видно произведения функций под интегралом. Однако одна из функций это $\ln x$, а в качестве другой можно считать единицу, на которую первая функция умножается. Иначе можно сказать, что вторая из функций — это $g(x) = x$, и она сразу внесена под знак дифференциала. Согласно формуле интегрирования по частям имеем

$$\int \ln x\,dx = (\ln x)\cdot x - \int x\,d\ln x = x\ln x - \int x\cdot\frac{1}{x}\,dx = x\ln x - x + C.$$

Теорема 20 (о замене и подстановке для первообразной). *Пусть даны функции $f : \langle c, d\rangle \to \mathbb{R}$, $\varphi : \langle a, b\rangle \to \langle c, d\rangle$, так что определена композиция $f(\varphi(x))$.*

(1) Замена.

Если

(*a*) φ *дифференцируема,*

(*b*) f *имеет первообразную* F,

то $f(\varphi(x))\varphi'(x)$ *тоже имеет первообразную и*

$$\int f(\varphi(x))\varphi'(x)\,dx = F(\varphi(x)).$$

(2) Подстановка.

Если

(*a*) φ *дифференцируема, строго монотонна и имеет дифференцируемую обратную* φ^{-1},

(*b*) $f(\varphi(x))\varphi'(x)$ *имеет первообразную* $G(x)$,

то $f(y)$ *тоже имеет первообразную и*

$$\int f(y)\,dy = G(\varphi^{-1}(y)).$$

Доказательство. (1) (Замена). Проверим, что производная $F(\varphi(x))$ равна $f(\varphi(x))\varphi'(x)$:

$$\frac{d}{dx}F(\varphi(x)) = F'(\varphi(x))\varphi'(x) = f(\varphi(x))\varphi'(x).$$

(2) (Подстановка). Проверим, что производная $G(\varphi^{-1}(y))$ равна $f(y)$:

$$\frac{d}{dy}G(\varphi^{-1}(y)) = G'(\varphi^{-1}(y))(\varphi^{-1}(y))'$$

(по теореме о производной обратной функции)

$$= \underbrace{f(\varphi(\varphi^{-1}(y)))}_{=f(y)}\,\underbrace{\varphi'(\varphi^{-1}(y))\,(\varphi^{-1}(y))'}_{=(\varphi\circ\varphi^{-1})'=1} = f(y).$$

Теорема доказана.

Надо обратить внимание на то, что во втором случае, т. е. при выполнении подстановки, функция должна быть обратимой.

Пример на замену. Замена осуществляется в том случае, если в результате преобразований подынтегральная функция приобрела вид

$$\int f(\varphi(x))\varphi'(x)\,dx.$$

В этот момент подынтегральное выражение надо представить в виде

$$\int f(\varphi(x))\,d\varphi(x),$$

после чего заменить выражение $\varphi(x)$ одной буквой, например $y = \varphi(x)$, и перейти к поиску первообразной $\int f(y)\,dy$. Найдя ее, надо вернуться к прежней переменной x.

Найдем первообразную $\displaystyle\int \frac{\ln x}{x}\,dx$. Смотрим, смотрим на функцию и понимаем, что если записать ее как произведение $\ln x \cdot \dfrac{1}{x}$, то правый сомножитель есть не что иное как производная от логарифма. Следовательно,

$$\int \frac{\ln x}{x}\,dx = \int \ln x \cdot \frac{1}{x}\,dx = \int (\ln x) \cdot (\ln x)'\,dx = \int \ln x\,d\ln x.$$

Полагая $\ln x = y$, перейдем к поиску первообразной:

$$\int y\,dy = \frac{1}{2}y^2 + C.$$

Возвращаясь к переменной x, получаем

$$\int \frac{\ln x}{x}\,dx = \frac{1}{2}(\ln x)^2 + C.$$

В данном ниже примере целесообразно применить подстановку с участием гиперболических функций.

Определение и свойства гиперболических функций. Функции *косинус гиперболический* $\operatorname{ch} x$ и *синус гиперболический* $\operatorname{sh} x$ определяются равенствами

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

Их графики строятся путем сложения и соответственно вычитания графиков двух экспонент. Легко проверяется, что $\operatorname{ch} x$ функция четная, а $\operatorname{sh} x$ нечетная.

Эти функции называют гиперболическими по той причине, что для них выполняется легко проверяемое тождество $\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1$, похожее на основное тригонометрическое тождество. Если на координатной плоскости изобразить линию, состоящую из точек, у которых абсцисса равна $\operatorname{ch} x$, а ордината $\operatorname{sh} x$, то получится гипербола.

Легко найти производные этих функций: $\operatorname{ch}' x = \operatorname{sh} x$, $\operatorname{sh}' x = \operatorname{ch} x$.

Пример на подстановку. Найдем первообразную $\displaystyle\int \frac{dy}{\sqrt{1+y^2}}$. Используем записанное выше тождество для гиперболических функции в виде $\operatorname{ch}^2 x = 1 + \operatorname{sh}^2 x$. Оно подсказывает, что если на место y в нашем интеграле подставить $\operatorname{sh} x$, т. е. положить $y = \operatorname{sh} x$ и учесть, что $dy = \operatorname{sh}' x\,dx = \operatorname{ch} x\,dx$, то получится значительно более простое выражение:

$$\int \frac{dy}{\sqrt{1+y^2}} == \int \frac{\operatorname{ch} x\,dx}{\sqrt{\operatorname{ch}^2 x}} = \int dx = x + C.$$

В найденную первообразную на место x надо подставить выражение для функции, обратной к функции $\operatorname{sh} x$. Найдем его (существование обратной гарантируется строгим возрастанием исходной функции). Для этого из равенства

$$y = \operatorname{sh} x = \frac{e^x - e^{-x}}{2} \tag{$*$}$$

выразим x через y. Записывая равенство $(*)$ в виде квадратного относительно e^x уравнения

$$e^{2x} - 2ye^x - 1 = 0,$$

находим, что

$$e^x = y + \sqrt{y^2+1}.$$

В итоге

$$x = \operatorname{sh}^{-1} y = \ln(y + \sqrt{y^2+1}).$$

Вспомним, что при взятии производных мы всегда действуем по правилам нахождения производных, и если функция элементарна, то получаемый результат окажется в этом же классе функции. С первообразными дело обстоит иначе. Оказывается, что не всегда первообразная элементарной функции выражается через элементарные функции. Примерами таких функций служат функции

$$\frac{\sin x}{x}, \quad \frac{e^x}{x}, \quad e^{-x^2}.$$

Как узнать, выражается ли первообразная той или иной функции через элементарные? Ответ на этот вопрос довольно сложен и здесь рассматриваться не будет. С практической точки зрения можно поступить так: взять какой-либо фундаментальный справочник по первообразным или обратиться к какому-либо программному средству, приспособленному для нахождения первообразных, и если в нем нет первообразной интересующей вас функции, то, скорее всего, она через элементарные не выражается.

Однако не все так плохо: есть большой класс функций, для которых первообразная выражается через элементарные — это класс рациональных функций.

Определение рациональной функции. *Рациональной* называют функцию вида $\dfrac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ — многочлены. Такие функции определены всюду на $\mathbb{R}$ за исключением тех точек, в которых знаменатель обращается в нуль.

Теорема 21 (о первообразной рациональной функции). *Любая рациональная функция на каждом из промежутков свой области определения имеет первообразную, причем эта первообразная выражается через рациональные функции, а также функции* $\ln x$ *и* $\operatorname{arctg} x$.

Доказательство. Шаг 1 (разложение на простейшие дроби). Пусть $R(x) = \dfrac{P(x)}{Q(x)}$ — рациональная функция. Из курса алгебры известно, что $R(x)$ представляется в виде

$$R(x) = p(x) + \sum_{j=1}^{l}\sum_{k=1}^{k_j} \frac{a_{jk}}{(x-x_j)^k} + \sum_{j=1}^{n}\sum_{k=1}^{m_j} \frac{b_{jk}x + c_{jk}}{(x^2+p_jx+q_j)^k},$$

где многочлен $p(x)$ степени $\deg P - \deg Q$ — частное от деления $P(x)$ на $Q(x)$, числа x_j, p_j, q_j определяются разложением $Q(x)$ над полем $\mathbb{R}$:

$$Q(x) = \alpha(x-x_1)^{k_1}\cdot\ldots\cdot(x-x_l)^{k_l}(x^2+p_1x+q_1)^{m_1}\cdot\ldots\cdot(x^2+p_nx+q_n)^{m_n},$$

а числа a_{jk}, b_{jk}, c_{jk} определяются методом неопределенных коэффициентов.

В итоге достаточно показать, как находить первообразную от каждого слагаемого.

Шаг 2 (первообразная дробей вида $\dfrac{a}{(x-x_1)^k}$). Первообразная находится явно по таблице первообразных:

$$\int (x - x_1)^{-k}\, dx = \begin{cases} \dfrac{1}{-k+1}(x - x_1)^{-k+1}, & k \neq 1, \\ \ln|x - x_1|, & k = 1. \end{cases}$$

Шаг 3 (первообразная дробей вида $\dfrac{bx + c}{(x^2 + px + q)^k}$). Выделим в знаменателе полный квадрат и введем новую переменную:

$$x^2 + px + q = \Big(\underbrace{x + \frac{p}{2}}_{=u}\Big)^2 + \underbrace{q - \frac{p^2}{4}}_{=a^2}.$$

(Заметим, что $q - \dfrac{p^2}{4} > 0$ по условию в силу неприводимости $x^2 + px + q$ над $\mathbb{R}$.) Используя подстановку $x = u - p/2$, получаем

$$\begin{aligned} \int \frac{bx + c}{(x^2 + px + q)^k}\, dx &= \int \frac{b(u - p/2) + c}{(u^2 + a^2)^k}\, du\Big|_{u=x+p/2} \\ &= b \int \frac{u\, du}{(u^2 + a^2)^k} + \left(c - \frac{bp}{2}\right) \int \frac{du}{(u^2 + a^2)^k}\Big|_{u=x+p/2}. \end{aligned}$$

Первый интеграл легко считается при помощи замены:

$$\int \frac{u\, du}{(u^2 + a^2)^k} = \frac{1}{2} \int \frac{du^2}{(u^2 + a^2)^k} = \frac{1}{2} \begin{cases} \dfrac{1}{-k+1}(u^2 + a^2)^{-k+1}, & k \neq 1, \\ \ln|u^2 + a^2|, & k = 1. \end{cases}$$

Второй интеграл считается по индукции. При $k = 1$ имеем

$$\int \frac{du}{u^2 + a^2} = \frac{1}{a^2} \int \frac{a\, d\frac{u}{a}}{\left(\frac{u}{a}\right)^2 + 1} = \frac{1}{a} \operatorname{arctg}\left(\frac{u}{a}\right).$$

При $k > 1$ первообразная находится при помощи рекуррентной формулы, которая получается при помощи интегрирования по частям:

$$\begin{aligned}
I_k(u) &= \int \frac{du}{(u^2+a^2)^k} \\
&= \frac{u}{(u^2+a^2)^k} - \int u\, d\left(\frac{1}{(u^2+a^2)^k}\right) \\
&= \frac{u}{(u^2+a^2)^k} - \int u \cdot (-k) \cdot \frac{1}{(u^2+a^2)^{k+1}} \cdot 2u\, du \\
&= \frac{u}{(u^2+a^2)^k} + 2k \int \frac{u^2+a^2-a^2}{(u^2+a^2)^{k+1}}\, du \\
&= \frac{u}{(u^2+a^2)^k} + 2k \int \frac{du}{(u^2+a^2)^k} - 2ka^2 \int \frac{du}{(u^2+a^2)^{k+1}} \\
&= \frac{u}{(u^2+a^2)^k} + 2kI_k(u) - 2ka^2 I_{k+1}(u).
\end{aligned}$$

В итоге

$$I_{k+1}(u) = \frac{1}{2ka^2}\left(\frac{u}{(u^2+a^2)^k} + (2k-1)I_k(u)\right).$$

Теорема доказана.

§ 6. Примеры и приложения

6.1. Минимизация.

Дифференциальное исчисление даже функции одной переменной — это мощный инструмент, позволяющий решать уже интересные задачи. Начнем с простого примера, в котором надо найти минимум или максимум некоторой функции, называемой в таком аспекте *целевой функцией*.

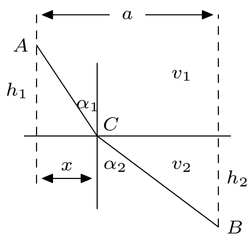

Рис. 6.1.

Пример (закон преломления). Предположим, что свет переходит из среды, в которой он распространяется с одной скоростью v_1, в среду со скоростью распространения v_2. Предполагается, что граница между средами плоская. Считается, что свет плоско-параллельный и его можно описывать лучами. Проходя через границу сред с разными скоростями распространения света, луч меняет свое направление. Траектория луча света определяется согласно принципу Ферма:

Траектория реализует минимальное время прохождения из точки одной среды в точку другой по сравнению со всеми другими траекториями.

Введем необходимые обозначения (рис. 6.1). Рассмотрим луч света, проходящий из точки A в точку B и преломляющийся в точке C на границе двух сред. Обозначим через α_1 угол между нормалью к плоскости перехода и лучом, называемый *углом падения*, через α_2 — аналогичный *угол преломления* (см. рис. 6.1). Для того чтобы установить закон отражения, нужно найти координату x точки C, которую будем отсчитывать, как показано на рис. 6.1. В наших обозначениях целевая функция, т. е. время, которое затратит рассматриваемый луч света, равна

$$T(x) = \frac{1}{v_1}\sqrt{h_1^2 + x^2} + \frac{1}{v_2}\sqrt{h_2^2 + (a-x)^2}.$$

Искомое положение x доставляет минимум функции $T(x)$. Найдем его. Для этого вычислим производную и приравняем ее к нулю:

$$T'(x) = \frac{1}{v_1}\underbrace{\frac{x}{\sqrt{h_1^2 + x^2}}}_{=\sin\alpha_1} + \frac{1}{v_2}\underbrace{\frac{-(a-x)}{\sqrt{h_2^2 + (a-x)^2}}}_{=\sin\alpha_2} = 0.$$

Нетрудно заметить, что первое из выражений в производной есть не что иное как отношение $\dfrac{\sin\alpha_1}{v_1}$, а второе — $\dfrac{\sin\alpha_2}{v_2}$, и уравнение $T'(x) = 0$ равносильно равенству

$$\frac{\sin\alpha_1}{v_1} = \frac{\sin\alpha_2}{v_2},$$

которой и является законом отражения.

Для того чтобы убедиться в том, что x действительно минимум (и вообще экстремум), найдет вторую производную и увидим, что она положительна:

$$T''(x) = \frac{1}{v_1}\frac{h_1^2}{\left(\sqrt{h_1^2 + x^2}\right)^3} + \frac{1}{v_2}\frac{h_2^2}{\left(\sqrt{h_2^2 + (a-x)^2}\right)^3} > 0.$$

Теорема о достаточном условии экстремума гарантирует, что найденная точка x действительно доставляет минимум.

6.2. Дифференциальные уравнения с разделяющимися переменными.

Дифференциальным уравнением называют уравнение, связывающее функцию и ее производные. В частности, одними из самых просто решаемых дифференциальных уравнений являются так называемые *уравнения с разделяющимися переменными.* Так называют уравнения, которые можно представить в виде

$$u'(x) = \frac{f(x)}{g(u(x))}. \tag{1}$$

Естественно предположить, что функции $f(x)$ и $g(u)$ имеют первообразные соответственно $F(x)$ и $G(u)$ и первообразная G обратима. Запишем уравнение (1) в виде

$$g(u(x))u'(x) = f(x). \tag{2}$$

По теореме 20 о замене переменной в неопределенном интеграле левая часть в (2) имеет первообразную, следовательно, из (2) получаем

$$G(u(x)) = F(x) + C. \tag{3}$$

В уравнении (3) функция u искомая. Ввиду предположения обратимости функции G из (3) имеем

$$u(x) = G^{-1}(F(x) + C). \tag{4}$$

Заметим, что в качестве решения получаем не одну функцию, а семейство функций, параметризованное константой c, т. е. для любой константы c функция u в (4) является решением. Для получения конкретного решения к уравнению необходимо добавить дополнительное условие. Так, если зафиксировать значение функции u в какой-либо точке, например, задав условие $u(x_0) = u_0$, называемое *начальным условием*, то можно найти константу c. Для этого надо подставить это условие в (4):

$$u_0 = u(x_0) = G^{-1}(F(x_0) + C),$$

или в (3):

$$G(u_0) = F(x_0) + C.$$

Из последнего соотношения C легко выражается.

Пример (охлаждение тела). Предположим, что есть небольшое тело в окружающей среде, температура которого может меняться во времени. Вместе с тем считаем, что тело настолько мало, что температура во всех его точках одинакова. Пусть $T(t)$ — температура (небольшого) тела в момент времени t, а T_0 — температура окружающей среды. Если они разные, то будет происходить передача энергии и температура тела будет изменяться. Наша цель — вывести закон, отражающий это изменение. Температура связана с количеством теплоты в этом теле. Количество теплоты в теле пропорционально температуре и равно $Q(t) = cmT(t)$, где m — масса, а c — теплоемкость тела. Количество теплоты меняется согласно физическому закону Ньютона, утверждающему, что скорость, с которой меняется количество теплоты, пропорциональна разности температур, так называемому температурному напору, т. е.

$$\frac{dQ(t)}{dt} = -\alpha(T(t) - T_0). \tag{5}$$

Будем предполагать, что $T(t) - T_0 > 0$, и тогда тело, естественно, теряет теплоту, поэтому в правой части поставлен минус. Коэффициент α эмпирический и зависит от разных факторов, например, от площади поверхности. Подставляя в соотношение (5) выражение для $Q(t)$, получаем уравнение

$$cm\frac{dT(t)}{dt} = -\alpha(T(t) - T_0). \tag{6}$$

Уравнение (5) полезно записать в дифференциальном виде, т. е. в виде равенства дифференциалов:

$$dQ(t) = -\alpha(T(t) - T_0)\,dt.$$

Здесь в соответствии с физической точкой зрения в правой части есть бесконечно малое приращение времени dt, т. е. такой промежуток времени, на котором можно пренебречь изменением температуры, и тогда количество теплоты, которое тело потеряет за этот промежуток времени, пропорционально величине этого промежутка и разности температур с некоторым коэффициентом. Здесь мы смотрим на дифференциалы как на малые изменения величин.

Уравнение (6) является уравнением с разделяющимися переменными. Записав его в виде

$$\frac{1}{T(t) - T_0}\frac{dT(t)}{dt} = -\frac{\alpha}{cm},$$

найдем первообразные от правой и левой частей:

$$\int \frac{1}{T(t)-T_0} T'(t)\,dt = -\frac{\alpha}{cm}t + C_1, \tag{7}$$

откуда

$$\ln(T(t)-T_0) = -\frac{\alpha t}{cm} + C_1.$$

Если требуется выразить температуру, то надо выразить

$$T(t) = T_0 + e^{(-\frac{\alpha t}{cm}+C_1)} = T_0 + C_2 e^{-\frac{\alpha t}{cm}}, \tag{8}$$

где $C_2 = e^{C_1}$. Получили решение.

Если задано начальное условие, например $T(0) = T_1$, то можно найти C_2. А именно, подставляя T_1 в (8), получаем $T_1 = T_0 + C_2$, т. е. $C_2 = T_1 - T_0$.

При намерении применить полученную информацию на практике появляется потребность найти еще одну из входящих в уравнение констант, это коэффициент пропорциональности α. Ее можно определить, например, так. Замерить температуру через одну минуту и получить еще одно условие, на этот раз для определения α.

6.3. Дифференциальные уравнения второго порядка с постоянными коэффициентами.

Определение. *Линейным уравнением второго порядка* называют уравнение

$$u''(x) + p(x)u'(x) + q(x)u(x) = f(x). \tag{1}$$

Здесь $u(x)$ — искомая функция, $p(x)$ и $q(x)$ — функции от x, называемые *коэффициентами*, а $f(x)$ называют *правой частью*. Если правая часть тождественно равна нулю, т. е. $f(x) = 0$, то уравнение называют *однородным*.

Мы будем рассматривать только частный случай однородного уравнения, у которого коэффициенты p и q постоянны, т. е. линейное однородное уравнение с постоянными коэффициентами. В этом случае можно получить явное решение и его проанализировать.

Будем искать решение в виде

$$u(x) = e^{\lambda x}. \tag{2}$$

Для такого рецепта есть естественное объяснение, но на нем мы останавливаться не будем — к указанному виду решения быстро привыкают и вопрос сам собой угасает. Подставим выражение (2) в уравнение (1):

$$\lambda^2 e^{\lambda x} + p\lambda e^{\lambda x} + q e^{\lambda x} = 0. \tag{3}$$

Мы хотим подобрать λ так, чтобы это уравнение выполнялось тождественно. Функция $e^{\lambda x}$ всюду отлична от нуля, поэтому уравнение (3) равносильно такому:

$$\lambda^2 + p\lambda + q = 0. \tag{4}$$

Способ решения уравнения (4) известен. Рассмотрим две возможности для дискриминанта: когда он положителен и когда отрицателен.

I. Предположим, что дискриминант уравнения (4) положителен. Тогда получаем два корня

$$\lambda_{1,2} = \frac{-p \pm \sqrt{p^2 - 4q}}{2}, \quad p^2 - 4q > 0.$$

При таком предположении получается два различных решения

$$u_1(x) = e^{\lambda_1 x}, \quad u_2(x) = e^{\lambda_2 x}.$$

Они различны в том смысле, что невозможно одно из них выразить через другое умножением на константу (т. е. они линейно независимы). Оказывается, что любая линейная комбинация этих решений

$$u(x) = C_1 u_1(x) + C_2 u_2(x) \tag{5}$$

тоже будет решением. Действительно, подставляем, опуская указания аргументов:

$$\begin{aligned}(C_1u_1 + C_2u_2)'' + p(C_1u_1 + C_2u_2)' + q(C_1u_1 + C_2u_2)\\ = C_1(u_1'' + pu_1' + qu_1) + C_2(u_2'' + pu_2' + qu_2) = 0.\end{aligned}$$

Если уравнение не было бы однородным, останется ли верным указанный факт? Ясно, что не будет.

Получили семейство решений, зависящих от двух параметров C_1, C_2. Для получения единственного решения надо поставить начальные условия. Добавим к уравнению начальные условия, иногда называемые *условиями Коши*, задав их для простоты в нуле:

$$u(0) = U_0, \quad u'(0) = U_1. \tag{6}$$

Добавление к уравнению условий (6) позволяет найти соответствующие константы и получить одно решение. Подставим информацию (6) в (5):

$$u(0) = C_1 e^{\lambda \cdot 0} + C_2 e^{\lambda_2 \cdot 0} = U_0,$$
$$u'(0) = C_1 \lambda_1 e^{\lambda \cdot 0} + C_2 \lambda_2 e^{\lambda_2 \cdot 0} = U_1.$$

Получили систему

$$\begin{aligned} C_1 + C_2 &= U_0 \\ \lambda_1 C_1 + \lambda_2 C_2 &= U_1. \end{aligned} \tag{7}$$

Поскольку числа λ_1, λ_2 различны, система (7) имеет единственное решение.

Заметим, что количество условий, налагаемых для получения одного решения, зависит от того, до какого порядка производные участвуют в уравнении. Так, в уравнении первого порядка, рассмотренном выше, такое условие одно, в рассматриваемом сейчас уравнении второго порядка — два. Будь уравнение третьего порядка, для обеспечения единственности решения понадобилось бы три условия, и т. д.

II. Мы рассмотрели случай, когда дискриминант уравнения (4) положителен. А что будет, если он отрицателен? Это означает, что вещественных корней уравнение не имеет. В этом случае уравнение также имеет решение, но искать его надо иным способом. Можно пойти одним из двух путей. Поскольку вещественных корней нет, есть два комплексных корня, можно подставить из в изложенную схему и объяснить содержание того, что значит e в комплексной степени. Можно пойти другим путем, без комплексных чисел, что мы и сделаем.

Предположим, что дискриминант отрицателен, т. е. $p^2 - 4q < 0$. Надо модифицировать вид, в котором ищутся решения. Будем искать решения в виде

$$u_1(x) = e^{\lambda x} \cos \mu x, \quad u_2(x) = e^{\lambda x} \sin \mu x. \tag{8}$$

Проделаем подробно все процедуры с решением $u_1(x)$, для второго их можно выполнить аналогично. Найдем производные:

$$u_1'(x) = \lambda e^{\lambda x} \cos \mu x - \mu e^{\lambda x} \sin \mu x,$$

$$\begin{aligned} u_1''(x) = \lambda^2 e^{\lambda x} \cos \mu x - \lambda \mu e^{\lambda x} \sin \mu x - \lambda \mu e^{\lambda x} \sin \mu x - \mu^2 e^{\lambda x} \cos \mu x \\ = (\lambda^2 - \mu^2) e^{\lambda x} \cos \mu x - 2\lambda \mu e^{\lambda x} \sin \mu x. \end{aligned}$$

Подставим полученную информацию в уравнение (1):

$$(\lambda^2-\mu^2)e^{\lambda x}\cos\mu x - 2\lambda\mu e^{\lambda x}\sin\mu x \\ + p(\lambda e^{\lambda x}\cos\mu x - \mu e^{\lambda x}\sin\mu x) + q e^{\lambda x}\cos\mu x = 0.$$

Это равенство должно выполняться тождественно, т. е. при всех значениях x. Для этого надо, чтобы коэффициенты при одинаковых функциях были нулевыми. Соответственно сгруппировав, получаем равенства

$$\lambda^2-\mu^2+p\lambda+q=0, \quad (2\lambda+p)\mu=0. \tag{9}$$

Получили систему двух уравнений относительно λ и μ. Из второго уравнения видно, что либо $\mu=0$, либо $2\lambda+p=0$. Если $\mu=0$, то из первого уравнения получаем $\lambda^2+p\lambda+q=0$, где слева расположен характеристический многочлен. У него вещественных корней по предположению нет, так что в этом случае вещественных корней, которые мы ищем, нет.

Следовательно, надо использовать возможность $2\lambda+p=0$, т. е. $\lambda=-\dfrac{p}{2}$. Подставим это выражение в первое из уравнений (9):

$$\left(-\frac{p}{2}\right)^2-\mu^2-\frac{p^2}{2}+q=0.$$

Отсюда

$$\mu=\pm\sqrt{q-\frac{p^2}{4}}.$$

По предположению подкоренное выражение положительно.

Можно заметить, что, проделывая те же действия со вторым решением, в итоге придем к такой же системе характеристических уравнений и будут те же самые решения.

В итоге оказалось, что есть два решения

$$u_1(x)=e^{\lambda x}\cos\mu x=e^{-\frac{p}{2}x}\cos\sqrt{q-\frac{p^2}{2}}x,$$

$$u_2(x)=e^{\lambda x}\sin\mu x=e^{-\frac{p}{2}x}\sin\sqrt{q-\frac{p^2}{2}}x.$$

Обратим внимание на то, что из двух знаков, с которыми было записано значение для μ, оставлен только плюс. Причина здесь такова. Для четной функции косинус, в аргументе которой стоит соответствующий корень, безразлично, с каким знаком брать корень. Для нечетной функции синус различия будут только в знаке, а поскольку в дальнейшем мы будем брать линейную комбинацию решений $u_1(x)$ и $u_2(x)$, безразлично, с каким знаком будет u_2. В итоге любое решение уравнения (1) в рассматриваемом случае выражается как линейная комбинация

$$u(x) = e^{\lambda x}(C_1 \cos \mu x + C_2 \sin \mu x).$$

Пример (масса на пружине с амортизатором). Рассмотрим простейшую механическую систему. Есть стенка, к которой крепится пружинка, к ней крепится грузик массой m и есть резервуар с жидкостью, в которую входит поршень, с трубочкой для перетекания жидкости (рис. 6.2). Задача состоит в описании движения этой механической системы. Описывать будем положением груза относительно его положения покоя. Иначе говоря, есть естественное состояние пружины, когда она не сжата и не растянута, и именно в этом положении считаем, что $u = 0$. Пружина характеризуется некоторым коэффициентом жесткости k, который входит в закон Гука, а относительно поршня будем предполагать, что его сопротивление пропорционально скорости. В некотором диапазоне небольших скоростей это довольно естественное предположение.

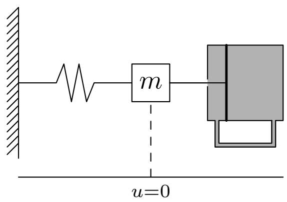

Рис. 6.2.

Для описания рассмотрим второй закон Ньютона согласно которому надо посчитать силы, действующие на тело, и приравнять их произведению массы на ускорение, которое запишется так:

$$m \cdot a = mu''(t).$$

Разберемся, как будут действовать силы. Если смещаемся вправо, т. е. $u(t) > 0$, сила будет направлена влево, потому что пружина сопротивляется растяжению. Тем самым на основании закона Гука появится выражение $-ku(t)$. Затем надо понять, в какую сторону будет действовать сила сопротивления поршня в цилиндре. Если двигаемся в положительном направлении, т. е. скорость положительна, сила сопротивления будет действовать влево, значит, ее также надо учесть со знаком минус, и добавится выражение $-\gamma u'(t)$, где γ — коэффициент сопротивления среды. Как было запланировано, по второму закону Ньютона имеем

$$mu''(t) = -ku(t) - \gamma u'(t). \tag{10}$$

Получили уравнение. Перепишем его в виде

$$u'' + \frac{\gamma}{m}u' + \frac{k}{m}u = 0.$$

Составим характеристическое уравнение:

$$\lambda^2 + \left(\frac{\gamma}{m}\right)\lambda + \frac{k}{m} = 0.$$

Найдем дискриминант и используем его знак для выбора вида искомого решения. Если

$$\left(\frac{\gamma}{m}\right)^2 - 4\frac{k}{m} > 0,$$

то решение ищется в виде $u_j(t) = e^{\lambda_j t}$. После подстановки в уравнение получим

$$\lambda_{1,2} = -\frac{\gamma}{2m} \pm \sqrt{\frac{\gamma^2}{4m^2} - \frac{k}{m}}. \tag{11}$$

Физики, получив выражение решения, проводят его анализ. Получилась экспонента, она может расти, а может убывать. Это зависит от знаков показателей. Получившиеся значения λ всегда отрицательны. Первый член в (11) отрицателен, а второй по модулю меньше первого, так что даже его добавление к первому дает в сумме отрицательное число. В качестве решений будут возникать комбинации двух затухающих экспонент, что вполне естественно. Когда реализуется ситуация с положительным дискриминантом? Когда γ большое, т. е. когда сопротивление в амортизаторе большое — тогда колебаний практически не будет, сразу произойдет затухание.

Пусть

$$\left(\frac{\gamma}{m}\right)^2 - 4\frac{k}{m} < 0.$$

Тогда решение ищем в виде $u_1(t) = e^{\lambda t}\cos\mu t$ либо $u_2(t) = e^{\lambda t}\sin\mu t$ и после подстановки получается

$$\lambda = -\frac{\gamma}{2m}, \quad \mu = \sqrt{\frac{k}{m} - \left(\frac{\gamma}{2m}\right)^2}.$$

Получается два решения:

$$u_1(t) = e^{-\frac{\gamma}{2m}t}\cos\sqrt{\frac{k}{m} - \left(\frac{\gamma}{2m}\right)^2}\cdot t,$$

$$u_2(t) = e^{-\frac{\gamma}{2m}t}\sin\sqrt{\frac{k}{m} - \left(\frac{\gamma}{2m}\right)^2}\cdot t.$$

Визуально решение выглядит как синусоида, поджатая к оси абсцисс графиком экспоненты.

В этот момент физик, во-первых, проверяет результат по размерности, и мы этим заниматься не будем. А во-вторых, рассматривает крайние случаи, когда, например, что-то отсутствует. Один из крайних случаев — отсутствие затухания, т. е. цилиндра с поршнем. Тогда $\gamma = 0$, значит, затухания не будет, и останется только выражение с корнем. Под корнем останется только одно первое слагаемое. Как известно, $\sqrt{\dfrac{k}{m}}$ — это частота, с которой система будет совершать колебания, причем бесконечно долго.

Оглавление

Основы математического анализа
для студентов-физиков. Лекции.
2. Дифференциальное исчисление
функций одной переменной

Дятлов Глеб Владимирович

Подписано в печать 5.11.2014. Формат $60 \times 84\,{}^{1}/_{16}$. Печать офсетная.
Усл. печ. л. 3.8. Уч.-изд. л. 3.8. Тираж 200 экз. Заказ №   .

Лицензия ЛР № 065614 от 8 января 1998 г.
Издательство Института математики,
пр. Академика Коптюга, 4, 630090 Новосибирск, Россия.

Отпечатано в ООО «Омега Принт»,
пр. Академика М. А. Лаврентьева, 6, 630090 Новосибирск, Россия.