В.А. Ильин, В.А. Садовничий,
Бл.Х. Сендов

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

НАЧАЛЬНЫЙ КУРС

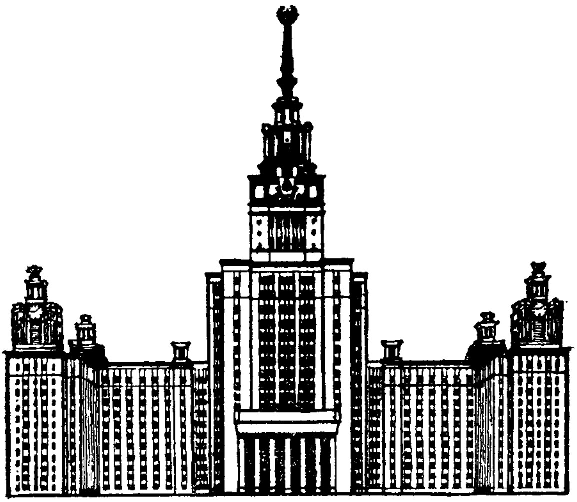
Под редакцией академика А.Н. Тихонова

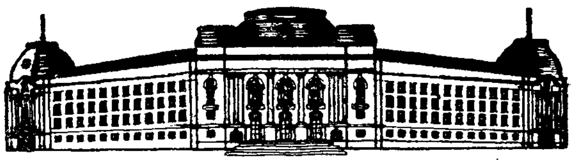
ИЗДАНИЕ ВТОРОЕ,
ПЕРЕРАБОТАННОЕ

Допущено Министерством высшего
и среднего специального образования СССР
в качестве учебника для студентов вузов,
обучающихся по специальностям „Математика",
„Прикладная математика"

ИЗДАТЕЛЬСТВО
МОСКОВСКОГО УНИВЕРСИТЕТА
1985

МГУ





СОВМЕСТНОЕ ИЗДАНИЕ
МОСКОВСКОГО
ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
ИМЕНИ М.В. ЛОМОНОСОВА
И СОФИЙСКОГО УНИВЕРСИТЕТА
ИМЕНИ КЛИМЕНТА ОХРИДСКОГО,
НАПИСАННОЕ В СООТВЕТСТВИИ
С ЕДИНОЙ ПРОГРАММОЙ
ПО МАТЕМАТИЧЕСКОМУ АНАЛИЗУ

УДК 517

Ильин В. А. и др. Математический анализ. Начальный курс/В. А. Ильин, В. А. Садовничий, Бл. Х. Сендов. Под ред. А. Н. Тихонова.— 2-е изд., перераб.— М.: Изд-во МГУ, 1985.— 662 с.

Учебник представляет собой первую часть трехтомного курса математического анализа для высших учебных заведений СССР, Болгарии и Венгрии, написанного в соответствии с соглашением о сотрудничестве между Московским, Софийским и Будапештским университетами. Книга включает в себя теорию вещественных чисел, теорию пределов, теорию непрерывности функций, дифференциальное и интегральное исчисления функций одной переменной и их приложения, дифференциальное исчисление функций многих переменных и теорию неявных функций.

Рецензент:

Кафедра математики МИФИ
(зав. кафедрой проф. **А. И. Прилепко**)

И $\frac{1702050000-150}{077(02)-85}$ 95–85

ПРЕДИСЛОВИЕ ТИТУЛЬНОГО РЕДАКТОРА

В настоящее время прогресс в математике в большой степени связан с развитием электронно-вычислительных средств. Математические методы исследования проникают во все области человеческой деятельности. Все это повышает интерес к математике со стороны смежных наук, использующих различный объем математических знаний, и ставит новые задачи в изучении самой математики. В связи с этим возникает потребность в написании учебника по математическому анализу, учитывающего указанные закономерности.

То обстоятельство, что решение математических задач реализуется на ЭВМ с помощью вычислительных алгоритмов, предъявляет повышенные требования к четкости алгоритмического уровня изложения математических дисциплин. Однако такое изложение должно базироваться на классических концепциях математики и не должно их затемнять.

Эти общие принципы вместе с задачей четкого, ясного и доступного изложения и положены в основу написания предлагаемой читателю книги. Книга написана с учетом согласованной между Московским и Софийским университетами программы преподавания первой части математического анализа.

В предлагаемом учебнике уделено большое внимание вопросам оптимизации, играющим в математике и ее приложениях большую роль. В частности, в книге впервые в учебной литературе в законченном виде излагается алгоритм отыскания как внутреннего, так и краевого экстремума функции. В учебнике уделено значительное внимание изучению вопроса об исходной информации, доступной при решении задачи. Так, например, для отыскания экстремума функции одной переменной авторы предлагают алгоритм, базирующийся на информации только о значениях функции в точках области ее задания. Предлагаемое решение не опирается на знание значений производной в точках области задания и пригодно для отыскания экстремума недифференцируемых функций. Такая постановка типична при решении задач об оптимизации производственных процессов.

При выборе метода изложения авторы отправляются от того, что выбор алгоритма решения задачи зависит от того, какая информация из постановки этой задачи может быть использована. Так, например, при введении понятия определенного интеграла

Римана авторы отправляются от концепции изложения, базирующейся на использовании значений функции в точках сегмента.

Эта концепция, несомненно, является более предпочтительной по сравнению с концепцией введения определенного интеграла Римана с помощью первообразной, ибо она отвечает идее численных методов вычисления определенного интеграла, используемых на ЭВМ.

В заключение хочу высказать уверенность, что предлагаемая книга будет способствовать повышению математической культуры читателей с различными запросами к объему математических знаний.

А. Тихонов

Москва,
сентябрь 1978 г.

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

Во втором издании книга подверглась существенной переработке и сокращению в целях максимального приближения ее материала к тому курсу, который реально может быть прочитан студентам первого года обучения.

Особенно существенной переработке были подвергнуты разделы, посвященные теории вещественных чисел, теории множеств и теории метрических, топологических и нормированных пространств.

В книге сохранены три уровня изложения (облегченный, основной и повышенный).

Так же, как и в первом издании, текст повышенного уровня выделен в книге двумя вертикальными чертами, текст основного уровня — одной вертикальной чертой, а остальной текст книги относится к облегченному уровню изложения.

Проведенные во втором издании переработки улучшили возможности использования книги на указанных трех различных уровнях изложения.

Авторы выражают благодарность сотрудникам кафедры математического анализа механико-математического факультета МГУ и сотрудникам кафедры общей математики факультета вычислительной математики и кибернетики МГУ за критические замечания по первому изданию этой книги. Авторы благодарят также В. М. Говорова, В. Н. Денисова, И. С. Ломова и В. В. Тихомирова за помощь при подготовке второго издания этой книги. Особую благодарность авторы приносят А. И. Прилепко, прочитавшему рукопись второго издания и сделавшему критические замечания, способствующие ее улучшению.

Москва,
февраль 1985 г.

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Настоящая книга является учебником по математическому анализу по согласованной между Московским и Софийским университетами единой программе первого года обучения. Она полностью охватывает материал первого года обучения, предусмотренный программой для студентов университетов СССР и НРБ, обучающихся по специальностям «математика», «механика» и «прикладная математика».

Особенностью этой книги является то, что она содержит три четко отделяемых друг от друга уровня изложения: облегченный, основной и повышенный, причем для понимания материала облегченного уровня не требуется чтения материалов основного и повышенного уровней, а для понимания материала основного уровня не требуется чтения материала повышенного уровня.

Облегченный уровень отвечает программе технических вузов СССР с углубленным изучением математического анализа; основной уровень изложения отвечает программе специальностей «прикладная математика» и «физика» университетов СССР; материал повышенного уровня дополняет материал основного уровня разделами, обычно излагаемыми на механико-математических факультетах университетов.

Текст, выделенный в книге двумя вертикальными чертами, относится к повышенному уровню изложения; текст, выделенный одной вертикальной чертой, — к основному уровню изложения; остальной текст книги составляет содержание облегченного уровня изложения.

Книга содержит вводную главу, разъясняющую возникновение основных понятий математического анализа и облегчающую восприятие последующего материала.

В книге нашла отражение возросшая роль вычислительных методов и содержится ряд примеров применения аппарата математического анализа для вычисления элементарных функций, интегралов и отыскания корней уравнений и точек экстремума.

В настоящее время в СССР и НРБ имеется целый ряд учебников по математическому анализу, среди которых особенно удачными, по нашему мнению, являются учебники, написанные Л. Д. Кудрявцевым и С. М. Никольским в СССР и Я. Тагамлицким в НРБ.

Авторы настоящей книги, несомненно, испытали влияние этих прекрасных учебников.

При написании этой книги авторы использовали часть материала книги В. А. Ильина и Э. Г. Позняка «Основы математического анализа», а также опыт преподавания математического анализа в университетах.

Авторы выражают глубокую благодарность титульному редактору этой книги академику А. Н. Тихонову за большое количество ценных советов и замечаний.

Авторы благодарят также Л. Д. Кудрявцева, И. И. Ляшко, В. Л. Макарова, Д. Б. Дойчинова и Т. Боянова, критические замечания которых способствовали улучшению этой книги.

Особой благодарностью авторы отмечают труд В. М. Говорова и Г. Христова, который намного превзошел рамки обычного редактирования.

София,
март 1978 г.

Глава 1

ОСНОВНЫЕ ПОНЯТИЯ МАТЕМАТИЧЕСКОГО АНАЛИЗА

В настоящей главе, не претендуя на точность формулировок и отправляясь от простейших задач механики, мы постараемся обрисовать основной круг понятий и проблем математического анализа.

1. Начнем наше рассмотрение с выяснения тех математических понятий, которые неизбежно возникают при описании самого простейшего вида движения — движения материальной точки вдоль прямой линии.

Если материальная точка движется вдоль оси Oy, а x обозначает время, отсчитываемое от некоторого начального момента, то для описания указанного движения необходимо знать правило, посредством которого каждому значению времени x ставится в соответствие координата y движущейся точки в момент времени x.

В механике такое правило называют законом движения. Абстрагируясь от конкретного механического смысла переменных x и y и рассматривая в качестве x и y две совершенно произвольные переменные величины, мы придем к понятию функции, являющемуся одним из важнейших понятий математического анализа.

Если известно правило, посредством которого каждому значению одной переменной x ставится в соответствие определенное значение другой переменной y, то говорят, что переменная y является функцией переменной x, и пишут $y=y(x)$ или $y=f(x)$.

При этом переменную x называют аргументом или независимой переменной, а переменную y — функцией аргумента x.

Букву f в записи $y=f(x)$ обычно называют характеристикой рассматриваемой функции, а значение $y=f(x)$ называют частным значением функции в точке x. Совокупность всех частных значений функции принято называть областью изменения этой функции.

Отметим сразу же, что приведенная формулировка понятия функции требует уточнения, ибо в этой формулировке ничего не говорится о том, из какого множества берутся значения независимой переменной x.

Множество, состоящее из тех и только тех чисел, которые являются значениями независимой переменной x, обычно называют

областью задания функции. Описание областей задания функции требует развития теории числовых множеств.

Отметим еще, что понятие функции (так же, как и понятие числа, множества и переменной величины) естественно считать **начальным понятием** (т. е. таким понятием, которое можно описать, но нельзя строго определить, ибо любая попытка дать строгое определение указанного понятия неизбежно сведется к замене определяемого понятия ему эквивалентным). Таким образом, вместо термина «определение функции» естественнее употреблять термин «понятие функции».

Отметим, наконец, что для обозначения аргумента функции и ее характеристики могут употребляться различные буквы. Так, например, запись $x=\varphi(t)$ обозначает, что переменная x является функцией аргумента t, причем характеристика этой функции обозначена через φ. При одновременном рассмотрении нескольких функций одного аргумента t для обозначения характеристик этих функций необходимо употреблять различные символы.

2. Часто приходится рассматривать такую функцию $y=f(x)$, аргумент x которой сам является функцией вида $x=\varphi(t)$ некоторой новой переменной t. В таком случае говорят, что переменная y представляет собой **сложную функцию** аргумента t, а переменную x называют **промежуточным** аргументом. Указанную сложную функцию называют также **суперпозицией** функций f и φ. Для обозначения указанной сложной функции естественно использовать символ $y=f[\varphi(t)]$.

Рассмотрим простой пример, иллюстрирующий возникновение понятия сложной функции. Предположим, что материальная точка M равномерно с постоянной угловой скоростью ω вращается по окружности радиуса R. Найдем закон движения проекции M' точки M на некоторую ось Oy, проходящую через центр O окружности и лежащую в ее плоскости (рис. 1.1). При этом естественно предположить, что в начальный момент времени $t=0$ движущаяся точка M находилась в точке M_0 пересечения окружности с осью Oy. Обозначим через y координату проекции M' точки M на ось Oy, а через x угол M_0OM, на который повернется точка M за время t. Очевидно, что $y=R\cos x$, $x=\omega t$, и мы получим, что координата y проекции M' представляет собой сложную функцию времени t вида $y=R\cos x$, где $x=\omega t$. Эту сложную функцию можно записать в виде $y=R\cos\omega t$. Отметим, что движение по закону $y=R\cos\omega t$ в механике принято называть **гармоническим колебанием**.

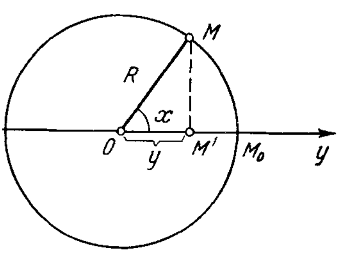


Рис. 1.1

3. Из курса физики известно, что важной характеристикой движения материальной точки является ее мгновенная скорость в каждый момент времени x. Если материальная точка движется вдоль оси Oy по закону $y=f(x)$, то, фиксировав произвольный момент времени x и какое угодно приращение времени Δx, мы можем утверждать, что в момент времени x движущаяся точка имеет координату $f(x)$, а в момент времени $x+\Delta x$ — координату $f(x+\Delta x)$.

Таким образом, число $\Delta y=f(x+\Delta x)-f(x)$ представляет собой путь, пройденный движущейся точкой за промежуток времени от x до $x+\Delta x$.

Отсюда вытекает, что отношение

$$\frac{\Delta y}{\Delta x}=\frac{f(x+\Delta x)-f(x)}{\Delta x}, \tag{1.1}$$

обычно называемое *разностным отношением*, представляет собой *среднюю скорость* движущейся точки за промежуток времени от x до $x+\Delta x$.

Мгновенной скоростью (или просто *скоростью*) движущейся точки называется предел, к которому стремится средняя скорость (1.1) при стремлении к нулю промежутка времени Δx.

Если использовать известный из курса средней школы символ предела, то можно записать следующее соотношение для мгновенной скорости $v(x)$ в момент времени x:

$$v(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}. \tag{1.2}$$

Физическое понятие мгновенной скорости приводит к фундаментальному математическому понятию производной. Абстрагируясь от механического смысла рассмотренной выше функции $y=f(x)$, мы *назовем производной произвольной функции* $y=f(x)$ *в данной фиксированной точке* x *предел, стоящий в правой части* (1.2) (*при условии, конечно, что этот предел существует*).

Используя для обозначения производной функции $y=f(x)$ в точке x символ $f'(x)$ или $y'(x)$, мы можем по определению записать следующее равенство:

$$f'(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

Операцию нахождения производной договоримся называть *дифференцированием*.

Наше рассмотрение показывает, что при вычислении производной фундаментальную роль играет понятие предела функции.

Предварительное представление о понятии предела функции (да и о самом понятии производной) дается в курсе средней школы. Однако строгое и последовательное изучение понятия предела возможно лишь на базе строгой теории вещественных чисел. Так, например, без строгой теории вещественных чисел невозможно установить существование двух следующих важных пределов:

$$\lim_{t\to 0}\frac{\sin t}{t} \text{ и } \lim_{t\to 0}(1+t)^{1/t},$$

неизбежно возникающих, как мы увидим ниже, при вычислении производных функций $y=\sin x$ и $y=\log_a x$.

Итак, проведенное нами рассмотрение показывает, что вопрос о существовании и вычислении производной упирается в необходимость развития строгой теории вещественных чисел и на ее базе теории пределов.

4. Займемся теперь вычислением производных двух конкретных элементарных функций $y=\sin x$ и $y=\log_a x$ и выясним, какие математические проблемы неизбежно возникают при этом.

Сначала вычислим производную функцию $y=\sin x$ в любой фиксированной точке x. Для этой функции разностное отношение (1.1), очевидно, имеет вид

$$\frac{\Delta y}{\Delta x}=\frac{\sin(x+\Delta x)-\sin x}{\Delta x}=\cos\left(x+\frac{\Delta x}{2}\right)\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}.$$

Таким образом, производная функции $y=\sin x$ в точке x по определению равна пределу

$$\lim_{\Delta x\to 0}\left\{\cos\left(x+\frac{\Delta x}{2}\right)\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\right\} \tag{1.3}$$

(при условии, что этот предел существует).

Можно ожидать, что

$$\lim_{\Delta x\to 0}\cos\left(x+\frac{\Delta x}{2}\right)=\cos x. \tag{1.4}$$

Заметим, однако, что не для всякой функции $f(x)$ справедливо равенство

$$\lim_{\Delta x\to 0} f\left(x+\frac{\Delta x}{2}\right)=f(x). \tag{1.5}$$

Функция $f(x)$, для которой в данной точке x справедливо равенство (1.5), *называется* ***непрерывной*** *(в точке x).* Понятие непрерывности функции является одним из важнейших математических понятий и будет основательно изучаться в систематическом курсе математического анализа. В частности, в систематическом курсе будет доказано, что функция $y=\cos x$ является непрерывной в каждой точке x, т. е. в каждой точке x справедливо равенство (1.4).

Заметим теперь, что для вычисления предела (1.3) недостаточно доказать справедливость соотношения (1.4). Для этого необходимо еще вычислить следующий предел:

$$\lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}=\lim_{t\to 0}\frac{\sin t}{t}\quad\left(t=\frac{\Delta x}{2}\right). \tag{1.6}$$

В систематическом курсе анализа будет строго доказано, что предел (1.6), часто называемый *первым замечательным пределом*, существует и равен единице.

Только после того, как будет установлена непрерывность функции $y=\cos x$ (т. е. равенство (1.4)) и вычислен первый замечательный предел (1.6), мы сможем, опираясь еще на то, что предел произведения равен произведению пределов сомножителей, строго утверждать, что предел (1.3) существует и равен $\cos x$ или, что то же самое, производная функции $y=\sin x$ существует и равна $\cos x$.

Перейдем теперь к вычислению производной функции $y=\log_a x$, считая, что $0<a\neq 1$, и фиксировав произвольную точку $x>0$. Для этой функции разностное отношение (1.1) имеет вид

$$\frac{\Delta y}{\Delta x}=\frac{\log_a(x+\Delta x)-\log_a x}{\Delta x}$$

($\Delta x\neq 0$ и выбирается так, что $x+\Delta x>0$). Таким образом, производная функции $y=\log_a x$ в любой точке $x>0$ по определению равна пределу

$$\lim_{\Delta x\to 0}\frac{\log_a(x+\Delta x)-\log_a x}{\Delta x} \tag{1.7}$$

(при условии, что этот предел существует). Преобразуем дробь, стоящую в (1.7), проделав следующие операции: 1) заменим разность логарифмов логарифмом частного; 2) произведем умножение и деление на одну и ту же величину $x>0$; 3) внесем множитель, стоящий перед логарифмом, под знак логарифма, сделав его показателем степени. В результате получим, что предел (1.7) равен

$$\lim_{\Delta x\to 0}\left\{\frac{1}{\Delta x}\log_a\left(1+\frac{\Delta x}{x}\right)\right\}=\lim_{\Delta x\to 0}\left\{\frac{1}{x}\frac{x}{\Delta x}\log_a\left(1+\frac{\Delta x}{x}\right)\right\}=$$

$$= \lim_{\Delta x \to 0} \left\{ \frac{1}{x} \log_a \left[\left(1 + \frac{\Delta x}{x} \right)^{\frac{x}{\Delta x}} \right] \right\} =$$

$$= \lim_{t \to 0} \left\{ \frac{1}{x} \log_a \left[(1+t)^{\frac{1}{t}} \right] \right\} \quad \left(t = \frac{\Delta x}{x} \to 0 \right). \tag{1.8}$$

Рассмотрим отдельно предел при $t \to 0$ выражения, заключенного в правой части последнего равенства в квадратные скобки:

$$\lim_{t \to 0} [(1+t)^{\frac{1}{t}}]. \tag{1.9}$$

Этот предел часто называют **вторым замечательным пределом**. В систематическом курсе анализа будет установлено, что этот предел равен иррациональному числу e, которое с точностью до пятнадцати знаков после запятой имеет вид $e = 2{,}718281828459045\ldots$

Кроме того, в систематическом курсе будет доказана непрерывность функции $y = \log_a x$ в каждой точке $x > 0$ и, в частности, в точке $x = e$. Но тогда из существования равного e предела (1.9) будет следовать, что

$$\lim_{t \to 0} \log_a [(1+t)^{\frac{1}{t}}] = \log_a e.$$

Последнее соотношение и соотношение (1.8) позволяют утверждать, что предел (1.7) равен

$$\lim_{\Delta x \to 0} \frac{\log_a (x + \Delta x) - \log_a x}{\Delta x} = \frac{1}{x} \log_a e.$$

Таким образом, после того как будет вычислен второй замечательный предел и установлена непрерывность функции $y = \log_a x$ в точке e, мы сможем строго утверждать, что логарифмическая функция имеет производную, причем

$$(\log_a x)' = \frac{1}{x} \log_a e \text{ при } 0 < a \neq 1, \ x > 0.$$

5. В курсе средней школы кроме двух рассмотренных нами функций $y = \sin x$ и $y = \log_a x$ изучались еще следующие функции: $y = \cos x$, $y = \operatorname{tg} x$, $y = \operatorname{ctg} x$, $y = x^\alpha$ (α — вещественное число), $y = a^x$ $(0 < a \neq 1)$, $y = \arcsin x$, $y = \arccos x$, $y = \operatorname{arctg} x$, $y = \operatorname{arcctg} x$.

Все перечисленные функции принято называть **простейшими элементарными**.

Замечательным является тот факт, что при вычислении производных всех простейших элементарных функций не возникает никаких новых трудностей, кроме тех, с которыми мы встретились при вычислении производных функций $y = \sin x$ и $y = \log_a x$. Нетрудно проверить, что для вычисления производных всех простейших элементарных функций требуются лишь арифметические свойства операции предельного перехода, два замечательных предела и факт непрерывности каждой из этих функций в точках областей их задания.

Отмеченное обстоятельство дает нам право без дальнейших разъяснений привести таблицу производных всех простейших элементарных функций.

1°. $(x^{\alpha})'=\alpha x^{\alpha-1}$ ($x>0$, α — вещественное число).

2°. $(\log_a x)'=\frac{1}{x}\log_a e$ $(0<a\neq 1,\ x>0)$.

В частности, при $a=e$

$$(\log_e x)'=\frac{1}{x}.$$

3°. $(a^x)'=a^x\log_e a$ $(0<a\neq 1)$.

В частности, при $a=e$

$$(e^x)'=e^x.$$

4°. $(\sin x)'=\cos x$.

5°. $(\cos x)'=-\sin x$.

6°. $(\operatorname{tg} x)'=\frac{1}{\cos^2 x}$ $\left(x\neq\frac{\pi}{2}+n\pi,\text{ где } n=0,\ \pm 1,\pm 2,\ldots\right)$.

7°. $(\operatorname{ctg} x)'=-\frac{1}{\sin^2 x}$ $(x\neq n\pi,\text{ где } n=0,\ \pm 1,\ \pm 2,\ldots)$.

8°. $(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}$ $(|x|<1)$.

9°. $(\arccos x)'=-\frac{1}{\sqrt{1-x^2}}$ $(|x|<1)$.

10°. $(\operatorname{arctg} x)'=\frac{1}{1+x^2}$.

11°. $(\operatorname{arcctg} x)'=-\frac{1}{1+x^2}$.

Строгое обоснование приведенной таблицы является одной из важных задач той части математического анализа, которую принято называть *дифференциальным исчислением*.

Традиционной задачей классического дифференциального исчисления является и несколько более общая задача — вычисление производной любой функции $f(x)$, которая получается из перечисленных выше простейших элементарных функций путем конечного числа суперпозиций и конечного числа четырех арифметических действий (сложения, умножения, вычитания и деления). Такую функцию $f(x)$ принято называть просто *элементарной*.

Итак, *элементарной* называется функция, которая получается из простейших элементарных функций путем конечного числа суперпозиций и четырех арифметических действий.

Примером элементарной функции может служить функция

$$f(x)=2^{\operatorname{arctg}\left(\frac{x+1}{x^2+2}\right)}.$$

Для вычисления производной любой элементарной функции следует присоединить к выписанной нами таблице производных простейших элементарных функций два правила: 1) правило дифференцирования сложной функции, 2) правило дифференцирования суммы, разности, произведения и частного функций.

Правило дифференцирования сложной функции $y=f(u)$, где $u=\varphi(x)$, имеет следующий вид: если функция $u=\varphi(x)$ имеет производную в данной точке x_0, а функция $y=f(u)$ имеет производную в соответствующей точке $u_0=\varphi(x_0)$, то сложная функция $y=f[\varphi(x)]$ имеет производную в точке x_0, причем эта производная (обозначим ее через y') равна

$$y'=f'[\varphi(x_0)]\varphi'(x_0), \tag{1.10}$$

т. е. равна произведению производной функции $y=f(u)$ в точке $u_0=\varphi(x_0)$ на производную функции $u=\varphi(x)$ в точке x_0.

Справедливость для производной сложной функции формулы (1.10) легко оправдать с помощью наводящих соображений, но строгий вывод формулы (1.10) не является простым и будет приведен в систематическом курсе математического анализа.

Гораздо проще устанавливаются правила дифференцирования суммы, разности, произведения и частного двух функций, которые имеют вид

$$[u(x)\pm v(x)]'=u'(x)\pm v'(x),$$

$$[u(x)v(x)]'=u'(x)v(x)+u(x)v'(x),$$

$$\left[\frac{u(x)}{v(x)}\right]'=\frac{u'(x)v(x)-u(x)v'(x)}{v^2(x)}$$

(в последней формуле требуется необращение в нуль функции $v(x)$ в рассматриваемой точке x).

Подводя итог, мы можем заключить, что одной из важных задач части математического анализа, называемой дифференциальным исчислением, является строгое обоснование таблицы производных простейших элементарных функций и правил дифференцирования сложной функции, а также суммы, разности, произведения и частного функций.

Это обоснование позволит вычислить производную любой элементарной функции $f(x)$, т. е. любой функции $f(x)$, получающейся из простейших элементарных путем конечного числа суперпозиций и четырех арифметических действий. При этом оказывается, что производная любой элементарной функции представляет собой также элементарную функцию, т. е. операция дифференцирования не выводит нас из класса элементарных функций.

Отмеченное обстоятельство оправдывает введение класса элементарных функций как традиционного объекта классического анализа.

6. Еще раз обратимся к рассмотрению механической задачи о движении материальной точки вдоль прямой линии — оси Oy, но на этот раз предположим, что для любого момента времени x нам задана мгновенная скорость $f(x)$ движущейся точки и требуется найти закон движения этой точки.

Поскольку мгновенная скорость $f(x)$ является производной функции $y=F(x)$, определяющей закон движения, то задача сводится к разысканию по данной функции $f(x)$ такой функции $F(x)$, производная $F'(x)$ которой равна $f(x)$. Отвлекаясь от механического смысла функций $f(x)$ и $F(x)$, мы придем к математическим понятиям первообразной и неопределенного интеграла.

Первообразной функции $f(x)$ называется такая функция $F(x)$, производная которой $F'(x)$ равна $f(x)$.

Это определение требует уточнения: следует четко оговорить, на каком множестве должно быть справедливо равенство $F'(x)=f(x)$. Отмеченное обстоятельство еще раз подчеркивает необходимость развития теории множеств. Уточнение понятия первообразной будет дано в систематическом курсе.

Заметим, что если функция $F(x)$ является первообразной функции $f(x)$, то и функция $F(x)+C$, где C — произвольная постоянная, также является первообразной функции $f(x)$ (в силу того, что производная постоянной C равна нулю).

Более трудным является обратное утверждение: любые две первообразные одной и той же функции $f(x)$ на интервале (a, b) могут отличаться лишь постоянным слагаемым. Доказательство этого утверждения требует развитого аппарата математического анализа и будет проведено в систематическом курсе анализа.

Опираясь на указанное утверждение, мы можем констатировать следующий факт: если функция $F(x)$ является одной из первообразных функции $f(x)$, то любая первообразная функции $f(x)$ имеет вид $F(x)+C$, где C — постоянная.

Совокупность всех первообразных данной функции $f(x)$ называется неопределенным интегралом от функции $f(x)$ и обозначается символом

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

Только что отмеченный нами факт позволяет утверждать, что если $F(x)$ — одна из первообразных функции $f(x)$, то неопределенный интеграл от функции $f(x)$ равен

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

где C — любая постоянная.

Возвратимся к поставленной нами задаче об отыскании закона движения материальной точки вдоль оси Oy по известной мгновенной скорости $f(x)$ этой точки. Мы теперь можем утверждать, что искомый закон движения определяется функцией $y=F(x)+C$, где $F(x)$ — любая первообразная функция $f(x)$, а C — постоянная. Как мы видим, без дополнительных условий закон движения по мгновенной скорости определяется неоднозначно: с точностью до постоянного слагаемого C. Для определения постоянной C должно быть привлечено дополнительное условие, обычно заключающееся в задании координаты y_0 движущейся точки в некоторый момент времени x_0. Используя это условие, мы получим соотношение $y_0=F(x_0)+C$, из которого $C=y_0-F(x_0)$, так что окончательно искомый закон движения имеет вид

$$y=F(x)+y_0-F(x_0).$$

7. Рассмотрим вопрос об отыскании первообразных и неопределенных интегралов от некоторых элементарных функций. Так как функция $f(x)=\cos x$ является производной функции $F(x)=\sin x$, то функция $F(x)=\sin x$ является одной из первообразных функции $f(x)=\cos x$, и потому любая первообразная функции $f(x)=\cos x$ имеет вид $\sin x+C$, где C — постоянная. Таким образом,

$$\int \cos x dx=\sin x+C.$$

Только что проведенное рассуждение имеет общий характер. Можно утверждать, что любая формула дифференциального исчисления $F'(x)=f(x)$, утверждающая, что функция $f(x)$ является производной функции $F(x)$, порождает эквивалентную ей формулу интегрального исчисления $\int f(x)dx=F(x)+C$, утверждающую, что неопределенный интеграл от функции $f(x)$ равен $F(x)+C$, где C — любая постоянная.

Таким образом, выписанная выше таблица производных простейших элементарных функций порождает эквивалентную ей таблицу важных неопределенных интегралов, которую мы приводим ниже.

1°. $\int x^\alpha dx=\frac{x^{\alpha+1}}{\alpha+1}+C \quad (a\neq 1).$

2°. $\int \frac{dx}{x}=\log_e x+C \quad (x>0).$

3°. $\int a^x dx=\frac{a^x}{\log_e a}+C \quad (0<a\neq 1).$

В частности, при $a=e$

$$\int e^x dx=e^x+C$$

4°. $\int \sin x dx=-\cos x+C.$

5°. $\int \cos x dx=\sin x+C.$

6^0. $\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C$ $\left(-\frac{\pi}{2} + \pi n < x < \frac{\pi}{2} + \pi n\right.$, где $n = 0, \pm 1,$ $\left.\pm 2, \ldots\right)$.

7^0. $\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C$ ($\pi n < x < \pi + \pi n$, где $n = 0, \pm 1, \pm 2, \ldots$).

8^0. $\int \frac{dx}{\sqrt{1 - x^2}} = \arcsin x + C$ ($|x| < 1$).

9^0. $\int \frac{dx}{1 + x^2} = \operatorname{arc\,tg} x + C$.

Приведенная таблица в систематическом курсе анализа будет дополнена двумя важнейшими правилами интегрирования (интегрированием посредством замены переменной и интегрированием по частям).

Здесь мы не будем приводить формулировку этих правил, а лишь отметим, что написанная таблица вместе с этими правилами составляет важный вычислительный аппарат той части математического анализа, которую принято называть *интегральным исчислением*.

Следует, однако, сразу же подчеркнуть, что для вычисления многих важных неопределенных интегралов этого аппарата оказывается недостаточно. Например, этого аппарата недостаточно для вычисления неопределенного интеграла

$$\int e^{-\frac{x^2}{2}}\, dx, \tag{1.11}$$

играющего важную роль в теории вероятностей и в других разделах точных наук.

Интеграл (1.11) служит примером интеграла от элементарной функции, не являющегося элементарной функцией. Таким образом, в отличие от операции дифференцирования, операция интегрирования выводит нас из класса элементарных функций. Это обстоятельство подчеркивает условность самого понятия элементарной функции как традиционного объекта классического анализа.

Недостаточность описанного нами аппарата ставит на повестку дня задачу о существовании и о вычислении первообразной и неопределенного интеграла от любой функции $f(x)$, только непрерывной в каждой точке x области своего задания.

Оказывается, такую задачу можно решить при помощи другого подхода к проблеме интегрирования функции, к выяснению которого мы сейчас и перейдем.

8. Снова предположим, что функция $f(x)$ представляет собой мгновенную скорость движущейся вдоль оси Oy материальной точки. Поставим цель — вычислить путь, пройденный этой точкой за промежуток времени от $x = a$ до $x = b$.

Для облегчения рассуждений будем считать, что скорость $f(x)$ неотрицательна для всех значений времени x.

Для решения поставленной задачи разобьем промежуток времени на малые промежутки, ограниченные моментами времени

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b.$$

Естественно считать, что на каждом малом промежутке времени от x_{k-1} до x_k ($k = 1, 2, \ldots, n$) скорость $f(x)$ меняется мало (что заведомо будет иметь место всякий раз, когда $f(x)$ является непрерывной в каждой точке x). Но тогда приближенно можно считать скорость $f(x)$ постоянной на каждом промежутке $[x_{k-1}, x_k]$ и равной значению $f(\xi_k)$, где ξ_k — некоторое значение времени из промежутка $[x_{k-1}, x_k]$.

Таким образом, путь $S[x_{k-1}, x_k]$, пройденный движущейся точкой за промежуток времени от x_{k-1} до x_k, приближенно можно считать равным произведению $f(\xi_k)$ на длину $\Delta x_k = x_k - x_{k-1}$ промежутка $[x_{k-1}, x_k]$. Итак,

$$S[x_{k-1}, x_k] \approx f(\xi_k) \Delta x_k.$$

В таком случае путь $S[a, b]$, пройденный материальной точкой за весь промежуток времени от $x = a$ до $x = b$, будет приближенно равен сумме

$$S[a, b] \approx f(\xi_1)\Delta x_1 + f(\xi_2)\Delta x_2 + \ldots + f(\xi_n)\Delta x_n. \tag{1.12}$$

Сумму, стоящую в правой части (1.12), принято называть **интегральной суммой**.

Естественно ожидать, что точное значение пути $S[a, b]$ мы получим, переходя в интегральной сумме, стоящей в правой части (1.12), к пределу при стремлении к нулю наибольшей из длин Δx_k (при этом, конечно, общее число n частичных промежутков будет неограниченно возрастать).

Используя символ предела и обозначая через d наибольшее из чисел $\Delta x_1, \Delta x_2, \ldots, \Delta x_n$, получим, что

$$S[a, b] = \lim_{d \to 0} \{f(\xi_1)\Delta x_1 + f(\xi_2)\Delta x_2 + \ldots + f(\xi_n)\Delta x_n\}. \tag{1.13}$$

Разумеется, требует уточнения вопрос о том, что мы понимаем под пределом интегральной суммы, стоящим в правой части (1.13). На этот раз операция предельного перехода встречается в новой и более сложной форме, чем при вычислении обычного предела функции $\lim\limits_{x \to x_0} f(x)$.

Строгое определение и изучение свойств предела вида (1.13) будет дано в систематическом курсе анализа. Здесь же мы укажем, что в математике *предел, стоящий в правой части* (1.13), *называется* ***определенным интегралом*** *от функции $f(x)$ в пределах от a до b и обозначается символом*

$$\int_a^b f(x)\,dx. \tag{1.14}$$

Итак, определенный интеграл (1.14) равен пути $S[a, b]$, пройденному движущейся со скоростью $f(x)$ материальной точкой за промежуток времени от $x=a$ до $x=b$.

Вместе с тем очевидно, что интегральная сумма, стоящая в правой части (1.12), геометрически представляет собой сумму площадей прямоугольников, основаниями которых служат отрезки Δx_k, а высотами — отрезки длины $f(\xi_k)$.

Иными словами, интегральная сумма, стоящая в правой части (1.12), равна площади ступенчатой фигуры, обведенной на

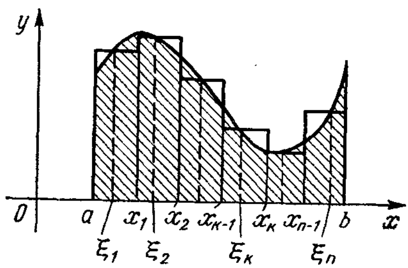


Рис. 1.2

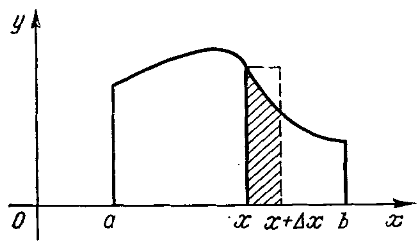


Рис. 1.3

рис. 1.2 жирной линией. Естественно ожидать, что при стремлении к нулю длины d наибольшего из чисел Δx_k площадь указанной ступенчатой фигуры будет стремиться к площади криволинейной фигуры, лежащей под графиком функции $y=f(x)$ на отрезке $a \leqslant x \leqslant b$ (на рис. 1.2 эта фигура заштрихована). Эту фигуру принято называть **криволинейной трапецией**.

Таким образом, определенный интеграл (1.14) равен площади указанной криволинейной трапеции.

Конечно, проведенные нами наглядные рассуждения требуют уточнения. В частности, в систематическом курсе анализа надлежит уточнить само понятие площади криволинейной трапеции и вообще площади плоской фигуры.

Итак, с понятием определенного интеграла (1.14) связаны две фундаментальные задачи: физическая задача о вычислении пути, пройденного движущейся со скоростью $f(x)$ материальной точкой за промежуток времени от $x=a$ до $x=b$, и геометрическая задача о вычислении площади криволинейиой трапеции.

9. Теперь настало время заняться вопросом о связи определенного интеграла (1.14) с введенным ранее неопределенным интегралом (или с первообразной), а также вопросом о способах вычисления определенного интеграла.

Обозначим через $F(x)$ определенный интеграл от функции $f(x)$ в пределах от a до x, где a — некоторое фиксированное значение аргумента, а x — переменное значение. Иными словами, положим *

$$F(x)=\int_a^x f(t)\,dt. \tag{1.15}$$

С геометрической точки зрения этот интеграл, как это показывает проведенное выше рассмотрение, равен площади криволинейной трапеции, лежащей под графиком функции $y=f(x)$ на сегменте $[a, x]$. На рис. 1.3 эта криволинейная трапеция обведена жирной линией.

Используя наглядные геометрические соображения, убедимся в том, что введенная нами функция (1.15) является одной из первообразных функции $f(x)$, т. е. убедимся в том, что $F'(x)=f(x)$.

Пусть Δx — некоторое достаточно малое приращение аргумента x. Очевидно, разность $F(x+\Delta x)-F(x)$ представляет собой площадь «узкой» криволинейной трапеции, заштрихованной на рис. 1.3. С другой стороны, если функция $f(x)$ непрерывна в каждой точке x, т. е. если значение этой функции при малом изменении аргумента меняется мало, то указанная площадь «узкой» криволинейной трапеции мало отличается от площади $f(x)\,\Delta x$ прямоугольника с основанием Δx и высотой $f(x)$.

Отсюда следует, что при малом Δx разностное отношение

$$\frac{F(x+\Delta x)-F(x)}{\Delta x} \tag{1.16}$$

мало отличается от высоты $f(x)$ указанного прямоугольника, т. е. предел при $\Delta x\to 0$ разностного отношения (1.16) обязан быть равен $f(x)$. Вместе с тем, по определению, указанный предел равен производной $F'(x)$.

Итак, мы убедились в том, что $F'(x)=f(x)$, т. е. функция (1.15) является одной из первообразных функции $f(x)$. Но тогда любая первообразная функции $f(x)$ равна

$$\Phi(x)=\int_a^x f(t)\,dt+C, \tag{1.17}$$

где C — постоянная.

Проведенные нами рассуждения имеют предварительный характер, но при наличии развитого аппарата математического ана-

* Переменную под знаком определенного интеграла мы обозначаем через t, чтобы не путать ее с верхним пределом интегрирования x.

лиза им легко придать строгий характер и строго доказать, что у любой функции $f(x)$, только непрерывной в каждой точке x, существует первообразная (а значит, и неопределенный интеграл), причем любая первообразная этой функции определяется равенством (1.17).

Равенство (1.17), в свою очередь, позволяет установить связь между определенным интегралом $\int\limits_a^b f(x)\,dx$ и любой первообразной $\Phi(x)$ функции $f(x)$.

Для установления такой связи возьмем в равенстве (1.17) в качестве верхнего предела интегрирования сначала число b, а затем число a. При этом получим

$$\Phi(b)=\int\limits_a^b f(t)\,dt+C=\int\limits_a^b f(x)\,dx+C, \tag{1.18}$$

$$\Phi(a)=\int\limits_a^a f(t)\,dt+C=C \tag{1.19}$$

(ибо интеграл $\int\limits_a^a f(t)\,dt$, очевидно, равен нулю).

Вычитая из равенства (1.18) равенство (1.19), мы получим знаменитую *формулу Ньютона—Лейбница*

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

сводящую вопрос о вычислении определенного интеграла $\int\limits_a^b f(x)\,dx$ к вычислению разности значений любой первообразной $\Phi(x)$ функции $f(x)$ в точках b и a.

Строгое обоснование формулы Ньютона — Лейбница является одной из важных задач математического анализа.

10. Заметим, однако, что точное аналитическое выражение для первообразной можно получить лишь для узкого класса функций. Поэтому наличие формулы Ньютона — Лейбница не снимает вопроса о приближенных способах вычисления определенного интеграла.

Простейший способ приближенного вычисления определенного интеграла (так называемый *метод прямоугольников*) основан на замене вычисляемого интеграла $\int\limits_a^b f(x)\,dx$ интегральной суммой, стоящей в правой части (1.12), у которой все точки

ξ_k являются серединами соответствующих сегментов $[x_{k-1}, x_k]$, а длины всех указанных сегментов, т. е. все числа $\Delta x_k = x_k - x_{k-1}$, равны друг другу.

В систематическом курсе анализа будет доказано, что при определенных требованиях на функцию $f(x)$ ошибка, совершаемая при замене интеграла $\int\limits_a^b f(x)\,dx$ указанной специальной интегральной суммой, имеет порядок n^{-3}, где n — число частичных сегментов.

Замечательным является то обстоятельство, что метод прямоугольников (как и многие другие методы приближенного вычисления определенного интеграла) допускает удобную реализацию на ЭВМ. Это обстоятельство и равенство (1.17) делают эти методы эффективным средством вычисления первообразных и неопределенных интегралов.

Ниже мы приводим результат вычисления на ЭВМ по методу прямоугольников так называемого **интеграла Пуассона**

$$F(x) = \int\limits_0^x e^{-\frac{t^2}{2}}\,dt$$

для значения x из сегмента $0 \leqslant x \leqslant 1$.

Результаты вычислений собраны нами в табл. 1, в которой в первой колонке стоит аргумент x интеграла Пуассона, во второй колонке указана длина h частичного сегмента (или шага), в третьей колонке приведен результат вычисления, а в четвертой колонке указано число n частичных сегментов *.

Таким образом, для интеграла Пуассона, не являющегося, как указано выше, элементарной функцией, с помощью ЭВМ и простейших приближенных методов без труда могут быть составлены таблицы его значений, делающие использование этого интеграла столь же доступным, как и использование любой элементарной функции.

11. Наряду с приближенными методами вычисления интегралов важную роль в современной математике играют приближенные методы отыскания корней различных уравнений.

Рассмотрим простейшее уравнение

$$f(x) = 0. \tag{1.20}$$

В систематическом курсе анализа будет доказано, что при опре-

* При этом следует учитывать ошибки округления, возникающие при переводе чисел, взятых в десятичной системе счисления, в двоичную систему ЭВМ и обратном переводе чисел, полученных в двоичной системе счисления, в десятичную. Вследствие указанных ошибок округления из десяти выписанных после запятой десятичных знаков можно гарантировать правильность первых шести знаков (остальные четыре знака в табл. 1 взяты в скобки).

Таблица 1

x	h	$F(x)$	n
0,0999999642	0,0099999964	0,039827 (9885)	10
0,1999999285	0,0099999964	0,079260 (0074)	20
0,2999998927	0,0099999964	0,117911 (8581)	30
0,3999998569	0,0099999964	0,155422 (3028)	40
0,4999998212	0,0099999964	0,191463 (1319)	50
0,5999997854	0,0099999964	0,225747 (6439)	60
0,6999997497	0,0099999964	0,258037 (1805)	70
0,7999997139	0,0099999964	0,288145 (4843)	80
0,8999996781	0,0099999964	0,315940 (7870)	90
0,9999996424	0,0099999964	0,341345 (6679)	100
0,0999999642	0,0009999996	0,039827 (8248)	100
0,1999999285	0,0009999996	0,079259 (6848)	200
0,2999998927	0,0009999996	0,117911 (3861)	300
0,3999998569	0,0009999996	0,155421 (6952)	400
0,4999998212	0,0009999996	0,191462 (4058)	500
0,5999997854	0,0009999996	0,225746 (8192)	600
0,6999997497	0,0009999996	0,258036 (2789)	700
0,7999997139	0,0009999996	0,288144 (5284)	800
0,8999996781	0,0009999996	0,315939 (7992)	900
0,9999996424	0,0009999996	0,341344 (6698)	1000

деленных требованиях, налагаемых на функцию $f(x)$, корень $x=c$ уравнения (1.20) может быть найден как предел последовательности итераций x_n ($n=1, 2, 3, \ldots$), первая из которых x_1 берется из некоторого достаточно широкого диапазона, а все последующие шаг за шагом определяются по формуле

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n = 1, 2, 3, \ldots). \tag{1.21}$$

Указанный метод приближенного вычисления корня уравнения (1.20) называется **методом Ньютона** (или **методом касательных**). Этот метод допускает очень удобную реализацию на ЭВМ.

В качестве конкретного примера рассмотрим уравнение (1.20) с функцией $f(x)$ вида $f(x)=x^k-a$, где a — положительное вещественное число, а $k\geqslant 2$ — целое положительное число. Для такой функции $f(x)$ положительным корнем уравнения (1.20) будет являться число $\sqrt[k]{a}$ (т. е. корень степени k из положительного вещественного числа a).

Формула (1.21), определяющая последовательные приближения метода Ньютона, на этот раз принимает вид

$$x_{n+1} = \frac{k-1}{k} x_n + \frac{a}{k x_n^{k-1}} \quad (n = 1, 2, 3, \ldots). \tag{1.22}$$

(Чтобы убедиться в этом, достаточно учесть, что $f'(x)=kx^{k-1}$.)

Формула (1.22) представляет собой эффективный легко реализуемый на ЭВМ алгоритм вычисления корня степени k из положительного числа a.

Приведем пример вычислений, проведенных на ЭВМ по этой формуле.

Всякое положительное вещественное число a представимо (и притом единственным способом) в виде $a=2^l x$, где l — целое число, а x удовлетворяет неравенствам $1/2 \leqslant x < 1$. Будем каждый раз выбирать за первое приближение x_1 число $x_1 = 2^{[l/k]}$, где k — степень извлекаемого корня, а символ $[l/k]$ обозначает целую часть числа l/k.

Результаты вычислений собраны нами в приводимую ниже табл. 2, в которой в первой колонке стоят числа a, из которых извлекается корень, во второй колонке указаны степени k извлекаемых корней, в третьей колонке приведен результат вычислений, а в четвертой колонке указано число сделанных итераций.

Таблица 2

a	k	$\sqrt[k]{a}$	n
2	2	1,41423181	4
3	2	1,732049942	5
4	2	1,999999046	5
2	5	1,148697853	5
3	5	1,245730400	5
4	5	1,319507599	6
2	10	1,071773529	5
3	10	1,116123199	6
4	10	1,148697853	6

12. Мы рассмотрели постановку важнейших задач математического анализа, отправляясь от простейшей механической модели — движения материальной точки вдоль прямой линии. Такая модель естественно привела нас к необходимости построения дифференциального и интегрального исчисления функции $f(x)$ одной независимой переменной.

При описании более сложных задач естествознания возникает понятие функции нескольких независимых переменных x_1, $x_2, \ldots, x_m$. Так, например, температура u нагреваемого тела представляет собой функцию четырех независимых переменных: трех координат x_1, x_2, x_3 точки этого тела и времени t. Эту функцию естественно обозначить символом $u=f(x_1, x_2, x_3, t)$.

Для функции нескольких независимых переменных естественно ввести понятие производной по каждой из переменных (такую производную называют частной производной по данной переменной).

Важной задачей для последующего развития математического анализа является построение дифференциального и интегрального исчислений функций нескольких переменных.

Наконец, математический анализ, понимаемый в совсем широком смысле, включает в себя теорию так называемых дифференциальных уравнений (т. е. уравнений, содержащих искомые функции под знаками производных).

В последние десятилетия широкое развитие получили теории, исходящие из обобщенной трактовки самого понятия функции, понятия производной и понятия решения дифференциального уравнения, связывающего производные функции.

Создание математического анализа является одним из величайших достижений человеческого разума. Оно позволило от рассмотрения отдельных разрозненных физических и геометрических задач (таких, как падение тела под действием силы тяжести, вычисление площади, лежащей под параболой) перейти к развитию общих методов решения больших классов задач. Развитие математического анализа, в свою очередь, оказало огромное влияние на прогресс науки и техники.

Классический математический анализ представляет собой очень удобную идеализированную модель, основанную на том, что мы располагаем точными значениями всех исходных величин и можем найти точные значения всех вычисляемых величин.

Заметим вместе с тем, что, отправляясь от этой модели, мы, как правило, можем оценить погрешность, возникающую вследствие того, что исходные величины заданы нам с некоторой ошибкой и все вычисления могут быть проведены лишь с определенной точностью.

Таким образом, аппарат математического анализа может быть использован для построения численных методов и оценки погрешностей.

Подводя итог, систематизируем первоочередные и наиболее важные проблемы, выявившиеся в результате проведенного нами предварительного рассмотрения.

1. Уточнение понятий вещественного числа, множества и функции.

2. Развитие теории пределов и связанного с этой теорией понятия непрерывности функции.

3. Построение аппарата дифференциального и интегрального исчислений.

4. Построение теории определенного интеграла как предела сумм специального вида.

5. Развитие приближенных методов вычисления определенных интегралов и приближенных методов решения уравнений.

6. Выяснение некоторых геометрических понятий (таких, как площадь плоской фигуры, длина дуги).

Глава 2

ВЕЩЕСТВЕННЫЕ ЧИСЛА

В предыдущей главе мы убедились в том, что развитие теории вещественных * чисел необходимо для строгого и последовательного изучения понятия предела, являющегося одним из важнейших понятий математического анализа.

Необходимая нам теория вещественных чисел, излагаемая в этой главе, включает в себя определение операций упорядочения сложения и умножения этих чисел и установление основных свойств указанных операций, а также доказательство существования точных граней у множеств чисел, ограниченных сверху или снизу.

В конце главы дается представление о дополнительных вопросах теории вещественных чисел, не являющихся необходимыми для построения теории пределов и вообще курса математического анализа (полнота множества вещественных чисел в смысле Гильберта, аксиоматическое построение теории вещественных чисел, связь между различными способами введения вещественных чисел).

Самый последний параграф главы посвящен элементарным вопросам теории множеств, близко примыкающих к теории вещественных чисел.

§ 1. МНОЖЕСТВО ЧИСЕЛ, ПРЕДСТАВИМЫХ БЕСКОНЕЧНЫМИ ДЕСЯТИЧНЫМИ ДРОБЯМИ, И ЕГО УПОРЯДОЧЕНИЕ

1. Свойства рациональных чисел. Понятие рационального числа и основные свойства рациональных чисел известны из курса средней школы. В настоящем пункте мы даем систематизацию хорошо известных из курса средней школы вопросов теории рациональных чисел.

Рациональным называется число, представимое (хотя бы одним способом) в виде отношения двух целых чисел, т. е. в виде дроби m/n, где m и n — целые числа и $n \neq 0$.

---

* Вместо термина «вещественное число» часто употребляют термин «действительное число».

Рациональные числа обладают следующими 16 основными свойствами *. (При формулировке этих свойств мы вместо термина «рациональное число» употребляем более краткий термин «число».)

1°. Любые два числа a и b связаны одним и только одним из трех знаков $>$, $<$ или $=$, причем если $a>b$, то $b<a$. Иными словами, существует правило, позволяющее установить, каким из указанных трех знаков связаны два данных числа. Это правило называется **правилом упорядочения** **.

2°. Существует правило, посредством которого любым числам a и b ставится в соответствие третье число c, называемое их **суммой** и обозначаемое символом $c=a+b$ ***. Операция нахождения суммы называется **сложением**.

3°. Существует правило, посредством которого любым числам a и b ставится в соответствие третье число c, называемое их **произведением** и обозначаемое символом $c=a\cdot b$ ****. Операция нахождения произведения называется **умножением**.

Правило упорядочения обладает следующим свойством:

4°. Из $a>b$ и $b>c$ вытекает, что $a>c$ (свойство транзитивности знака $>$); из $a=b$ и $b=c$ вытекает, что $a=c$ (свойство транзитивности знака $=$).

Операция сложения обладает следующими четырьмя свойствами:

5°. $a+b=b+a$ (коммутативность или перестановочное свойство).

6°. $(a+b)+c=a+(b+c)$ (ассоциативность или сочетательное свойство).

7°. Существует число 0 такое, что $a+0=a$ для любого числа a (особая роль нуля).

8°. Для каждого числа a существует противоположное ему число a' такое, что $a+a'=0$.

Аналогичными четырьмя свойствами обладает операция умножения:

---

* Все приводимые нами свойства рациональных чисел могут быть получены из свойств целых чисел.

** Правило упорядочения рациональных чисел формулируется так: два неотрицательных числа $a=m_1/n_1$ и $b=m_2/n_2$, у которых $n_1>0$ и $n_2>0$, связаны тем же знаком, что и два целых числа $m_1\cdot n_2$ и $m_2\cdot n_1$; два неположительных числа a и b связаны тем же знаком, что и два неотрицательных числа $|b|$ и $|a|$; если a неотрицательно, а b отрицательно, то $a>b$.

*** Правило образования суммы двух рациональных чисел $a=m_1/n_1$ и $b=m_2/n_2$ определяется равенством $\frac{m_1}{n_1}+\frac{m_2}{n_2}=\frac{m_1\cdot n_2+m_2\cdot n_1}{n_1\cdot n_2}$, которое получается с помощью известного приема приведения к общему знаменателю.

**** Правило образования произведения двух рациональных чисел $a=m_1/n_1$ и $b=m_2/n_2$ определяется равенством $\frac{m_1}{n_1}\cdot\frac{m_2}{n_2}=\frac{m_1\cdot m_2}{n_1\cdot n_2}$.

9°. $a \cdot b = b \cdot a$ (коммутативность).

10°. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (ассоциативность).

11°. Существует число 1 такое, что $a \cdot 1 = a$ для любого числа a (особая роль единицы).

12°. Для каждого числа $a \neq 0$ существует обратное ему число a' такое, что $a \cdot a' = 1$.

Операции сложения и умножения связаны следующим свойством:

13°. $(a+b) \cdot c = a \cdot c + b \cdot c$ (дистрибутивность или распределительное свойство умножения относительно суммы).

Следующие два свойства связывают операцию упорядочения с операцией сложения или соответственно умножения:

14°. Из $a > b$ вытекает, что $a + c > b + c$.

15°. Из $a > b$ и $c > 0$ вытекает, что $a \cdot c > b \cdot c$.

Особая роль принадлежит последнему свойству.

16°. Каково бы ни было число a, можно число 1 повторить слагаемым столь раз, что сумма превзойдет a *.

Перечисленные 16 свойств называют **основными** потому, что все другие алгебраические свойства, относящиеся к операциям сложения и умножения и к сочетанию равенств и неравенств, могут быть извлечены как логические следствия из указанных 16 свойств.

Так, например, из основных свойств вытекает следующее часто используемое в дальнейшем свойство, позволяющее почленно складывать неравенства одного знака:

$$\text{если } a > b \text{ и } c > d, \text{ то } a + c > b + d.$$

В самом деле, из неравенств $a > b$ и $c > d$ и из свойств 14° и 5° вытекает, что $a + c > b + c$ и $b + c > b + d$, а из двух последних неравенств и свойства 4° вытекает, что $a + c > b + d$.

2. Недостаточность рациональных чисел для измерения отрезков числовой оси. *Договоримся называть числовой осью прямую, на которой выбраны определенная точка O (начало отсчета), масштабный отрезок OE, длину которого мы принимаем равной единице, и положительное направление (обычно от O к E).* Очевидно, каждому рациональному числу соответствует на числовой оси определенная точка. В самом деле, из курса средней школы известно, как построить отрезок, длина которого составляет $1/n$ часть длины масштабного отрезка OE (n — любое целое положительное число). Следовательно, мы можем построить и отрезок, длина которого относится к длине масштабного отрезка как m/n, где m и n — любые целые положительные числа. Отложив такой отрезок вправо (влево) от точки O, мы получим точку M_1 (M_2), соответствующую рациональному числу m/n ($-m/n$) (рис. 2.1).

* Это свойство часто называют аксиомой Архимеда.

Заметим теперь, что не каждой точке M числовой оси соответствует рациональное число. Так, например, если точка M выбрана так, что длина отрезка OM равна диагонали квадрата, стороной которого служит масштабный отрезок OE, то поскольку длина масштабного отрезка OE равна единице, по теореме Пифагора длина x отрезка OM является корнем уравнения $x^2=2$ и, как показано в курсах средней школы, не является рациональным числом. Но это и означает, что указанной точке M не соответствует рациональное число.

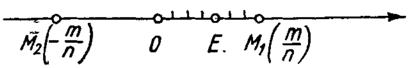


Рис. 2.1

Рис. 2.2

Естественно, возникает потребность расширить множество рациональных чисел и ввести в рассмотрение более широкое множество чисел так, чтобы каждой точке числовой оси соответствовало некоторое число из этого более широкого множества (или, что то же самое, чтобы с помощью этого более широкого множества чисел можно было выразить длину любого отрезка OM числовой оси).

Убедимся в том, что посредством измерения отрезка OM каждой точке M числовой оси можно поставить в соответствие вполне определенную бесконечную десятичную дробь.

Пусть M — любая точка числовой оси. Ради определенности предположим, что точка M лежит направо от O. Проведем процесс измерения отрезка OM при помощи масштабного отрезка OE.

Сначала выясним, сколько раз целый масштабный отрезок уложится в отрезке OM *. Могут представиться два случая:

1) Отрезок OE укладывается в отрезке OM целое число a_0 раз с некоторым остатком NM, меньшим OE (рис. 2.2). В этом случае целое число a_0 представляет собой результат измерения по недостатку с точностью до числа 1.

2) Отрезок OE укладывается в отрезке OM целое число a_0 раз **без остатка**. В этом случае процесс измерения можно считать законченным и целое рациональное число a_0 считать длиной отрезка OM. Формально мы можем утверждать, что в этом случае точке M соответствует бесконечная десятичная дробь $a_0{,}000\ldots$, которая отождествляется с целым рациональным числом a_0.

* В силу аксиомы Архимеда для отрезка, каковы бы ни были два отрезка AB и CD, повторив один из этих отрезков слагаемым достаточно большое число раз, мы получим отрезок, длина которого превосходит длину второго отрезка.

В первом случае процесс измерения следует продолжить и выяснить, сколько раз 1/10 часть масштабного отрезка OE укладывается в отрезке NM (являющемся остатком измерения с помощью целого отрезка OE). Снова могут представиться два случая:

1) 1/10 часть OE укладывается в отрезке NM a_1 раз с некоторым остатком PM, меньшим 1/10 части OE (см. рис. 2.2). В этом случае рациональное число a_0, a_1 представляет собой результат измерения OM по недостатку с точностью до числа 1/10.

2) 1/10 часть OE укладывается в отрезке NM целое число a_1 раз без остатка. В этом случае процесс измерения можно считать законченным и рациональное число a_0, a_1 считать длиной отрезка OM. Формально мы можем утверждать, что в этом случае точке M соответствует бесконечная десятичная дробь $a_0,a_1000...$, отождествляемая с рациональным число a_0,a_1.

Продолжая указанные рассуждения далее, мы придем к двум возможностям:

1) либо описанный процесс измерения оборвется на n-м шаге вследствие того, что точке M соответствует рациональное число $a_0,a_1a_2 \ldots a_n$ (в этом случае точке M соответствует бесконечная десятичная дробь $a_0,a_1a_2 \ldots a_n000 \ldots$, которую мы отождествляем с рациональным числом $a_0,a_1a_2 \ldots a_n$);

2) либо описанный процесс измерения никогда не оборвется и мы получим бесконечную последовательность рациональных чисел

$$a_0;\ a_0, a_1;\ \ldots;\ a_0, a_1a_2 \ldots a_n;\ \ldots, \tag{2.1}$$

представляющих собой результат измерения по недостатку отрезка OM с точностью до $1, \frac{1}{10}, \ldots, \frac{1}{10^n}, \ldots$.

Каждое из чисел последовательности (2.1) может быть получено обрыванием на соответствующем знаке бесконечной десятичной дроби

$$a_0,\ a_1a_2 \ldots a_n \ldots. \tag{2.2}$$

Таким образом, в случае 2) точке M числовой оси отвечает вполне определенная бесконечная десятичная дробь (2.2). Можно сказать, что и в случае 1) точке M отвечает бесконечная десятичная дробь (2.2), но в этом случае у этой дроби все десятичные знаки с номером, большим n, равны нулю, т. е. указанная дробь в случае 1) имеет вид $a_0, a_1a_2 \ldots a_n000 \ldots$.

Приведенные нами рассуждения применимы и для случая, когда точка M лежит левее точки O, только в этом случае естественно считать, что все элементы последовательности (2.1) и бесконечная дробь имеют отрицательный знак.

Итак, мы убедились, что описанный нами процесс измерения позволяет поставить в соответствие каждой точке M числовой оси вполне определенную бесконечную десятичную дробь. Это обстоятельство естественно приводит нас к необходимости рассмотрения чисел, представимых бесконечными десятичными дробями.

Замечание. Конечно, описанный нами процесс измерения отрезка OM можно видоизменить так, что он будет приводить к рассмотрению не бесконечных десятичных, а, например, бесконечных двоичных или бесконечных троичных дробей. Желание рассматривать бесконечные десятичные дроби вызвано лишь той особой ролью, которую традиционно играет десятичная система счисления. Развитие электронной вычислительной техники повысило роль двоичной и троичной систем счисления, ибо (в силу конструктивных особенностей ЭВМ) эти системы счисления более удобны в практике использования ЭВМ.

3. Упорядочение множества бесконечных десятичных дробей. Во вводной главе мы уже отмечали, что понятие числа относится к так называемым начальным понятиям (т. е. к понятиям, которые могут быть разъяснены, но не могут быть строго определены, ибо всякая попытка дать строгое определение такого понятия неизбежно сведется к замене определяемого понятия ему эквивалентным). Мы введем понятие вещественных чисел, отправляясь от множества бесконечных десятичных дробей.

Рассмотрим множество всевозможных бесконечных десятичных дробей (как положительных, т. е. взятых со знаком $+$, так и отрицательных, т. е. взятых со знаком $-$).

Мы будем придерживаться следующего плана.

Для множества всех чисел, представимых бесконечными десятичными дробями, мы введем операцию упорядочения. После этого мы убедимся, что для введенной нами операции упорядочения остается справедливым то же самое свойство 4°, которое сформулировано в п. 1 для рациональных чисел (т. е. свойство транзитивности знаков $>$ и $=$).

Наличие только одного этого свойства позволит нам доказать замечательную теорему о том, что у *множества чисел, представимых бесконечными десятичными дробями и ограниченных сверху (или соответственно снизу), существует число, представимое бесконечной десятичной дробью и являющееся точной верхней (или соответственно точной нижней) гранью указанного множества чисел.*

После этого вводятся операции сложения и умножения чисел, представимых бесконечными десятичными дробями. Это дает нам возможность ввести *вещественные числа как такие числа, которые представимы бесконечными десятичными дробями и для которых указанным нами способом определены операции упорядочения, сложения и умножения.* Доказанная нами теорема о существовании точных граней позволит доказать *существование суммы и произведения* двух любых вещественных чисел, а также справедливость для этих чисел тех же самых *16 основных свойств,* которые сформулированы в п. 1 для рациональных чисел.

Приступим к реализации указанного плана.

В этом пункте мы введем для чисел, представимых бесконечными десятичными дробями, операцию упорядочения и установим, что эта операция обладает свойством 4°, сформулированном в п. 1 для рациональных чисел (т. е. свойством транзитивности знаков $>$ и $=$).

Рассмотрим произвольное число, представимое бесконечной десятичной дробью, отличной от 0,000... . *Это число мы будем называть* ***положительным***, *если оно представимо бесконечной десятичной дробью, взятой со знаком* $+$, *и* ***отрицательным***, *если оно представимо бесконечной десятичной дробью, взятой со знаком* $-$.

Числа, не являющиеся положительными, мы будем называть ***неположительными***, *а числа, не являющиеся отрицательными,—* ***неотрицательными***.

Сразу же отметим, что все рациональные числа относятся к множеству чисел, представимых бесконечными десятичными дробями. Представление данного рационального числа бесконечной десятичной дробью можно получить двумя способами:

1) взяв точку M, отвечающую данному рациональному числу на числовой оси, и произведя измерение отрезка OM с помощью масштабного отрезка способом, указанным в п. 2;

2) взяв обыкновенную дробь m/n, представляющую данное рациональное число, и поделив числитель m на знаменатель n «столбиком» *.

Мы представляем читателю убедиться в том, что оба эти способа эквивалентны друг другу. Так, при любом из указанных способов рациональному числу 1/2 ставится в соответствие бесконечная десятичная дробь 0,5000..., рациональному числу 4/3 — бесконечная десятичная дробь 1,333... .

Прежде чем перейти к формулировке правила упорядочения чисел, представимых бесконечными десятичными дробями, рассмотрим вопрос о представлении в виде бесконечных десятичных дробей тех рациональных чисел, которые представимы в виде конечной десятичной дроби.

Заметим, что такие рациональные числа допускают **два представления** в виде бесконечных десятичных дробей. Например, рациональное число $1/2=0{,}5$ можно представить в виде двух бесконечных десятичных дробей:

$$1)\ 1/2=0{,}5000\ldots,\quad 2)\ 1/2=0{,}4999\ldots .$$

Вообще, рациональное число $a=a_0, a_1a_2\ldots a_n$, где $a_n\neq 0$, можно записать в виде двух бесконечных десятичных дробей:

$$1)\ a=a_0, a_1a_2\ldots a_n000\ldots,\qquad 2)\ a=a_0, a_1a_2\ldots a_{n-1}(a_n-1)999\ldots .$$

* В курсе средней школы доказывается, что при таком делении получается обязательно **периодическая** бесконечная десятичная дробь.

Естественно, мы должны отождествить указанные две бесконечные десятичные дроби (т. е. считать, что они представляют одно и то же вещественное число).

Рассмотрим теперь два произвольных вещественных числа a и b и предположим, что эти числа представляются бесконечными десятичными дробями

$$a = \pm a_0, a_1 a_2 \dots a_n \dots, \quad b = \pm b_0, b_1 b_2 \dots b_n \dots, \tag{2.3}$$

где из двух знаков $+$ и $-$ в каждом представлении берется какой-то один.

Исключим уже рассмотренный выше случай, когда обе бесконечные десятичные дроби в (2.3) имеют одинаковые знаки и служат двумя различными представлениями одного и того же рационального числа, представимого конечной десятичной дробью. После исключения этого случая договоримся называть два числа a и b равными, если их представления в виде бесконечных десятичных дробей (2.3) имеют одинаковые знаки и если справедлива бесконечная цепочка равенств

$$a_0 = b_0, \quad a_1 = b_1, \quad a_2 = b_2, \dots, a_n = b_n, \dots . \tag{2.4}$$

Итак, *мы называем два числа a и b равными, если их представления в виде бесконечных десятичных дробей* (2.3) *имеют одинаковые знаки и если либо справедлива цепочка равенств* (2.4), *либо бесконечные десятичные дроби в* (2.3) *служат двумя представлениями одного и того же рационального числа, представимого конечной десятичной дробью.*

Пусть даны два неравных числа a и b, представимых бесконечными десятичными дробями. Установим правило, позволяющее заключить, каким из двух знаков, $>$ или $<$, связаны эти числа.

Договоримся называть **модулем** *числа a, представимого бесконечной десятичной дробью, число, представимое той же самой бесконечной десятичной дробью, что и число a, но всегда взятой со знаком $+$.*

Модуль числа a будем обозначать символом $|a|$. Число $|a|$ всегда является неотрицательным.

Рассмотрим отдельно три возможных случая: 1) случай, когда a и b оба неотрицательны; 2) случай, когда оба числа a и b отрицательны; 3) случай, когда одно из чисел a и b неотрицательно, а другое отрицательно.

1) Пусть сначала a и b **оба неотрицательны** и имеют представления $a = a_0, a_1 a_2 \dots$; $b = b_0, b_1 b_2 \dots$. Так как числа a и b не являются равными, то нарушается хотя бы одно из равенств (2.4).

Обозначим через k наименьший из номеров n, для которого нарушается равенство $a_n = b_n$, т. е. предположим, что

$$a_0 = b_0, \; a_1 = b_1, \dots, a_{k-1} = b_{k-1}, \; a_k \neq b_k.$$

Тогда мы будем считать, что $a>b$, если $a_k>b_k$, и будем считать, что $a<b$, если $a_k<b_k$.

2) Пусть теперь оба числа a и b отрицательны. Тогда мы будем считать, что $a>b$, если $|b|>|a|$, и $a<b$, если $|b|<<|a|$ *.

3) Пусть, наконец, одно число (например, a) неотрицательно, а другое число (b) отрицательно. Тогда, естественно, мы будем считать, что $a>b$.

Итак, мы полностью сформулировали правило упорядочения чисел, представимых бесконечными десятичными дробями.

Чтобы сделать сформулированное правило безупречным с логической точки зрения (или, как говорят в математике, корректным), докажем следующую лемму.

Лемма. *Если* $a=a_0, a_1a_2\ldots a_n\ldots$ — *произвольное неотрицательное число, а* $b'=b_0, b_1b_2\ldots b_{n-1}b_n000\ldots$ и $b''=b_0, b_1b_2\ldots$ $\ldots b_{n-1}(b_n-1)999\ldots$ *при* $b_n>0$ — *два различных представления одного и того же рационального числа* $b_0, b_1b_2\ldots b_n$, *то условие* $a<b'$ *эквивалентно условию* $a<b''$, *а условие* $a>b'$ *эквивалентно условию* $a>b''$.

Эта лемма позволяет при упорядочении двух неравных чисел не заботиться о том, какое из двух возможных представлений в виде бесконечной десятичной дроби взято для числа, представимого конечной десятичной дробью.

Доказательство. Для полного доказательства леммы следует доказать четыре утверждения: 1) из $a<b'$ вытекает $a<b''$; 2) из $a<b''$ вытекает $a<b'$; 3) из $a>b'$ вытекает $a>b''$; 4) из $a>b''$ вытекает $a>b'$.

Мы ограничимся доказательством утверждений 1) и 2), ибо утверждения 3) и 4) доказываются аналогично.

Пусть $a<b'$. Тогда по правилу упорядочения найдется номер k такой, что

$$a_0=b_0,\ a_1=b_1, \ldots, a_{k-1}=b_{k-1},\ a_k<b_k \tag{2.5}$$

(в этих соотношениях следует считать все b_{n+1}, $b_{n+2}, \ldots$ равными нулю).

Сразу же заметим, что $k\leqslant n$, ибо при $k>n$ неравенство $a_k<b_k$ не может выполняться, так как $0\leqslant a_k\leqslant 9$, а $b_k=0$.

Если при этом $k<n$, то, поскольку при $k<n$ все десятичные знаки до порядка k у b' и b'' совпадают, условия $a<b'$ и $a<b''$, очевидно, эквивалентны.

Остается рассмотреть случай $k=n$. В этом случае соотношения (2.5) принимают вид $a_0=b_0$, $a_1=b_1, \ldots, a_{n-1}=b_{n-1}$, $a_n<b_n$. Самое последнее неравенство эквивалентно неравенству $a_n\leqslant b_n-1$. Если при этом $a_n<b_n-1$, то по правилу упорядочения $a<b''$.

* При этом мы учитываем, что для неотрицательных a и b правило упорядочения уже определено (см. случай 1)).

Если же в указанном последнем неравенстве $a_n = b_n - 1$, то все десятичные знаки у чисел a и b'' до порядка n совпадают. Поскольку у числа b'' все десятичные знаки порядка, бо́льшего n, равны девяти, то и в этом случае $a < b''$, ибо у числа a все десятичные знаки порядка, бо́льшего n, не могут быть равны девяти (в силу того, что a не равно b').

Итак, утверждение 1) доказано.

Перейдем к доказательству утверждения 2). Предположим, что $a < b''$. Договоримся о следующих обозначениях бесконечных десятичных дробей, представляющих числа b' и b''.

$$b' = b_0', b_1'b_2' \dots b_n' \dots, \quad b'' = b_0'', b_1''b_2'' \dots b_n'' \dots .$$

В этих представлениях

$$b_0' = b_0'' = b_0,\ b_1' = b_1'' = b_1, \dots, b'_{n-1} = b''_{n-1} = b_{n-1},\ b_n' = b_n,\ b_n'' = b_n - 1.$$

Иными словами, справедлива цепочка соотношений

$$b_0' = b_0'',\ b_1' = b_1'', \dots, b'_{n-1} = b''_{n-1},\quad b_n' > b_n''.$$

С другой стороны, поскольку $a < b''$, найдется номер k такой, что справедлива цепочка соотношений

$$a_0 = b_0'',\ a_1 = b_1'', \dots, a_{k-1} = b''_{k-1},\ a_k < b_k''.$$

Обозначим через m н а и м е н ь ш и й из двух номеров n и k и сопоставим между собой две последние цепочки соотношений. Используя свойства транзитивности знаков $>$ и $=$ для целых чисел, мы получим при этом следующую цепочку соотношений:

$$a_0 = b_0',\ a_1 = b_1', \dots, a_{m-1} = b'_{m-1},\ a_m < b'_m.$$

Полученные соотношения на основании правила упорядочения вещественных чисел устанавливают справедливость неравенства $a < b'$. Тем самым утверждение 2) также доказано.

Еще раз подчеркнем, что доказанная лемма позволяет при упорядочении двух чисел, представимых бесконечными десятичными дробями, пользоваться любым из двух представлений в виде бесконечной десятичной дроби для рациональных чисел, представимых конечной десятичной дробью.

Легко убедиться в том, что сформулированное правило упорядочения в применении к двум рациональным числам, представленным в виде бесконечных десятичных дробей, приводит к тому же результату, что и прежнее правило упорядочения рациональных чисел, представленных в виде отношения двух целых чисел.

В самом деле, достаточно рассмотреть случай двух н е о т р и ц а т е л ь н ы х рациональных чисел a и b. Пусть $a > b$ согласно прежнему правилу упорядочения рациональных чисел, и пусть $a = a_0, a_1a_2 \dots a_n \dots$; $b = b_0, b_1b_2 \dots b_k \dots$. Отложив рациональные числа a и b на числовой оси, мы получим отвечающие им точки M_1 и M_2, причем, поскольку $a>b$, отрезок OM_1 больше отрезка OM_2. Из описанного в п. 2 процесса измерения отрезка числовой оси вытекает, что целое число $a_0a_1a_2\ldots a_k$ показывает, сколько раз 10^{-k} часть масштабного отрезка OE укладывается в отрезке OM_1, а целое число $b_0b_1b_2\ldots b_n$ показывает, сколько раз 10^{-k} часть OE укладывается в отрезке OM_2. Поскольку отрезок OM_1 больше отрезка OM_2, то найдется номер k такой, что $a_0a_1\ldots a_{k-1}=b_0b_1\ldots b_{k-1}$, а $a_0a_1\ldots a_k>b_0b_1\ldots b_k$, но это и означает, что $a>b$ согласно правилу упорядочения чисел, представимых бесконечными десятичными дробями.

Докажем теперь, что для сформулированного нами правила упорядочения чисел, представимых бесконечными десятичными дробями, остается справедливым свойство 4°, приведенное в п. 1 для рациональных чисел, т. е. докажем, что *для любых трех чисел a, b и c, представимых бесконечными десятичными дробями, из справедливости неравенств $a>b$ и $b>c$ вытекает справедливость неравенства $a>c$ (свойство транзитивности знака $>$), а из справедливости равенств $a=b$ и $b=c$ вытекает справедливость равенства $a=c$ (свойство транзитивности знака $=$).*

Свойство транзитивности знака $=$ сразу же вытекает из справедливости соответствующего свойства для целых чисел.

Докажем свойство транзитивности знака $>$. Пусть $a>b$, $b>c$. Требуется доказать, что $a>c$.

Рассмотрим три возможных случая: 1) c неотрицательно; 2) c отрицательно, a неотрицательно; 3) c отрицательно и a отрицательно.

1) Пусть сначала c неотрицательно. Тогда b также неотрицательно, ибо если бы b было отрицательно, то в силу правила упорядочения мы получили бы, что $c>b$, и это противоречило бы условию $b>c$. Далее, повторяя те же рассуждения, мы получим, что и a неотрицательно (ибо в противном случае мы получили бы, что $b>a$, и это противоречило бы условию $a>b$).

Итак, в рассматриваемом случае все три числа a, b и c неотрицательны. Записав представления этих чисел бесконечными десятичными дробями
$a=a_0,a_1a_2\ldots$; $b=b_0,b_1b_2\ldots$; $c=c_0,c_1c_2\ldots$, мы получим, что в силу условия $a>b$ найдется номер k такой, что

$$a_0=b_0,\quad a_1=b_1,\ldots,a_{k-1}=b_{k-1},\quad a_k>b_k. \tag{2.6}$$

Аналогично в силу условия $b>c$ найдется номер p такой, что

$$b_0=c_0,\quad b_1=c_1,\ldots,b_{p-1}=c_{p-1},\quad b_p>c_p. \tag{2.7}$$

Обозначим через m наименьший из двух номеров k и p. Тогда, очевидно, из соотношений (2.6) и (2.7) и из справедливости свойства транзитивности знаков $>$ и $=$ для целых чисел вытекает, что $a_0=c_0$, $a_1=c_1$, ..., $a_{m-1}=c_{m-1}$, $a_m>c_m$, а это и означает (по правилу упорядочения), что $a>c$.

2) Пусть теперь c отрицательно, a неотрицательно. Тогда (независимое от знака числа b) неравенство $a>c$ справедливо в силу правила упорядочения.

3) Рассмотрим, наконец, случай, когда оба числа a и c отрицательны. Заметим, что в этом случае и b отрицательно (ибо в противном случае мы получили бы из правила упорядочения, что $b>a$, и это противоречило бы условию $a>b$).

Итак, в рассматриваемом случае все три числа a, b и c отрицательны. Но в таком случае (в силу правила упорядочения) неравенства $a>b$, $b>c$ эквивалентны неравенствам $|b|>|a|$ и $|c|>|b|$. Из последних двух неравенств (в силу свойства транзитивности знака $>$, уже доказанного нами в случае 1) для неотрицательных чисел) вытекает, что $|c|>|a|$, а это и означает (в силу правила упорядочения отрицательных чисел a и c), что $a>c$. Тем самым доказательство свойства транзитивности знака $>$ полностью завершено.

§ 2. ОГРАНИЧЕННЫЕ СВЕРХУ (ИЛИ СНИЗУ) МНОЖЕСТВА ЧИСЕЛ, ПРЕДСТАВИМЫХ БЕСКОНЕЧНЫМИ ДЕСЯТИЧНЫМИ ДРОБЯМИ

1. Основные понятия. Рассмотрим совершенно произвольное множество $\{x\}$ чисел, представимых бесконечными десятичными дробями.

Отдельные числа, входящие в состав множества $\{x\}$, мы будем называть элементами этого множества.

Всюду в этом параграфе мы будем требовать, чтобы рассматриваемое множество $\{x\}$ содержало хотя бы один элемент (такое множество принято называть н е п у с т ы м).

Введем важное понятие ограниченности множества сверху (или соответственно снизу).

О п р е д е л е н и е 1. *Множество $\{x\}$ чисел, представимых бесконечными десятичными дробями, называется* ***о г р а н и ч е н н ы м с в е р х у*** *(соответственно* ***о г р а н и ч е н н ы м с н и з у****), если существует такое представимое бесконечной десятичной дробью число M (соответственно такое представимое бесконечной десятичной дробью число m), что каждый элемент x множества $\{x\}$ удовлетворяет неравенству*

$$x \leqslant M \quad (\textit{соответственно } x \geqslant m). \tag{2.8}$$

При этом число M (число m) называется ***в е р х н е й г р а н ь ю*** *(****н и ж н е й г р а н ь ю****) множества $\{x\}$.*

Конечно, любое ограниченное сверху множество $\{x\}$ имеет бесконечно много верхних граней. В самом деле, если число M — одна из верхних граней множества $\{x\}$, то любое число M', бо́льшее числа M, также является верхней гранью множества $\{x\}$ (ибо из справедливости неравенства (2.8) будет следовать, что $x \leqslant M'$). Аналогичное замечание можно сделать в отношении нижних граней ограниченного снизу множества $\{x\}$.

Так, например, множество всех представимых бесконечными десятичными дробями отрицательных чисел ограничено сверху. В качестве верхней грани M такого множества можно взять любое неотрицательное число. Множество всех целых положительных чисел 1, 2, 3, ... ограничено снизу. В качестве нижней грани этого множества можно взять любое число m, удовлетворяющее неравенству $m \leqslant 1$.

Естественно, возникает вопрос о существовании наименьшей из верхних граней ограниченного сверху множества и наибольшей из нижних граней ограниченного снизу множества.

Определение 2. *Наименьшая из всех верхних граней ограниченного сверху множества $\{x\}$ называется* ***точной верхней гранью*** *этого множества и обозначается символом $\bar{x} = \sup\{x\}$* *.

Наибольшая из всех нижних граней ограниченного снизу множества $\{x\}$ называется ***точной нижней гранью*** *этого множества и обозначается символом $\underline{x} = \inf\{x\}$* **.

Определение 2 можно сформулировать и по-другому, а именно:

Число $\bar{x}$ (число $\underline{x}$) называется ***точной верхней (точной нижней)*** *гранью ограниченного сверху (снизу) множества $\{x\}$, если выполнены следующие два требования: 1) каждый элемент x множества $\{x\}$ удовлетворяет неравенству $x \leqslant \bar{x}$ ($x \geqslant \underline{x}$); 2) каково бы ни было число x', меньшее $\bar{x}$ (бо́льшее $\underline{x}$), найдется хотя бы один элемент x множества $\{x\}$, удовлетворяющий неравенству $x > x'$ ($x < x'$).*

В этом определении требование 1) утверждает, что число $\bar{x}$ (число $\underline{x}$) является одной из верхних (нижних) граней, а требование 2) говорит о том, что эта грань является наименьшей (наибольшей) и уменьшена (увеличена) быть не может.

2. Существование точных граней. Существование у любого ограниченного сверху (снизу) множества точной верхней (точной нижней) грани не является очевидным и требует доказательства. Докажем следующую основную теорему.

Основная теорема 2.1. *Если множество $\{x\}$ чисел, представимых бесконечными десятичными дробями, ограничено сверху (соответственно снизу) и содержит хотя бы один элемент, то у этого множества существует точная верхняя (соответственно точная нижняя) грань.*

---

* sup — первые три буквы латинского слова supremum («супремум»), которое переводится как «наивысшее».

** inf — первые три буквы латинского слова infimum («инфимум»), которое переводится как «наинизшее».

Доказательство. Мы остановимся лишь на доказательстве существования точной верхней грани у любого ограниченного сверху множества, ибо существование точной нижней грани у любого ограниченного снизу множества доказывается совершенно аналогично.

Итак, пусть множество $\{x\}$ ограничено сверху, т. е. существует такое число M, что каждый элемент x множества $\{x\}$ удовлетворяет неравенству $x \leqslant M$.

Могут представиться два случая:

1°. Среди элементов множества $\{x\}$ есть хотя бы одно неотрицательное число. 2°. Все элементы множества $\{x\}$ являются отрицательными числами. Эти случаи мы рассмотрим отдельно.

1°. Рассмотрим лишь неотрицательные числа, входящие в состав множества $\{x\}$. Каждое из этих чисел представим в виде бесконечной десятичной дроби и рассмотрим целые части этих десятичных дробей. В силу неравенства $x \leqslant M$ все целые части не превосходят числа M, а поэтому найдется наибольшая из целых частей, которую мы обозначим через $\bar{x}_0$. Сохраним среди неотрицательных чисел множества $\{x\}$ те, у которых целая часть равна $\bar{x}_0$, и отбросим все остальные числа. У сохраненных чисел рассмотрим первые десятичные знаки после запятой. Наибольший из этих знаков обозначим через $\bar{x}_1$. Сохраним среди неотрицательных чисел множества $\{x\}$ те, у которых целая часть равна $\bar{x}_0$, а первый десятичный знак равен $\bar{x}_1$, и отбросим все остальные числа. У сохраненных чисел рассмотрим вторые десятичные знаки после запятой. Наибольший из этих знаков обозначим через $\bar{x}_2$. Продолжая аналогичные рассуждения далее, мы последовательно определим десятичные знаки некоторого числа $\bar{x}$:

$$\bar{x} = \bar{x}_0,\ \bar{x}_1 \bar{x}_2 \ldots \bar{x}_n \ldots . \tag{2.6}$$

Докажем, что это число $\bar{x}$ и является точной верхней гранью множества $\{x\}$. Для этого достаточно доказать два утверждения: 1) каждый элемент x множества $\{x\}$ удовлетворяет неравенству $x \leqslant \bar{x}$; 2) каково бы ни было число x', меньшее $\bar{x}$, найдется хотя бы один элемент x множества $\{x\}$, удовлетворяющий неравенству $x > x'$.

Докажем сначала утверждение 1). Так как $\bar{x}$ по построению является неотрицательным числом, то любой отрицательный элемент x множества $\{x\}$ заведомо удовлетворяет неравенству $x < \bar{x}$.

Поэтому нам достаточно доказать, что любой неотрицательный элемент x множества $\{x\}$ удовлетворяет неравенству $x \leqslant \bar{x}$.

Предположим, что некоторый неотрицательный элемент $x = x_0, x_1 x_2 \ldots x_n \ldots$ не удовлетворяет неравенству $x \leqslant \bar{x}$. Тогда $x > \bar{x}$ и по правилу упорядочения найдется номер k такой, что $x_0 = \bar{x}_0$, $x_1 = \bar{x}_1, \ldots, x_{k-1} = \bar{x}_{k-1}$, $x_k > \bar{x}_k$. Но последние соотношения противоречат тому, что в качестве $\bar{x}_k$ берется наибольший из десятичных знаков x_k тех элементов x, у которых целая часть и первые $k-1$ знаков после запятой соответственно равны $\bar{x}_0, \bar{x}_1, \ldots$ $\ldots, \bar{x}_{k-1}$.

Полученное противоречие доказывает утверждение 1).

Докажем теперь утверждение 2). Пусть x' — любое число, удовлетворяющее условию $x'<\bar{x}$. Требуется доказать, что существует хотя бы один элемент x множества $\{x\}$, удовлетворяющий неравенству $x>x'$.

Если число x' является отрицательным, то неравенству $x>x'$ заведомо удовлетворяет неотрицательный элемент x множества $\{x\}$ (по предположению хотя бы один такой элемент существует).

Остается рассмотреть случай, когда число x', удовлетворяющее условию $x'<\bar{x}$, является неотрицательным. Пусть $x'=x_0', x_1' \ldots$ $\ldots x_n' \ldots$. Из условия $x'<\bar{x}$ и из правила упорядочения вытекает, что найдется номер m такой, что

$$x_0' = \bar{x}_0, \quad x_1' = \bar{x}_1, \quad \ldots, \quad x_{m-1}' = \bar{x}_{m-1}, \quad x_m' < \bar{x}_m. \tag{2.10}$$

С другой стороны, из построения числа (2.9) вытекает, что для любого номера m найдется неотрицательный элемент $x=x_0, x_1 x_2 \ldots$ $\ldots x_n \ldots$ множества $\{x\}$ такой, у которого целая часть и все первые m знаков после запятой те же, что у числа $\bar{x}$. Иными словами, для номера m найдется элемент x такой, для которого

$$x_0 = \bar{x}_0, \quad x_1 = \bar{x}_1, \quad \ldots, \quad x_{m-1} = \bar{x}_{m-1}, \quad x_m = \bar{x}_m. \tag{2.11}$$

Сопоставляя (2.10) и (2.11), мы получим, что

$$x_0 = x_0', \quad x_1 = x_1', \quad \ldots, \quad x_{m-1} = x_{m-1}', \quad x_m > x_m',$$

а это и означает (в силу правила упорядочения), что $x>x'$. Утверждение 2), а с ним и вся теорема для случая 1° доказаны.

2°. Аналогично доказывается существование точной верхней грани и во втором случае, когда все элементы x множества $\{x\}$ являются отрицательными числами.

В этом случае мы представим все элементы x отрицательными бесконечными десятичными дробями и обозначим через $\bar{x}_0$ наименьшую из целых частей этих дробей, через $\bar{x}_1$ — наименьший из первых десятичных знаков тех дробей, целая часть которых равна $\bar{x}_0$, через $\bar{x}_2$ — наименьший из вторых десятичных знаков тех дробей, целая часть и первый десятичный знак которых соответственно равны $\bar{x}_0$ и $\bar{x}_1$ и т. д.

Таким образом мы определим неположительное число $\bar{x} = -\bar{x}_0, \bar{x}_1 \bar{x}_2 \ldots \bar{x}_n \ldots$.

В полной аналогии со случаем 1° доказывается, что это число $\bar{x}$ является точной верхней гранью множества $\{x\}$, т. е. доказывается справедливость утверждений 1) и 2), сформулированных при рассмотрении случая 1°. Теорема доказана.

§ 3. ПРИБЛИЖЕНИЕ ЧИСЕЛ, ПРЕДСТАВИМЫХ БЕСКОНЕЧНЫМИ ДЕСЯТИЧНЫМИ ДРОБЯМИ, РАЦИОНАЛЬНЫМИ ЧИСЛАМИ

Докажем три леммы о приближении чисел, представимых бесконечными десятичными дробями, рациональными числами.

Сначала убедимся в том, что произвольное число a, представимое бесконечной десятичной дробью, можно с наперед заданной точностью приблизить рациональными числами.

Ради определенности будем считать a неотрицательным и представим его дробью $a=a_0, a_1a_2 \dots a_n \dots$.

Обрывая указанную дробь на n-м знаке после запятой, мы получим рациональное число $a_0, a_1a_2 \dots a_n$, причем из правила упорядочения чисел, представимых бесконечными десятичными дробями, сразу же вытекает, что

$$a_0, a_1a_2 \dots a_n \leqslant a.$$

Увеличив указанное рациональное число на 10^{-n}, мы получим другое рациональное число $a_0, a_1a_2 \dots a_n + 10^{-n}$, которое (в силу правила упорядочения) обязано удовлетворять неравенству

$$a \leqslant a_0, a_1a_2 \dots a_n + 10^{-n}.$$

Итак, для любого номера n мы нашли два рациональных числа $\alpha_1 = a_0, a_1a_2 \dots a_n$ и $\alpha_2 = a_0, a_1a_2 \dots a_n + 10^{-n}$ такие, что $\alpha_1 \leqslant a \leqslant \alpha_2$ и $\alpha_2 - \alpha_1 = 10^{-n}$.

Убедимся в том, что для любого наперед взятого положительного рационального числа ε, начиная с некоторого номера n, справедливо неравенство $10^{-n} < \varepsilon$.

В самом деле, в силу аксиомы Архимеда найдется лишь конечное число натуральных чисел, не превосходящих чисел $1/\varepsilon$. Значит, лишь для конечного числа номеров n справедливо неравенство $10^n \leqslant 1/\varepsilon$, или $10^{-n} \geqslant \varepsilon$. Для всех остальных номеров n справедливо противоположное неравенство $10^{-n} < \varepsilon$, что и требовалось доказать.

Мы приходим к следующему утверждению.

Лемма 1. *Для любого представимого бесконечной десятичной дробью числа a и любого наперед взятого положительного рационального числа ε найдутся два рациональных числа α_1 и α_2 такие, что $\alpha_1 \leqslant a \leqslant \alpha_2$ и $\alpha_2 - \alpha_1 < \varepsilon$.*

Докажем еще две леммы, характеризующие густоту распределения рациональных чисел среди произвольных чисел, представимых бесконечными десятичными дробями.

Лемма 2. *Каковы бы ни были два представимых бесконечными десятичными дробями числа a и b такие, что $a > b$, найдется рациональное число α, заключенное между ними, т. е. такое, что $a > \alpha > b$ (а следовательно, найдется и бесконечное множество различных рациональных чисел, заключенных между a и b).*

Доказательство. Достаточно рассмотреть случай, когда оба числа a и b неотрицательны, ибо случай, когда a и b неположительны, сводится к указанному случаю посредством перехода к модулям, а случай, когда b отрицательно, а a положительно, тривиален (в качестве α можно взять нуль).

Итак, пусть $a>b$ и оба числа a и b неотрицательны. Предположим, что $a=a_0, a_1a_2\ldots a_n\ldots$; $b=b_0, b_1b_2\ldots b_n\ldots$, причем в случае, если a является рациональным числом, представимым конечной десятичной дробью, договоримся брать представление a десятичной дробью, заканчивающейся бесконечным числом девяток.

Так как $a>b$, то найдется номер k такой, что $a_0=b_0$, $a_1=b_1$, ..., $a_{k-1}=b_{k-1}$, $a_k>b_k$.

В силу принятой нами договоренности все десятичные знаки a_n при $n>k$ не могут быть равны нулю.

Обозначим через p наименьший из номеров n, бóльших k, для которых $a_n\neq0$. Тогда число a можно записать в виде

$$a=a_0, a_1a_2\ldots a_k00\ldots0a_p\ldots \qquad (a_p>0).$$

С помощью правила упорядочения легко проверить, что рациональное число

$$\alpha=a_0, a_1a_2\ldots a_k00\ldots0(a_p-1)999\ldots$$

удовлетворяет неравенствам $a>\alpha>b$. Лемма доказана.

Лемма 3. *Пусть x_1 и x_2 — два заданных числа, представимых бесконечными десятичными дробями.*

Пусть далее для любого положительного рационального числа ε найдутся два рациональных числа γ_1 и γ_2 такие, что

$$\gamma_1\leqslant x_1\leqslant\gamma_2;\qquad \gamma_1\leqslant x_2\leqslant\gamma_2;\qquad \gamma_2-\gamma_1<\varepsilon.$$

Тогда числа x_1 и x_2 равны.

Доказательство. Допустим противное, т. е. предположим, что $x_1\neq x_2$. Не ограничивая общности, будем считать, что $x_1<x_2$. В силу леммы 2 найдутся два рациональных числа α_1 и α_2 такие, что

$$x_1<\alpha_1<\alpha_2<x_2.$$

Пусть теперь γ_1 и γ_2 — какие угодно рациональные числа, удовлетворяющие неравенствам $\gamma_1\leqslant x_1\leqslant\gamma_2$, $\gamma_1\leqslant x_2\leqslant\gamma_2$.

Из написанных выше неравенств и свойства транзитивности знаков $>$ и $=$ получим $\gamma_1<\alpha_1<\alpha_2<\gamma_2$. Но тогда $\gamma_2-\gamma_1>\alpha_2-\alpha_1$, что противоречит тому, что разность $\gamma_2-\gamma_1$ может быть сделана меньше любого наперед взятого положительного рационального числа ε. Лемма доказана.

§ 4. ОПЕРАЦИИ СЛОЖЕНИЯ И УМНОЖЕНИЯ. ОПИСАНИЕ МНОЖЕСТВА ВЕЩЕСТВЕННЫХ ЧИСЕЛ

1. Определение операций сложения и умножения. Описание понятия вещественных чисел. Хорошо известно, как складывают два числа, представимых бесконечными десятичными дробями, когда требуется вычислить их сумму на практике.

Для того чтобы сложить два таких числа a и b, заменяют их с требуемой точностью рациональными числами и за приближенное значение суммы чисел a и b берут сумму указанных рациональных чисел. При этом совершенно не заботятся о том, с какой стороны (по недостатку или по избытку) взятые рациональные числа приближают a и b.

Фактически указанный практический способ сложения чисел, представимых бесконечными десятичными дробями, предполагает, что чем точнее рациональные числа α и β приближают (с любой стороны) числа a и b соответственно, тем точнее сумма $\alpha+\beta$ приближает то представимое бесконечной десятичной дробью число, которое должно являться суммой чисел a и b.

Желание оправдать указанный практический способ сложения естественно приводит к следующему определению.

Определение 1. *Суммой двух представимых бесконечными десятичными дробями чисел a и b называется такое представимое бесконечной десятичной дробью число x, которое для любых рациональных чисел α_1, α_2, β_1, β_2, удовлетворяющих соотношениям $\alpha_1 \leqslant a \leqslant \alpha_2$, $\beta_1 \leqslant b \leqslant \beta_2$, удовлетворяет неравенствам*

$$\alpha_1+\beta_1 \leqslant x \leqslant \alpha_2+\beta_2.$$

Это число x обозначают символом $a+b$.

В п. 2 будет доказано, что такое число x существует и притом только одно. Там же будет установлено, что таким числом является точная верхняя грань множества $\{\alpha_1+\beta_1\}$ сумм всех рациональных чисел α_1 и β_1, удовлетворяющих неравенствам $\alpha_1 \leqslant a$, $\beta_1 \leqslant b$, или точная нижняя грань множества $\{\alpha_2+\beta_2\}$ сумм всех рациональных чисел α_2 и β_2, удовлетворяющих неравенствам $a \leqslant \alpha_2$, $b \leqslant \beta_2$.

В п. 2 будет доказано также, что в применении к двум рациональным числам данное нами определение приводит к тому же результату, что и старое определение суммы рациональных чисел.

Перейдем теперь к определению произведения двух чисел, представимых бесконечными десятичными дробями. Сначала определим произведение двух положительных чисел a и b.

Определение 2. *Произведением двух представимых положительными бесконечными десятичными дробями чисел a и b называется такое представимое бесконечной десятичной дробью число x, которое для любых рациональных чисел α_1, α_2, β_1, β_2,* *удовлетворяющих соотношениям* $0<\alpha_1\leqslant a\leqslant\alpha_2$, $0<\beta_1\leqslant b\leqslant\beta_2$, *удовлетворяет неравенствам* $\alpha_1\cdot\beta_1\leqslant x\leqslant\alpha_2\cdot\beta_2$.

Это число x обозначают символом $a\cdot b$.

В п. 2 будет установлено, что такое число x существует и притом только одно. Таким числом x является точная верхняя грань множества $\{\alpha_1\cdot\beta_1\}$ произведений всех рациональных чисел α_1 и β_1, удовлетворяющих неравенствам $0<\alpha_1\leqslant a$, $0<\beta_1\leqslant b$, или точная нижняя грань множества $\{\alpha_2\cdot\beta_2\}$ произведений всех рациональных чисел α_2 и β_2, удовлетворяющих неравенствам $a\leqslant\alpha_2$, $b\leqslant\beta_2$.

Произведение чисел любого знака определяется по следующему правилу:

1) для любого представимого бесконечной десятичной дробью числа a полагают, что

$$a\cdot 0=0\cdot a=0;$$

2) для произвольных отличных от нуля и представимых бесконечными десятичными дробями чисел a и b полагают

$$a\cdot b=\begin{cases} \ \ |a|\cdot|b|, & \text{если } a \text{ и } b \text{ одного знака,}\\ -|a|\cdot|b|, & \text{если } a \text{ и } b \text{ разных знаков.}\end{cases}$$

В п. 2 будет установлено, что в применении к двум рациональным числам данное нами определение произведения приводит к тому же результату, что и прежнее определение произведения рациональных чисел.

Теперь мы располагаем всем тем, что необходимо для описания понятия вещественных чисел.

Договоримся называть ***вещественными*** *числа, представимые бесконечными десятичными дробями, при условии, что для этих чисел указанным выше способом определены три операции: упорядочения, сложения и умножения.*

Так как все изложенное в § 2 и 3 (и, в частности, основная теорема 2.1 и леммы 1—3) справедливо для произвольных чисел, представимых бесконечными дробями, для которых определена только одна операция упорядочения, то все изложенное в этих параграфах справедливо и для произвольных вещественных чисел.

В дальнейшем будут рассматриваться числа, представимые бесконечными десятичными дробями, для которых кроме операции упорядочения определены также и операции сложения и умножения.

Такие числа в соответствии со сформулированным нами понятием мы в дальнейшем будем называть вещественными.

2. Существование и единственность суммы и произведения вещественных чисел.

Теорема о существовании суммы вещественных чисел. *Для любых вещественных чисел a и b существует вещественное число x, являющееся их суммой.*

Доказательство. Фиксируем произвольные рациональные числа α_2 и β_2, удовлетворяющие неравенствам $a \leqslant \alpha_2$, $b \leqslant \beta_2$, и рассмотрим всевозможные рациональные числа α_1 и β_1, удовлетворяющие неравенствам $\alpha_1 \leqslant a$, $\beta_1 \leqslant b$.

Убедимся в том, что множество $\{\alpha_1+\beta_1\}$ всех сумм $\alpha_1+\beta_1$, отвечающих указанным выше всевозможным рациональным α_1 и β_1, ограничено сверху.

В силу свойства транзитивности знаков $>$ и $=$ из неравенств $a \leqslant \alpha_2$ и $\alpha_1 \leqslant a$ вытекает, что $\alpha_1 \leqslant \alpha_2$, а из неравенств $b \leqslant \beta_2$ и $\beta_1 \leqslant b$ вытекает, что $\beta_1 \leqslant \beta_2$.

Но два неравенства $\alpha_1 \leqslant \alpha_2$ и $\beta_1 \leqslant \beta_2$ одного знака, связывающие рациональные числа, можно складывать почленно (см. конец п. 1 § 1). Значит, справедливо неравенство

$$\alpha_1+\beta_1 \leqslant \alpha_2+\beta_2,$$

которое и доказывает ограниченность множества $\{\alpha_1+\beta_1\}$ сверху и тот факт, что число $\alpha_2+\beta_2$ является одной из верхних граней этого множества.

По основной теореме 2.1 (см. § 2) у множества $\{\alpha_1+\beta_1\}$ существует точная верхняя грань, которую мы обозначим через x. Остается убедиться в том, что это вещественное число x и является суммой чисел a и b, т. е. удовлетворяет неравенствам $\alpha_1+\beta_1 \leqslant x \leqslant \alpha_2+\beta_2$. Справедливость левого неравенства $\alpha_1+\beta_1 \leqslant x$ вытекает из того, что x является верхней гранью множества $\{\alpha_1+\beta_1\}$, а справедливость правого неравенства $x \leqslant \alpha_2+\beta_2$ вытекает из того, что число $\alpha_2+\beta_2$ является одной из верхних граней множества $\{\alpha_1+\beta_1\}$, а число x является точной, т. е. наименьшей, верхней гранью этого множества. Теорема доказана.

Аналогично можно было бы доказать, что в качестве x можно взять точную нижнюю грань множества $\{\alpha_2+\beta_2\}$ сумм $\alpha_2+\beta_2$ всевозможных рациональных чисел α_2 и β_2, удовлетворяющих неравенствам $a \leqslant \alpha_2$, $b \leqslant \beta_2$.

Теорема единственности суммы двух вещественных чисел. *Может существовать только одно вещественное число x, являющееся суммой двух данных вещественных чисел a и b.*

Доказательство. Предположим, что существуют два вещественных числа x_1 и x_2, удовлетворяющие неравенствам

$$\begin{cases} \alpha_1+\beta_1 \leqslant x_1 \leqslant \alpha_2+\beta_2, \\ \alpha_1+\beta_2 \leqslant x_2 \leqslant \alpha_2+\beta_2 \end{cases} \tag{2.12}$$

для всевозможных рациональных чисел α_1, α_2, β_1, β_2, удовлетворяющих неравенствам

$$\alpha_1 \leqslant a \leqslant \alpha_2, \qquad \beta_1 \leqslant b \leqslant \beta_2. \tag{2.13}$$

Фиксируем произвольное положительное рациональное число ε. В силу леммы 1 из § 3 для положительного рационального числа

$\varepsilon/2$ и для данного вещественного числа a найдутся такие рациональные числа α_1 и α_2, что $\alpha_1 \leqslant a \leqslant \alpha_2$, причем $\alpha_2 - \alpha_1 < \varepsilon/2$.

Аналогично для указанного $\varepsilon/2$ и для данного вещественного числа b найдутся такие рациональные числа β_1 и β_2, что $\beta_1 \leqslant b \leqslant \beta_2$, причем $\beta_2 - \beta_1 < \varepsilon/2$.

Если взять в неравенствах (2.13) указанные α_1, α_2, β_1 и β_2, то мы получим, что оба числа x_1 и x_2 удовлетворяют неравенствам (2.12), которые можно переписать в виде

$$\gamma_1 \leqslant x_1 \leqslant \gamma_2, \qquad \gamma_1 \leqslant x_2 \leqslant \gamma_2,$$

положив $\gamma_1 = \alpha_1 + \beta_1$, $\gamma_2 = \alpha_2 + \beta_2$.

Учитывая, что

$$\gamma_2 - \gamma_1 = (\alpha_2 + \beta_2) - (\alpha_1 + \beta_1) = (\alpha_2 - \alpha_1) + (\beta_2 - \beta_1) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

мы получим, что оба числа x_1 и x_2 заключены между рациональными числами γ_1 и γ_2, разность между которыми меньше наперед взятого положительного рационального ε.

В силу леммы 3 из § 3 мы получим, что $x_1 = x_2$.

Теорема доказана.

С л е д с т в и е. *В применении к двум рациональным числам a и b данное нами определение суммы вещественных чисел приводит к тому же результату, что и прежнее определение суммы рациональных чисел.*

В самом деле, пусть a и b — два рациональных числа, $a+b$ — их сумма согласно прежнему определению, α_1, α_2, β_1 и β_2 — какие угодно рациональные числа, удовлетворяющие неравенствам (2.13). Тогда, очевидно, справедливы неравенства *

$$\alpha_1 + \beta_1 \leqslant a + b \leqslant \alpha_2 + \beta_2, \tag{2.14}$$

причем согласно теореме единственности число $a+b$ является единственным вещественным числом, удовлетворяющим неравенствам (2.14).

Совершенно аналогично доказывается *существование и единственность произведения двух данных вещественных чисел.*

Ясно, что достаточно доказать существование и единственность произведения двух п о л о ж и т е л ь н ы х чисел a и b.

Для доказательства существования произведения фиксируем произвольные рациональные числа α_2 и β_2, удовлетворяющие неравенствам $a \leqslant \alpha_2$, $b \leqslant \beta_2$, и рассмотрим всевозможные рациональные числа α_1 и β_1, удовлетворяющие неравенствам $0 < \alpha_1 \leqslant a$, $0 < \beta_1 \leqslant b$. Легко убедиться в том, что множество $\{\alpha_1 \cdot \beta_1\}$ всех произведений $\alpha_1 \cdot \beta_1$ ограничено сверху, причем число $\alpha_2 \cdot \beta_2$ является одной из верхних граней этого множества.

* Ибо для рациональных чисел неравенства одного знака можно складывать почленно (см. конец п. 1 § 1).

По основной теореме 2.1 существует точная верхняя грань этого множества x, которая, как легко проверить, удовлетворяет неравенствам $\alpha_1 \cdot \beta_1 \leqslant x \leqslant \alpha_2 \cdot \beta_2$, т. е. является произведением чисел a и b.

Аналогично можно было бы доказать, что *произведением положительных чисел a и b является точная нижняя грань множества $\{\alpha_2 \cdot \beta_2\}$ произведений $\alpha_2 \cdot \beta_2$ всевозможных рациональных чисел α_2 и β_2, удовлетворяющих неравенствам $a \leqslant \alpha_2$, $b \leqslant \beta_2$.*

Для доказательства единственности произведения двух положительных вещественных чисел a и b предположим, что существуют два вещественных числа x_1 и x_2, удовлетворяющие неравенствам

$$\alpha_1 \cdot \beta_1 \leqslant x_1 \leqslant \alpha_2 \cdot \beta_2, \qquad \alpha_1 \cdot \beta_1 \leqslant x_2 \leqslant \alpha_2 \cdot \beta_2 \tag{2.15}$$

для всевозможных рациональных α_1, α_2, β_1 и β_2 таких, что*

$$0 < \alpha_1 \leqslant a \leqslant \alpha_2 \leqslant M, \qquad 0 < \beta_1 \leqslant b \leqslant \beta_2 \leqslant M. \tag{2.16}$$

Фиксировав любое положительное рациональное число ε, мы с помощью леммы 1 найдем для данных вещественных чисел a и b такие рациональные числа α_1, α_2, β_1 и β_2, удовлетворяющие неравенствам (2.16), для которых $\alpha_2 - \alpha_1 < \varepsilon/2M$ и $\beta_2 - \beta_1 < \varepsilon/2M$.

Но тогда в силу (2.15) оба числа x_1 и x_2 будут заключены между рациональными числами $\alpha_2 \cdot \beta_2$ и $\alpha_1 \cdot \beta_1$, разность между которыми

$$\alpha_2 \cdot \beta_2 - \alpha_1 \cdot \beta_1 = \alpha_2 (\beta_2 - \beta_1) + \beta_1 (\alpha_2 - \alpha_1) < 2M \cdot \frac{\varepsilon}{2M} = \varepsilon.$$

В силу леммы 3 из § 3 получаем, что $x_1 = x_2$.

С помощью теоремы единственности так же, как и для суммы, доказывается, что *в применении к двум рациональным числам данное нами определение произведения вещественных чисел приводит к тому же самому результату, что и прежнее определение произведения рациональных чисел.*

§ 5. СВОЙСТВА ВЕЩЕСТВЕННЫХ ЧИСЕЛ

1. Свойства вещественных чисел. В этом пункте мы убедимся в справедливости для произвольных вещественных чисел всех основных свойств, перечисленных в п. 1 § 1 для рациональных чисел. Справедливость для вещественных чисел свойств 1°—4° уже установлена выше. Таким образом, нужно выяснить лишь вопрос о справедливости для вещественных чисел свойств 5°—16°. Легко убедиться в справедливости для вещественных чисел свойств 5°—8° и 14°, связанных с понятием суммы. Справедливость свойств 5°—8° непосредственно вытекает из определения

* В качестве M можно взять, например, число $M = 2(a+b)$.

суммы вещественных чисел и из справедливости указанных свойств для рациональных чисел.

Остановимся на доказательстве свойств 14°, т. е. *докажем, что если a, b* и *c — любые три вещественных числа* и $a>b$, *то* $a+c>b+c$.

Так как $a>b$, то в силу леммы 2 из § 3 найдутся рациональные числа α_1 и β_2 такие, что $a>\alpha_1>\beta_2>b$. Для вещественного числа c и для положительного рационального числа $\varepsilon=\alpha_1-\beta_2$ найдутся рациональные числа γ_1 и γ_2 такие, что $\gamma_1\leqslant c\leqslant\gamma_2$, причем $\gamma_2-\gamma_1<\varepsilon=\alpha_1-\beta_2$ (см. лемму 1 § 3). Пусть, далее, α_2 и β_1 — любые рациональные числа, удовлетворяющие неравенствам $\alpha_2\geqslant a$, $b\geqslant\beta_1$. Тогда по определению суммы вещественных чисел

$$\alpha_2+\gamma_2\geqslant a+c\geqslant\alpha_1+\gamma_1, \qquad \beta_2+\gamma_2\geqslant b+c\geqslant\beta_1+\gamma_1.$$

Для доказательства того, что $a+c>b+c$, в силу транзитивности знака $>$ достаточно доказать, что $\alpha_1+\gamma_1>\beta_2+\gamma_2$, но это непосредственно вытекает из неравенства $\gamma_2-\gamma_1<\alpha_1-\beta_2$.

Заметим, что вопрос о вычитании вещественных чисел как о действии, обратном сложению, полностью исчерпывается на основании свойств 5°—8°. *Назовем разностью вещественных чисел a и b вещественное число c такое, что $c+b=a$.*

Убедимся в том, что таковой разностью является число $c=a+b'$, где b' — число, противоположное b.

В самом деле, используя свойства 5°—8°, можем записать

$$c+b=(a+b')+b=a+(b'+b)=a+0=a.$$

Убедимся в том, что существует только одно вещественное число, являющееся разностью двух данных вещественных чисел.

Предположим, что кроме указанного выше числа $c=a+b'$ существует еще одно число d такое, что $d+b=a$. Тогда, с одной стороны, $(d+b)+b'=a+b'=c$, с другой стороны, $(d+b)+b'=d+(b+b')=d+0=d$, т. е. $c=d$.

Из определения разности и из свойства 8° вытекает, что число a', противоположное a, равно разности $0-a$. Это число обычно записывают в виде $-a$. Не вызывает затруднения перенесение на случай вещественных чисел свойств 9°, 10°, 11°, 12°, 13° и 15°, связанных с понятием произведения. Отметим лишь в отношении свойства 12°, что если a — положительное вещественное число, а α_1 и α_2 — какие угодно рациональные числа, удовлетворяющие неравенствам $0<\alpha_1\leqslant a\leqslant\alpha_2$, то число a', обратное для a, определяется как единственное вещественное число, удовлетворяющее неравенствам $\frac{1}{\alpha_2}\leqslant a'\leqslant\frac{1}{\alpha_1}$ *.

* В качестве числа a' может быть взята точная верхняя грань множества всех рациональных чисел $\{1/\alpha_2\}$.

Свойства 9°—12° позволяют заключить, что *для любых двух вещественных чисел a и b ($b \neq 0$) существует и притом только одно вещественное число c, удовлетворяющее условию $c \cdot b = a$. Это число c называется* ***частным*** *чисел a и b.* Из определения частного и из свойства 12° вытекает, что число a', обратное числу a, равно частному $1/a$.

Заметим, наконец, что на случай вещественных чисел переносится и последнее, 16-е, свойство рациональных чисел, а именно: *Каково бы ни было вещественное число a, можно число 1 повторить слагаемым столько раз, что полученная сумма превзойдет a**.

Докажем это свойство. В случае $a<0$ доказательства не требуется, ибо $1>a$. Пусть $a \geqslant 0$, $a = a_0,\ a_1 a_2 \ldots$. В силу того, что определение суммы вещественных чисел в применении к сумме рациональных чисел совпадает с определением суммы рациональных чисел, повторив число 1 слагаемым n раз, получим целое число n. Таким образом, достаточно доказать, что для числа a найдется целое число n такое, что $n>a$. Но это очевидно: достаточно взять $n = a_0 + 2$.

Таким образом, на случай вещественных чисел переносятся все основные свойства, сформулированные для рациональных чисел в п. 1 настоящего параграфа. Следовательно, для вещественных чисел сохраняют силу все правила алгебры, относящиеся к арифметическим действиям и к сочетанию равенств и неравенств.

2. Некоторые часто употребляемые соотношения. Докажем справедливость для любых вещественных чисел a и b следующих двух соотношений:

$$|a \cdot b| = |a| \cdot |b|, \tag{2.17}$$

$$|a+b| \leqslant |a| + |b|. \tag{2.18}$$

Соотношение (2.17) непосредственно вытекает из определения произведения двух вещественных чисел. Докажем соотношение (2.18). На основании определения модуля и правила упорядочения для любых вещественных чисел a и b справедливы неравенства

$$-|a| \leqslant a \leqslant |a|, \qquad -|b| \leqslant b \leqslant |b|.$$

В силу справедливости основных свойств для вещественных чисел можно почленно складывать неравенства одного знака (это доказано в конце п. 1 § 1). Поэтому

$$-(|a| + |b|) \leqslant a+b \leqslant (|a| + |b|).$$

Используя в случае $a+b>0$ правое, а в случае $a+b \leqslant 0$ левое из последних неравенств, мы получим неравенство (2.18).

3. Некоторые конкретные множества вещественных чисел. В дальнейшем нам часто придется иметь дело с различными мно-

* Заметим, что это свойство называют **аксиомой Архимеда**.

жествами вещественных чисел. Будем обозначать произвольное множество вещественных чисел символом $\{x\}$, а числа, входящие в состав этого множества, будем называть элементами или точками этого множества. Мы будем говорить, что точка x_1 множества $\{x\}$ отлична от точки x_2 этого множества, если вещественные числа x_1 и x_2 не равны друг другу. Если при этом справедливо неравенство $x_1>x_2$ $(x_1<x_2)$, то мы будем говорить, что точка x_1 лежит правее (левее) точки x_2.

Рассмотрим некоторые наиболее употребляемые частные виды множеств вещественных чисел.

1°. Множество вещественных чисел x, удовлетворяющих неравенствам $a\leqslant x\leqslant b$, где $a<b$, будем называть сегментом и обозначать символом $[a, b]$. При этом числа a и b мы будем называть граничными точками или концами сегмента $[a, b]$, а любое число x, удовлетворяющее неравенствам $a<x<b$, будем называть внутренней точкой сегмента $[a, b]$.

2°. Множество всех вещественных чисел x, удовлетворяющих неравенствам $a<x<b$, будем называть интервалом и обозначать символом (a, b).

3°. Интервал $(a-\varepsilon, a+\varepsilon)$, где $\varepsilon>0$, будем называть ε-окрестностью точки a.

4°. Любой интервал, содержащий точку a, будем называть окрестностью точки a.

5°. Множество всех вещественных чисел x, удовлетворяющих неравенствам $a\leqslant x<b$ (или $a<x\leqslant b$), будем называть полусегментом и обозначать символом $[a, b)$ (или $(a, b]$).

6°. Множество всех вещественных чисел будем называть числовой (бесконечной) прямой и обозначать символом $(-\infty, +\infty)$.

7°. Множество вещественных чисел x, удовлетворяющих неравенству $x\geqslant a$ (или $x\leqslant b$), будем называть полупрямой и обозначать символом $[a, +\infty)$ (или $(-\infty, b]$).

8°. Множество всех вещественных чисел x, удовлетворяющих неравенству $x>a$ (или $x<b$), будем называть открытой полупрямой и обозначать символом $(a, +\infty)$ (или $(-\infty, a)$).

Произвольное множество $\{x\}$ будем называть плотным в себе, если в любой окрестности каждой точки x этого множества содержится хотя бы одна точка множества, отличная от x.

Примером плотного в себе множества может служить любое из определенных выше множеств 1°—8°. Другим примером плотного множества может служить множество всех рациональных чисел, входящих в состав любого из множеств 1°—8°.

В изложенном нами материале содержатся сведения, необходимые для построения аппарата математического анализа. В следующих параграфах этой главы будут рассмотрены некоторые дополнительные вопросы теории вещественных чисел и элементы теории множеств.

§ 6. ДОПОЛНИТЕЛЬНЫЕ ВОПРОСЫ ТЕОРИИ ВЕЩЕСТВЕННЫХ ЧИСЕЛ

Выше, для того чтобы ввести вещественные числа, были использованы бесконечные десятичные дроби. Для множества бесконечных десятичных дробей были определены правила упорядочения, сложения и умножения и было установлено, что эти правила удовлетворяют 16 основным свойствам (перечисленным в п. 1 § 1 для рациональных чисел). Описанный метод введения вещественных чисел, хотя и обладает несомненными эвристическими и методическими достоинствами, не является единственно возможным. Вещественные числа можно было бы ввести с помощью бесконечных двоичных дробей, с помощью так называемых дедекиндовых сечений в области рациональных чисел *, с помощью последовательностей рациональных чисел ** и другими способами.

Чтобы выяснить взаимосвязь между различными методами введения вещественных чисел, привлечем некоторые новые понятия и установим еще одно важное свойство множества изученных выше вещественных чисел.

1. Полнота множества вещественных чисел. Пусть A и B — два произвольных множества. Будем говорить, что *между множествами A и B установлено* ***взаимно однозначное соответствие****, если каждому элементу множества A отвечает единственный элемент множества B, каждый элемент множества B сопоставлен некоторому элементу множества A и разным элементам множества A отвечают разные элементы множества B.*

Назовем два множества, для элементов каждого из которых определены правила упорядочения, сложения и умножения, ***изоморфными*** *друг другу относительно этих правил, если между элементами этих множеств можно установить взаимно однозначное соответствие так, что если элементам a и b первого множества соответствуют элементы a' и b' второго множества, то* 1) *элементы a' и b' связаны тем же знаком* ($>$, $<$ или $=$), *что и элементы a и b;* 2) *элементу $a+b$ соответствует элемент $a'+b'$;* 3) *элементу $a \cdot b$ соответствует элемент $a' \cdot b'$.*

Аналогично можно было бы говорить не о правилах упорядочения, сложения и умножения, а о каких-либо других правилах, характеризующих соотношения между элементами, и вве-

* Введение вещественных чисел с помощью дедекиндовых сечений изложено, например, в гл. 1 книги Ф. Франклина «Математический анализ» или в гл. 1 книги Г. М. Фихтенгольца «Основы математического анализа».

** Этот способ введения вещественных чисел принадлежит Кантору. Его изложение можно, например, найти в книге В. В. Немыцкого, М. И. Слудской и А. Н. Черкасова «Курс математического анализа», т. I, гл. II, а также в книге Я. Тагамлицкого «Диференсиално смятане» (София, 1971).

сти понятие множеств, изоморфных друг другу относительно указанных правил.

Примером двух множеств, изоморфных друг другу относительно правил упорядочения, сложения и умножения, служит множество рациональных чисел, введенных в виде отношения целых чисел, с соответствующими (см. п. 1 § 1) правилами упорядочения, сложения и умножения и множество рациональных чисел, записанных в виде бесконечных дробей с обычными правилами упорядочения, сложения и умножения вещественных чисел.

Рассмотрим более внимательно два множества: множество всех рациональных чисел и множество всех вещественных чисел. Для каждого из этих множеств определены правила упорядочения, сложения и умножения и справедливы остальные из 16 основных свойств. Вместе с тем ясно, что множество всех вещественных чисел является более «широким», чем множество всех рациональных чисел, ибо в целом множество всех вещественных чисел не изоморфно относительно правил упорядочения, сложения и умножения множеству всех рациональных чисел *, но в множестве вещественных чисел можно выделить часть, изоморфную относительно указанных правил множеству рациональных чисел.

Естественно, возникает вопрос, нельзя ли и для множества всех вещественных чисел построить более «широкое» множество объектов, обладающее такими свойствами: 1) в этом более «широком» множестве определены правила упорядочения, сложения и умножения и справедливы остальные из 16 основных свойств; 2) в целом более «широкое» множество не изоморфно относительно указанных правил множеству всех вещественных чисел; 3) в более «широком» множестве можно выделить часть, изоморфную относительно указанных правил множеству всех вещественных чисел. Мы докажем, что такого более «широкого» множества не существует, т. е. множество всех вещественных чисел является полным относительно правил упорядочения, сложения и умножения и остальных 16 основных свойств.

Вообще, *произвольное множество объектов, для которого определены некоторые правила и справедливы некоторые свойства, называется* ***полным*** *относительно этих правил и свойств, если нельзя построить более «широкое» множество объектов такое, чтобы* 1) *в этом более «широком» множестве были определены те же правила и справедливы те же свойства;* 2) *в целом это более «широкое» множество не было изоморфно данному относительно указанных правил*; 3) *в этом более «широком»*

* Это вытекает из того, что между множеством всех рациональных чисел и всех вещественных чисел нельзя установить взаимно однозначное соответствие. В п. 3 § 7 будет доказано, что такого соответствия нет между рациональными числами и вещественными числами сегмента [0, 1]. Отсюда и вытекает требуемое утверждение.

множестве существовала часть, изоморфная данному множеству относительно указанных правил.

Можно утверждать, что множество всех рациональных чисел не является полным относительно правил упорядочения, сложения и умножения и остальных 16 основных свойств, ибо существует более «широкое» множество (множество вещественных чисел), удовлетворяющее требованиям 1), 2) 3) из только что сформулированного определения.

Докажем теперь, что *множество всех вещественных чисел является полным относительно правил упорядочения, сложения и умножения и остальных 16 основных свойств.*

Предположим противное, т. е. предположим, что существует более «широкое» множество объектов $\{x'\}$ такое, что выполнены требования 1), 2), 3) из сформулированного выше определения, и обозначим через $\{\bar{x}'\}$ ту часть множества $\{x'\}$, которая изморфна относительно правил упорядочения, сложения и умножения множеству $\{x\}$ всех вещественных чисел.

Заметим прежде всего, что у множества $\{x'\}$ существует единственная пара элементов $0'$ и $1'$, играющих особую роль нуля и единицы *. Далее можно утверждать, что элементы $0'$ и $1'$ входят в состав множества $\{\bar{x}'\}$ и находятся во взаимно однозначном соответствии с вещественными числами 0 и 1 **. Пусть α' — какой-либо элемент множества $\{x'\}$, не принадлежащий множеству $\{\bar{x}'\}$.

В силу правила упорядочения мы можем разбить все элементы множества $\{\bar{x}'\}$ на два класса — верхний и нижний, отнеся к верхнему классу все элементы $\bar{x}'$, удовлетворяющие неравенству $\bar{x}'>\alpha'$, а к нижнему классу все элементы $\bar{x}'$, удовлетворяющие неравенству $\bar{x}'<\alpha'$. Оба эти класса не являются пустыми. В самом деле, докажем, например, что верхний класс не пуст. Повторив элемент $1'$ слагаемым достаточное число раз, мы, в силу свойства 16°, получим элемент n' множества $\{\bar{x}'\}$, удовлетворяющий неравенству $n'>\alpha'$, т. е. принадлежащий верхнему классу. Из свойства 4° вытекает, что каждый элемент нижнего класса меньше любого элемента верхнего класса.

* Если бы нашлись два элемента $0'_1$ и $0_2'$, играющие особую роль нуля, то в силу свойства суммы мы получили бы $0_1'=0_1'+0_2'=0_2'+0_1'=0_2'$, т. е. $0_2'=0_2'$. Аналогично доказывается единственность элемента $1'$, играющего особую роль единицы.

** Докажем, например, что нулевой элемент $0'$ множества $\{x'\}$ принадлежит множеству $\{x'\}$ и находится в соответствии с вещественным числом 0. Обозначим через θ тот элемент множества $\{\overline{x'}\}$, который находится в соответствии с вещественным числом 0, и заметим, что сумма $\theta+\theta$ отвечает вещественному числу $0+0=0$, и потому $\theta+\theta=\theta$. С другой стороны, $0'+\theta=\theta$ (по определению нулевого элемента $0'$). Из двух последних равенств заключаем, что $\theta+\theta=0'+\theta$. Прибавляя к обеим частям полученного равенства элемент θ', противоположный θ, и учитывая, что $\theta+\theta'=0'$, получим $\theta+0'=0'+0'$, или (в силу свойства нулевого элемента) $\theta=0'$. Аналогично проводятся рассуждения для единичного элемента.

В силу изоморфизма множества $\{\bar{x}'\}$ и множества $\{x\}$ всех вещественных чисел можно утверждать, что множество всех вещественных чисел разбивается на два класса, причем каждое число из нижнего класса меньше любого числа из верхнего класса. Но это означает, что нижний класс вещественных чисел ограничен сверху и имеет (в силу теоремы 2.1) точную верхнюю грань M, а верхний класс имеет точную нижнюю грань m. Из определения точных граней вытекает, что обе грани m и M заключены между вещественными числами, как угодно близкими между собой, а поэтому $m=M$. Так как число $m=M$ является одним из вещественных чисел, то оно принадлежит одному из классов, т. е. существует либо наименьший элемент в верхнем классе, либо наибольший элемент в нижнем классе. Докажем, что оба эти утверждения абсурдны. Пусть, например, существует наименьший элемент в верхнем классе вещественных чисел. Тогда существует наименьший элемент m' и в верхнем классе, отвечающем разбиению множества $\{\bar{x}'\}$. По определению верхнего класса $m'>\alpha'$. Согласно свойствам суммы существует разность $m'-\alpha'$, причем согласно этим свойствам $m'-\alpha'>0'$. Но тогда в силу свойства 12° для элемента $m'-\alpha'$ существует обратный, который в силу свойств произведения равен частному $1'/(m'-\alpha')$. Согласно свойству 16° элемент $1'$ можно повторить слагаемым столько раз, что полученный при этом «целый» элемент n' будет принадлежать $\{\bar{x}'\}$ и удовлетворять неравенству $n'>1'/(m'-\alpha')$. Из последнего неравенства в силу свойств произведения и суммы получим *

$$m'-\frac{1'}{n'}>\alpha'. \tag{2.19}$$

Так как элементы m', $1'$ и n' принадлежат множеству $\{\bar{x}'\}$, то и элемент $\left(m'-\frac{1'}{n'}\right)$ также принадлежит этому множеству и, очевидно, удовлетворяет неравенству $m'-\frac{1'}{n'}<m'$. Но тогда неравенство (2.19) означает, что в верхнем классе имеется элемент, меньший m', т. е. m' не является наименьшим элементом. Полученное противоречие доказывает полноту множества вещественных чисел **.

2. Аксиоматическое введение множества вещественных чисел. Для введения вещественных чисел мы использовали множество бесконечных десятичных дробей. Определив для множества этих дробей правило упорядочения и операции сложения и ум-

* Эти свойства обеспечивают применимость всех правил алгебры.

** При доказательстве этой теоремы использовалась идея так называемого дедекиндова сечения. Дедекиндовым сечением в области рациональных чисел называется разбиение множества всех рациональных чисел на два непустых подмножества A и B таких, что любой элемент A меньше любого элемента B.

ножения, мы установили, что элементы этого множества обладают 16 основными свойствами и, кроме того, свойством полноты относительно 16 основных свойств.

Описанный способ введения вещественных чисел, хотя и обладает несомненными эвристическими и методическими достоинствами, не является единственно возможным и целесообразным с научной точки зрения. Для окончательного оформления и полного логического завершения наших представлений о вещественных числах более предпочтительным является аксиоматический метод введения этих чисел.

Этот метод заключается в следующем.

Множество вещественных чисел вводится как совокупность объектов *, удовлетворяющих 17 аксиомам, в качестве которых берутся 16 основных свойств и аксиома о полноте относительно 16 указанных свойств. Впредь мы будем называть упомянутые 17 аксиом **аксиомами вещественного числа**. Конкретной реализацией совокупности объектов, удовлетворяющих 17 аксиомам вещественного числа, и является изученное нами выше множество бесконечных десятичных дробей. Возможны и другие реализации указанной совокупности объектов.

Имеет место следующее замечательное утверждение:

Любая реализация совокупности объектов $\{x'\}$, удовлетворяющих 17 аксиомам вещественного числа, изоморфна изученному выше множеству $\{x\}$ бесконечных десятичных дробей.

Доказательство этого утверждения можно найти в книге В. А. Ильина и Э. Г. Позняка: «Основы математического анализа», ч. I (М., Наука, 1982, с. 608—612), а также в книге В. А. Ильина, В. А. Садовничего и Бл. Х. Сендова «Математический анализ» (М., Наука, 1979, с. 65—69).

Подчеркнем, что аксиоматический метод и понятие изоморфных (в различных смыслах) совокупностей объектов широко используются в разнообразных разделах современной математики и физики (при построении геометрии, теории вероятностей, классической механики, статистической физики, квантовой механики ** и др. разделов).

В заключение заметим, что в геометрии множество точек прямой вводится как совокупность объектов, удовлетворяющих некоторым аксиомам, среди которых фундаментальную роль играет аксиома о полноте этой совокупности относительно остальных аксиом. Упомянутые аксиомы позволяют установить

* При этом ничего не предполагается о природе этих объектов.

** Так, квантовая механика первоначально возникла в виде двух внешне различных теорий: «матричной механики» Гейзенберга и «волновой механики» Шредингера. Позже было доказано, что эти две теории используют две изоморфные друг другу конкретные реализации одной общей совокупности объектов, вводимой аксиоматически и называемой абстрактным гильбертовым пространством.

взаимно однозначное соответствие между множеством точек прямой и множеством всех вещественных чисел *. Это соответствие позволяет изображать вещественные числа точками на прямой (числовой оси), чем мы будем широко пользоваться в иллюстративных целях.

§ 7. ЭЛЕМЕНТЫ ТЕОРИИ МНОЖЕСТВ

1. Понятие множества. В предыдущих параграфах при изучении теории вещественных чисел важным понятием являлось понятие множества. Подчеркнем, что множество мы рассматривали как начальное понятие, неопределяемое через другие. В этом параграфе мы будем изучать множества произвольной природы, или, как говорят, абстрактные множества. Это означает, что объекты, составляющие данное множество, или, как говорят, элементы данного множества, уже не обязаны быть вещественными числами. Элементами абстрактного множества могут быть, например, функции, буквы алфавита, фигуры на плоскости и т. д.

В математике обычно вводят множество как совокупность объектов любой природы, обладающих определенным свойством.

Множества мы будем обозначать прописными буквами A, B, ... или X, Y, ... и т. п., их элементы — малыми буквами a, b, ... или x, y, ... и т. п. Утверждение «элемент a принадлежит множеству A» будем записывать в виде $a \in A$, если же элемент a не принадлежит множеству A, то будем писать, что $\overline{a \in A}$ или $a \notin A$. Если рассматриваются два произвольных множества A и B и известно, что все элементы множества B содержатся в множестве A, то B называется подмножеством множества A и обозначается этот факт так: $B \subset A$. При этом говорят, что множество B включается в множество A. (Заметим, что при этом возможен случай $B = A$, т. е. случай, когда каждый элемент множества B принадлежит множеству A и, наоборот, каждый элемент множества A принадлежит множеству B.)

В дальнейшем удобно будет рассматривать множества, являющиеся подмножествами некоторого фиксированного множества E.

Если множество вводится как совокупность объектов, обладающих некоторым свойством, причем оказывается, что объектов, обладающих указанным свойством, не существует, то множество называется пустым и обозначается символом $\varnothing$.

Таким образом, пустое множество — это множество, не содержащее ни одного элемента. Пустое множество является подмножеством любого множества.

Заметим, что когда речь идет о некотором выборе элементов, для которых ранее было введено обозначение, скажем, о наборе

* См. книгу В. А. Ильина и Э. Г. Позняка «Аналитическая геометрия» (М., Наука, 1981; приложение в конце книги).

элементов x, то данная совокупность или данное множество может обозначаться и так $\{x\}$ (говорят — «множество элементов «икс»). Далее, если X — какое-то множество, а P — определенное свойство, то запись $\{x \in X : P(x)\}$ или $\{x \in X \mid P(x)\}$ обозначает множество элементов x, обладающих свойством P. Например, если обозначить через $N=\{x\}$ множество натуральных чисел: 1, 2, 3, ..., то запись $\{x \in N : x^2-4=0\}$ означает множество корней уравнения $x^2-4=0$, являющихся натуральными числами. В данном случае это множество состоит из одного элемента: 2. Таким образом, $\{x \in N : x^2-4=0\}=2$.

Множество всех тех вещественных чисел $\{x\}$, которые одновременно удовлетворяют двум условиям: $x<1$ и $2<x$, является пустым. Пустым является и множество $\{x \in E : x \neq x\}$.

2. Операции над множествами. *Суммой (или объединением) двух множеств A и B называется третье множество C, состоящее из элементов, принадлежащих хотя бы одному из множеств A и B.* Сумма двух множеств обозначается так: $C=A \cup B$. Аналогично определяется сумма A любого числа множеств A_α. В этом случае пишут $A=\bigcup\limits_{\alpha} A_\alpha$, что и означает, что множество A состоит из элементов, принадлежащих хотя бы одному A_α.

Заметим, что не следует путать понятие суммы двух множеств с понятием суммы двух вещественных чисел. Например, если мы рассматриваем множества $A=\{1\}$, $B=\{2\}$, т. е. множества, состоящие всего из одного элемента: в первом случае из единицы, во втором из числа два, то $C=A \cup B=\{1; 2\}$ есть множество, состоящее из двух элементов — чисел 1 и 2. Ясно, что при этом $1+2=3$ не является даже элементом множества C.

Пересечением двух множеств A и B называется третье множество C, состоящее из элементов, принадлежащих как множеству A, так и множеству B, т. е. из элементов, общих для множеств A и B. Пересечение C двух множеств A и B обозначается так: $C=A \cap B$. Аналогично определяется пересечение C произвольного числа множеств $A_\alpha : C=\bigcap\limits_{\alpha} A_\alpha$, т. е. множество C, состоящее из элементов, принадлежащих каждому множеству A_α.

Разностью $C=A \setminus B$ двух множеств A и B называется множество, состоящее из элементов A, не принадлежащих B. Заметим, что если рассматриваемые множества являются подмножествами некоторого фиксированного множества E, то *разность $A'=E \setminus A$ называется дополнением множества A или дополнением до E множества A.*

Подчеркнем также, что понятие разности двух множеств также не следует путать с понятием разности двух вещественных чисел.

Дополнительные сведения о свойствах операций над множествами и понятие отображения множеств будут даны в п. 4 в конце настоящего параграфа.

3. Счетные и несчетные множества. Несчетность сегмента [0, 1]. Мощность множества. Важным вопросом при изучении множеств является вопрос о том, как сравнивать между собой два множества, имея в виду «количество» элементов, в них содержащихся. Если мы имеем два множества, каждое из которых содержит конечное число элементов, то элементы в этих множествах мы можем просто каким-нибудь способом занумеровать. При этом может оказаться, что в первом и втором множествах содержится одинаковое число элементов. Назовем такие два множества, содержащие конечное и одинаковое число элементов, эквивалентными. Если в одном из рассматриваемых множеств элементов окажется больше, то мы будем говорить, что оно имеет большую мощность, чем другое из рассматриваемых множеств.

Обратимся теперь к множествам, состоящим из, вообще говоря, бесконечного числа элементов. Примерами таких множеств являются множество рациональных чисел или множество вещественных чисел, лежащих на сегменте [0, 1].

Назовем два множества A и B эквивалентными, если между ними существует взаимно однозначное соответствие, т. е. каждому элементу $a \in A$ отвечает единственный элемент $b \in B$, каждый элемент $b \in B$ сопоставлен некоторому элементу $a \in A$ и разным элементам множества A отвечают разные элементы множества B.

Взаимно однозначное соответствие называют иногда биективным соответствием.

В частности, множества, содержащие конечное число элементов, эквивалентны тогда и только тогда, когда они содержат одинаковое число элементов. Эквивалентность множеств A и B обозначается так: $A \sim B$.

Покажем, например, что *множество $R = \{r\}$ рациональных чисел и множество $N = \{n\}$ натуральных чисел эквивалентны.* Заметим сначала, что для любого целого $k \neq 0$ два рациональных числа m/n и mk/nk являются одинаковыми (здесь $n \neq 0$). Поэтому всякое рациональное число r можно записать в виде $r = \frac{p}{q}$ $(q > 0)$ и дробь считать несократимой. Число 0 будем считать записанным одним способом: $0 = \frac{0}{1}$.

Назовем число $h = |p| + q$ высотой рационального числа p/q. Ясно, что рациональных чисел r, имеющих данную высоту, конечное число. Будем нумеровать натуральными числами 1, 2, 3, ... рациональные числа по возрастанию высоты, т. е. сперва занумеруем все рациональные числа высоты $h = 1$. Такое число только одно: 0. Этому рациональному числу припишем индекс 1, т. е. поставим ему в соответствие натуральное число 1. Затем занумеруем рациональные числа высоты $h = 2$. Таких чисел два: $1 = \frac{1}{}$

и $1 = \frac{-1}{1}$. Первому из них поставим в соответствие натуральное число 2 (т. е. занумеруем его индексом 2), второму — число 3. После этого занумеруем рациональные числа высоты 3 и т. д. Ясно, что при этом мы установим взаимно однозначное соответствие между всеми рациональными числами и всеми натуральными числами, т. е. $R \sim N$.

Введем понятие счетного множества.

Определение 1. *Множество называется* счетным, *если оно эквивалентно множеству натуральных чисел.*

Согласно этому определению и рассуждениям, проведенным выше, мы получаем, что множество рациональных чисел является счетным множеством.

Докажем следующие два простых утверждения о счетных множествах.

Утверждение 1. *Всякое непустое подмножество счетного множества является или множеством, состоящим из конечного числа элементов, или множеством счетным.*

Доказательство. Пусть A — исходное счетное множество, т. е. $A \sim N$ — множеству натуральных чисел. Это означает, что элементы множества A можно занумеровать каким-нибудь способом. Расположим элементы множества A в виде последовательности: $a_1, a_2, \ldots, a_n, \ldots$. Пусть B — непустое подмножество множества A. Рассмотрим последовательно элементы $a_1, a_2, a_3, \ldots$ множества A. Если $a_1 \in B$, то этот элемент мы обозначим через b_1; если $a_1 \overline{\in} B$, мы переходим к рассмотрению элемента a_2. При рассмотрении элемента a_2 могут представиться две возможности: а) элемент $a_2 \in B$; если при этом было выполнено, что и $a_1 \in B$, то элемент a_2 мы обозначим через b_2; если же $a_1 \overline{\in} B$, то элемент a_2 обозначается через b_1; б) элемент $a_2 \overline{\in} B$, тогда переходим к рассмотрению элемента a_3 и т. д. Ясно, что при этом может случиться, что все элементы множества B будут расположены в виде конечной последовательности: $b_1, b_2, \ldots, b_M$ $(M < \infty)$. В этом случае множество B состоит из конечного числа элементов. Если этого не случится, то мы выпишем все элементы множества B в виде бесконечной последовательности элементов $b_1, b_2, \ldots, b_n, \ldots$, откуда следует, что множество B счетное. Утверждение доказано.

Утверждение 2. *Сумма любой конечной или счетной совокупности счетных множеств есть множество счетное.*

Доказательство. Рассмотрим, например, случай, когда имеется счетная совокупность счетных множеств. Пусть $A_1, A_2, A_3, \ldots$ —совокупность множеств, каждое из которых счетно. Расположим элементы множеств $A_1, A_2, A_3, \ldots$ в виде последовательностей:

$$\begin{array}{lcl}
A_1 & = & \{a_{11}, \ a_{12}, \ a_{13}, \ \ldots\}, \\
A_2 & = & \{a_{21}, \ a_{22}, \ a_{23}, \ \ldots\}, \\
A_3 & = & \{a_{31}, \ a_{32}, \ a_{33}, \ \ldots\}, \\
\cdot & \cdot & \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot \ \cdot
\end{array}$$

Пусть $A=\bigcup\limits_{n=1}^{\infty} A_n$. Произведем нумерацию элементов a множества $A=\{a\}$ следующим образом*:

$$a_1=a_{11}, \ a_2=a_{21}, \ a_3=a_{12}, \ a_4=a_{31}, \ a_5=a_{22}, \ a_6=a_{13} \ \text{и т. д.}$$

У некоторых множеств A_i и A_j могут оказаться общие элементы (при $i \neq j$). В этом случае мы их учитываем только один раз.

Таким образом, элементы множества A можно занумеровать, т. е. поставить во взаимно однозначное соответствие с множеством натуральных чисел N, т. е. A счетно. Утверждение доказано.

Возникает вопрос: существуют ли бесконечные несчетные множества, т. е. такие бесконечные множества, которые нельзя поставить во взаимно однозначное соответствие с множеством натуральных чисел? Ответ содержится в доказываемой ниже теореме.

Теорема 2.2. *Множество всех точек сегмента* [0, 1] *несчетно.*

Доказательство. Рассмотрим интервал (0, 1). Очевидно, что если мы докажем, что интервал (0, 1) несчетен, то и сегмент [0, 1] будет несчетен, так как множество точек сегмента [0, 1] отличается от множества точек интервала (0,1) всего двумя точками: 0 и 1. Итак, докажем, что множество точек интервала (0, 1) несчетно. Допустим противное, т. е. предположим, что все вещественные числа интервала (0, 1) можно занумеровать.

Записывая все числа интервала (0, 1) в виде бесконечных десятичных дробей, получим, что

$$\begin{gathered}
x_1=0,\ a_{11}a_{12}\ldots a_{1n}\ldots, \\
x_2=0,\ a_{21}a_{22}\ldots a_{2n}\ldots, \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \ , \\
x_n=0,\ a_{n1}a_{n2}\ldots a_{nn}\ldots, \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot
\end{gathered}$$

Рассмотрим на интервале (0,1) вещественное число $x=0, b_1b_2\ldots$ $\ldots b_n\ldots$, где b_1 — любая цифра, отличная от a_{11}, 0 и 9; b_2 — любая цифра, отличная от a_{22}, 0 и 9; и т. д.; b_n — любая цифра, отличная от a_{nn}, 0 и 9. Достаточно доказать, что число x не совпадает ни с одним из чисел $x_1, x_2, \ldots, x_n, \ldots$. Число x не содержит после запятой нулей и девяток, т. е. это число не принадлежит классу ра-

---

* В записи всех элементов множеств $A_1, A_2, \ldots$, приведенной выше, стрелки указывают порядок, в котором мы производим нумерацию.

циональных чисел, представимых двумя способами в виде бесконечных десятичных дробей. В таком случае число x допускает единственное представление в виде бесконечной десятичной дроби и оно отлично от всех чисел $x_1, x_2, \dots, x_n, \dots$, ибо совпадение числа x с каким-либо числом x_n означало бы совпадение b_n и a_{nn}. Таким образом, интервал $(0, 1)$, а вместе с тем и сегмент $[0, 1]$ несчетен. Теорема доказана.

Определение 2. *Множество, эквивалентное множеству точек сегмента* $[0, 1]$, *называется множеством **мощности континуума**.*

Из доказанной теоремы 2.2 следует, что множества мощности континуума и счетные множества не являются эквивалентными между собой множествами. В частности, из теоремы 2.2 следует, что существуют иррациональные числа, так как уже на сегменте $[0, 1]$ не все числа рациональны: в противном случае их можно было бы перенумеровать. Из теоремы 2.2 также следует, что иррациональных чисел несчетное множество, так как если бы их было счетное множество или конечное число, то по утверждению 2 и всех чисел — рациональных и иррациональных — было бы счетное множество.

*Рассмотрим два произвольных множества A и B. Если эти множества являются эквивалентными, то мы будем говорить, что они имеют **одинаковую мощность** или являются **равномощными**.*

Для обозначения равномощности множеств A и B используют следующую символику:

$$m(A) = m(B)\ *.$$

*Если множество A эквивалентно некоторому подмножеству множества B и при этом множество A не содержит подмножества, эквивалентного множеству B, то будем говорить, что **мощность множества A меньше мощности множества B**.*

Для обозначения того, что мощность множества A меньше мощности B, используют следующую символику:

$$m(A) < m(B).$$

Например, из данного выше определения множества мощности континуума, из теоремы 2.2 и из утверждения 1 о счетных множествах следует, что мощность счетного множества меньше мощности множества сегмента $[0, 1]$, т. е. мощности континуума.

Итак, нами введено сравнение мощностей двух множеств. Логически возможны еще два случая:

---

* Величину $m(A)$, представляющую собой общую характеристику класса всех эквивалентных множеству A множеств, принято называть **кардинальным числом**. В частности, если A состоит из конечного числа элементов, то $m(A)$ равно количеству элементов этого множества.

а) Множество A содержит подмножество, эквивалентное множеству B, а множество B содержит подмножество, эквивалентное A.

б) Множества A и B не эквивалентны, и ни одно из них не содержит подмножества, эквивалентного другому множеству. Нетрудно доказать, что в случае а) множества A и B будут эквивалентны. Случай же б) на самом деле невозможен.

Заметим еще, что трудной проблемой оказался вопрос о существовании множества промежуточной мощности между мощностью счетных множеств и мощностью континуума. Оказалось, что утверждение как о существовании, так и об отсутствии множества промежуточной мощности не противоречит аксиомам теории множеств и не может быть выведено из них. Тем самым это утверждение является одной из аксиом аксиоматической теории множеств.

В заключение докажем, что сегмент $[0, 1]$ и интервал $(0, 1)$ — эквивалентные или, что то же, равномощные множества. Для этого установим взаимно однозначное соответствие между их элементами. Выберем на сегменте $[0, 1]$ и интервале $(0, 1)$ последовательность точек $\left\{\frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right\}$.

Точке 0 сегмента $[0, 1]$ поставим в соответствие точку $\frac{1}{2}$ интервала $(0, 1)$, точке 1 сегмента $[0, 1]$ поставим в соответствие точку $\frac{1}{3}$ интервала $(0, 1)$, далее точке $\frac{1}{2}$ сегмента $[0, 1]$ поставим в соответствие точку $\frac{1}{4}$ интервала $(0, 1)$ и т. д., точке $\frac{1}{n}$ сегмента $[0, 1]$ поставим в соответствие точку $\frac{1}{n+2}$ интервала $(0, 1)$, $n \geqslant 2$. Всем остальным точкам сегмента (т. е. точкам, отличным от 0,1 и не принадлежащим выбранной последовательности) ставятся в соответствие те же точки интервала, т. е. точки, имеющие те же абсциссы. Таким образом, взаимно однозначное соответствие между сегментом $[0, 1]$ и интервалом $(0, 1)$ установлено.

4. Свойства операций над множествами. Отображение множеств. Отметим ряд свойств, введенных выше операций над множествами. Отношение в к л ю ч е н и я двух множеств обладает следующими свойствами:

1°) $A \subset A$;

2°) если $A \subset B$ и $B \subset A$, то $A = B$;

3°) если $B \subset A$ и $A \subset C$, то $B \subset C$;

4°) $\varnothing \subset A$ для любого множества A.

Операции суммы (объединения) и пересечения множеств обладают следующими, непосредственно проверяемыми свойствами:

5°) $(\bigcup\limits_{\alpha} A_\alpha)\cap B=\bigcup\limits_{\alpha}(A_\alpha\cap B)$ (дистрибутивность пересечения);

6°) $(\bigcap\limits_{\alpha} A_\alpha)\cup B=\bigcap\limits_{\alpha}(A_\alpha\cup B)$ (дистрибутивность объединения);

7°) $A\subset B$ эквивалентно условиям $A\cup B=B$ или $A\cap B=A$.

Напомним, что для подмножеств $\{A\}$ некоторого фиксированного множества E мы ввели операцию дополнения $A'=E\setminus A$. Очевидно эта операция удовлетворяет следующим свойствам:

8°) $A\cup A'=E$, $A\cap A'=\varnothing$;

9°) $\varnothing'=E$, $E'=\varnothing$;

10°) $(\bigcup\limits_{\alpha} A_\alpha)'=\bigcap\limits_{\alpha} A_\alpha'$;

11°) $(\bigcap\limits_{\alpha} A_\alpha)'=\bigcup\limits_{\alpha} A_\alpha$.

Последние два свойства суть правила де Моргана*.

Симметрической разностью двух множеств A и *B назовем множество $C(A\cup B)\setminus(A\cap B)$*. Симметрическая разность множеств A и B обозначается символом $A\triangle B$. Легко видеть, что $A\triangle B=(A\setminus B)\cup(B\setminus A)$.

Важнейшим понятием в анализе является понятие отображения одного множества в другое. Пусть X и Y — какие-то множества. *Если в силу некоторого закона f каждому элементу $x\in X$ соответствует элемент $y=f(x)\in Y$, то говорят, что задано отображение f множества X в множество Y*. Записывают этот факт в виде

$$f{:}X\to Y \text{ или } X\xrightarrow{f}Y.$$

В этом случае элемент $y=f(x)$ называют образом элемента x или значением f на элементе x, а элемент x — прообразом или одним из прообразов элемента y. Часто элемент $x\in X$ называют переменным или аргументом отображения f.

Образом множества $A\subset X$ при отображении $f:X\to Y$ называют множество всех таких элементов из Y, которые являются образами элементов $x\in A$. Это множество обозначается символом $f(A)$. Если $B\subset Y$, то прообразом (или полным прообразом) множества B называют совокупность всех элементов $x\in X$ таких, что $f(x)\in B$. Прообраз множества B обозначается символом $f^{-1}(B)$.

Отображение $f:X\to Y$ иногда удобно называть функцией с областью определения X и областью (или множеством) значений $f(X)\subset Y$. В некоторых разделах математики в зависимости от природы множеств X и Y и свойств f отображение f называется оператором, функционалом и т. д.

* А. де Морган — шотландский математик (1806—1871).

Про отображение $f: X \to Y$ говорят, что оно **сюръективно** (или является отображением X на Y), если $f(X) = Y$; **инъективно** (или является **вложением**), если для любых элементов x_1, x_2 множества X из условия $f(x_1) = f(x_2)$ вытекает, что $x_1 = x_2$, т. е. различные элементы имеют различные образы; **биективно** (или **взаимно однозначно**), если оно сюръективно и инъективно одновременно. Если отображение $f: X \to Y$ биективно, то, как мы отмечали в п. 3, множества X и Y называются эквивалентными (или равномощными). В случае биекции $f: X \to Y$ можно определить **обратное отображение** $f^{-1}: Y \to X$ по правилу: если при отображении f элементу $x \in X$ соответствует элемент $y \in Y$, то $f^{-1}(y)$ полагается равным элементу x. Для любого $y \in Y$ в силу сюръективности отображения f элемент $f^{-1}(y)$ всегда существует, а ввиду инъективности отображения f этот элемент $f^{-1}(y)$ единственен.

Глава 3

ТЕОРИЯ ПРЕДЕЛОВ

В гл. 1 уже указывалось, что одной из основных операций математического анализа является операция предельного перехода и что эта операция встречается в курсе анализа в различных формах.

В настоящей главе изучаются простейшие формы операции предельного перехода. Мы начинаем с изучения самой простейшей формы операции предельного перехода, основанной на понятии предела так называемой числовой последовательности.

Понятие предела числовой последовательности облегчит нам введение и другой весьма важной формы операции предельного перехода, основанной на понятии предельного значения (или, короче, предела) функции.

В конце главы дается общее определение предела функции по базе.

§ 1. ПОСЛЕДОВАТЕЛЬНОСТЬ И ЕЕ ПРЕДЕЛ

1. Понятие последовательности. Арифметические операции над последовательностями. Понятие числовой последовательности известно из курса средней школы. Примерами числовых последовательностей могут служить: 1) последовательность всех элементов арифметической или геометрической прогрессии; 2) последовательность периметров правильных n-угольников, вписанных в данную окружность; 3) последовательность рациональных чисел $x_1=0{,}3$, $x_2=0{,}33$, $x_3=0{,}333$, ..., приближающих число 1/3.

Если каждому значению n из натурального ряда чисел 1, 2, ..., n, ... *ставится в соответствие по определенному закону некоторое вещественное число x_n, то множество занумерованных вещественных чисел*

$$x_1,\ x_2,\ \ldots,\ x_n,\ \ldots \tag{3.1}$$

мы и будем называть ***числовой последовательностью*** *или просто* ***последовательностью.***

Отдельные числа x_n мы будем называть элементами или членами последовательности (3.1). Для сокращенной записи последовательности (3.1) будем использовать символ $\{x_n\}$.

Так, например, символ $\left\{\frac{1}{n^2}\right\}$ обозначает последовательность 1, $1/2^2$, $1/3^2$, ..., $1/n^2$, ..., а символ $\{1+(-1)^n\}$ обозначает последовательность 0, 2, 0, 2, ...

Рассмотрим наряду с последовательностью (3.1) еще одну последовательность

$$y_1, y_2, \ldots, y_n, \ldots . \qquad (3.2)$$

Назовем последовательность $x_1+y_1, x_2+y_2, \ldots, x_n+y_n, \ldots$ *суммой последовательностей* (3.1) *и* (3.2), *последовательность* $x_1-y_1, x_2-y_2, \ldots, x_n-y_n, \ldots$ — *разностью последовательностей* (3.1) *и* (3.2), *последовательность* $x_1 \cdot y_1, x_2 \cdot y_2, \ldots, x_n \cdot y_n, \ldots$ — *произведением последовательностей* (3.1) *и* (3.2) *и, наконец, последовательность* $\frac{x_1}{y_1}, \frac{x_2}{y_2}, \ldots, \frac{x_n}{y_n}, \ldots$ — *частным последовательностей* (3.1) *и* (3.2).

Конечно, при определении частного последовательностей (3.1) и (3.2) необходимо требовать, чтобы все элементы последовательности (3.2) были отличны от нуля. Заметим, однако, что если у последовательности $\{y_n\}$ обращается в нуль лишь конечное число элементов, то частное $\left\{\frac{x_n}{y_n}\right\}$ можно определить с того номера, начиная с которого все элементы y_n отличны от нуля.

2. Ограниченные, неограниченные, бесконечно малые и бесконечно большие последовательности. Совокупность всех элементов произвольной последовательности $\{x_n\}$ образует некоторое числовое множество. Отправляясь от понятий ограниченного сверху, снизу или с обеих сторон множества, мы приходим к следующим определениям.

Определение 1. *Последовательность* $\{x_n\}$ *называется ограниченной сверху (снизу), если существует вещественное число* M *(вещественное число* m*) такое, что каждый элемент этой последовательности* x_n *удовлетворяет неравенству*

$$x_n \leqslant M \qquad (x_n \geqslant m).$$

При этом число M (число m) называется верхней гранью (нижней гранью) последовательности $\{x_n\}$, а неравенство $x_n \leqslant M$ $(x_n \geqslant m)$ называется условием ограниченности этой последовательности сверху (снизу).

Отметим, что любая ограниченная сверху последовательность имеет бесконечное множество верхних граней* и что в условии ограниченности последовательности сверху $x_n \leqslant M$ в качестве M может браться любая из верхних граней. Аналогичное замечание относится и к нижним граням ограниченной снизу последовательности.

Определение 2. *Последовательность* $\{x_n\}$ *называется ограниченной с обеих сторон (или просто ограниченной), если она ограничена и сверху, и снизу, т. е. если существуют два*

* В самом деле, если число M — одна из верхних граней, то в силу свойства транзитивности знаков $>$ и $=$ и любое число M^*, бо́льшее M, является верхней гранью.

вещественных числа M и m такие, что каждый элемент этой последовательности x_n удовлетворяет неравенствам

$$m \leqslant x_n \leqslant M. \qquad (3.3)$$

При этом числа m и M называются соответственно нижней и верхней гранями последовательности $\{x_n\}$, а неравенства (3.3) называются условием ограниченности последовательности $\{x_n\}$.

Подчеркнем, что в условии ограниченности (3.3) могут фигурировать любая нижняя и любая верхняя грани последовательности. Определение ограниченности последовательности требует существования хотя бы одной пары вещественных чисел m и M таких, что для любого элемента последовательности x_n справедливы неравенства (3.3).

Заметим, что условие ограниченности последовательности можно записать не только в форме удовлетворения неравенствам (3.3), но и в другой эквивалентной форме: *последовательность $\{x_n\}$ является ограниченной тогда и только тогда, когда существует положительное вещественное число A такое, что каждый элемент последовательности x_n удовлетворяет неравенству*

$$|x_n| \leqslant A. \qquad (3.4)$$

В самом деле, если каждый элемент x_n удовлетворяет неравенству (3.4), то, положив $m=-A$, $M=+A$, мы получим, что x_n удовлетворяет неравенствам (3.3). Если, наоборот, каждый элемент x_n удовлетворяет неравенствам (3.3), то, обозначив через A наибольшее из двух чисел $|m|$ и $|M|$, мы можем утверждать, что x_n удовлетворяет неравенству (3.4).

В соответствии с определением 2 ограниченной последовательности и условием ограниченности, взятым в форме (3.4), мы можем определить понятие неограниченной последовательности.

Последовательность $\{x_n\}$ называется неограниченной, если для любого положительного вещественного числа A найдется хотя бы один элемент последовательности x_n, удовлетворяющий неравенству*

$$|x_n| > A. \qquad (3.5)$$

С точки зрения этого определения всякая последовательность, которая ограничена только сверху или только снизу, является неограниченной.

Так, например, последовательность 1, 2, 1, 4, ..., 1, $2n$, ... ограничена только снизу и является неограниченной: какое бы положительное вещественное число A мы ни взяли, найдется элемент этой последовательности с четным номером, удовлетворяющий неравенству (3.5).

* Сколь бы большим мы ни взяли это число.

Последовательность $\left\{\frac{1}{n}\right\}$, очевидно, является ограниченной: каждый элемент этой последовательности удовлетворяет неравенствам (3.3) при любых $m \leqslant 0$ и $M \geqslant 1$.

Введем теперь понятия бесконечно большой и бесконечно малой последовательностей.

Определение 3. *Последовательность $\{x_n\}$ называется бесконечно большой, если для любого положительного вещественного числа A* найдется номер N такой**, что при всех $n \geqslant N$ элементы x_n этой последовательности удовлетворяют неравенству* (3.5).

Очевидно, что всякая бесконечно большая последовательность является неограниченной, ибо определение бесконечно большой последовательности требует, чтобы для любого $A>0$ неравенству (3.5) удовлетворяли все элементы последовательности, начиная с некоторого номера N, а определение неограниченной последовательности требует, чтобы для любого $A>0$ неравенству (3.5) удовлетворял хотя бы один элемент последовательности.

Вместе с тем не всякая неограниченная последовательность является бесконечно большой. Так, например, рассмотренная выше последовательность 1, 2, 1, 4, ..., 1, $2n$, ..., будучи неограниченной, не является бесконечно большой, ибо для любого $A>1$ неравенство (3.5) не имеет места для элементов x_n со сколь угодно большими нечетными номерами n.

Определение 4. *Последовательность $\{\alpha_n\}$*** называется бесконечно малой, если для любого положительного вещественного числа ε**** найдется номер N такой*****, что при всех $n \geqslant N$ элементы α_n этой последовательности удовлетворяют неравенству*

$$|\alpha_n| < \varepsilon. \tag{3.6}$$

Докажем, что последовательность $q, q^2, \ldots, q^n, \ldots$ является бесконечно большой при $|q|>1$ и бесконечно малой при $|q|<1$.

Пусть сначала $|q|>1$. Тогда $|q|=1+\delta$, где $\delta>0$. Используя формулу бинома Ньютона, можем записать

$$|q|^N = (1+\delta)^N = 1 + N\delta + \text{(положительные члены)}.$$

Отсюда следует неравенство

$$|q|^N > \delta N. \tag{3.7}$$

* Сколь бы большим мы ни взяли это число.

** Так как этот номер N, вообще говоря, зависит от A, то иногда пишут: $N=N(A)$.

*** Элементы бесконечно малых последовательностей мы будем стремиться обозначать греческими буквами.

**** Сколь бы малым мы ни взяли это число.

***** Так как этот номер N, вообще говоря, зависит от ε, то иногда пишут: $N=N(\varepsilon)$.

Фиксируем произвольное положительное число A и выберем по нему номер N такой, чтобы было справедливо неравенство

$$\delta N > A. \tag{3.8}$$

Убедимся в том, что по любому $A>0$ можно выбрать номер N, удовлетворяющий неравенству (3.8). *Договоримся обозначать символом* $[x]$ *целую часть положительного вещественного числа* x. Поскольку неравенство (3.8) эквивалентно неравенству $N > \frac{A}{\delta}$, то этому неравенству заведомо будет удовлетворять номер N, выбранный из условия $N = \left[\frac{A}{\delta}\right] + 1 = \left[\frac{A}{|q|-1}\right] + 1$.

Заметим теперь, что поскольку $|q|>1$, то из свойств произведения вещественных чисел мы получим, что при всех $n \geqslant N$

$$|q|^n \geqslant |q|^N. \tag{3.9}$$

Сопоставляя неравенства (3.7), (3.8) и (3.9), мы получим, что для любого $A>0$ найдется номер $N = \left[\frac{A}{|q|-1}\right] + 1$ такой, что при всех $n \geqslant N$

$$|q^n| = |q|^n > A.$$

Это и доказывает, что при $|q|>1$ последовательность $\{q^n\}$ является бесконечно большой.

Рассмотрим теперь случай $|q|<1$. Мы должны доказать, что в этом случае последовательность $\{q_n\}$ является бесконечно малой. Исключая тривиальный случай $q=0$, положим $\frac{1}{|q|} = 1 + \delta$, где $\delta>0$. Используя, как и выше, бином Ньютона, мы вместо (3.7) получим неравенство

$$\frac{1}{|q|^N} > \delta N, \text{ или } |q|^N < \frac{1}{qN}. \tag{3.7$'$}$$

Фиксируем произвольное положительное число ε и выберем по нему номер N такой, чтобы было справедливо неравенство

$$\frac{1}{\delta N} < \varepsilon. \tag{3.8$'$}$$

В силу того, что неравенство (3.8′) эквивалентно неравенству $N > \frac{1}{\varepsilon\delta}$, для выбора указанного номера достаточно положить $N = \left[\frac{1}{\varepsilon\delta}\right] + 1 = \left[\frac{|q|}{\varepsilon(1-|q|)}\right] + 1$. Далее, поскольку, в силу свойств произведения вещественных чисел, при $|q|<1$ для всех $n \geqslant N$ справедливо неравенство

$$|q|^n \leqslant |q|^N, \tag{3.9$'$}$$

то из сопоставления неравенств (3.7′), (3.8′) и (3.9′) мы получим, что для любого $\varepsilon>0$ найдется номер $N=\left[\frac{|q|}{\varepsilon(1-|q|)}\right]+1$ такой, что при всех $n\geqslant N$ справедливо неравенство

$$|q^n|=|q|^n<\varepsilon.$$

Это и доказывает, что при $|q|<1$ последовательность является бесконечно малой.

3. Основные свойства бесконечно малых последовательностей.

Теорема 3.1. *Сумма* $\{\alpha_n+\beta_n\}$ *двух бесконечно малых последовательностей* $\{\alpha_n\}$ и $\{\beta_n\}$ *представляет собой бесконечно малую последовательность.*

Доказательство. Фиксируем произвольное положительное число ε. Так как последовательность $\{\alpha_n\}$ является бесконечно малой, то для положительного числа $\varepsilon/2$ найдется номер N_1 такой, что при $n\geqslant N_1$ справедливо неравенство

$$|\alpha_n|<\varepsilon/2. \tag{3.10}$$

Аналогично, так как последовательность $\{\beta_n\}$ является бесконечно малой, то для положительного числа $\varepsilon/2$ найдется номер N_2 такой, что при $n\geqslant N_2$ справедливо неравенство

$$|\beta_n|<\varepsilon/2. \tag{3.11}$$

Обозначим через N наибольший из двух номеров N_1 и N_2. Тогда при $n\geqslant N$ будут справедливы оба неравенства (3.10) и (3.11).

Учитывая, что модуль суммы двух чисел не превосходит суммы их модулей, мы получим, что для всех номеров $n\geqslant N$

$$|\alpha_n+\beta_n|\leqslant|\alpha_n|+|\beta_n|. \tag{3.12}$$

Из соотношений (3.12), (3.10) и (3.11) вытекает, что при $n\geqslant N$ справедливо неравенство

$$|\alpha_n+\beta_n|<\varepsilon.$$

Это и означает, что последовательность $\{\alpha_n+\beta_n\}$ является бесконечно малой. Теорема доказана.

Теорема 3.2. *Разность* $\{\alpha_n-\beta_n\}$ *двух бесконечно малых последовательностей* $\{\alpha_n\}$ *и* $\{\beta_n\}$ *представляет собой бесконечно малую последовательность.*

Доказательство этой теоремы отличается от доказательства теоремы 3.1 только тем, что вместо неравенства (3.12) следует взять неравенство

$$|\alpha_n-\beta_n|\leqslant|\alpha_n|+|\beta_n|.$$

Следствие. *Алгебраическая сумма любого конечного числа бесконечно малых последовательностей представляет собой бесконечно малую последовательность.*

Теорема 3.3. *Произведение ограниченной последовательности на бесконечно малую последовательность представляет собой бесконечно малую последовательность.*

Доказательство. Пусть $\{x_n\}$ — ограниченная и $\{\alpha_n\}$ — бесконечно малая последовательности. По определению ограниченной последовательности найдется вещественное число $A>0$ такое, что для всех элементов x_n справедливо неравенство

$$|x_n| \leqslant A. \tag{3.13}$$

Фиксируем произвольное положительное число ε. Так как последовательность $\{\alpha_n\}$ является бесконечно малой, то для положительного числа ε/A найдется номер N такой, что при $n \geqslant N$ справедливо неравенство

$$|\alpha_n| < \varepsilon/A. \tag{3.14}$$

Учитывая, что модуль произведения двух чисел равен произведению модулей этих чисел, мы получим с помощью неравенств (3.13) и (3.14), что для всех $n \geqslant N$

$$|x_n \cdot \alpha_n| = |x_n| \cdot |\alpha_n| < A \cdot \frac{e}{A} = \varepsilon.$$

Это и означает, что последовательность $\{x_n \cdot \alpha_n\}$ является бесконечно малой. Теорема доказана.

Теорема 3.4. *Всякая бесконечно малая последовательность является ограниченной.*

Доказательство. Пусть $\{\alpha_n\}$ — бесконечно малая последовательность. Фиксируем некоторое положительное число ε. По определению бесконечно малой последовательности найдется отвечающий этому ε номер N такой, что $|\alpha_n| < \varepsilon$ для всех номеров $n \geqslant N$. Обозначим через A наибольшее из следующих N чисел: ε, $|\alpha_1|$, $|\alpha_2|$, ..., $|\alpha_{N-1}|$. Тогда очевидно, что $|\alpha_n| \leqslant A$ для всех номеров n, а это и означает ограниченность последовательности $\{\alpha_n\}$. Теорема доказана.

Следствие из теорем *3.3 и 3.4. Произведение двух (и любого конечного числа) бесконечно малых последовательностей представляет собой бесконечно малую последовательность.*

Теорема 3.5. *Если все элементы бесконечно малой последовательности $\{\alpha_n\}$ равны одному и тому же числу c, то $c=0$.*

Доказательство. Допустим, что $c \neq 0$. Обозначим через ε положительное число $\varepsilon = |c|$. По определению бесконечно малой последовательности для указанного $\varepsilon = |c|$ найдется номер N такой, что $|\alpha_n| < |c|$ при всех $n \geqslant N$. Но неравенство $|\alpha_n| < |c|$ (в силу того, что все α_n равны c) превращается в заведомо абсурдное неравенство $|c| < |c|$. Следовательно, наше допущение $c \neq 0$ не имеет места, и теорема доказана.

Теорема 3.6. *Если $\{x_n\}$ — бесконечно большая последовательность, то, начиная с некоторого номера n, определено частное*

$\left\{\frac{1}{x_n}\right\}$ *двух последовательностей* $\{1\}$* *и* $\{x_n\}$, *которое представляет собой бесконечно малую последовательность. Если все элементы бесконечно малой последовательности* $\{\alpha_n\}$ *отличны от нуля, то частное* $\left\{\frac{1}{\alpha_n}\right\}$ *двух последовательностей* $\{1\}$ *и* $\{\alpha_n\}$ *представляет собой бесконечно большую последовательность.*

Доказательство. Докажем сначала первую часть теоремы. Заметим, что у бесконечно большой последовательности $\{x_n\}$ лишь конечное число элементов может быть равно нулю. В самом деле, по определению бесконечно большой последовательности для числа $A=1$ найдется номер N^* такой, что $|x_n|>A=1$ для всех $n>N^*$. Значит, при $n\geqslant N^*$ все элементы x_n не обращаются в нуль, и мы можем, начиная с номера N^*, рассматривать частное $\left\{\frac{1}{x_n}\right\}$ последовательностей $\{1\}$ и $\{x_n\}$. Докажем, что это частное является бесконечно малой последовательностью. Фиксируем произвольное положительное число ε. По определению бесконечно большой последовательности для положительного числа $\frac{1}{\varepsilon}$ найдется номер N (этот номер мы возьмем таким, чтобы он превосходил N^*) такой, что $|x_n|>\frac{1}{\varepsilon}$ при $n\geqslant N$ или, что то же самое, $\left|\frac{1}{x_n}\right| = \frac{1}{|x_n|}<\varepsilon$ при $n\geqslant N$. Это и означает, что последовательность $\left\{\frac{1}{x_n}\right\}$ является бесконечно малой.

Для доказательства второй части теоремы предположим, что все элементы бесконечно малой последовательности $\{\alpha_n\}$ отличны от нуля. Фиксируем произвольное положительное число A. Так как $\{\alpha_n\}$ является бесконечно малой последовательностью, то для положительного числа $\frac{1}{A}$ найдется номер N такой, что $|\alpha_n|<\frac{1}{A}$ при $n\geqslant N$ или, что то же самое, $\left|\frac{1}{\alpha_n}\right|=\frac{1}{|\alpha_n|}>A$ при $n\geqslant N$. Это и означает, что последовательность $\left\{\frac{1}{\alpha_n}\right\}$ является бесконечно большой. Теорема доказана.

4. Сходящиеся последовательности и их свойства. Введем фундаментальное понятие сходящейся последовательности и ее предела.

Определение 1. *Последовательность* $\{x_n\}$ *называется* ***сходящейся***, *если существует такое вещественное число* a, *что последовательность* $\{x_n-a\}$ *является бесконечно малой. При этом ве-*

* Символ $\{1\}$ обозначает последовательность, все элементы которой равны 1.

щественное число a называется **пределом** *последовательности* $\{x_n\}$ *.

Если последовательность $\{x_n\}$ является сходящейся и имеет своим пределом число a, то символически это записывают так:

$$\lim_{n\to\infty} x_n = a \quad \text{или} \quad x_n \to a \text{ при } n\to\infty.$$

Используя определение бесконечно малой последовательности, мы приходим к другому определению сходящейся последовательности, эквивалентному определению 1.

Определение 2. *Последовательность $\{x_n\}$ называется* **сходящейся**, *если существует такое вещественное число a, что для любого положительного вещественного числа ε найдется номер N такой* **, *что при всех $n \geqslant N$ элементы x_n этой последовательности удовлетворяют неравенству*

$$|x_n - a| < \varepsilon. \tag{3.15}$$

При этом число a называется **пределом** *последовательности* $\{x_n\}$.

Неравенство (3.15) можно записать в эквивалентной форме

$$-\varepsilon < x_n - a < +\varepsilon$$

или, что то же самое,

$$a - \varepsilon < x_n < a + \varepsilon. \tag{3.15'}$$

На геометрическом языке неравенства (3.15′) означают, что элементы x_n при $n \geqslant N$ лежат в интервале $(a-\varepsilon,\ a+\varepsilon)$, который мы договорились называть ε-окрестностью точки a.

Это позволяет сформулировать еще одно определение сходящейся последовательности, эквивалентное определениям 1 и 2.

Определение 3. *Последовательность $\{x_n\}$ называется* **сходящейся**, *если существует такое число a, что в любой ε-окрестности точки a находятся все элементы последовательности $\{x_n\}$, начиная с некоторого номера (зависящего, конечно, от ε).*

Установим специальное представление для элементов любой сходящейся последовательности $\{x_n\}$. В силу определения 1 разность $x_n - a = \alpha_n$ является элементом бесконечно малой последовательности. Следовательно, элемент x_n сходящейся последовательности, имеющей своим пределом вещественное число a, может быть представлен в следующем специальном виде:

$$x_n = a + \alpha_n, \tag{3.16}$$

где α_n — элемент некоторой бесконечно малой последовательности $\{\alpha_n\}$.

---

* В соответствии с этим определением всякая бесконечно малая последовательность является сходящейся и имеет своим пределом число $a=0$.

** Так как этот номер N, вообще говоря, зависит от ε, то иногда пишут: $N = N(\varepsilon)$.

З а м е ч а н и е 1. *Из определения сходящейся последовательности и ее предела сразу же вытекает, что удаление любого конечного числа элементов последовательности не влияет на сходимость этой последовательности и величину ее предела.*

З а м е ч а н и е 2. *Последовательности, не являющиеся сходящимися, принято называть р а с х о д я щ и м и с я.*

З а м е ч а н и е 3. Иногда формально договариваются трактовать бесконечно большие последовательности как последовательности, сходящиеся к пределу ∞. Такая формализация позволяет использовать для бесконечно большой последовательности $\{x_n\}$ следующую символику

$$\lim_{n\to\infty} x_n = \infty.$$

Если при этом элементы бесконечно большой последовательности, начиная с некоторого номера, имеют положительный [отрицательный] знак, то используют следующую символику*:

$$\lim_{n\to\infty} x_n = +\infty \ [\lim_{n\to\infty} x_n = -\infty].$$

В качестве примера *докажем, что последовательность* $\{x_n\}$ *с элементами*

$$x_n = \underbrace{0{,}33\ldots3}_{n \text{ раз}}$$

сходится к пределу $a=1/3$. Фиксируем произвольное положительное число ε и докажем возможность выбора по этому ε такого номера N, что $\left|x_n - \frac{1}{3}\right| < \varepsilon$ при всех $n \geqslant N$. Так как число $1/3$ представимо бесконечной десятичной дробью $0{,}333\ldots$, то из правила упорядочения вещественных чисел вытекают следующие неравенства:

$$\underbrace{0{,}33\ldots3}_{n \text{ раз}} \leqslant \frac{1}{3} \leqslant \underbrace{0{,}33\ldots3}_{n \text{ раз}} + \frac{1}{10^n},$$

справедливые для любого номера n.

Из последних неравенств мы получим для числа $x_n = \underbrace{0{,}33\ldots3}_{n \text{ раз}}$ следующее соотношение:

$$\left|x_n - \frac{1}{3}\right| \leqslant \frac{1}{10^n}.$$

Так как $\frac{1}{10^n} \leqslant \frac{1}{10^N}$ для всех $n \geqslant N$, то для нахождения по данному $\varepsilon > 0$ номера N такого, что $\left|x_n - \frac{1}{3}\right| < \varepsilon$ при всех $n \geqslant N$, достаточно выбрать этот номер N из условия $\frac{1}{10^N} < \varepsilon$.

---

* Иными словами, $\lim_{n\to\infty} x_n = +\infty [= -\infty]$, если для любого $A > 0$ найдется отвечающий этому A номер N такой, что $x_n > A \, [x_n < -A]$ для всех $x \geqslant N$.

Напомним, что в п. 2 мы установили возможность выбора номера N из условия $|q|^N<\varepsilon$ для любого $|q|<1$. Там доказано, что такой номер N можно взять равным

$$N=\left[\frac{|q|}{\varepsilon(1-|q|)}\right]+1.$$

В нашем случае $|q|=0,1$, так что

$$N=\left[\frac{1}{9\varepsilon}\right]+1.$$

Перейдем к установлению свойств произвольных сходящихся последовательностей.

Теорема 3.7. *Сходящаяся последовательность имеет только один предел.*

Доказательство. Предположим, что два вещественных числа a и b являются пределами сходящейся последовательности $\{x_n\}$. Тогда в силу специального представления элементов сходящейся последовательности (3.16) мы получим, что $x_n=a+\alpha_n$ и $x_n=b+\beta_n$, где $\{\alpha_n\}$ и $\{\beta_n\}$ — некоторые бесконечно малые последовательности. Из последних двух равенств получим, что $\alpha_n-\beta_n=b-a$. В силу теоремы 3.2 последовательность $\{\alpha_n-\beta_n\}$ является бесконечно малой, а в силу равенства $\alpha_n-\beta_n=b-a$ все элементы этой бесконечно малой последовательности равны одному и тому же вещественному числу $b-a$. На основании теоремы 3.5 это число $b-a$ равно нулю, т. е. $b=a$. Теорема доказана.

Теорема 3.8. *Всякая сходящаяся последовательность является ограниченной.*

Доказательство. Пусть $\{x_n\}$ — сходящаяся последовательность и a — ее предел. Фиксируем некоторое положительное число ε и по нему номер N такой, что $|x_n-a|<\varepsilon$ при $n\geqslant N$ или, что то же самое, $a-\varepsilon<x_n<a+\varepsilon$ при $n\geqslant N$. Обозначим через A наибольшее из следующих $(N+1)$ чисел: $|a-\varepsilon|$, $|a+\varepsilon|$, $|x_1|, |x_2|, \ldots$ $\ldots, |x_{N-1}|$. Тогда, очевидно, $|x_n|\leqslant A$ для всех номеров n, а это и доказывает ограниченность последовательности $\{x_n\}$. Теорема доказана.

Замечание 4. *Не всякая ограниченная последовательность является сходящейся.* Так, например, последовательность 0, 1, 0, 1, ..., 0, 1, ... является ограниченной, но не является сходящейся. В самом деле, обозначим *n-й* член этой последовательности символом x_n и предположим, что эта последовательность сходится к некоторому пределу a. Но тогда каждая из последовательностей $\{x_{n+1}-a\}$ и $\{x_n-a\}$ являлась бы бесконечно малой. Стало быть, являлась бы бесконечно малой и разность этих последовательностей $\{x_{n+1}-x_n\}$, а этого быть не может в силу того, что $|x_{n+1}-x_n|=1$ для всех номеров n.

Следующие четыре теоремы показывают, что четыре арифметические операции над элементами сходящихся последовательностей приводят к аналогичным операциям над их пределами.

Теорема 3.9. *Сумма сходящихся последовательностей* $\{x_n\}$ *и* $\{y_n\}$ *представляет собой сходящуюся последовательность, предел которой равен сумме пределов последовательностей* $\{x_n\}$ *и* $\{y_n\}$.

Доказательство. Предположим, что последовательности $\{x_n\}$ и $\{y_n\}$ сходятся к пределам a и b соответственно. Тогда в силу специального представления элементов сходящейся последовательности (3.16) будут справедливы соотношения

$$x_n = a + \alpha_n, \qquad y_n = b + \beta_n, \tag{3.17}$$

в которых α_n и β_n представляют собой элементы некоторых бесконечно малых последовательностей $\{\alpha_n\}$ и $\{\beta_n\}$. Из соотношений (3.17) вытекает, что

$$(x_n + y_n) - (a + b) = \alpha_n + \beta_n. \tag{3.18}$$

Так как сумма $\{\alpha_n + \beta_n\}$ двух бесконечно малых последовательностей $\{\alpha_n\}$ и $\{\beta_n\}$ представляет собой бесконечно малую последовательность (теорема 3.1), то из соотношения (3.18) вытекает в силу определения 1, что последовательность $\{x_n + y_n\}$ сходится и вещественное число $a + b$ является ее пределом. Теорема доказана.

Теорема 3.10. *Разность сходящихся последовательностей* $\{x_n\}$ *и* $\{y_n\}$ *представляет собой сходящуюся последовательность, предел которой равен разности пределов последовательностей* $\{x_n\}$ *и* $\{y_n\}$.

Доказательство этой теоремы аналогично доказательству теоремы 3.9, только вместо соотношения (3.18) мы получим соотношение

$$(x_n - y_n) - (a - b) = \alpha_n - \beta_n.$$

Теорема 3.11. *Произведение сходящихся последовательностей* $\{x_n\}$ *и* $\{y_n\}$ *представляет собой сходящуюся последовательность, предел которой равен произведению пределов последовательностей* $\{x_n\}$ *и* $\{y_n\}$.

Доказательство. Предположим, что последовательности $\{x_n\}$ и $\{y_n\}$ сходятся к пределам a и b соответственно. Тогда для элементов этих последовательностей справедливы специальные представления (3.17), перемножая которые, мы получим

$$x_n \cdot y_n = a \cdot b + a \cdot \beta_n + b \cdot \alpha_n + \alpha_n \cdot \beta_n$$

или, что то же самое,

$$x_n y_n - a \cdot b = a\beta_n + b\alpha_n + \alpha_n \cdot \beta_n. \tag{3.19}$$

Для доказательства теоремы в силу определения 1 остается убедиться в том, что в правой части (3.19) стоит элемент бесконечно малой последовательности, но это сразу же вытекает из теоремы

3.3 (согласно этой теореме последовательности $\{a\cdot\beta_n\}$ и $\{b\cdot\alpha_n\}$ являются бесконечно малыми), из следствия из теорем 3.3 и 3.4 (согласно этому следствию последовательность $\{\alpha_n\cdot\beta_n\}$ является бесконечно малой) и из теоремы 3.1 (согласно этой теореме сумма трех бесконечно малых последовательностей $\{a\cdot\beta_n\}$, $\{b\cdot\alpha_n\}$ и $\{\alpha_n\cdot\beta_n\}$ является бесконечно малой последовательностью). Теорема доказана.

Теореме о частном двух сходящихся последовательностей предпошлем следующую лемму.

Лемма 1. *Если последовательность $\{y_n\}$ сходится к отличному от нуля пределу b, то, начиная с некоторого номера, определено частное $\left\{\frac{1}{y_n}\right\}$ последовательностей $\{1\}$ и $\{y_n\}$, которое представляет собой ограниченную последовательность.*

Доказательство. Учитывая, что $b\neq 0$, обозначим через ε положительное число $\varepsilon=\frac{|b|}{2}$. Для этого ε найдется номер N такой, что при $n\geqslant N$ справедливо неравенство $|y_n-b|<\varepsilon$ или, что то же самое,

$$|y_n-b|<\frac{|b|}{2}. \tag{3.20}$$

Итак, для всех номеров n, начиная с номера N, выполняется неравенство (3.20). Убедимся в том, что из неравенства (3.20) вытекает следующее неравенство:

$$|y_n|>\frac{|b|}{2}, \tag{3.21}$$

которое тем самым оказывается справедливым также для всех номеров n, начиная с номера N. В самом деле, так как модуль суммы двух чисел не превосходит суммы модулей этих чисел, то, исходя из тождества $b\equiv(b-y_n)+y_n$ и используя неравенство (3.20), мы получим

$$|b|\leqslant|b-y_n|+|y_n|<\frac{|b|}{2}+|y_n|.$$

Из последнего неравенства сразу же вытекает неравенство (3.21), справедливость которого, начиная с номера N, установлена.

Неравенство (3.21) позволяет утверждать, что при $n\geqslant N$ элементы y_n не обращаются в нуль и, начиная с номера N, можно рассматривать частное $\left\{\frac{1}{y_n}\right\}$.

Из (3.21), в свою очередь, вытекает, что для всех $n\geqslant N$ справедливо неравенство

$$\left|\frac{1}{y_n}\right|<\frac{2}{|b|}.$$

Это последнее неравенство и доказывает, что последовательность $\left\{\frac{1}{y_n}\right\}$, если ее рассматривать, начиная с номера N, является ограниченной. Лемма доказана.

Теорема 3.12. *Частное двух сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$, предел второй из которых отличен от нуля, определено, начиная с некоторого номера, и представляет собой сходящуюся последовательность, предел которой равен частному пределов последовательностей $\{x_n\}$ и $\{y_n\}$.*

Доказательство. Предположим, что последовательности $\{x_n\}$ и $\{y_n\}$ сходятся к пределам a и $b\neq 0$ соответственно. В силу леммы 1 найдется номер N такой, что при $n\geqslant N$ элементы y_n не обращаются в нуль, определена последовательность $\left\{\frac{1}{y_n}\right\}$ и эта последовательность является ограниченной. Начиная с указанного номера N, мы и будем рассматривать частное $\left\{\frac{x_n}{y_n}\right\}$. В силу определения 1 достаточно доказать, что последовательность $\left\{\frac{x_n}{y_n}-\frac{a}{b}\right\}$ является бесконечно малой. Будем исходить из тождества

$$\frac{x_n}{y_n}-\frac{a}{b}=\frac{x_n\cdot b-y_n\cdot a}{y_n\cdot b}. \tag{3.22}$$

Так как для элементов x_n и y_n справедливы специальные представления (3.17), то

$$x_n\cdot b-y_n\cdot a=(a+\alpha_n)\cdot b-(b+\beta_n)\cdot a=\alpha_n b-\beta_n\cdot a. \tag{3.23}$$

Подставляя (3.23) в (3.22), получим

$$\frac{x_n}{y_n}-\frac{a}{b}=\frac{1}{y_n}\left(\alpha_n-\frac{a}{b}\beta_n\right). \tag{3.24}$$

Остается доказать, что в правой части (3.24) стоит элемент бесконечно малой последовательности, но это сразу вытекает из теоремы 3.3 и из того, что последовательность $\left\{\frac{1}{y_n}\right\}$ (в силу леммы 1) является ограниченной, а последовательность $\left\{\alpha_n-\frac{a}{b}\beta_n\right\}$ (как разность двух бесконечно малых) является бесконечно малой последовательностью. Теорема доказана.

Убедимся теперь в том, что неравенства, которым удовлетворяют элементы сходящихся последовательностей, приводят к аналогичным неравенствам для пределов этих последовательностей.

Теорема 3.13. *Если все элементы сходящейся последовательности $\{x_n\}$, по крайней мере начиная с некоторого номера, удовлетворяют неравенству $x_n\geqslant b\,[x_n\leqslant b]$, то и предел x этой последовательности удовлетворяет неравенству $x\geqslant b\,[x\leqslant b]$.*

Доказательство. Предположим, что все элементы x_n, по крайней мере начиная с некоторого номера N^*, удовлетворяют неравенству $x_n \geqslant b$. Докажем, что и предел x этой последовательности удовлетворяет неравенству $x \geqslant b$. Допустим, что это не так, т. е. справедливо неравенство $x < b$.

Тогда по определению сходящейся последовательности для положительного числа $\varepsilon = b - x$ найдется такой номер N (этот номер мы возьмем еще и таким, чтобы он превосходил N^*), что при $n \geqslant N$ будет справедливо неравенство $|x_n - x| < \varepsilon$ или $|x_n - x| < b - x$. Последнее неравенство эквивалентно неравенствам $-(b-x) < x_n - x < b - x$, правое из которых означает, что $x_n < b$ при всех $n \geqslant N$, а это противоречит условию теоремы. Полученное противоречие означает, что наше предположение о том, что $x < b$, неверно, т. е. $x \geqslant b$.

Случай $x_n \leqslant b$ рассматривается аналогично. Теорема доказана.

Замечание 5. *Если все элементы сходящейся последовательности $\{x_n\}$ удовлетворяют строгому неравенству $x_n > b$, то отсюда, вообще говоря, не вытекает, что и предел x этой последовательности удовлетворяет строгому неравенству $x > b$.* (Можно лишь утверждать, что $x \geqslant b$.)

Например, если $x_n = \frac{1}{n}$, то для всех номеров $x_n > 0$, однако предел $\lim\limits_{n\to\infty} \frac{1}{n} = x = 0$ не удовлетворяет неравенству $x > 0$.

Следствие 1. *Если все элементы двух сходящихся последовательностей $\{x_n\}$ и $\{y_n\}$, по крайней мере начиная с некоторого номера, удовлетворяют неравенствам $x_n \leqslant y_n$, то и пределы этих последовательностей удовлетворяют такому же неравенству*

$$\lim_{n\to\infty} x_n \leqslant \lim_{n\to\infty} y_n. \tag{3.25}$$

В самом деле, начиная с указанного номера, все элементы последовательности $\{y_n - x_n\}$ неотрицательны. В силу теоремы 3.13 и предел указанной последовательности $\lim\limits_{n\to\infty}(y_n - x_n)$ неотрицателен. В силу теоремы 3.10 $\lim\limits_{n\to\infty}(y_n - x_n) = \lim\limits_{n\to\infty} y_n - \lim\limits_{n\to\infty} x_n$ и мы получим, что $\lim\limits_{n\to\infty} y_n - \lim\limits_{n\to\infty} x_n \geqslant 0$. Из последнего неравенства вытекает (3.25).

Следствие 2. *Если все элементы сходящейся последовательности $\{x_n\}$ находятся на сегменте $[a, b]$, то и предел x этой последовательности лежит на сегменте $[a, b]$.*

В самом деле, так как $a \leqslant x_n \leqslant b$ для всех номеров n, то (в силу теоремы 3.13) $a \leqslant x \leqslant b$.

Последнюю теорему, к доказательству которой мы сейчас переходим, можно назвать принципом двустороннего ограничения.

Теорема 3.14. *Пусть $\{x_n\}$ и $\{y_n\}$ — две сходящиеся последовательности, имеющие общий предел a. Пусть, кроме того, все элементы третьей последовательности $\{z_n\}$, по крайней мере начиная с некоторого номера, удовлетворяют неравенствам*

$$x_n \leqslant z_n \leqslant y_n. \tag{3.26}$$

Тогда последовательность $\{z_n\}$ сходится к тому же самому пределу a.

Доказательство. Предположим, что неравенства (3.26) справедливы, начиная с номера N^*. Тогда, начиная с того же самого номера N^*, справедливы и неравенства

$$x_n - a \leqslant z_n - a \leqslant y_n - a. \tag{3.27}$$

Из неравенств (3.27) вытекает, что для каждого номера n, превосходящего N^*,

$$|z_n - a| \leqslant \max\{|x_n - a|, \quad |y_n - a|\}. \tag{3.28}$$

(Эта запись означает, что $|z_n - a|$ не превосходит наибольшего из двух чисел $|x_n - a|$ и $|y_n - a|$.)

Фиксируем произвольное положительное число ε. Тогда в силу сходимости последовательностей $\{x_n\}$ и $\{y_n\}$ к пределу a найдутся номера N_1 и N_2 такие, что

$$\begin{cases} |x_n - a| < \varepsilon & \text{при } n \geqslant N_1, \\ |y_n - a| < \varepsilon & \text{при } n \geqslant N_2. \end{cases} \tag{3.29}$$

Если мы теперь обозначим через N наибольший из трех номеров N^*, N_1 и N_2, то при $n \geqslant N$ будут справедливы оба неравенства в (3.29) и мы получим в силу (3.28), что при $n \geqslant N$ справедливо неравенство

$$|z_n - a| < \varepsilon.$$

Это и доказывает сходимость последовательности $\{z_n\}$ к пределу a. Теорема доказана.

§ 2. МОНОТОННЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

1. Понятие монотонной последовательности.

Определение 1. *Последовательность $\{x_n\}$ называется неубывающей [невозрастающей], если каждый элемент этой последовательности, начиная со второго, не меньше [не больше] предыдущего ее элемента, т. е. если для всех номеров n ($n = 1, 2, \ldots$) справедливо неравенство*

$$x_n \leqslant x_{n+1} \qquad [x_n \geqslant x_{n+1}]. \tag{3.30}$$

Определение 2. *Последовательность $\{x_n\}$ называется монотонной, если она является либо неубывающей, либо невозрастающей.*

Если элементы неубывающей последовательности для всех номеров n удовлетворяют строгому неравенству $x_n < x_{n+1}$, то эту последовательность называют возрастающей.

Аналогично, если элементы невозрастающей последовательности для всех номеров n удовлетворяют строгому неравенству $x_n > x_{n+1}$, то эту последовательность называют убывающей.

Заметим, что всякая монотонная последовательность заведомо ограничена с одной стороны (либо сверху, либо снизу). В самом деле, всякая неубывающая последовательность ограничена снизу (в качестве нижней грани можно взять величину ее первого элемента), а всякая невозрастающая последовательность ограничена сверху (в качестве верхней грани также можно взять величину ее первого элемента).

Отсюда следует, что неубывающая последовательность будет ограниченной с обеих сторон, или просто ограниченной, тогда и только тогда, когда она ограничена сверху, а невозрастающая последовательность будет ограниченной тогда и только тогда, когда она ограничена снизу.

Рассмотрим примеры монотонных последовательностей.

1. Последовательность 1, 1, 2, 2, ... является неубывающей. Она ограничена снизу величиной своего первого элемента, а сверху не ограничена.

2. Последовательность $\frac{2}{1}, \frac{3}{2}, \frac{4}{3}, \ldots, \frac{n+1}{n}, \ldots$ является убывающей. Она ограничена с обеих сторон: сверху величиной своего первого элемента 2, а снизу, например, числом 1.

2. Теорема о сходимости монотонной ограниченной последовательности. Справедливо следующее фундаментальное утверждение.

Основная теорема 3.15. *Если неубывающая (невозрастающая) последовательность $\{x_n\}$ ограничена сверху (снизу), то она сходится.*

Доказательство. Рассмотрим случай неубывающей и ограниченной сверху последовательности $\{x_n\}$. Множество всех элементов такой последовательности ограничено сверху, а потому по основной теореме 2.1 гл. 2 у этого множества существует точная верхняя грань, которую мы обозначим символом $\bar{x}$. Докажем, что это число $\bar{x}$ и является пределом последовательности $\{x_n\}$. Во-первых, заметим, что по определению верхней грани любой элемент x_n последовательности $\{x_n\}$ удовлетворяет неравенству

$$x_n \leqslant \bar{x}. \tag{3.31}$$

Далее фиксируем произвольное положительное число ε и заметим, что по определению точной верхней грани найдется хотя бы один элемент последовательности x_N, удовлетворяющий неравенству

$$\bar{x} - \varepsilon < x_N. \tag{3.32}$$

Учтем теперь, что последовательность $\{x_n\}$ является неубывающей и вследствие этого $x_N \leqslant x_n$ для всех номеров n, удовлетворяющих неравенству $n \geqslant N$. Сопоставляя неравенство $x_N \leqslant x_n$ с неравенством (3.32), мы получим, что для всех $n \geqslant N$

$$\bar{x}-\varepsilon < x_n. \tag{3.33}$$

Объединяя неравенства (3.31) и (3.33), мы получим, что для всех $n \geqslant N$ справедливы неравенства $\bar{x}-\varepsilon < x_n \leqslant \bar{x}$.

Следовательно, для всех $n \geqslant N$ справедливо неравенство $|x_n-\bar{x}|<\varepsilon$, которое и доказывает, что последовательность $\{x_n\}$ сходится к пределу $\bar{x}=\sup\{x_n\}$.

Если последовательность $\{x_n\}$ является невозрастающей и ограничена снизу, то совершенно аналогично доказывается, что она сходится к пределу $\underline{x}=\inf\{x_n\}$. Теорема доказана.

Замечание 1. Теорему 3.15 можно сформулировать в другом виде. Во-первых, заметим, что в силу сказанного в п. 1 последовательность $\{x_n\}$, удовлетворяющая условию теоремы 3.15, является ограниченной с обеих сторон, или просто ограниченной. Поэтому теорему 3.15 можно переформулировать так: *для того чтобы монотонная последовательность $\{x_n\}$ сходилась, достаточно, чтобы она была ограничена.*

Легко убедиться в том, что эта формулировка может быть заменена более «сильной»: *для того чтобы монотонная последовательность $\{x_n\}$ сходилась, необходимо и достаточно, чтобы она была ограничена* (необходимость вытекает из теоремы 3.8).

Замечание 2. Конечно, не всякая сходящаяся последовательность является монотонной. Например, заведомо сходящаяся к нулю последовательность $\frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \ldots, \frac{1}{n}, -\frac{1}{n}, \ldots$ не является монотонной, так как знаки ее элементов чередуются.

Замечание 3. Из приведенного выше доказательства теоремы 3.15 вытекает, что *все элементы неубывающей, ограниченной сверху последовательности $\{x_n\}$ не больше ее предела $\bar{x}$ $(x_n \leqslant \bar{x})$.* Аналогично легко убедиться в том, что *все элементы невозрастающей, ограниченной снизу последовательности не меньше ее предела $\underline{x}$.*

Извлечем важное следствие из теоремы 3.15.

Договоримся называть бесконечную последовательность сегментов $[a_1, b_1], [a_2, b_2], \ldots, [a_n, b_n], \ldots$ стягивающейся системой сегментов, если выполнены два требования: 1) каждый следующий сегмент содержится в предыдущем, т. е. $a_n \leqslant a_{n+1}$, $b_{n+1} \leqslant b_n$ для любого $n=1, 2, \ldots$; 2) длина n-го сегмента $[a_n, b_n]$, т. е. разность b_n-a_n, стремится к нулю при $n \to \infty$.

Следствие из теоремы 3.15. *У всякой стягивающейся системы сегментов $\{[a_n, b_n]\}$ существует и притом единственная точка c, принадлежащая всем сегментам этой системы.*

Доказательство. Прежде всего заметим, что точка c, принадлежащая всем сегментам, может быть только одна. В самом деле, если бы нашлась еще одна точка d, отличная от c и принадлежащая всем сегментам, то, предположив ради определенности, что $d>c$, мы получили бы, что сегмент $[c, d]$ принадлежит всем сегментам $[a_n, b_n]$. Но тогда для любого номера n выполнялись бы неравенства $b_n-a_n \geqslant d-c>0$, что невозможно в силу того, что $b_n-a_n \to 0$ при $n \to \infty$.

Докажем теперь, что существует точка c, принадлежащая всем сегментам. Так как система сегментов является стягивающейся, то последовательность левых концов $\{a_n\}$ не убывает, а последовательность правых концов $\{b_n\}$ не возрастает. Поскольку обе эти последовательности ограничены (все их элементы находятся на сегменте $[a_1, b_1]$), то обе они сходятся (в силу теоремы 3.15). Из того, что разность b_n-a_n является бесконечно малой, вытекает, что эти две последовательности $\{a_n\}$ и $\{b_n\}$ сходятся к общему пределу, который мы обозначим через c. В силу замечания 3 для любого номера n справедливы неравенства $a_n \leqslant c \leqslant b_n$, т. е. c принадлежит всем сегментам $[a_n, b_n]$. Следствие доказано.

3. Число e. Применим теорему 3.15 для доказательства сходимости последовательности $\{x_n\}$, элемент x_n которой определяется равенством

$$x_n = \left(1+\frac{1}{n}\right)^n.$$

В силу теоремы 3.15 достаточно доказать, что эта последовательность 1) является возрастающей; 2) ограничена сверху.

Применяя формулу бинома Ньютона, получим для x_n следующее выражение:

$$x_n = 1 + n\cdot\frac{1}{n} + \frac{n(n-1)}{2!}\cdot\frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!}\cdot\frac{1}{n^3} + \ldots$$
$$\ldots + \frac{n(n-1)(n-2)\ldots[n-(n-1)]}{n!}\cdot\frac{1}{n^n}.$$

Это выражение перепишем в следующем виде:

$$x_n = 2 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \ldots$$
$$\ldots + \frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-1}{n}\right). \tag{3.34}$$

Для следующего элемента последовательности $\{x_n\}$ в полной аналогии с (3.34) получится следующее выражение:

$$x_{n+1} = 2 + \frac{1}{2!}\left(1-\frac{1}{n+1}\right) + \frac{1}{3!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right) + \ldots$$

$$\ldots + \frac{1}{n!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\ldots\left(1-\frac{n-1}{n+1}\right)+$$
$$+\frac{1}{(n+1)!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\ldots\left(1-\frac{n}{n+1}\right). \quad (3.35)$$

Для того чтобы убедиться в том, что последовательность $\{x_n\}$ является возрастающей, сравним между собой выражения (3.34) и (3.35). Во-первых, заметим, что правая часть (3.34) состоит из n слагаемых, а правая часть (3.35) — из $n+1$ слагаемых, причем последнее $(n+1)$-е слагаемое в правой части (3.35) является строго положительным.

Сопоставим теперь между собой любое из остальных n слагаемых в правой части (3.35) с соответствующим слагаемым в правой части (3.34). Легко видеть, что для любого номера k, равного 2, 3, ..., n, справедливо неравенство

$$\frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)<$$
$$<\frac{1}{k!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\ldots\left(1-\frac{k-1}{n+1}\right).$$

Это последнее неравенство означает, что k-е слагаемое в правой части (3.34) меньше соответствующего k-го слагаемого в правой части (3.35).

Итак, мы доказали, что $x_n<x_{n+1}$, т. е. последовательность $\{x_n\}$ является возрастающей.

Докажем теперь, что эта последовательность ограничена сверху. Заметим, что если каждую круглую скобку в правой части (3.34) заменить единицей, то указанная правая часть возрастет. Поэтому

$$x_n < 2 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!}. \quad (3.36)$$

Заметим далее, что для любого номера $k\geqslant 2$ справедливо неравенство $k!=2\cdot 3\ldots k\geqslant 2^{k-1}$.

Поэтому неравенство (3.36) дает право утверждать, что

$$x_n < 2 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}} = 3 - \frac{1}{2^{n-1}} < 3. \quad (3.37)$$

(Мы воспользовались формулой для суммы членов геометрической прогрессии.) Неравенство (3.37) доказывает ограниченность последовательности $\{x_n\}$.

По основной теореме 3.15 последовательность $\{x_n\}$ имеет предел, который, следуя Л. Эйлеру *, мы обозначим через e.

* Леонард Эйлер (1707—1783) — великий математик, член Петербургской Академии наук, бо́льшую часть жизни провел в России, по происхождению швейцарец.

Убедимся в том, что число e удовлетворяет неравенствам $2 \leqslant e \leqslant 3$. Для этого (в силу следствия 2 из теоремы 3.13) достаточно доказать, что каждый элемент x_n последовательности $\{x_n\}$ удовлетворяет неравенствам $2 \leqslant e \leqslant 3$.

Неравенство $x_n \leqslant 3$ вытекает из (3.37), а неравенство $2 \leqslant x_n$ вытекает из (3.34), если отбросить в (3.34) все члены, кроме первого.

В дальнейшем выяснится, что число e играет важную роль в математике, и будет указан способ вычисления этого числа с любой степенью точности.

4. Примеры сходящихся монотонных последовательностей. Начнем с рассмотрения последовательности, которая широко используется в современной вычислительной математике для приближенного нахождения корня квадратного из положительного вещественного числа a. Эта последовательность определяется следующей рекуррентной формулой *:

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \ (n = 1, 2, \ldots), \tag{3.38}$$

где в качестве первого приближения x_1 берется любое положительное число.

Прежде всего докажем, что такая последовательность $\{x_n\}$ сходится, для чего в силу теоремы 3.15 достаточно доказать, что она ограничена снизу и, начиная со второго номера, является невозрастающей.

Начнем с доказательства ограниченности снизу. По условию $x_1 > 0$. Но тогда из рекуррентной формулы (3.38), взятой при $n = 1$, вытекает, что $x_2 > 0$, далее из той же формулы (3.38), взятой при $n = 2$, вытекает, что $x_3 > 0, \ldots$. Продолжая эти рассуждения далее, мы последовательно докажем, что все $x_n > 0$. Итак, рассматриваемая последовательность ограничена снизу.

Докажем теперь, что при $n \geqslant 2$ все элементы x_n удовлетворяют неравенству $x_n \geqslant \sqrt{a}$. Переписав формулу (3.38) в виде

$$x_{n+1} = \frac{\sqrt{a}}{2}\left(\frac{x_n}{\sqrt{a}} + \frac{\sqrt{a}}{x_n}\right), \tag{3.39}$$

воспользуемся тривиальным неравенством $t + \frac{1}{t} \geqslant 2$, справедливым для всех $t > 0$ **.

---

* Рекуррентная формула (от латинского recurrens — возвращающийся) — формула, выражающая $(n+1)$-й элемент последовательности через значения ее первых n элементов.

** Это неравенство для всех $t > 0$ эквивалентно неравенству $t^2 + 1 \geqslant 2t$, вытекающему из того, что $(t-1)^2 \geqslant 0$.

Взяв в этом неравенстве $t=\frac{x_n}{\sqrt{a}}>0$, мы получим, что $\frac{x_n}{\sqrt{a}}+$ $+\frac{\sqrt{a}}{x_n}\geqslant 2$, и поэтому соотношение (3.39) дает $x_{n+1}\geqslant\sqrt{a}$ при $n=1$, $2,\ldots$. Это и означает, что $x_n\geqslant\sqrt{a}$ при $n=2, 3,\ldots$.

Докажем, наконец, что последовательность $\{x_n\}$, если ее рассматривать с номера $n=2$, является невозрастающей. Из рекуррентной формулы (3.38) вытекает, что

$$\frac{x_{n+1}}{x_n}=\frac{1}{2}\left(1+\frac{a}{x_n^2}\right).$$

Из последнего соотношения, учитывая, что $x_n\geqslant\sqrt{a}$ при $n\geqslant 2$, получим, что $\frac{x_{n+1}}{x_n}\leqslant 1$ при $n>2$ или $x_{n+1}\leqslant x_n$ при $n\geqslant 2$. Итак, при $n\geqslant 2$ последовательность $\{x_n\}$ является невозрастающей.

По теореме 3.15 последовательность $\{x_n\}$ сходится к некоторому пределу x. Остается найти этот предел. Учитывая, что $x_n\geqslant\sqrt{a}$ при $n\geqslant 2$, мы получим (в силу теоремы 3.13), что $x\geqslant\sqrt{a}>0$.

Принимая во внимание, что $x>0$, перейдем к пределу при $n\to\infty$ в рекуррентном соотношении (3.38). Мы получим в пределе при $n\to\infty$ из (3.38) следующее равенство:

$$x=\frac{1}{2}\left(x+\frac{a}{x}\right).$$

Это равенство представляет собой уравнение для определения предела x. Единственный положительный корень этого уравнения есть $x=\sqrt{a}$.

Итак, мы окончательно доказали, что последовательность $\{x_n\}$, определяемая рекуррентной формулой (3.38) при любом выборе $x_1>0$, сходится к пределу $\sqrt{a}$.

В качестве другого применения теоремы 3.15 рассмотрим вопрос о вычислении предела последовательности $\{x_n\}$, элементы которой имеют вид

$$x_n=\frac{t^n}{n!},\tag{3.40}$$

где t — любое фиксированное вещественное число. Для любого фиксированного t найдется номер N такой, что $|t|<n+1$ при всех $n\geqslant N$. Но тогда, поскольку $\frac{x_{n+1}}{x_n}=\frac{t}{n+1}$, мы получим, что $|x_{n+1}|<|x_n|$ при всех $n\geqslant N$, т. е., начиная с номера N, последовательность $\{|x_n|\}$ является убывающей. Так как, кроме того, эта последовательность ограничена снизу (например, числом нуль), то по теореме 3.15 она сходится к некоторому пределу x. Для нахождения x запишем соотношение (3.40) для номера $n+1$:

$$x_{n+1} = \frac{t^{n+1}}{(n+1)!} = \frac{t^n}{n!} \cdot \frac{t}{n+1} = x_n \cdot \frac{t}{n+1}.$$

Таким образом, $|x_{n+1}| = |x_n| \frac{|t|}{n+1}$, и, перейдя в последнем равенстве к пределу при $n \to \infty$, мы получим соотношение

$$|x| = |x| \cdot \lim_{n\to\infty} \frac{|t|}{n+1} = 0.$$

Итак, последовательность $|x_n|$, а вместе с ней и последовательность (3.40), сходится к пределу $x=0$.

В обоих рассмотренных примерах мы применили часто употребляемый прием: сначала с помощью теоремы 3.15 доказали существование предела последовательности, а затем нашли этот предел, устремив номер n к бесконечности в рекуррентном соотношении, выражающем $(n+1)$-й элемент последовательности через ее n-й элемент.

В качестве третьего примера изучим вопрос о сходимости последовательности $\{x_n\}$, элемент x_n которой имеет вид

$$x_n = \sqrt{a + \sqrt{a + \sqrt{a + \ldots + \sqrt{a}}}} \tag{3.41}$$

при условии, что $a>0$ и общее число извлекаемых корней равно n.

Указанную последовательность $\{x_n\}$ можно задать рекуррентным соотношением

$$x_{n+1} = \sqrt{a + x_n} \quad (n = 1, 2, \ldots) \tag{3.42}$$

при условии, что

$$x_1 = \sqrt{a}. \tag{3.43}$$

Для доказательства сходимости рассматриваемой последовательности достаточно (в силу теоремы 3.15) доказать, что она возрастает и ограничена.

Сначала докажем возрастание последовательности (3.41), т. е. докажем, что для любого номера

$$x_n < x_{n+1}. \tag{3.44}$$

Доказательство этого неравенства проведем по индукции. Достаточно доказать два утверждения: 1) неравенство (3.44) справедливо для номера $n=1$, т. е. справедливо неравенство

$$x_1 \leqslant x_2; \tag{3.45}$$

2) из справедливости неравенства (3.44) для данного номера n

вытекает справедливость этого неравенства и для номера $n+1$, т. е. вытекает справедливость неравенства

$$x_{n+1}<x_{n+2}. \tag{3.46}$$

Справедливость неравенства (3.45) сразу вытекает из равенства (3.43) и из соотношения $x_2=\sqrt{a+\sqrt{a}}>\sqrt{a}$.

Докажем, что из справедливости неравенства (3.44) вытекает справедливость неравенства (3.46). Из неравенства (3.44) и рекуррентной формулы (3.42) вытекает, что

$$x_{n+1}=\sqrt{a+x_n}<\sqrt{a+x_{n+1}}. \tag{3.47}$$

С другой стороны, записывая рекуррентное соотношение (3.42) для номера $n+1$, мы получим равенство

$$x_{n+2}=\sqrt{a+x_{n+1}}. \tag{3.48}$$

Из сопоставления равенства (3.48) с неравенством (3.47) и вытекает неравенство (3.46). Тем самым индукция завершена и возрастание последовательности (3.41) доказано.

Докажем теперь, что эта последовательность ограничена сверху. Снова пользуясь методом математической индукции, мы докажем, что для всех номеров n

$$x_n\leqslant M, \tag{3.49}$$

где M — наибольшее из двух чисел a и 2.

Сначала проверим, что неравенство (3.49) справедливо для номера $n=1$. Пользуясь равенством (3.43) и рассматривая отдельно случаи $0<a\leqslant 2$ и $a>2$, мы получим

$$\begin{aligned} x_1&=\sqrt{a}\leqslant\sqrt{2}<2 \quad \text{при } 0<a\leqslant 2,\\ x_1&=\sqrt{a}<a \qquad\quad\ \text{при } a>2. \end{aligned} \tag{3.50}$$

Из (3.50) вытекает, что $x_1\leqslant M$, где $M=\max\{a, 2\}$.

Пусть теперь неравенство (3.49) справедливо для данного номера n. Пользуясь рекуррентным соотношением (3.42) и рассматривая отдельно случаи $0<a\leqslant 2$ и $a>2$, мы получим, что

$$\begin{aligned} x_{n+1}&=\sqrt{a+x_n}\leqslant\sqrt{2+2}=2 \quad \text{при } 0<a\leqslant 2,\\ x_{n+1}&=\sqrt{a+x_n}\leqslant\sqrt{a+a}=\sqrt{2a}<a \quad \text{при } a>2. \end{aligned} \tag{3.51}$$

Из (3.51) вытекает справедливость неравенства $x_{n+1}\leqslant M$, где $M=\max\{a, 2\}$.

Таким образом, индукция завершена, и ограниченность последовательности (3.41) доказана.

По теореме 3.15 последовательность (3.41) сходится к некоторому пределу x.

Остается найти этот предел.

Из соотношения (3.41) очевидно, что все элементы рассматриваемой последовательности неотрицательны. Следовательно, в силу теоремы 3.13 и искомый предел x этой последовательности неотрицателен.

Возводя в квадрат рекуррентное соотношение (3.42), мы получим равенство

$$x^2_{n+1}=a+x_n. \tag{3.52}$$

Так как последовательность $\{x_n\}$ сходится к пределу x, то, переходя в равенстве (3.52) к пределу при $n\to\infty$ и пользуясь теоремой о пределе суммы и произведения двух сходящихся последовательностей, мы получим следующее соотношение для определения искомого предела x:

$$x^2=a+x$$

или, что то же самое,

$$x^2-x-a=0. \tag{3.53}$$

Соотношение (3.53) представляет собой квадратное уравнение для определения искомого предела x. Это уравнение имеет два корня:

$$x=\frac{1+\sqrt{1+4a}}{2}>0 \text{ и } x=\frac{1-\sqrt{1+4a}}{2}<0.$$

Так как искомый предел, как уже указано выше, является неотрицательным числом, то мы окончательно получим, что он совпадает с положительным корнем уравнения (3.53), т. е. равен

$$x=\frac{1+\sqrt{1+4a}}{2}.$$

§ 3. ПРОИЗВОЛЬНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ

1. Предельные точки, верхний и нижний пределы последовательности. Рассмотрим некоторую последовательность $x_1, x_2, \ldots, x_n, \ldots$ и произвольную в о з р а с т а ю щ у ю последовательность целых положительных чисел $k_1, k_2, \ldots, k_n, \ldots$. Выберем из последовательности $\{x_n\}$ элементы с номерами $k_1, k_2, \ldots, k_n, \ldots$ и расположим их в порядке возрастания указанных номеров. Мы получим при этом новую последовательность

$$x_{k_1}, x_{k_2}, \ldots, x_{k_n}, \ldots,$$

которую принято называть п о д п о с л е д о в а т е л ь н о с т ь ю исходной последовательности $\{x_n\}$.

В частности, и сама последовательность $\{x_n\}$ может рассматриваться как подпоследовательность с номерами $k_n=n$.

Заметим сразу же, что всегда $k_n \geqslant n$, ибо любая подпоследовательность, не совпадающая со всей последовательностью, получается путем некоторого прорежения элементов последовательности.

Справедливы два тривиальных утверждения:

1°. *Если последовательность $\{x_n\}$ сходится к пределу a, то и любая ее подпоследовательность сходится к тому же самому пределу a.*

2°. *Если все подпоследовательности некоторой последовательности $\{x_n\}$ сходятся, то все они сходятся к одному и тому же пределу a (к этому же пределу a сходится и вся последовательность).*

Докажем сначала утверждение 1°.

Фиксируем произвольное положительное число ε и, пользуясь сходимостью последовательности $\{x_n\}$ к пределу a, выберем по этому ε номер N такой, что $|x_n - a| < \varepsilon$ при всех $n \geqslant N$. Пусть $\{x_{k_n}\}$ — произвольная подпоследовательность последовательности $\{x_n\}$. Так как $k_N \geqslant N$, то для всех номеров $n \geqslant N$ элементы подпоследовательности $\{x_{k_n}\}$ удовлетворяют неравенству $|x_{k_n} - a| < \varepsilon$, а это и означает, что подпоследовательность $\{x_{k_n}\}$ сходится к пределу a.

Для доказательства утверждения 2° достаточно учесть, что так как сама последовательность $\{x_n\}$ (как частный случай подпоследовательности) сходится к некоторому пределу a, то и любая ее подпоследовательность сходится к тому же пределу a (в силу утверждения 1°).

В полной аналогии с утверждением 1° доказывается, что *любая подпоследовательность бесконечно большой последовательности представляет собой также бесконечно большую последовательность.*

Введем фундаментальное понятие предельной точки последовательности.

Определение 1. *Точка x бесконечной прямой $(-\infty, +\infty)$ называется* ***предельной точкой*** *последовательности $\{x_n\}$, если в любой ε-окрестности точки x содержится бесконечно много элементов этой последовательности.*

Определение 2. *Точка x бесконечной прямой $(-\infty, +\infty)$ называется* ***предельной точкой*** *последовательности $\{x_n\}$, если из этой последовательности можно выделить подпоследовательность, сходящуюся к пределу x.*

Убедимся в том, что определения 1 и 2 эквивалентны.

1) Пусть в любой ε-окрестности x содержится бесконечно много элементов последовательности $\{x_n\}$. Рассмотрим совокупность ε-окрестностей точки x, для которых ε последовательно равно $1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots$.

В первой из этих окрестностей выберем элемент последовательности x_{k_1} с некоторым номером k_1, во второй из указанных окрестностей выберем элемент последовательности x_{k_2} с номером k_2, удовлетворяющим условию $k_2 > k_1$, в третьей из указанных окрестностей выберем элемент последовательности x_{k_3} с номером k_3, удовлетворяющим условию $k_3 > k_2 \ldots$. Этот процесс можно продолжать неограниченно, так как в любой ε-окрестности точки x содержится бесконечно много элементов последовательности $\{x_n\}$. В результате мы получим подпоследовательность $x_{k_1}, x_{k_2}, x_{k_3}, \ldots, x_{k_n}, \ldots$ последовательности $\{x_n\}$, которая сходится к пределу x, ибо $|x_{k_n} - x| < \frac{1}{n}$.

2) Предположим, что из последовательности $\{x_n\}$ можно выделить подпоследовательность, сходящуюся к пределу x. Тогда в любой ε-окрестности точки x лежит бесконечно много элементов подпоследовательности (все, начиная с некоторого номера). Так как каждый элемент подпоследовательности является элементом и всей последовательности, то в любой ε-окрестности x лежит бесконечно много элементов последовательности.

Эквивалентность определений 1 и 2 доказана.

Выясним вопрос о наличии предельных точек у сходящейся последовательности.

Лемма 1. *Каждая сходящаяся последовательность имеет только одну предельную точку, совпадающую с пределом этой последовательности.*

Доказательство. Пусть последовательность $\{x_n\}$ сходится к пределу x. Тогда в любой ε-окрестности x лежит бесконечно много элементов последовательности $\{x_n\}$ (все, начиная с некоторого номера), а поэтому x является предельной точкой последовательности $\{x_n\}$.

Остается доказать, что ни одно число x', отличное от x, не является предельной точкой последовательности $\{x_n\}$, но это непосредственно вытекает из доказанного выше утверждения 1, согласно которому из сходимости всей последовательности к пределу x вытекает сходимость любой ее подпоследовательности к тому же пределу x.

Приведем пример *ограниченной последовательности* $\{x_n\}$, *имеющей две предельные точки.* Докажем, что последовательность $\frac{1}{2}, 1-\frac{1}{2}, \frac{1}{3}, 1-\frac{1}{3}, \ldots, \frac{1}{n}, 1-\frac{1}{n}, \ldots$ имеет только две предельные точки $x=0$ и $x=1$. Тот факт, что эти две точки $x=0$ и $x=1$ являются предельными, вытекает из того, что подпоследовательность всех нечетных элементов рассматриваемой последовательности сходится к пределу $x=0$, а подпоследовательность всех четных элементов рассматриваемой последовательности сходится к пределу $x=1$. Остается доказать, что ни одно число x_0, отличное от 0 и 1, не является предельной точкой нашей последовательности. Так как $x_0 \neq 0$ и $x_0 \neq 1$, то заведомо можно указать столь малое положительное число ε, что ε-окрестности трех точек 0, 1 и x_0 не будут иметь общих точек (рис. 3.1).

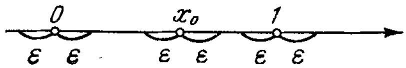


Рис. 3.1

Но все нечетные элементы нашей последовательности, начиная с некоторого номера, находятся в ε-окрестности числа 0, а все четные элементы нашей последовательности, начиная с некоторого номера, находятся в ε-окрестности числа 1. Поэтому за пределами ε-окрестностей чисел 0 и 1 (и, в частности, в ε-окрестности числа x_0) может лежать лишь конечное число элементов нашей последовательности. Это и означает, что x_0 не является предельной точкой последовательности.

Приведем теперь пример *ограниченной последовательности* $\{x_n\}$, *имеющей бесконечно много предельных точек.* Выше (в п. 3 § 7 гл. 2) мы установили, что множество всех рациональных чисел из сегмента $[0, 1]$ можно занумеровать в последовательность $\{x_n\}$. Докажем, что любое вещественное число x из сегмента $[0, 1]$ является предельной точкой указанной последовательности $\{x_n\}$. Заметим, что, каково бы ни было число x из сегмента $[0, 1]$ для любого $0<\varepsilon<1/2$ хотя бы одно из двух чисел $x-\varepsilon$ и $x+\varepsilon$ также принадлежит сегменту $[0, 1]$.

Предположим ради определенности, что число $x+\varepsilon$ принадлежит сегменту $[0, 1]$. Между двумя не равными друг другу вещественными числами x и $x+\varepsilon$, в силу леммы 2 § 3 гл. 2, лежит бесконечно много различных рациональных чисел. Это означает, что при любом $0<\varepsilon<1/2$ в ε-окрестности точки x лежит бесконечно много элементов последовательности $\{x_n\}$, т. е. x является предельной точкой этой последовательности.

Естественно, возникает идея рассмотрения наибольшей и наименьшей предельных точек последовательности.

Определение 3. *Наибольшая предельная точка последовательности $\{x_n\}$ называется* **верхним пределом** *этой последовательности и обозначается символом*

$$\overline{x} = \overline{\lim_{n\to\infty}}\, x_n.$$

Определение 4. *Наименьшая предельная точка последовательности $\{x_n\}$ называется* **нижним пределом** *этой последовательности и обозначается символом*

$$\underline{x} = \underline{\lim_{n\to\infty}}\, x_n.$$

Возникает вопрос о существовании хотя бы одной предельной точки и верхнего и нижнего пределов у любой ограниченной последовательности.

Справедлива следующая замечательная теорема.

Основная теорема 3.16. *У всякой ограниченной последовательности существуют верхний и нижний пределы и, в. частности, существует хотя бы одна предельная точка.*

Доказательство. Остановимся на доказательстве существования у любой ограниченной последовательности хотя бы одной предельной точки и верхнего предела. (Существование нижнего предела доказывается аналогично.)

Пусть $\{x_n\}$ — произвольная ограниченная последовательность. По условию ограниченности найдутся два вещественных числа m и M такие, что любой элемент x_n последовательности $\{x_n\}$ удовлетворяет неравенствам $m \leqslant x_n \leqslant M$.

Рассмотрим множество $\{x\}$ всех вещественных чисел x таких, что правее * каждого из этих чисел либо вовсе нет элементов последовательности $\{x_n\}$, либо таких элементов лишь конечное число.

Иными словами, вещественное число x принадлежит множеству $\{x\}$, если правее x лежит не более чем конечное число элементов последовательности $\{x_n\}$, и не принадлежит множеству $\{x\}$, если правее этого числа x лежит бесконечно много элементов последовательности $\{x_n\}$.

Заметим, что множество $\{x\}$ заведомо не является пустым: ему принадлежит любое вещественное число x, удовлетворяющее неравенству $x \geqslant M$ (ибо правее такого x нет элементов последовательности $\{x_n\}$). Кроме того, очевидно, что множество $\{x\}$ ограничено снизу и в качестве его нижней грани может быть взято любое число, меньшее m (правее такого числа лежат все элементы последовательности $\{x_n\}$, а их бесконечно много).

По основной теореме 2.1 гл. 2 у множества $\{x\}$ существует точная нижняя грань, которую мы обозначим символом $\bar{x}$. Докажем, что это число $\bar{x} = \inf\{x\}$ и является верхним пределом последовательности $\{x_n\}$.

Достаточно доказать два утверждения:

1°. Число $\bar{x}$ является предельной точкой последовательности $\{x_n\}$ (т. е. в любой ε-окрестности $\bar{x}$ лежит бесконечно много элементов последовательности $\{x_n\}$).

2°. Ни одно число $\overline{\overline{x}}$, большее $\bar{x}$, уже не является предельной точкой последовательности $\{x_n\}$ (это и будет означать, что $\bar{x}$ является наибольшей предельной точкой, т. е. верхним пределом $\{x_n\}$).

Для доказательства утверждения 1° фиксируем произвольное положительное число ε. По определению нижней грани любое число, меньшее $\bar{x}$ (и, в частности, число $\bar{x} - \varepsilon$), не принадлежит введенному нами множеству $\{x\}$. Значит, правее $\bar{x} - \varepsilon$ лежит бесконечно много элементов последовательности $\{x_n\}$.

---

* Напомним, что термин «y лежит правее x» означает, что числа x и y связаны неравенством $y > x$.

Далее, из того, что число $\bar{x}$ является точной нижней гранью $\{x\}$, и из неравенства $\bar{x}<\bar{x}+\varepsilon$ вытекает, что найдется хотя бы один элемент x' множества $\{x\}$, удовлетворяющий неравенствам $\bar{x}\leqslant x'<\bar{x}+\varepsilon$, т. е. лежащий левее $\bar{x}+\varepsilon$ (рис. 3.2). В силу определения множества $\{x\}$ правее этого числа x' лежит не более чем конечное число элементов последовательности $\{x_n\}$.

На рис. 3.2 условно указано, что правее числа $\bar{x}-\varepsilon$ лежит бесконечно много элементов последовательности $\{x_n\}$, а правее числа x' лежит не более чем конечное число элементов этой последовательности.

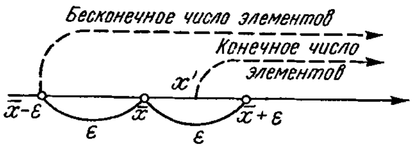


Рис. 3.2

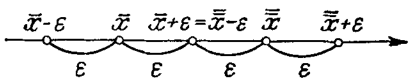


Рис. 3.3

Так как правее $\bar{x}-\varepsilon$ лежит бесконечно много, а правее x' — лишь конечное число элементов последовательности $\{x_n\}$, то мы приходим к выводу, что на полусегменте $(\bar{x}-\varepsilon,\ x']$ (а значит, и на интервале $(\bar{x}-\varepsilon,\ \bar{x}+\varepsilon)$) лежит бесконечно много элементов последовательности $\{x_n\}$.

Итак, мы доказали, что для любого $\varepsilon>0$ в ε-окрестности точки $\bar{x}$ лежит бесконечно много элементов последовательности $\{x_n\}$. Это и означает, что $\bar{x}$ является предельной точкой последовательности $\{x_n\}$. Утверждение 1° доказано.

Подчеркнем, что попутно мы доказали, что для любого $\varepsilon>0$ правее числа $\bar{x}+\varepsilon$ лежит не более чем конечное число элементов последовательности $\{x_n\}$.

Это последнее утверждение используем для доказательства утверждения 2° о том, что $\bar{x}$ является наибольшей предельной точкой.

Пусть $\bar{\bar{x}}$ — любое число, большее $\bar{x}$. Обозначим через ε положительное число $\varepsilon=(\bar{\bar{x}}-\bar{x})/2$. При таком выборе ε интервалы $(\bar{x}-\varepsilon,\ \bar{x}+\varepsilon)$ и $(\bar{\bar{x}}-\varepsilon,\ \bar{\bar{x}}+\varepsilon)$, т. е. ε-окрестности точек $\bar{x}$ и $\bar{\bar{x}}$, не будут иметь общих точек, а точнее, вся ε-окрестность точки $\bar{\bar{x}}$ будет лежать правее числа $\bar{x}+\varepsilon$, т. е. правой границы ε-окрестности точки $\bar{x}$ (рис. 3.3).

Выше мы установили, что для любого $\varepsilon>0$ правее $\bar{x}+\varepsilon$ лежит не более чем конечное число элементов последовательности $\{x_n\}$. Значит, в рассматриваемой нами ε-окрестности точки $\bar{\bar{x}}$ лежит не более чем конечное число элементов последовательности $\{x_n\}$, а это и означает, что $\bar{\bar{x}}$ не является предельной точкой последовательности $\{x_n\}$. Утверждение 2° доказано.

Мы доказали существование у ограниченной последовательности $\{x_n\}$ верхнего предела (т. е. наибольшей предельной точки). Совершенно аналогично доказывается, что у такой последовательности существует нижний предел, являющийся точной верхней гранью того множества вещественных чисел $\{x\}$, левее каждого из которых лежит не более чем конечное число элементов последовательности $\{x_n\}$. Теорема 3.16 доказана.

Следствие 1 из теоремы 3.16. *Если $\{x_n\}$ — ограниченная последовательность, $\underline{x}$ и $\overline{x}$ — ее нижний и верхний пределы, ε — любое положительное число, то на интервале $(\underline{x}-\varepsilon, \overline{x}+\varepsilon)$ лежат все элементы этой последовательности, начиная с некоторого номера (зависящего, конечно, от ε).*

Достаточно доказать, что для любого $\varepsilon>0$ вне интервала $(\underline{x}-\varepsilon, \overline{x}+\varepsilon)$ лежит не более чем конечное число элементов последовательности $\{x_n\}$. Тем более достаточно доказать, что правее $\overline{x}+\frac{\varepsilon}{2}$ и левее $\underline{x}-\frac{\varepsilon}{2}$ лежит не более чем конечное число элементов последовательности $\{x_n\}$. Тот факт, что для любого $\varepsilon>0$ правее $\overline{x}+\frac{\varepsilon}{2}$ лежит не более чем конечное число элементов $\{x_n\}$, уже установлен в процессе доказательства теоремы 3.16. Совершенно аналогично доказывается, что для любого $\varepsilon>0$ левее $\underline{x}-\frac{\varepsilon}{2}$ лежит не более чем конечное число элементов последовательности $\{x_n\}$.

Следствие 2 из теоремы 3.16. *Пусть $\{x_n\}$ — ограниченная последовательность, $\underline{x}$ и $\overline{x}$ — ее нижний и верхний пределы, (a, b) — интервал, вне которого лежит не более чем конечное число элементов последовательности $\{x_n\}$. Тогда интервал $(\underline{x}, \overline{x})$ содержится в интервале (a, b) и, в частности, $\overline{x}-\underline{x}\leqslant b-a$.*

Доказательство. Достаточно доказать два неравенства $\overline{x}\leqslant b$ и $a\leqslant\underline{x}$. Первое из этих неравенств вытекает из того, что точка b, правее которой лежит не более чем конечное число элементов последовательности $\{x_n\}$, принадлежит множеству $\{x\}$, рассмотренному при доказательстве теоремы 3.16, а $\overline{x}$ является точной нижней гранью этого множества. Второе неравенство $a\leqslant\underline{x}$ устанавливается аналогично.

Следствие 3 из теоремы 3.16 (теорема Больцано—Вейерштрасса*). *Из всякой ограниченной последовательности можно выделить сходящуюся подпоследовательность.*

Эта теорема является непосредственным следствием теоремы 3.16 и определения 2 предельной точки.

Теорема 3.16 проливает свет на то, как устроено множество всех предельных точек любой ограниченной последовательности.

---

* Бернгард Больцано — чешский философ и математик (1781—1848), Карл Вейерштрасс — немецкий математик (1815—1897).

Если, как и выше, обозначить через $\underline{x}$ и $\bar{x}$ нижний и верхний пределы этой последовательности, то можно утверждать, что все ее предельные точки лежат на сегменте $[\underline{x}, \bar{x}]$, причем если указанная последовательность не является сходящейся, то она имеет по крайней мере две предельные точки $\underline{x}$ и $\bar{x}$. Рассмотренная нами выше последовательность $\frac{1}{2}, 1-\frac{1}{2}, \frac{1}{3}, 1-\frac{1}{3}, \ldots, \frac{1}{n}, 1-\frac{1}{n}, \ldots$ представляет собой пример последовательности, имеющей только две предельные точки $\underline{x}=0$ и $\bar{x}=1$.

Другая рассмотренная выше последовательность $\{x_n\}$, содержащая все рациональные числа из сегмента $[0, 1]$, представляет собой пример последовательности, предельные точки которой покрывают весь сегмент $[\underline{x}, \bar{x}]$, у которого $\underline{x}=0$, $\bar{x}=1$.

Легко построить пример последовательности, предельными точками которой служат: 1) наперед заданное конечное множество точек $a_1, a_2, \ldots, a_k$; 2) наперед взятая бесконечная последовательность точек $a_1, a_2, \ldots, a_n, \ldots$ * (во втором случае каждая предельная точка последовательности предельных точек $\{a_n\}$ будет являться предельной точкой исходной последовательности $\{x_n\}$).

2. Расширение понятий предельной точки и верхнего и нижнего пределов. Аналогом теоремы Больцано — Вейерштрасса для неограниченной последовательности является следующее утверждение.

Лемма 2. *Из всякой неограниченной последовательности можно выделить бесконечно большую подпоследовательность (и, в частности, бесконечно большую подпоследовательность, все элементы которой имеют определенный знак).*

Доказательство. Прежде всего заметим, что если у неограниченной последовательности отбросить любое конечное число первых ее элементов, то после такого отбрасывания получится снова неограниченная последовательность **. Пусть $\{x_n\}$ — произвольная неограниченная последовательность. Тогда найдется элемент x_{k_1} этой последовательности, удовлетворяющий неравенству $|x_{k_1}|>1$. Учитывая, что последовательность $\{x_n\}$, рассматриваемая с номера k_1+1, также является неограниченной, мы получим, что найдется элемент этой последовательности x_{k_2}, удовлетворяющий неравенству $|x_{k_2}|>2$ при $k_2>k_1$. Продолжая эти рассуждения далее, мы получим, что для любого номера n найдется элемент x_{k_n}, удовлетворяющий неравенству $|x_{k_n}|>n$ при $k_n>k_{n-1}$.

---

* Таковой является последовательность

$$\underbrace{a_1}, \underbrace{a_1, a_2}, \underbrace{a_1, a_2, a_3}, \ldots .$$

** Ибо предположение о том, что это не так, приводит к противоречию с требованием неограниченности исходной последовательности.

Очевидно, что построенная нами подпоследовательность x_{k_1}, $x_{k_2}, \ldots, x_{k_n}, \ldots$ является бесконечно большой. Замечая, что эта подпоследовательность заведомо содержит бесконечно много либо положительных, либо отрицательных членов, мы может выделить из нее бесконечно большую подпоследовательность, все элементы которой имеют определенный знак. Лемма доказана.

Из леммы 2 и из теоремы Больцано — Вейерштрасса вытекает следующее утверждение.

Лемма 3. *Из совершенно произвольной последовательности можно выделить либо сходящуюся подпоследовательность, либо бесконечно большую подпоследовательность, все элементы которой имеют определенный знак.*

Лемма 3 естественно приводит к идее расширения понятия предельной точки последовательности. Договоримся формально дополнить введенные выше конечные предельные точки последовательности еще двумя возможными предельными точками $+\infty$ и $-\infty$.

Будем говорить, что $+\infty$ $[-\infty]$ является предельной точкой последовательности $\{x_n\}$, если из этой последовательности можно выделить бесконечно большую подпоследовательность, все элементы которой положительны [отрицательны].

При таком расширении понятия предельной точки из леммы 3 вытекает следующее утверждение: *у совершенно произвольной последовательности существует хотя бы одна предельная точка* *.

Естественно, считая, что $+\infty$ и $-\infty$ связаны с любым конечным вещественным числом x соотношением $-\infty < x < +\infty$, убедимся в том, что *у совершенно произвольной последовательности $\{x_n\}$ существуют верхний и нижний пределы (т. е. существуют наибольшая и наименьшая предельные точки).*

Ради определенности остановимся на доказательстве существования верхнего предела.

В силу теоремы 3.16 достаточно рассмотреть лишь случай, когда последовательность $\{x_n\}$ не является ограниченной.

Если при этом последовательность $\{x_n\}$ не является ограниченной сверху, то из нее можно выделить бесконечно большую последовательность, все элементы которой положительны, и поэтому $+\infty$ является предельной точкой, а значит, и верхним пределом последовательности $\{x_n\}$.

Остается рассмотреть случай, когда неограниченная последовательность $\{x_n\}$ является ограниченной сверху, т. е. когда существует вещественное число M такое, что все элементы x_n последовательности удовлетворяют неравенству $x_n \leqslant M$. Так как при этом последовательность $\{x_n\}$ не является ограниченной снизу, то из нее можно выделить бесконечно большую подпоследовательность,

* Либо конечная, либо равная $+\infty$ или $-\infty$.

все элементы которой отрицательны, т. е. $-\infty$ является предельной точкой такой последовательности.

Если при этом указанная последовательность $\{x_n\}$ не имеет ни одной конечной предельной точки, то единственная предельная точка $-\infty$ и является верхним пределом этой последовательности.

Если же при этом у указанной последовательности есть хотя бы одна конечная предельная точка x_0, то, фиксировав некоторое $\varepsilon>0$, мы выделим из этой последовательности подпоследовательность тех ее элементов x_n, которые удовлетворяют неравенствам* $x_0-\varepsilon<x_n\leqslant M$.

Выделенная подпоследовательность ограничена, и по теореме 3.16 у нее существует наибольшая предельная точка, которая является наибольшей предельной точкой (т. е. верхним пределом) и всей последовательности $\{x_n\}$. Существование у совершенно произвольной последовательности верхнего предела доказано. Аналогично доказывается, что у совершенно произвольной последовательности существует нижний предел.

В заключение заметим, что почти все понятия и утверждения, установленные нами в этом и в предыдущем пунктах, переносятся на случай *произвольного числового множества* $\{x\}$, *имеющего бесконечное число элементов.*

*Точку a бесконечной прямой $(-\infty, +\infty)$ назовем **предельной точкой** такого множества, если в любой ε-окрестности точки a содержится бесконечно много элементов этого множества.*

Наибольшую и наименьшую предельные точки множества $\{x\}$ назовем соответственно **верхней** и **нижней предельными точками** этого множества.

Повторяя рассуждения теоремы 3.16 с заменой термина «последовательность $\{x_n\}$» термином «множество $\{x\}$, содержащее бесконечное число элементов», мы придем к следующему утверждению: *у всякого ограниченного множества $\{x\}$, имеющего бесконечное число элементов, существуют верхняя и нижняя предельные точки (и, в частности, существует хотя бы одна предельная точка).*

Из этого утверждения вытекает следующее обобщение теоремы Больцано — Вейерштрасса: *из элементов всякого ограниченного множества $\{x\}$, имеющего бесконечное число элементов, можно выделить сходящуюся подпоследовательность***.

Как и для случая последовательности, удобно расширить понятие предельной точки и считать, что $+\infty$ $[-\infty]$ *является **предельной точкой** множества $\{x\}$, если из элементов этого мно-*

* Заметим, что $x_0\leqslant M$, ибо все элементы x_n последовательности $\{x_n\}$ удовлетворяют неравенству $x_n\leqslant M$. Далее заметим, что в силу того, что x_0 — предельная точка, существует бесконечно много элементов x_n последовательности $\{x_n\}$, удовлетворяющих неравенству $x_0-\varepsilon\leqslant x_n\leqslant M$.

** Любые два элемента которой являются различными элементами множества $\{x\}$.

жества можно выделить бесконечно большую последовательность, состоящую из положительных [отрицательных] чисел.

Эта формализация позволяет нам утверждать, что *у совершенно произвольного числового множества* $\{x\}$, *имеющего бесконечное число элементов, существуют хотя бы одна предельная точка, а также верхняя и нижняя предельные точки.*

3. Критерий Коши * сходимости последовательности. При изучении вопроса о сходимости последовательности $\{x_n\}$ с помощью определения сходящейся последовательности приходится оценивать разность $x_n - a$ элементов последовательности и ее предполагаемого предела a.

Иными словами, приходится предугадывать величину предела a этой последовательности.

В этом пункте мы установим «внутренний» критерий сходимости последовательности, позволяющий сделать заключение о ее сходимости лишь по величине ее элементов и не использующий величины предполагаемого предела этой последовательности. Для установления такого критерия введем понятие фундаментальной последовательности.

Определение. *Последовательность* $\{x_n\}$ *называется фундаментальной, если для любого положительного числа* ε *найдется номер* N *такой, что для всех номеров* n, *удовлетворяющих условию* $n \geqslant N$, *и для любого натурального* p $(p=1, 2, \ldots)$ *справедливо неравенство*

$$|x_{n+p} - x_n| < \varepsilon.$$

Установим два важных свойства любой фундаментальной последовательности.

Свойство 1. *Для любого положительного числа* ε *найдется элемент фундаментальной последовательности* x_N *такой, что в* ε*-окрестности этого элемента* x_N *находятся все элементы* x_n *этой последовательности с номерами* n, *удовлетворяющими условию* $n \geqslant N$.

Другими словами, для любого $\varepsilon > 0$ найдется элемент фундаментальной последовательности x_N, вне ε-окрестности которого лежит не более чем конечное число элементов этой последовательности.

Для доказательства этого свойства следует фиксировать произвольное положительное число ε и взять в определении фундаментальной последовательности номер n равным N. Мы получим при этом, что для любого натурального p $(p=1, 2, \ldots)$ элементы фундаментальной последовательности удовлетворяют неравенству

$$|x_{N+p} - x_N| < \varepsilon$$

или, что то же самое, неравенству

$$x_N - \varepsilon < x_{N+p} < x_N + \varepsilon.$$

* Огюстен Луи Коши — французский математик (1789—1857).

Так как p — любое натуральное число, то последние неравенства и означают, что все элементы фундаментальной последовательности, номер которых не меньше N, находятся в интервале $(x_N-\varepsilon,\ x_N+\varepsilon)$, т. е. в ε-окрестности x_N. Свойство 1 доказано.

Свойство 2. *Фундаментальная последовательность является ограниченной.*

Доказательство. Фиксируем некоторое $\varepsilon>0$. Так как последовательность $\{x_n\}$ является фундаментальной, то для этого ε (в силу свойства 1) найдется элемент x_N такой, что все элементы x_n с номерами $n\geqslant N$ удовлетворяют неравенству

$$x_N-\varepsilon<x_n<x_N+\varepsilon.$$

Обозначим теперь через A наибольшее из следующих $(N+1)$ чисел: $|x_1|$, $|x_2|, \ldots, |x_{N-1}|$, $|x_N-\varepsilon|$, $|x_N+\varepsilon|$. Тогда, очевидно, для всех номеров n будет справедливо неравенство $|x_n|\leqslant A$, которое и означает ограниченность последовательности $\{x_n\}$. Свойство 2 доказано.

Докажем теперь следующую вспомогательную теорему.

Теорема 3.17. *Для того чтобы последовательность $\{x_n\}$ была сходящейся, необходимо и достаточно, чтобы она была ограниченной и ее верхний и нижний пределы $\overline{x}$ и $\underline{x}$ совпадали между собой.*

Доказательство. 1) *Необходимость.* Пусть последовательность $\{x_n\}$ сходится. Тогда она ограничена (в силу теоремы 3.8) и имеет единственную предельную точку (в силу леммы 1 из п. 1 этого параграфа). Это и означает, что ее верхний и нижний пределы $\overline{x}$ и $\underline{x}$ совпадают между собой.

2) *Достаточность.* Пусть последовательность $\{x_n\}$ ограничена (при этом она в силу теоремы 3.16 имеет верхний предел $\overline{x}$ и нижний предел $\underline{x}$), и пусть $\overline{x}=\underline{x}$. Положим $x=\overline{x}=\underline{x}$. В силу следствия 1 из теоремы 3.16 для любого $\varepsilon>0$ в интервале $(x-\varepsilon,\ x+\varepsilon)$ лежат элементы последовательности $\{x_n\}$, начиная с некоторого номера. В силу определения 3 сходящейся последовательности (см. п. 4 § 1) это и означает, что последовательность $\{x_n\}$ сходится к пределу x. Теорема 3.17 доказана.

Докажем теперь следующую важнейшую теорему.

Основная теорема 3.18 (критерий Коши сходимости последовательности). *Для того чтобы последовательность $\{x_n\}$ была сходящейся, необходимо и достаточно, чтобы она была фундаментальной.*

Доказательство. 1) *Необходимость.* Пусть последовательность $\{x_n\}$ сходится к некоторому пределу x. Докажем, что эта последовательность является фундаментальной. Фиксируем произвольное положительное число ε. Так как последовательность $\{x_n\}$ сходится к пределу x, то для положительного числа $\varepsilon/2$ найдется номер N такой, что при всех $n\geqslant N$

$$|x_n - x| < \varepsilon/2. \tag{3.54}$$

Если p — любое натуральное число, то при всех $n \geqslant N$ и подавно будет справедливо неравенство

$$|x_{n+p} - x| < \varepsilon/2 \tag{3.55}$$

(ибо при $n \geqslant N$ заведомо будет справедливо неравенство $n+p \geqslant N$).

Так как модуль суммы двух чисел не превосходит суммы их модулей, то из неравенств (3.54) и (3.55) мы получим, что при всех $n \geqslant N$ и для любого натурального p

$$|x_{n+p} - x_n| = |[x_{n+p} - x] + [x - x_n]| \leqslant |x_{n+p} - x| + |x_n - x| < \varepsilon,$$

а это и означает фундаментальность последовательности $\{x_n\}$.

2) *Достаточность.* Пусть последовательность $\{x_n\}$ является фундаментальной. Требуется доказать, что эта последовательность является сходящейся. В силу теоремы 3.17 достаточно доказать, что последовательность $\{x_n\}$ ограничена и что ее верхний и нижний пределы $\bar{x}$ и $\underline{x}$ совпадают между собой.

Ограниченность любой фундаментальной последовательности уже установлена нами выше (см. свойство 2). Остается доказать, что для любой фундаментальной последовательности $\{x_n\}$ верхний предел $\bar{x}$ и нижний предел $\underline{x}$ совпадают. Фиксируем произвольное положительное число ε. В силу свойства 1 фундаментальной последовательности найдется элемент этой последовательности x_N такой, что вне ε-окрестности этого элемента, т. е. вне интервала $(x_N - \varepsilon,\ x_N + \varepsilon)$ лежит не более чем конечное число элементов последовательности $\{x_n\}$. Но тогда в силу следствия 2 из теоремы 3.16 интервал $(\underline{x}, \bar{x})$ обязан содержаться в интервале $(x_N - \varepsilon,\ x_N + \varepsilon)$ и, в частности, должно быть справедливо неравенство $\bar{x} - \underline{x} \leqslant (x_N + \varepsilon) - (x_N - \varepsilon) = 2\varepsilon$.

Так как, кроме того, $\bar{x} \geqslant \underline{x}$, то для любого $\varepsilon > 0$ будут справедливы неравенства $0 \leqslant \bar{x} - \underline{x} \leqslant 2\varepsilon$. В силу произвольности ε из этих неравенств вытекает, что $\bar{x} - \underline{x} = 0$ *, т. е. $\bar{x} = \underline{x}$. Критерий Коши полностью доказан.

Применим критерий Коши для установления расходимости последовательности $\{x_n\}$ с элементами $x_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$.

Заметим, что если для любого номера n натуральное число p взять равным n, то мы получим, что для всех номеров n

$$|x_{n+p} - x| = |x_{2n} - x_n| = \left(1 + \frac{1}{2} + \ldots + \frac{1}{2n}\right) -$$

* В самом деле, если бы разность $\bar{x} - \underline{x}$ равнялась положительному числу α, то, взяв в качестве ε число $\alpha/3$, мы бы получили противоречие с неравенством $\bar{x} - \underline{x} \leqslant 2\varepsilon$.

$$-\left(1+\frac{1}{2}+\ldots+\frac{1}{n}\right)=\underline{\frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{2n}}\geqslant$$
$$\geqslant n\cdot\frac{1}{2n}=\frac{1}{2},$$

ибо подчеркнутая сумма содержит n слагаемых, наименьшее из которых равно $\frac{1}{2n}$.

Таким образом, для положительного числа $\varepsilon=\frac{1}{2}$ не существует номера N такого, что при всех $n\geqslant N$ и для любого натурального p справедливо неравенство $|x_{n+p}-x_n|<\varepsilon$. Это означает, что рассматриваемая последовательность не является фундаментальной и (в силу критерия Коши) расходится.

В качестве второго примера применим критерий Коши для установления сходимости последовательности $\{x_n\}$ с элементами $x_n=1+q+\ldots+q^n$, где q — любое число из интервала $0<q<1$.

Для любого номера n и любого натурального числа p ($p=1, 2, \ldots$) справедливо неравенство

$$|x_{n+p}-x_n|=(1+q+\ldots+q^{n+p})-(1+q+\ldots+q^n)=$$
$$=q^{n+1}+q^{n+2}+\ldots+q^{n+p}=\frac{q^{n+1}-q^{n+1+p}}{1-q}<\frac{q^{n+1}}{1-q}. \tag{3.56}$$

Фиксируем произвольное положительное число ε. Так как при $0<q<1$ последовательность $\{q^n\}$ является бесконечно малой, то для положительного числа $\varepsilon(1-q)$ найдется номер N такой, что при всех $n\geqslant N$ справедливо неравенство

$$q^n<\varepsilon(1-q). \tag{3.57}$$

Из неравенств (3.56) и (3.57) вытекает, что при всех $n\geqslant N$ и для любого натурального p

$$|x_{n+p}-x_n|<\frac{\varepsilon(1-q)}{1-q}=\varepsilon,$$

а это означает, что рассматриваемая последовательность является фундаментальной и (в силу критерия Коши) сходится.

§ 4. ПРЕДЕЛ (ИЛИ ПРЕДЕЛЬНОЕ ЗНАЧЕНИЕ) ФУНКЦИИ

Перейдем теперь к изучению другой более сложной формы операции предельного перехода, основанной на понятии предела (или предельного значения) функции. Но прежде всего мы должны уточнить сами понятия переменной величины и функции.

1. Понятия переменной величины и функции. Как мы уже видели в гл. 1, к понятию функции приводит изучение двух переменных величин, изменение которых взаимообусловлено. Поэтому естественно начать с уточнения понятия переменной величины.

Рассмотрение реальных физических переменных величин приводит нас к выводу, что эти величины не всегда могут принимать произвольные значения. Так, скорость материальной точки не может быть больше $3\cdot10^{10}$ см/с (т. е. скорости света в пустоте), температура тела не может быть меньше $-273°$, смещение материальной точки, совершающей гармонические колебания по закону $y=A\cos(\omega t+\delta)$, может принимать значения только из сегмента $[-A, +A]$.

Отвлекаясь от конкретных физических свойств наблюдаемых в природе переменных величин, мы приходим к понятию математической переменной величины, характеризуемой только численными значениями, которые она может принимать *.

Множество $\{x\}$ всех значений, которые может принимать данная переменная величина x, называется областью изменения данной переменной величины. Переменная величина считается заданной, если задана область ее изменения.

В дальнейшем мы, как правило, будем обозначать переменные величины малыми латинскими буквами x, y, t, ... , а области изменения этих переменных величин соответственно символами $\{x\}$, $\{y\}$, $\{t\}$, ...

Перейдем теперь к уточнению понятия функции.

Пусть задана переменная величина x, имеющая областью изменения некоторое множество $\{x\}$.

Если каждому значению переменной x из множества $\{x\}$ ставится в соответствие по известному закону некоторое число y, то говорят, что на множестве $\{x\}$ задана функция $y=y(x)$ или $y=f(x)$.

При этом переменная x называется аргументом или независимой переменной, множество $\{x\}$ называется областью задания функции, а то число y, которое соответствует данному значению x, называется частным значением функции в точке x. Совокупность всех частных значений образует вполне определенное множество $\{y\}$, которое называют либо областью изменения функции, либо множеством всех значений функции.

В обозначении $y=f(x)$ букву f часто называют характеристикой функции.

Для обозначения аргумента, функции и ее характеристики могут употребляться различные символы.

Остановимся на примерах функций.

1°. $y=\sqrt{4-x^2}$. Эта функция задана на сегменте $-2\leqslant x\leqslant 2$ (при этом выражение под знаком корня является неотрицатель-

* Понятие переменной величины также относится к числу начальных математических понятий.

ным), а множеством всех ее значений является сегмент $0 \leqslant y \leqslant 2$ (рис. 3.4).

2°. Так называемая **функция Дирихле***, которая определяется так:

$$y = D(x) = \begin{cases} 0, & \text{если } x \text{ — иррациональное число,} \\ 1, & \text{если } x \text{ — рациональное число.} \end{cases}$$

Эта функция задана на бесконечной прямой $-\infty < x < +\infty$, а множество всех ее значений состоит из двух точек 0 и 1.

3°.

$$y = \operatorname{sgn} x = \begin{cases} +1, & \text{если } x > 0, \\ 0, & \text{если } x = 0, \\ -1, & \text{если } x < 0. \end{cases}$$

(Термин «sgn» происходит от латинского слова signum — знак.) Читается: «y равно сигнум x». Эта функция задана на всей бесконечной прямой $-\infty < x < +\infty$, а множество всех ее значений состоит из трех точек $y = -1$, $y = 0$ и $y = 1$ (рис. 3.5).

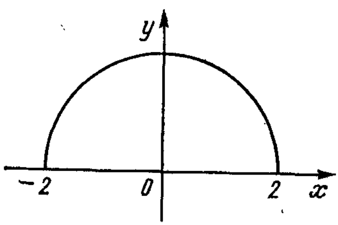


Рис. 3.4

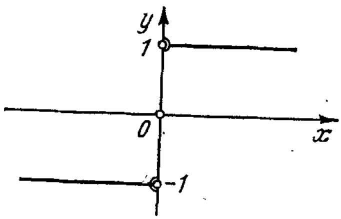


Рис. 3.5

4°. $y = [x]$, или $y = E(x)$, где символ $[x]$ или $E(x)$ обозначает целую часть числа x или, точнее, наибольшее целое число, не превосходящее x. Читается: «y равно антье x» (от французского слова entier — целый). Эта функция задана на всей бесконечной прямой $-\infty < x < +\infty$, а множеством всех ее значений является множество всех целых чисел (рис. 3.6).

5°. $y = n!$. Эта функция задана на множестве всех натуральных чисел $n = 1, 2, 3, \ldots$. Множеством всех значений этой функции является множество натуральных чисел вида $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$, где $n = 1, 2, 3, \ldots$ (рис. 3.7).

Часто закон, устанавливающий соответствие между множеством всех значений аргумента и множеством всех значений функции, задается посредством формул. Такой способ задания функции называется **аналитическим**.

---

* Петер Густав Лежен-Дирихле — немецкий математик (1805—1859).

При этом следует подчеркнуть, что функция может задаваться разными формулами на разных участках области своего задания.

Например, функция

$$y = \begin{cases} -x & \text{при } x < 0, \\ x^2 & \text{при } x \geqslant 0 \end{cases}$$

задана аналитическим способом на всей бесконечной прямой $-\infty < x < +\infty$ (рис. 3.8).

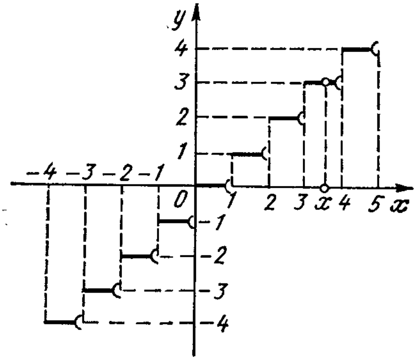


Рис. 3.6

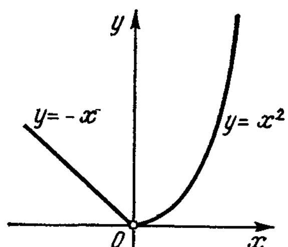


Рис. 3.8

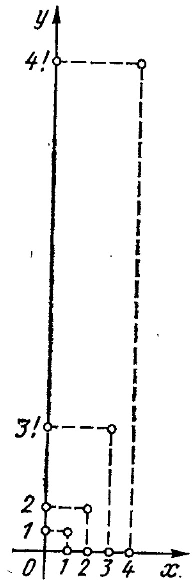


Рис. 3.7

Весьма распространенным способом задания функции является так называемый табличный способ, заключающийся в задании таблицы отдельных значений аргумента и соответствующих им значений функции. При таком способе задания можно приближенно вычислить не содержащиеся в таблице значения функции, отвечающие промежуточным значениям аргумента. Для этого применяется метод интерполяции, заключающийся в замене функции между ее соседними табличными значениями какой-либо функцией простой природы (например, линейной или квадратичной). Примером табличного способа задания функции может служить расписание движения поезда, которое определяет местоположение поезда в отдельные моменты времени. Интерполяция позволяет приближенно определить местоположение поезда в любой промежуточный момент времени.

В практике физических измерений весьма распространенным является и еще один способ задания функции — так называемый графический способ, при котором соответствие между аргументом и функцией задается посредством графика (снимаемого, например, на осциллографе).

2. Предел функции по Гейне и по Коши. Пусть функция $y=f(x)$ определена на некотором бесконечном множестве $\{x\}$, и пусть a — точка бесконечной прямой $(-\infty, +\infty)$, быть может и не принадлежащая множеству $\{x\}$, но обладающая тем свойством, что в любой δ-окрестности этой точки a имеются точки множества $\{x\}$, отличные от a *.

Например, множеством $\{x\}$ может служить интервал (a, b); в этом случае точка a, являясь граничной точкой интервала, ему не принадлежит, но в любой δ-окрестности a содержатся точки указанного интервала.

Другим примером множества $\{x\}$, на котором задана функция $f(x)$, может служить множество всех рациональных чисел, принадлежащих интервалу $(a-\delta, a+\delta)$ с выкинутой точкой a.

Заметим, кстати, что *при любом $\delta>0$ интервал $(a-\delta, a+\delta)$, из которого выкинута точка a, принято называть проколотой δ-окрестностью точки a.*

Определение 1 (предел функции по Гейне**). *Число b называется пределом (или предельным значением) функции $y=f(x)$ в точке a (или при $x\to a$), если для любой последовательности значений аргумента $x_1, x_2, \ldots, x_n, \ldots$, сходящейся к a и состоящей из чисел x_n, отличных от a, соответствующая последовательность значений функции $f(x_1), f(x_2), \ldots, f(x_n) \ldots$ сходится к числу b.*

Определение 1* (предел функции по Коши). *Число b называется пределом (или предельным значением) функции $y=f(x)$ в точке a (или при $x\to a$), если для любого положительного числа ε найдется отвечающее ему положительное число δ такое***, что для всех значений аргумента x, удовлетворяющих условию $0<|x-a|<\delta$, справедливо неравенство*

$$|f(x)-b|<\varepsilon. \qquad (3.58)$$

Для обозначения предельного значения функции $y=f(x)$ в точке a используют следующую символику:

$$\lim_{x\to a} f(x) = b \quad \text{или} \quad f(x)\to b \quad \text{при } x\to a.$$

---

* Это означает, что a является предельной точкой множества $\{x\}$.

** Генрих Эдуард Гейне — немецкий математик (1821—1881).

*** Так как δ зависит от ε, то иногда пишут: $\delta=\delta(\varepsilon)$.

Прежде чем доказывать эквивалентность определений 1 и 1*, сделаем несколько замечаний, разъясняющих смысл этих определений.

Замечание 1. Подчеркнем важность фигурирующего в определении 1 требования, обязывающего элементы последовательности значений аргумента x_n быть отличными от a, и аналогичного требования в определении 1*, обязывающего брать значения аргумента x, удовлетворяющие условию $0<|x-a|$, т. е. отличные от a. Это требование вызвано уже тем, что функция $y=f(x)$ может быть не определена в точке a. Отсутствие этого требования сделало бы невозможным использование предела функции для определения производной функции. В самом деле, из гл. 1 нам известно, что производная $f'(a)$ функции $f(x)$ в точке a представляет собой предел при $x \to a$ следующей функции:

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

Очевидно, что эта функция $F(x)$ не определена в точке a и это вызвано существом дела.

Замечание 2. Особо подчеркнем, что множество $\{x\}$, на котором задана функция $f(x)$, вовсе не обязано сплошь покрывать некоторую проколотую δ-окрестность точки a. От этого множества $\{x\}$ требуется только, чтобы оно имело хотя бы один элемент в любой проколотой δ-окрестности точки a. Примером множества $\{x\}$ может служить множество всех элементов последовательности $\left\{\frac{1}{n}\right\}$, лежащих в фиксированной δ-окрестности точки $a=0$.

Замечание 3. Заметим, что фигурирующее в определении 1* условие $0<|x-a|<\delta$ эквивалентно соотношениям $a-\delta<x<a+\delta$, $x\neq a$, т. е. означает, что x принадлежит проколотой δ-окрестности точки a. Аналогично, фигурирующее в определении 1* неравенство (3.58) эквивалентно неравенствам $b-\varepsilon<f(x)<b+\varepsilon$, т. е. означает, что $f(x)$ принадлежит ε-окрестности b.

Замечание 4. Привлекая идею приближения функции $f(x)$ в окрестности точки a с наперед заданной точностью ε, мы можем следующим образом перефразировать определение 1* предела функции по Коши: *число b называется предельным значением функции $f(x)$ в точке a, если для любой наперед заданной точности $\varepsilon>0$ можно указать такую δ-окрестность точки a, что для всех значений аргумента x, отличных от a и принадлежащих указанной δ-окрестности точки a, число b приближает значение функции $f(x)$ с точностью* ε (рис. 3.9).

Замечание 5. Отметим, что функция $f(x)$ может иметь в точке a только один предел. В самом деле, для определения предела функции по Гейне это вытекает из единственности предела последовательности $\{f(x_n)\}$, а для определения предела функции по Коши это вытекает из устанавливаемой ниже эквивалентности этого предела пределу функции по Гейне.

Докажем теперь следующую важную теорему.

Теорема 3.19. *Определения 1 и 1* предела функции по Гейне и по Коши являются эквивалентными.*

Доказательство. 1) Пусть сначала число b является пределом функции $y=f(x)$ в точке a по Коши. Докажем, что это же число b является пределом функции $y=f(x)$ в точке a и по Гейне. Пусть $\{x_n\}$ — любая сходящаяся к a последовательность значений аргумента, все элементы которой отличны от a. Требуется доказать, что соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к b.

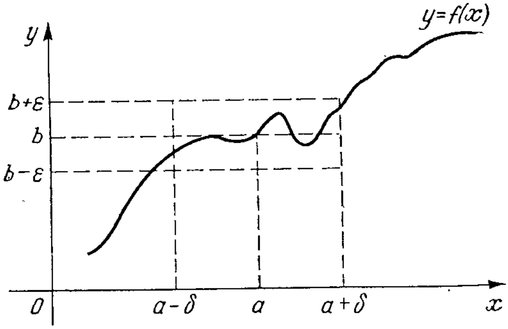


Рис. 3.9

Фиксируем произвольное положительное число ε и по нему положительное число δ, которое в силу определения предела функции по Коши гарантирует справедливость неравенства (3.58) для всех значений x, для которых $0<|x-a|<\delta$.

В силу сходимости последовательности $\{x_n\}$ к a для указанного положительного числа δ найдется номер N такой, что при всех $n\geqslant N$ справедливо неравенство $|x_n-a|<\delta$. Поскольку $x_n\neq a$ для всех номеров n, то при всех $n\geqslant N$ справедливы неравенства $0<|x_n-a|<\delta$ и, значит, в силу определения предела функции по Коши при всех $n\geqslant N$ справедливо неравенство $|f(x_n)-b|<\varepsilon$. Это и означает, что последовательность $\{f(x_n)\}$ сходится к числу b.

2) Пусть теперь число b является пределом функции $y=f(x)$ в точке a по Гейне. Докажем, что это же число b является пределом функции $y=f(x)$ в точке a и по Коши. Предположим, что это не так. Тогда для некоторого положительного числа ε и для сколь угодно малого положительного числа δ найдется хотя бы одно значение аргумента x такое, что $0<|x-a|<\delta$, но $|f(x)-b|\geqslant\varepsilon$.

Таким образом, мы можем взять последовательность $\delta_n=\dfrac{1}{n}$ $(n=1, 2, \ldots)$ и утверждать, что для каждого ее элемента $\delta_n=\dfrac{1}{n}$ найдется хотя бы одно значение аргумента x_n такое, что

$$0<|x_n-a|<\frac{1}{n}, \quad \text{но } |f(x_n)-b|\geqslant\varepsilon. \tag{3.59}$$

Левое из неравенств (3.59) означает, что последовательность $\{x_n\}$ сходится к a и состоит из чисел, отличных от a. Но в таком случае согласно определению предела по Гейне соответствующая последовательность значений функции $\{f(x_n)\}$ обязана сходиться к числу b, а этому противоречит правое из неравенств (3.59), справедливое для всех номеров n. Полученное противоречие доказывает теорему.

Приведем примеры функций, как обладающих, так и не обладающих в данной точке a предельным значением.

1°. Функция $f(x)=c=\text{const}$ имеет равный c предел в каждой точке a бесконечной прямой. В самом деле, для любого значения аргумента x разность $f(x)-c$ равна нулю, и поэтому $|f(x)-c|<\varepsilon$ для любого $\varepsilon>0$ и для всех значений аргумента (в данном случае для любого $\varepsilon>0$ в определении предела по Коши можно брать в качестве δ любое положительное число).

2°. Функция $f(x)=x$ в любой точке a бесконечной прямой имеет предел, равный a. В самом деле, для этой функции последовательности значений аргумента и соответствующих значений функции тождественны, и поэтому, если последовательность $\{x_n\}$ сходится к a, то и последовательность $\{f(x_n)\}$ также сходится к a.

3°. Функция Дирихле $D(x)$, значения которой в рациональных точках равны единице, а в иррациональных точках — нулю, не имеет предела ни в одной точке a бесконечной прямой. Это вытекает из того, что для сходящейся к a последовательности рациональных значений аргумента предел последовательности соответствующих значений функции равен единице, в то время как для сходящейся к a последовательности иррациональных значений аргумента предел последовательности соответствующих значений функции равен нулю.

Введем теперь понятие одностороннего (т. е. правого или левого) предела функции в данной точке a. Для этого нам прежде всего следует уточнить характер того множества $\{x\}$, на котором задана функция $f(x)$. Мы теперь потребуем, чтобы это *множество $\{x\}$ для любого $\delta>0$ имело хотя бы один элемент, принадлежащий интервалу $(a, a+\delta)$ [интервалу $(a-\delta, a)$]*.

Определение 2 (правый [левый] предел функции по Гейне). *Число b называется правым пределом [левым пределом] функции $y=f(x)$ в точке a, если для любой последовательности значений аргумента $\{x_n\}$, сходящейся к a и состоящей из чисел, бо́льших a [меньших a], соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к числу b.*

Определение 2* (правый [левый] предел функции по Коши). *Число b называется правым пределом [левым пределом] функции $y=f(x)$ в точке a, если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех значений аргумента x,* *удовлетворяющих условию* $a<x<a+\delta$ [*условию* $a-\delta<x<a$], *справедливо неравенство* (3.58).

Для обозначения правого [левого] предела функции $f(x)$ в точке a используют следующую символику:

$$\lim_{x\to a+0} f(x)=b \quad [\lim_{x\to a-0} f(x)=b]$$

или более краткую символику

$$f(a+0)=b \quad [f(a-0)=b].$$

В полной аналогии с теоремой 3.19 доказывается *эквивалентность определений* 2 *и* 2*: следует лишь во всех проведенных при доказательстве этой теоремы рассуждениях брать значения аргумента x и элементы последовательности $\{x_n\}$ бо́льшими числа a [меньшими числа a].

В качестве примера рассмотрим функцию

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

Эта функция имеет в точке $a=0$ как правый, так и левый пределы, причем $\operatorname{sgn}(0+0)=+1$, $\operatorname{sgn}(0-0)=-1$. В самом деле, для любой сходящейся к $a=0$ последовательности $\{x_n\}$, состоящей из чисел, больших нуля, соответствующая последовательность $\{\operatorname{sgn} x_n\}$ сходится к $+1$, а для любой сходящейся к $a=0$ последовательности $\{x_n\}$, состоящей из чисел, меньших нуля, соответствующая последовательность $\{\operatorname{sgn} x_n\}$ сходится к -1.

Из проведенных рассуждений вытекает, что у рассматриваемой функции $y=\operatorname{sgn} x$ не существует в точке $a=0$ предела.

Итак, функция $y=\operatorname{sgn} x$ не имеет в точке $a=0$ предела, но имеет в этой точке правый предел, равный $+1$, и левый предел, равный -1. Тот факт, что правый и левый пределы этой функции не равны друг другу, не является случайным, ибо справедливо следующее утверждение: *если функция $f(x)$ имеет в точке a как правый, так и левый пределы и если эти односторонние пределы равны одному и тому же числу b, то эта функция имеет в точке a предел, равный b**.

Для доказательства этого утверждения достаточно воспользоваться определениями 1* и 2* и учесть, что если неравенство (3.58) справедливо для значений аргумента x, удовлетворяющих условиям $a<x<a+\delta$ и $a-\delta<x<a$, то неравенство (3.58) справедливо и для всех значений аргумента x, удовлетворяющих условию $0<|x-a|<\delta$.

* Очевидно, справедливо и обратное утверждение: если функция $f(x)$ имеет в точке a равный b предел, то как правый, так и левый пределы $f(x)$ в точке a существуют и оба равны b.

Сформулируем теперь понятие предела функции при $x \to \infty$. Для введения этого понятия следует потребовать, чтобы *множество $\{x\}$, на котором задана функция $y=f(x)$, для любого $\delta>0$ имело хотя бы один элемент, лежащий вне сегмента* $[-\delta, +\delta]$.

Определение 3 (предел функции при $x \to \infty$ по Гейне). *Число b называется **пределом** (или **предельным значением**) **функции** $y=f(x)$ **при** $x \to \infty$, если для любой бесконечно большой последовательности значений аргумента $\{x_n\}$ соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к числу b.*

Определение 3* (предел функции при $x \to \infty$ по Коши). *Число b называется **пределом** (или **предельным значением**) **функции** $y=f(x)$ **при** $x \to \infty$, если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех значений аргумента x, удовлетворяющих условию $|x|>\delta$, справедливо неравенство* (3.58).

Для обозначения предела функции $y=f(x)$ при $x \to \infty$ используют следующий символ:

$$\lim_{x \to \infty} f(x) = b.$$

В полной аналогии с теоремой 3.19 доказывается *эквивалентность определений* 3 *и* 3*. Следует лишь в рассуждениях, использованных при доказательстве этой теоремы, всюду заменить сходящуюся последовательность значений аргумента $\{x_n\}$ бесконечно большой последовательностью значений аргумента $\{x_n\}$, а неравенство $0<|x-a|<\delta$ заменить неравенством $|x|>\delta$.

Примером функции, имеющей предел при $x \to \infty$, может служить функция $f(x)=\dfrac{1}{x}$ $(x \neq 0)$, В самом деле, для любой бесконечно большой последовательности значений аргумента $\{x_n\}$ соответствующая последовательность значений функции $f(x_n)=1/x_n$ (в силу теоремы 3.6) является бесконечно малой, т. е. имеет своим пределом число $b=0$. Значит, в силу определения 3

$$\lim_{x \to \infty} \frac{1}{x} = 0.$$

Сформулируем, наконец, понятие предела функции при стремлении x к бесконечности определенного знака. Для введения такого понятия потребуем, чтобы *функция $y=f(x)$ была задана на таком множестве $\{x\}$, которое для любого $\delta>0$ имеет хотя бы один элемент, лежащий правее δ [левее $-\delta$].*

Определение 4 (предел функции при $x \to +\infty$ [$x \to -\infty$] по Гейне). *Число b называется **пределом** (или **предельным значением**) **функции** $y=f(x)$ **при** $x \to +\infty$ [при $x \to -\infty$], если для любой бесконечно большой последовательности значений аргумента $\{x_n\}$, все элементы которой* *положительны [отрицательны], соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к числу b.*

Определение 4* (предел функции при $x\to+\infty$ [$x\to-\infty$] по Коши). *Число b называется пределом (или предельным значением) функции $y=f(x)$ при $x\to+\infty$ [при $x\to-\infty$], если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех значений аргумента x, удовлетворяющих условию $x>\delta$ [$x<-\delta$], справедливо неравенство* (3.58).

Для обозначения введенных понятий используется следующая символика:

$$\lim_{x\to+\infty} f(x)=b \quad [\lim_{x\to-\infty} f(x)=b].$$

*Эквивалентность определений 4 и 4** доказывается по схеме доказательства теоремы 3.19: следует только во всех рассуждениях заменить сходящуюся последовательность значений аргумента $\{x_n\}$ на бесконечно большую последовательность значений аргумента $\{x_n\}$, состоящую из положительных [отрицательных] чисел, а неравенство $0<|x-a|<\delta$ заменить неравенством $x>\delta$ [$x<-\delta$].

Замечание 6. Отметим, что изученное нами в § 1—3 понятие предела числовой последовательности $\{x_n\}$ можно рассматривать как частный случай предела функции при $x\to+\infty$. В самом деле, если взять в качестве $\{x\}$ множество всех натуральных чисел 1, 2, ..., n, ..., а в качестве функции $f(x)$, заданной на этом множестве, ту функцию, которая каждому значению аргумента n ставит в соответствие n-й член последовательности x_n, то определение 4* предела такой функции при $x\to+\infty$ в точности совпадет с определением предела числовой последовательности $\{x_n\}$.

Замечание 7. Естественно, возникает идея связать воедино все введенные нами понятия пределов функции и предел числовой последовательности. В § 5 настоящей главы вводится понятие общего предела функции по базе, включающее в себя как частный случай все введенные нами понятия пределов функции и понятие предела числовой последовательности.

3. Критерий Коши существования предела функции. Ради определенности рассмотрим подробно случай предела функции $y=f(x)$ в точке a, введенного определениями 1 и 1*.

Определение. *Будем говорить, что функция $y=f(x)$ удовлетворяет в точке a условию Коши, если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для любых двух значений аргумента x' и x'', удовлетворяющих условиям*

$$0<|x'-a|<\delta,\ 0<|x''-a|<\delta, \tag{3.60}$$

справедливо неравенство

$$|f(x')-f(x'')|<\varepsilon. \tag{3.61}$$

Теорема 3.20 (критерий Коши существования предела функции в точке a). *Для того чтобы функция $y=f(x)$ имела в точке a конечный предел, необходимо и достаточно, чтобы функция $y=f(x)$ удовлетворяла в точке a условию Коши.*

Доказательство. 1) *Необходимость.* Пусть существует конечный предел $\lim\limits_{x\to a} f(x)=b$. Фиксируем произвольное положительное число ε. В силу определения 1* предела функции по Коши для положительного числа $\varepsilon/2$ найдется положительное число δ такое, что, каковы бы ни были два значения аргумента x' и x'', удовлетворяющие условиям $0<|x'-a|<\delta$, $0<|x''-a|<\delta$, для соответствующих значений функции справедливы неравенства

$$|f(x')-b|<\frac{\varepsilon}{2}, \quad |f(x'')-b|<\frac{\varepsilon}{2} \tag{3.62}$$

Так как модуль суммы двух величин не превосходит суммы их модулей, то в силу неравенств (3.62) мы получим, что

$$|f(x')-f(x'')|=|[f(x')-b]+[b-f(x'')]|\leqslant$$
$$\leqslant|f'(x')-b|+|f(x'')-b|<\varepsilon,$$

а это и означает, что функция $y=f(x)$ удовлетворяет в точке a условию Коши.

2) *Достаточность.* Пусть функция $f(x)$ удовлетворяет в точке a условию Коши. Требуется доказать, что функция $f(x)$ имеет в точке a предел. Пусть $\{x_n\}$ — произвольная последовательность значений аргумента, сходящаяся к a и состоящая из чисел, отличных от a. В силу определения 1 предела по Гейне достаточно доказать, что соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к некоторому числу b и что это число b одно и то же для всех сходящихся к a последовательностей $\{x_n\}$, состоящих из чисел, отличных от a.

Докажем сначала, что для каждой сходящейся к a последовательности $\{x_n\}$ значений аргумента, отличных от a, соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к некоторому пределу. Фиксируем произвольное положительное число ε и по нему отвечающее ему, согласно условию Коши, положительное число δ. В силу сходимости последовательности $\{x_n\}$ к a и в силу условия $x_n\neq a$ для этого $\delta>0$ найдется номер N такой, что $0<|x_n-a|<\delta$ при $n\geqslant N$. Если теперь p — любое натуральное число ($p=1, 2, 3, \ldots$), то тем более $0<|x_{n+p}-a|<\delta$ при $n\geqslant N$ *.

Таким образом, при $n\geqslant N$ и для любого натурального p справедливы два неравенства:

$$0<|x_{n+p}-a|<\delta, \quad 0<|x_n-a|<\delta.$$

* Ибо если $n\geqslant N$, то и подавно $n+p\geqslant N$.

Из этих двух неравенств и из условия Коши вытекает, что при $n \geqslant N$ и для любого натурального p

$$|f(x_{n+p})-f(x_n)|<\varepsilon,$$

а это означает фундаментальность последовательности $\{f(x_n)\}$. В силу критерия Коши сходимости числовой последовательности (см. теорему 3.18) последовательность $\{f(x_n)\}$ сходится к некоторому числу b.

Остается доказать, что для любых двух сходящихся к a последовательностей значений аргумента $\{x_n\}$ и $\{x_n'\}$, все элементы которых отличны от a, соответствующие последовательности значений функции $\{f(x_n)\}$ и $\{f(x_n')\}$ сходятся к одному и тому же пределу. Предположим, что последовательности $\{f(x_n)\}$ и $\{f(x_n')\}$ сходятся к пределам b и b' соответственно. Рассмотрим новую последовательность значений аргумента $x_1, x_1', x_2, x_2', x_3, x_3', \ldots, x_n, x_n', \ldots$, также сходящуюся к a и состоящую из чисел, отличных от a. В силу доказанного выше соответствующая последовательность значений функции $f(x_1), f(x_1'), f(x_2), f(x_2'), \ldots, f(x_n), f(x_n'), \ldots$ обязана сходиться к некоторому пределу b''. Но тогда в силу утверждения, доказанного в начале п. 1 § 3, и любая подпоследовательность этой последовательности обязана сходиться к тому же самому пределу b''. Значит, как подпоследовательность нечетных элементов $f(x_1), f(x_2), \ldots, f(x_n), \ldots$, так и подпоследовательность четных элементов $f(x_1'), f(x_2'), \ldots, f(x_n'), \ldots$ обе сходятся к b''. Отсюда вытекает, что $b=b'=b''$. Теорема полностью доказана.

Аналогично формулируется условие Коши и доказывается критерий Коши и для случаев правого [левого] предела в точке a, предела при $x \to \infty$ и предела при $x \to +\infty$ $[x \to -\infty]$.

При формулировке условия Коши достаточно в приведенном выше определении заменить условия (3.60) для случая правого [левого] предела в точке a условиями

$$a<x'<a+\delta,\ a<x''<a+\delta \quad [a-\delta<x'<a,\ a-\delta<x''<a],$$

для случая предела при $x \to \infty$ условиями

$$|x'|>\delta,\ |x''|>\delta$$

и, наконец, для случая предела при $x \to +\infty$ $[x \to -\infty]$ условиями

$$x'>\delta,\ x''>\delta \quad [x'<-\delta,\ x''<-\delta].$$

Соответствующие критерии Коши доказываются по схеме доказательства теоремы 3.20: следует только во всех рассуждениях понимать под последовательностями значений аргумента $\{x_n'\}$ и $\{x_n''\}$ в случае правого [левого] предела в точке a последовательности, сходящиеся к a и состоящие из чисел, бóльших a [меньших a], в случае предела при $x \to \infty$ бесконечно большие последовательности и, наконец, в случае предела при $x \to +\infty$ $[x \to -\infty]$ бесконечно большие последовательности, состоящие из положительных [отрицательных] чисел.

4. Арифметические операции над функциями, имеющими предел. Справедлива следующая фундаментальная теорема.

Основная теорема 3.21. *Пусть две функции $f(x)$ и $g(x)$ заданы на одном и том же множестве $\{x\}$ и имеют в точке a пределы, соответственно равные b и c. Тогда функции $f(x)+g(x)$, $f(x)-g(x)$, $f(x)\cdot g(x)$ и $f(x)/g(x)$ имеют в точке a пределы, соответственно равные $b+c$, $b-c$, $b\cdot c$, b/c (в случае частного нужно дополнительно требовать, чтобы c было отлично от нуля).*

Доказательство. Пусть $\{x_n\}$ — произвольная сходящаяся к a последовательность значений аргумента, все элементы которой отличны от a. В силу определения 1 предела по Гейне соответствующие последовательности значений функции $\{f(x_n)\}$ и $\{g(x_n)\}$ сходятся к пределам b и c соответственно. Но тогда в силу теорем 3.9—3.12 последовательности $\{f(x_n)+g(x_n)\}$, $\{f(x_n)-g(x_n)\}$, $\{f(x_n)\cdot g(x_n)\}$ и $\left\{\frac{f(x_n)}{g(x_n)}\right\}$ сходятся к пределам $b+c$, $b-c$, $b\cdot c$ и b/c соответственно. Это последнее в силу произвольности последовательности значений аргумента $\{x_n\}$, сходящейся к a, и в силу определения 1 предела по Гейне означает, что функции $f(x)+g(x)$, $f(x)-g(x)$, $f(x)\cdot g(x)$ и $f(x)/g(x)$ имеют в точке a пределы, соответственно равные $b+c$, $b-c$, $b\cdot c$ и b/c. Теорема доказана.

Доказательство соответствующей теоремы для случаев правого [левого] предела в точке a, предела при $x\to\infty$ и предела при $x\to+\infty$ [$x\to-\infty$] проводится по той же схеме. Все отличие состоит в том, что в качестве последовательности значений аргумента $\{x_n\}$ следует взять в случае правого [левого] предела в точке a последовательность, сходящуюся к a и состоящую из чисел, больших a [меньших a], в случае предела при $x\to\infty$ — бесконечно большую последовательность и, наконец, в случае предела при $x\to+\infty$ [$x\to-\infty$] бесконечно большую последовательность, состоящую из положительных [отрицательных] чисел.

Рассмотрим примеры применения теоремы 3.21. Выше в п. 2 мы убедились в том, что для любой точки a бесконечной прямой $\lim\limits_{x\to a} x = a$. Используя теорему 3.21, мы можем утверждать, что

$$\lim_{x\to a} x^2 = \lim_{x\to a} x\cdot\lim_{x\to a} x = a\cdot a = a^2$$

и, вообще, для любого номера n

$$\lim_{x\to a} x^n = a^n.$$

Пусть теперь $P_n(x)=b_0+b_1x+b_2x^2+\ldots+b_nx^n$, где b_0, b_1, ..., b_{n-1}, $b_n\neq 0$ — некоторые постоянные числа. Такая функция $P_n(x)$ называется многочленом степени n. В силу той же теоремы 3.21

$$\lim_{x\to a} P_n(x) = \lim_{x\to a} [b_0 + b_1x + \ldots + b_nx^n] = b_0+b_1a + \ldots + b_na^n = P_n(a)$$

для любой точки a бесконечной прямой.

Итак, *многочлен $P_n(x)$ имеет предел в любой точке a бесконечной прямой, и этот предел равен частному значению этого многочлена в точке a.*

Пусть, наконец, $P_n(x)$ и $Q_m(x)$ — два произвольных многочлена степеней n и m соответственно. Частное $R(x) = \frac{P_n(x)}{Q_m(x)}$ принято называть рациональной дробью. В силу теоремы 3.21 для случая частного

$$\lim_{x\to a} R(x) = \lim_{x\to a} \frac{P_n(x)}{Q_m(x)} \frac{\lim\limits_{x\to a} P_n(x)}{\lim\limits_{x\to a} Q_m(x)} = \frac{P_n(a)}{Q_m(a)} = R(a)$$

в любой точке a, не являющейся корнем многочлена $Q_m(x)$. Таким образом, *рациональная дробь имеет предел в каждой точке a бесконечной прямой, не являющейся корнем ее знаменателя, и этот предел равен частному значению этой дроби в указанной точке a.*

5. Бесконечно малые и бесконечно большие функции. Ради определенности будем рассматривать предел функции в точке a.

Функция $\alpha(x)$ называется бесконечно малой в точке a, если предел этой функции в точке a равен нулю.

Примером бесконечно малой в точке a функции может служить функция $\alpha(x) = (x-a)^n$, где n — любое целое положительное число.

В самом деле, в конце предыдущего пункта мы установили, что многочлен $(x-a)^n$ имеет предел в каждой точке a, причем этот предел равен частному значению этого многочлена в точке $x=a$, т. е. равен нулю.

Заметим, что *если функция $f(x)$ имеет предел в точке a, равный числу b, то функция $\alpha(x)=f(x)-b$ является бесконечно малой в точке a.*

Это вытекает из того, что пределы каждой из функций $f(x)$ и $g(x) \equiv b$ в точке a равны числу b, и из теоремы 3.21 для случая разности $f(x)-g(x)$.

Сформулированное утверждение приводит нас к следующему специальному представлению для функции $f(x)$, имеющей равный b предел в точке a:

$$f(x) = b + \alpha(x), \tag{3.63}$$

где $\alpha(x)$ — некоторая бесконечно малая в точке a функция. Представление (3.63) весьма удобно в различных приложениях теории пределов.

Введем теперь понятие бесконечно большой в данной точке a справа [или слева] функции.

Функция $A(x)$ называется бесконечно большой в точке a справа [слева] функцией, если для любой сходящейся к a последовательности $\{x_n\}$ значений аргумента, все элементы которой больше a [меньше a], соответствующая последовательность значений функции $\{A(x_n)\}$ является бесконечно большой последовательностью, все элементы которой, начиная с некоторого номера, либо положительны, либо отрицательны.

Для бесконечно больших в точке a справа [слева] функций используется следующая символика:

$$\lim_{x\to a+0} A(x) = +\infty \; [\lim_{x\to a-0} A(x) = +\infty]$$

или

$$\lim_{x\to a+0} A(x) = -\infty \; [\lim_{x\to a-0} A(x) = -\infty].$$

Иногда употребляют более лаконичную символику:

$$A(a+0) = +\infty \; [A(a-0) = +\infty]$$

или

$$A(a+0) = -\infty \; [A(a-0) = -\infty].$$

Остановимся на методике сравнения двух бесконечно малых в данной точке a функций. *Пусть $\alpha(x)$ и $\beta(x)$ — две функции, заданные для одних и тех же значений аргумента и обе являющиеся бесконечно малыми в точке a.*

1°. Говорят, что $\alpha(x)$ *является в точке a бесконечно малой более высокого порядка, чем $\beta(x)$ (имеет в точке a более высокий порядок малости, чем $\beta(x)$), если*

$$\lim_{x\to a} \frac{\alpha(x)}{\beta(x)} = 0. \tag{3.64}$$

2°. Говорят, что $\alpha(x)$ *и $\beta(x)$ являются в точке a бесконечно малыми одного порядка (имеют в точке a одинаковый порядок малости), если предел, стоящий в левой части* (3.64), *равен конечному числу, отличному от нуля.*

3°. Говорят, что $\alpha(x)$ *и $\beta(x)$ являются в точке a эквивалентными бесконечно малыми, если предел, стоящий в левой части* (3.64), *равен единице.*

Для обозначения того, что $\alpha(x)$ является в данной точке бесконечно малой более высокого порядка, чем $\beta(x)$, используют следующую запись:

$$\alpha = o(\beta)$$

(читается: «α равно o малому от β»).

Итак, *символ $o(\beta)$ обозначает любую бесконечно малую в данной точке a функцию, имеющую в этой точке более высокий порядок малости, чем бесконечно малая в той же точке функция*

$\beta(x)$. Из этого определения символа «o малое» вытекают следующие его свойства:

1) $o(\beta)+o(\beta)=o(\beta)$, $o(\beta)-o(\beta)=o(\beta)$;

2) если $\gamma=o(\beta)$, то $o(\beta)+o(\gamma)=o(\beta)$;

3) если α и β — любые две бесконечно малые в данной точке функции, то $\alpha\cdot\beta=o(\alpha)$ и $\alpha\cdot\beta=o(\beta)$.

Аналогично сравниваются две бесконечно большие в данной точке a справа (или слева) функции.

Пусть $A(x)$ и $B(x)$ определены для одних и тех же значений аргумента и для определенности

$$\lim_{x\to a+0} A(x)=+\infty, \quad \lim_{x\to a+0} B(x)=+\infty.$$

1°. Говорят, что $A(x)$ *имеет в точке a справа более высокий порядок роста, чем $B(x)$, если функция $\frac{A(x)}{B(x)}$ является бесконечно большой в точке a справа.*

2°. Говорят, что $A(x)$ *и $B(x)$ имеют в точке a справа одинаковый порядок роста, если предел функции $\frac{A(x)}{B(x)}$ при $x\to a+0$ равен конечному числу, отличному от нуля.*

Приведем примеры сравнения бесконечно малых и бесконечно больших функций.

1. Функции $\alpha(x)=x^3-x^5$ и $\beta(x)=5x^3+x^4$ являются в точке $x=0$ бесконечно малыми одного порядка, ибо

$$\lim_{x\to 0}\frac{\alpha(x)}{\beta(x)}\quad \lim_{x\to 0}\frac{x^3-x^5}{5x^3+x^4}=\lim_{x\to 0}\frac{1-x^2}{5+x}=\frac{1}{5}.$$

2. Функции $\alpha(x)=(x-2)^2\cdot(x-1)$ и $\beta(x)=(x-2)^2$ являются в точке $x=2$ эквивалентными бесконечно малыми, ибо

$$\lim_{x\to 2}\frac{\alpha(x)}{\beta(x)}=\lim_{x\to 2}\frac{(x-2)^2(x-1)}{(x-2)^2}=\lim_{x\to 2}(x-1)=1.$$

3. Функции $A(x)=\frac{2+x}{x}$ и $B(x)=\frac{1}{x}$ являются бесконечно большими одинакового порядка роста в точке $x=0$ как справа, так и слева, ибо

$$\lim_{x\to 0}\frac{A(x)}{B(x)}=\lim_{x\to 0}(2+x)=2.$$

Аналогично определяются и сравниваются функции, бесконечно малые или бесконечно большие при $x\to\infty$, а также при $x\to+\infty$ (соответственно при $x\to-\infty$).

§ 5. ОБЩЕЕ ОПРЕДЕЛЕНИЕ ПРЕДЕЛА ФУНКЦИИ ПО БАЗЕ

Анализируя определения различных видов предела функции $f(x)$ по Коши, мы легко можем заметить, что во всех этих определениях требуется, чтобы для любого $\varepsilon>0$ все значения этой функции, отвечающие значениям аргумента x, принадлежащим некоторому множеству C_δ, удовлетворяли неравенству (3.58), т. е. принадлежали ε-окрестности b.

При этом множество C_δ, определенное для всех $\delta>0$, имеет разный вид при определении различных видов предела. При определении предела в точке a множество C_δ представляет собой проколотую δ-окрестность точки a, при определении правого [левого] предела в точке a множество C_δ представляет собой интервал $(a, a+\delta)$ [соответственно $(a-\delta, a)$], при определении предела при $x\to\infty$ множество C_δ представляет собой внешнюю часть сегмента $[-\delta, +\delta]$ и, наконец, при определении предела при $x\to\infty$ [при $x\to-\infty$] множество C_δ представляет собой открытую полупрямую $(\delta, +\infty)$ [соответственно $(-\infty, -\delta)$].

Если функция $f(x)$ задана на множестве $\{x\}$, то во всех определениях пределов по Коши требуется, чтобы неравенство (3.58) было справедливо для тех элементов множества $\{x\}$, которые принадлежат соответствующему множеству C_δ. Договоримся обозначать символом B_δ подмножество тех элементов $\{x\}$, которые принадлежат C_δ, т. е. положим

$$B_\delta=\{x\}\cap C_\delta.$$

Естественно, возникает вопрос, какими общими свойствами обладает совокупность всех подмножеств B_δ множества $\{x\}$.

Анализ условий, при которых формулируются определения 1*—4* пределов функции по Коши, приводит нас к выводу, что множество $\{x\}$ задания функции $f(x)$ всякий раз имеет хотя бы один элемент, принадлежащий C_δ, т. е. множество B_δ *всегда не является пустым.*

Далее легко убедиться в том, что для всех видов пределов пересечение двух любых множеств совокупности $\{B_\delta\}$ представляет собой некоторое множество той же совокупности.

Так, например, пересечение двух множеств B_δ и B_δ', первое из которых состоит из значений аргумента, принадлежащих проколотой δ-окрестности точки a, а второе — из значений аргумента, принадлежащих проколотой δ'-окрестности точки a, представляет собой совокупность значений аргумента, принадлежащих проколотой δ''-окрестности точки a, где δ'' — наименьшее из двух положительных чисел δ и δ', т. е. представляет собой множество B_δ'' той же совокупности $\{B_\delta\}$.

В более общей ситуации, которая может встретиться, например, при изучении функции нескольких переменных, *пересече-*

ние двух любых множеств совокупности $\{B_\delta\}$ само может не являться элементом этой совокупности, но обязательно содержит элемент этой совокупности.

Проведенное рассмотрение, естественно, приводит нас к фундаментальному понятию базы множества $\{x\}$ задания функции.

Определение 1. *Будем говорить, что бесконечная совокупность $B=\{B_\delta\}$ подмножеств B_δ множества $\{x\}$ образует базу (или базис фильтра) множества $\{x\}$, если для элементов этой совокупности выполнены два требования: 1) каждый элемент B_δ является непустым подмножеством множества $\{x\}$; 2) в пересечении любых двух элементов совокупности $\{B_\delta\}$ обязательно содержится некоторый элемент этой же совокупности.*

Приведем примеры наиболее употребительных баз (базисов фильтра).

1°. Пусть функция $f(x)$ задана на множестве $\{x\}$, имеющем хотя бы один элемент в любой проколотой δ-окрестности точки a. Указанную проколотую δ-окрестность точки a обозначим символом C_δ и положим $B_\delta=\{x\}\cap C_\delta$. Очевидно, совокупность $B=\{B_\delta\}$ множеств B_δ при всех $\delta>0$ образует базу множества $\{x\}$, ибо каждое множество B_δ при любом $\delta>0$ не является пустым и пересечение любых двух множеств совокупности $\{B_\delta\}$, как уже отмечалось выше, представляет собой множество из той же совокупности.

Рассмотренную базу $\{B_\delta\}$ принято обозначать символом $x\to a$.

2°. Пусть функция $f(x)$ задана на множестве $\{x\}$, имеющем при любом $\delta>0$ хотя бы один элемент, принадлежащий интервалу $(a, a+\delta)$ [соответственно $(a-\delta, a)$]. Обозначив указанный интервал символом C_δ, положим $B_\delta=\{x\}\cap C_\delta$. Тривиально проверяется, что совокупность $B=\{B_\delta\}$ множеств B_δ, отвечающих всевозможным $\delta>0$, образует базу множества $\{x\}$.

Указанную базу принято обозначать символом $x\to a+0$ [соответственно $x\to a-0$].

3°. Пусть функция $f(x)$ задана на множестве $\{x\}$, имеющем хотя бы один элемент вне сегмента $[-\delta, +\delta]$ при любом $\delta>0$. Положим $C_\delta=(-\infty, +\infty)\setminus[-\delta, +\delta]$, $B_\delta=\{x\}\cap C_\delta$. Легко проверить, что совокупность $B=\{B_\delta\}$ образует базу множества $\{x\}$.

Эту базу принято обозначать символом $x\to\infty$.

4°. Пусть функция $f(x)$ задана на множестве $\{x\}$, имеющем при любом $\delta>0$ хотя бы один элемент на полупрямой $(+\delta, +\infty)$ [соответственно $(-\infty, -\delta)$]. Обозначим указанную полупрямую символом C_δ и положим $B_\delta=\{x\}\cap C_\delta$. Легко убедиться в том, что совокупность $B=\{B_\delta\}$ образует базу множества $\{x\}$.

Эту базу обозначают символом $x\to+\infty$ [соответственно $x\to-\infty$].

5°. Пусть, наконец, множество $\{x\}$ представляет собой множество всех натуральных чисел 1, 2, 3, ..., n, ... Положив $B_\delta = \{x\} \cap (+\delta, +\infty)$ для любого $\delta > 0$, мы легко убедимся и в том, что совокупность $B = \{B_\delta\}$ образует базу множества $\{x\}$.

Эту базу принято обозначать символом $n \to \infty$.

Сформулируем теперь фундаментальное определение предела функции $f(x)$ по базе B множества ее задания, содержащее в себе как все рассмотренные выше виды предела функции, так и предел числовой последовательности.

Предположим, что функция $f(x)$ задана на множестве $\{x\}$ и что совокупность $B = \{B_\delta\}$ подмножеств B_δ множества $\{x\}$ образует базу множества $\{x\}$.

Множество всех значений, которые принимает функция $f(x)$, когда ее аргумент x пробегает множество B_δ, договоримся называть **образом множества** B_δ *и обозначать символом $f(B_\delta)$.*

Определение 2. *Число b называется* **пределом функции** $f(x)$ **по базе** B **множества ее задания,** *если для любого $\varepsilon > 0$ существует такой элемент B_δ базы B, образ $f(B_\delta)$ которого принадлежит ε-окрестности точки b, т. е. принадлежит интервалу $(b - \varepsilon, b + \varepsilon)$.*

Для обозначения предела функции $f(x)$ по базе B множества ее задания будем использовать символ

$$\lim_{B} f(x) = b.$$

Читатель без труда проверит, что это общее определение предела по базе содержит в себе как частные случаи изученные выше виды пределов, отвечающие базам $x \to a$, $x \to a + 0$, $x \to a - 0$, $x \to \infty$, $x \to +\infty$, $x \to -\infty$ и $n \to \infty$.

Легко проверить также, что для общего определения предела по базе остаются справедливыми основные свойства предела, отвечающего простейшей базе $x \to a$.

Мы ограничимся тем, что докажем критерий Коши существования общего предела функции $f(x)$ по базе B множества ее задания.

Теорема 3.22. *Для существования предела функции $f(x)$ по базе $B = \{B_\delta\}$ множества ее задания необходимо и достаточно, чтобы для любого $\varepsilon > 0$ нашелся элемент B_δ базы B, образ $f(B_\delta)$ которого содержится в некотором интервале длины 2ε.*

Доказательство. 1) *Необходимость* очевидна: если существует предел b функции $f(x)$ по базе B, то для любого $\varepsilon > 0$ найдется элемент этой базы B_δ, образ которого $f(B_\delta)$ содержится в интервале $(b - \varepsilon, b + \varepsilon)$, имеющем длину 2ε.

2) *Достаточность.* Пусть для любого $\varepsilon > 0$ существует элемент B_δ базы B, образ которого $f(B_\delta)$ содержится в некотором интервале длины 2ε. Рассмотрим бесконечно малую последова-

тельность положительных чисел $\varepsilon_n = \frac{1}{n}$ ($n=1, 2, 3, \ldots$). Для каждого ε_n найдется элемент базы B_{δ_n}, образ которого $f(B_{\delta_n})$ содержится в некотором интервале длины $2\varepsilon_n$.

По определению базы в пересечении элементов B_{δ_1} и B_{δ_2}, обязательно лежит некоторый элемент базы, который мы обозначим символом $\widehat{B}_{\delta_2}$. Образ этого элемента $f(\widehat{B}_{\delta_2})$ лежит как в некотором интервале I_1 длины $2\varepsilon_1$, так и в некотором интервале I_2' длины $2\varepsilon_2$. Пересечение интервалов I_1 и I_2' представляет собой интервал I_2 длины, не большей $2\varepsilon_2$, содержащийся в интервале I_1. Далее, по определению базы в пересечении элементов $\widehat{B}_{\delta_2}$ и B_{δ_3} обязательно лежит некоторый элемент базы, который мы обозначим символом $\widehat{B}_{\delta_3}$. Образ этого элемента $f(\widehat{B}_{\delta_3})$ лежит как в интервале I_2 длины, не большей $2\varepsilon_2$, так и в некотором интервале I_3' длины $2\varepsilon_3$. Пересечение интервалов I_2 и I_3' представляет собой интервал I_3 длины, не большей $2\varepsilon_3$, содержащийся в интервале I_2.

Продолжая эти рассуждения далее, мы построим последовательность элементов базы $\widehat{B}_{\delta_2}, \widehat{B}_{\delta_3}, \ldots, \widehat{B}_{\delta_n}, \ldots$ таких, что образ $f(\widehat{B}_{\delta_n})$ каждого элемента $\widehat{B}_{\delta_n}$ содержится в некотором интервале I_n длины, не большей $2\varepsilon_n$, причем в последовательности интервалов $I_2, I_3, \ldots, I_n, \ldots$ каждый следующий интервал содержится в предыдущем. Обозначим символом $\bar{I}_n$ сегмент, получающийся добавлением к интервалу I_n его концов. Так как последовательность $\bar{I}_2, \bar{I}_3, \ldots, \bar{I}_n, \ldots$ представляет собой стягивающуюся систему сегментов (см. п. 2 § 2), то в силу следствия из теоремы 3.15 существует, и притом единственная, точка b, принадлежащая всем сегментам.

Остается доказать, что b является пределом функции $f(x)$ по базе B, т. е. убедиться в том, что для любого $\varepsilon>0$ найдется элемент базы B, образ которого содержится в интервале $(b-\varepsilon, b+\varepsilon)$.

В силу того, что система сегментов $\{\bar{I}_n\}$ является стягивающейся и b является общей точкой всех сегментов, мы можем утверждать, что для любого $\varepsilon>0$ найдется сегмент $\bar{I}_n$ с достаточно большим номером n, содержащийся в интервале $(b-\varepsilon, b+\varepsilon)$. Это означает, что при соответствующем номере n элемент базы $\widehat{B}_{\delta_n}$ имеет образ $f(\widehat{B}_{\delta_n})$, содержащийся в интервале $(b-\varepsilon, b+\varepsilon)$. Теорема доказана.

Подчеркнем, что доказанная теорема содержит в качестве частных случаев как критерий Коши сходимости числовой последовательности, так и критерии Коши существования всех рассмотренных выше видов предела функции.

В качестве возможных обобщений изложенной теории можно рассматривать функции, заданные на подмножествах произвольного метрического пространства (см. по этому поводу дополнение 2 к гл. 12).

З а м е ч а н и е. Базы B и D множества $\{x\}$ называются э к в и в а л е н т н ы м и, если для любого элемента B_{δ_1} базы B найдется такой элемент D_{δ_2} базы D, что $D_{\delta_2} \subset B_{\delta_1}$, и для любого элемента D_{δ_1} базы D найдется такой элемент B_{δ_2} базы B, что $B_{\delta_2} \subset D_{\delta_1}$. Совокупность всевозможных эквивалентных между собой баз B множества $\{x\}$ называется ф и л ь т р о м множества $\{x\}$.

Нетрудно убедиться, что утверждения о пределах функции по *эквивалентным базам B и D справедливы одновременно.*

Глава 4

НЕПРЕРЫВНОСТЬ ФУНКЦИИ

В настоящей главе будет всесторонне изучаться важнейшее понятие математического анализа — понятие непрерывности функции.

В дополнении 2 к гл. 12 понятие непрерывности будет введено в общей ситуации, когда задано отображение одного метрического пространства в другое.

§ 1. ПОНЯТИЕ НЕПРЕРЫВНОСТИ ФУНКЦИИ

1. Определение непрерывности функции. Пусть точка a принадлежит области задания функции $f\{x\}$* и любая ε-окрестность точки a содержит отличные от a точки области задания функции $f\{x\}$**.

Формальное определение непрерывности в точке a. *Функция $f\{x\}$ называется непрерывной в точке a, если функция $f(x)$ имеет в точке a предел и этот предел равен частному значению $f(a)$ функции $f(x)$ в точке a.*

Используя определения предела функции $y=f(x)$ в точке a по Гейне и по Коши, мы приходим к определению непрерывности функции в данной точке a по Гейне и по Коши.

Определение 1 (непрерывность в точке a по Гейне). *Функция $y=f(x)$ называется непрерывной в точке a, если для любой сходящейся к a последовательности значений аргумента $x_1, x_2, \ldots, x_n$, соответствующая последовательность значений функции $f(x_1), f(x_2), \ldots, f(x_n), \ldots$ сходится к числу $f(a)$.*

Замечание 1. По сравнению с определением 1 предела функции по Гейне (см. п. 2 § 4, гл. 3) в определении непрерывности по Гейне мы опустили требование, обязывающее все элементы последовательности $\{x_n\}$ быть отличными от a. Это можно сделать в силу того, что добавление к элементам последовательности $\{f(x_n)\}$, сходящейся к числу $f(a)$, любого числа новых элементов, равных $f(a)$, не нарушит сходимости этой последовательности к $f(a)$.

---

* Заметим, что этого не требовалось, когда мы рассматривали предел функции $f(x)$ в точке a.

** Т. е. точка a является предельной точкой множества $\{x\}$, на котором задана функция $f(x)$.

Определение 1* (непрерывность в точке a по Коши). *Функция $f(x)$ называется непрерывной в точке a, если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех значений аргумента x, удовлетворяющих условию $|x-a|<\delta$, справедливо неравенство $|f(x)-f(a)|<\varepsilon$.*

Замечание 2. По сравнению с определением 1* предела функции по Коши (см. п. 2 § 4, гл. 3) в определении непрерывности по Коши мы опустили требование, обязывающее все значения аргумента x удовлетворять неравенству $|x-a|>0$, т. е. быть отличными от a. Это можно сделать в силу того, что для значений $x=a$ разность $f(x)-f(a)$ равна нулю и удовлетворяет неравенству $|f(x)-f(a)|<\varepsilon$ при любом $\varepsilon>0$.

Условие непрерывности функции $f(x)$ в точке a символически можно выразить следующим равенством:

$$\lim_{x\to a} f(x)=f(a).$$

Так как $a=\lim\limits_{x\to a} x$, то этому равенству можно придать следующую форму:

$$\lim_{x\to a} f(x)=f(\lim_{x\to a} x).$$

Следовательно, для непрерывной в точке a функции символ $\lim\limits_{x\to a}$ предельного перехода и символ f характеристики функции можно менять местами.

Из теоремы об эквивалентности определений предельного значения по Гейне и по Коши (см. теорему 3.19 из п. 2 § 4 гл. 3) следует, что *определения непрерывности функции по Гейне и по Коши (определения 1 и 1*) эквивалентны.*

Сформулируем теперь определение односторонней непрерывности функции $f(x)$ в точке a, т. е. непрерывности в точке a либо только справа, либо только слева.

От множества $\{x\}$ задания функции $f(x)$ мы на этот раз должны потребовать, чтобы это *множество включало точку a и для любого $\delta>0$ имело хотя бы один элемент, лежащий на интервале $(a, a+\delta)$ [соответственно $(a-\delta, a)$].*

Формальное определение непрерывности в точке a справа [слева]. *Функция $f(x)$ называется непрерывной в точке a справа [слева], если правый [левый] предел этой функции в точке a существует и равен частному значению $f(a)$ функции $f(x)$ в точке a.*

Используя определения правого [левого] предела функции $f(x)$ в точке a по Гейне и по Коши, мы придем к определениям непрерывности функции $f(x)$ в точке a справа [слева] по Гейне и по Коши.

Определение 2 (непрерывность функции в точке a справа [слева] по Гейне). *Функция $f(x)$ называется непрерывной в точке a справа [слева], если для любой сходящейся к a последовательности значений аргумента $\{x_n\}$, все элементы которой удовлетворяют условию $x_n>a$ $[x_n<a]$, соответствующая последовательность значений функции $\{f(x_n)\}$ сходится к числу $f(a)$.*

Заметим, что в этом определении условие $x_n>a$ $[x_n<a]$ можно заменить менее жестким условием $x_n\geqslant a$ $[x_n\leqslant a]$, ибо добавление к последовательности $\{f(x_n)\}$, сходящейся к $f(a)$, какого угодно числа новых элементов, равных $f(a)$, не нарушит сходимости этой последовательности к $f(a)$.

В применениях более эффективно условие $x_n>a$ $[x_n<a]$.

Определение 2* (непрерывность функции в точке a справа [слева] по Коши). *Функция $f(x)$ называется непрерывной в точке a справа [слева], если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех значений аргумента x, удовлетворяющих условию $a<x<a+\delta$ $[a-\delta<x<a]$, справедливо неравенство*

$$|f(x)-f(a)|<\varepsilon.$$

Заметим, что и в этом определении условие $a<x<a+\delta$ $[a-\delta<x<a]$ можно было бы заменить менее жестким условием $a\leqslant x<a+\delta$ $[a-\delta<x\leqslant a]$.

Эквивалентность определений 2 и 2 вытекает из эквивалентности соответствующих определений предела функции.*

Тот факт, что функция $f(x)$ непрерывна в точке a справа [слева], записывают так:

$$\lim_{x\to a+0} f(x)=f(a) \text{ или } f(a+0)=f(a)$$

$$\Big[\lim_{x\to a-0} f(x)=f(a) \text{ или } f(a-0)=f(a)\Big].$$

Замечание 3. *Если функция $f(x)$ непрерывна в точке a и слева, и справа, то она непрерывна в этой точке.* Действительно, в силу утверждения, доказанного в п. 2 § 4 гл. 3, в этом случае существует предел функции в точке a, равный $f(a)$.

Точки, в которых функция не обладает свойством непрерывности, называются точками разрыва этой функции.

Рассмотрим примеры.

1) Степенная функция $f(x)=x^n$, где n — натуральное число, непрерывна в каждой точке a бесконечной прямой $(-\infty, +\infty)$. Действительно, в гл. 3 было установлено, что предельное значение этой функции в любой точке a бесконечной прямой равно частному значению a^n.

2) Многочлены и рациональные дроби имеют в каждой точке области задания предельное значение, равное частному значению (см. п. 3 § 4 гл. 3). Поэтому они являются непрерывными функциями в каждой точке области задания.

3) Функция $f(x) = \operatorname{sgn} x$ имеет разрыв в точке $x=0$ и непрерывна во всех остальных точках числовой оси. Действительно, в точке $x=0$, как было показано в гл. 3, существуют правый (равный $+1$) и левый (равный -1) пределы функции $\operatorname{sgn} x$. Поскольку эти односторонние пределы не равны друг другу, функция $\operatorname{sgn} x$ в точке 0 разрывна (не является непрерывной). В остальных точках оси она обладает предельным значением, равным частному значению, и непрерывна.

4) Функция Дирихле $D(x)$ (см. § 4 гл. 3) разрывна в каждой точке числовой оси, поскольку она не имеет предельного значения ни в одной точке.

Заметим, однако, что функция $f(x) = x \cdot D(x)$, где $D(x)$ — функция Дирихле, является непрерывной в точке $x=0$ и разрывной во всех остальных точках бесконечной прямой. Разрывность $f(x)$ в любой точке $x_0 \neq 0$ устанавливается точно так же, как для функции $D(x)$ (для любой сходящейся к x_0 последовательности $\{x_n\}$ рациональных точек соответствующая последовательность $\{f(x_n)\}$ сходится к числу $x_0 \neq 0$, а для любой сходящейся к x_0 последовательности $\{x_n\}$ иррациональных точек соответствующая последовательность $\{f(x_n)\}$ сходится к нулю).

Убедимся в том, что функция $f(x) = x \cdot D(x)$ непрерывна в точке $x=0$. Для любой бесконечно малой последовательности значений аргумента $\{x_n\}$ последовательность $\{D(x_n)\}$ ограничена, а потому (в силу теоремы 3.3 из гл. 3) последовательность $f(x_n) = x_n \cdot D(x_n)$ является бесконечно малой, т. е. имеет предел нуль, равный частному значению $f(0)$.

Мы будем говорить, что *функция непрерывна на множестве $\{x\}$, если она непрерывна в каждой точке этого множества*.

Например, функция, непрерывная в каждой точке интервала, называется непрерывной на интервале.

Особо *договоримся называть функцию $f(x)$ непрерывной на сегменте $[a, b]$, если она непрерывна в каждой внутренней точке этого сегмента и, кроме того, непрерывна справа в точке a и непрерывна слева в точке b.*

Выше, давая определение непрерывности функции $f(x)$ в точке a, мы предположили, что точка a обладает тем свойством, что в любой ее ε-окрестности содержатся точки области задания, отличные от a. Формально этого предположения можно бы было и не делать и допустить, что точка a обладает ε-окрестностью, свободной от точек области задания функции, а в самой точке a функция определена. В этом случае формально функцию $f(x)$ можно считать непрерывной в точке a. Однако вся содержательная часть понятия непрерывности функции относится как раз к случаю, когда a — предельная точка области определения функции.

Определение непрерывности функции можно дать и в следующей, эквивалентной форме.

Определение 1**. *Функция $f(x)$ называется непрерывной в точке a, если для любой окрестности точки $f(a)$ найдется такая окрестность точки a, что образ всех точек множества задания функции, лежащих в этой окрестности точки a, при отображении, осуществляемом функцией $f(x)$, целиком лежит в указанной окрестности точки $f(a)$.*

В дополнении 2 к гл. 12 будет показано (даже в более общей ситуации), что последнее определение непрерывности эквивалентно предыдущим. Предлагается в качестве упражнения проверить это.

Используя введенное в § 5 гл. 3 общее определение предела функции $f(x)$ по базе B множества ее задания, мы можем объединить в одной формулировке понятие непрерывности в точке a, в точке a справа и в точке a слева.

Пусть функция $f(x)$ задана на множестве $\{x\}$, которое включает точку a и допускает базу B одного из видов $x \to a$, $x \to a+0$, $x \to a-0$*.

Функция $f(x)$ называется непрерывной в точке a, если ее предел по базе B множества ее задания существует и равен $f(a)$.

2. Арифметические операции над непрерывными функциями. Убедимся в том, что арифметические операции над непрерывными функциями приводят снова к непрерывным функциям.

Справедлива следующая теорема.

Основная теорема 4.1. *Пусть на одном и том же множестве заданы функции $f(x)$ и $g(x)$, непрерывные в точке a. Тогда функции $f(x)+g(x)$, $f(x)-g(x)$, $f(x)\cdot g(x)$ и $\frac{f(x)}{g(x)}$ непрерывны в точке a (в случае частного нужно дополнительно требовать $g(a)\neq 0$).*

Доказательство. Так как непрерывные в точке a функции $f(x)$ и $g(x)$ имеют в точке a пределы, соответственно равны $f(a)$ и $g(a)$, то в силу теоремы 3.21 из гл. 3 пределы функций $f(x)+g(x)$, $f(x)-g(x)$, $f(x)\cdot g(x)$ и $\frac{f(x)}{g(x)}$ существуют и равны соответственно $f(a)+g(a)$, $f(a)-g(a)$, $f(a)\cdot g(a)$ и $\frac{f(a)}{g(a)}$. Но как раз эти величины равны частным значениям перечисленных функций в точке a. По определению эти функции непрерывны в точке a, что и требовалось доказать.

---

* См. § 5 гл. 3.

3. Сложная функция и ее непрерывность. *Функции, полученные в результате суперпозиции двух или нескольких функций, мы будем называть* ***сложными***. Под суперпозицией двух функций мы понимаем функцию, полученную в результате наложения или последовательного применения указанных двух функций в определенном порядке. Ясно, что достаточно определить сложную функцию, полученную в результате суперпозиции только двух функций. Указанный алгоритм можно будет применять, беря суперпозицию трех и большего конечного числа функций.

Пусть функция $x=\varphi(t)$ задана на множестве $\{t\}$, и пусть $\{x\}$ — множество ее значений. Допустим, что на множестве $\{x\}$ задана функция $y=f(x)$. Тогда говорят, что на множестве $\{t\}$ задана сложная функция $y=f[\varphi(t)]=F(t)$ или $y=f(x)$, где $x=\varphi(t)$.

Справедлива следующая теорема.

Теорема 4.2. *Пусть функция $x=\varphi(t)$ непрерывна в точке a, а функция $y=f(x)$ непрерывна в точке $b=\varphi(a)$. Тогда сложная функция $y=f[\varphi(t)]=F(t)$ непрерывна в точке a.*

Доказательство. Пусть $\{t_n\}$ — произвольная последовательность значений аргумента сложной функции, сходящаяся к a. Так как функция $x=\varphi(t)$ непрерывна в точке a, то (в силу определения 1 непрерывности по Гейне) соответствующая последовательность значений функции $x_n=\varphi(t_n)$ сходится к числу $b=\varphi(a)$. Далее, поскольку функция $y=f(x)$ непрерывна в точке $b=\varphi(a)$ и для нее указанная выше последовательность $\{x_n\}$, сходящаяся к $b=\varphi(a)$, является последовательностью значений аргумента, то (в силу того же определения 1 непрерывности по Гейне) соответствующая последовательность значений функции $f(x_n)=$ $=f[\varphi(t_n)]=F(t_n)$ сходится к числу $f(b)=f[\varphi(a)]=F(a)$.

Итак, для любой последовательности $\{t_n\}$ значений аргумента сложной функции, сходящейся к a, соответствующая последовательность значений самой сложной функции $\{F(t_n)\}=f[\varphi(t_n)]$ сходится к числу $F(a)=f[\varphi(a)]$. В силу определения 1 непрерывности по Гейне сложная функция непрерывна в точке a. Теорема доказана.

§ 2. СВОЙСТВА МОНОТОННЫХ ФУНКЦИЙ

1. Монотонные функции. Введем понятие монотонной функции.

Определение 1. *Функция $f(x)$ называется* ***неубывающей*** [***невозрастающей***] *на множестве $\{x\}$, если для любых x_1 и x_2 из этого множества таких, что $x_1<x_2$, справедливо неравенство $f(x_1)\leqslant f(x_2)$ $[f(x_1)\geqslant f(x_2)]$.*

Неубывающие и невозрастающие функции называют **монотонными** функциями.

Определение 2. *Функция называется* ***возрастающей*** [***убывающей***] *на множестве $\{x\}$, если для любых x_1 и x_2 из* *этого множества таких, что* $x_1<x_2$, *справедливо неравенство* $f(x_1)<f(x_2)$ $[f(x_1)>f(x_2)]$.

Возрастающие и убывающие функции называются строго монотонными.

Приведем примеры монотонных функций.

1. Функция $f(x)=x^3$ — строго монотонна, а именно возрастает на всей числовой оси.

2. Функция $y=x^2$ — возрастает на полуоси $x\geqslant 0$ и убывает на полуоси $x\leqslant 0$.

3. Функция $y=\operatorname{sgn} x$ — неубывающая на всей числовой оси.

4. Функция $f(x)=\frac{1}{x}$ — убывает на множествах $x<0$ и $x>0$.

2. Понятие обратной функции. Пусть функция $y=f(x)$ задана на сегменте $[a, b]$, и пусть сегмент $[\alpha, \beta]$ является множеством значений этой функции. Пусть, кроме того, каждому y из сегмента $[\alpha, \beta]$ соответствует только одно значение x из сегмента $[a, b]$, для которого $f(x)=y$. Тогда на сегменте $[\alpha, \beta]$ определена функция, которая каждому y из $[\alpha, \beta]$ ставит в соответствие то значение x из $[a, b]$, для которого $f(x)=y$. Эта функция обозначается символом $x=f^{-1}(y)$ и называется обратной для функции $y=f(x)$.

В проведенных выше рассуждениях вместо сегментов $[a, b]$ и $[\alpha, \beta]$ можно было бы рассматривать интервалы (a, b) и (α, β) или, например, случай, когда один или оба из этих интервалов превращаются в бесконечную прямую или открытую полупрямую.

Можно рассматривать и самый общий случай, когда задано отображение f одного множества $\{x\}$ на другое множество $\{y\}$, причем отображение f устанавливает взаимно однозначное соответствие между элементами этих множеств. Тогда можно определить обратное отображение f^{-1} множества $\{y\}$ на множество $\{x\}$. В этом случае уравнение $y=f(x)$ можно разрешить относительно x, т. е. можно однозначно определить x, зная элемент y, и мы имеем $x=f^{-1}(y)$.

Отметим, что если $x=f^{-1}(y)$ — обратная функция для $y=f(x)$, то, очевидно, функция $y=f(x)$ является обратной для функции $x=f^{-1}(y)$. Поэтому функции $y=f(x)$ и $x=f^{-1}(y)$ называются взаимно обратными. Очевидно, что $f[f^{-1}(y)]=y$, $f^{-1}[f(x)]=x$.

Приведем примеры взаимно обратных функций.

1. Пусть на сегменте $[a, b]$ задана функция $y=2x$. Множеством значений этой функции будет сегмент $[2a, 2b]$. Функция $x=f^{-1}(y)=\frac{y}{2}$, определенная на $[2a, 2b]$, будет обратной к заданной функции $y=2x$.

2. Рассмотрим на сегменте $[0, 2]$ функцию $y=x^2$. Множество значений этой функции есть сегмент $[0, 4]$. На этом сегменте определена обратная к заданной функции функция $x=\sqrt{y}$.

3. Рассмотрим на сегменте $[0, 1]$ функцию

$$y = \begin{cases} x, & \text{если } x \text{ — рациональное число,} \\ 1-x, & \text{если } x \text{ — иррациональное число.} \end{cases}$$

Нетрудно убедиться, что заданная на сегменте $[0, 1]$ функция

$$x = \begin{cases} y, & \text{если } y \text{ — рациональное число,} \\ 1-y, & \text{если } y \text{ — иррациональное число,} \end{cases}$$

будет обратной к заданной функции.

Докажем несколько утверждений о монотонных функциях.

Начнем с доказательства леммы, справедливой для любой монотонной (не обязательно строго монотонной) функции.

Лемма. *Если функция $f(x)$ является монотонной на сегменте $[a, b]$, то у нее существуют правый и левый пределы в любой внутренней точке сегмента $[a, b]$ и, кроме того, существуют правый предел в точке a и левый предел в точке b.*

Доказательство. Для полного доказательства леммы достаточно доказать два факта: 1) существование правого предела в любой точке c, удовлетворяющей неравенствам $a \leqslant c < b$; 2) существование левого предела в любой точке c, удовлетворяющей неравенствам $a < c \leqslant b$.

Мы установим только первый из указанных двух фактов, ибо второй устанавливается аналогично.

При этом мы проведем все рассуждения для функции $f(x)$, неубывающей на сегменте $[a, b]$ (ибо для невозрастающей функции они проводятся аналогично).

Итак, пусть функция $f(x)$ не убывает на $[a, b]$, c — любая точка, удовлетворяющая неравенствам $a \leqslant c < b$. Рассмотрим множество $\{f(x)\}$ всех значений функции $f(x)$ для значений аргумента x, удовлетворяющих неравенствам $c < x \leqslant b$. Это множество $\{f(x)\}$ непусто (в силу того, что $c < b$) и ограничено снизу (в силу неубывания функции $f(x)$ для всех x из полусегмента $c < x \leqslant b$ справедливо неравенство $f(c) \leqslant f(x)$, которое означает, что $f(c)$ является нижней гранью рассматриваемого множества). По основной теореме 2.1 гл. 2 у рассматриваемого множества существует точная нижняя грань, которую мы обозначим символом γ. Докажем, что это число γ и является правым пределом функции $f(x)$ в точке c, т. е. докажем, что $\gamma = f(c+0)$.

Фиксируем произвольное $\varepsilon > 0$. По определению точной нижней грани найдется положительное число δ, не превосходящее $b-c$ и такое, что значение функции $f(c+\delta)$ удовлетворяет неравенству $f(c+\delta) < \gamma + \varepsilon$.

Но тогда в силу неубывания функции $f(x)$ для всех x из интервала $c < x < c+\delta$ и подавно будет справедливо неравенство $f(x) < \gamma + \varepsilon$. Так как, кроме того, для всех x из указанного интервала справедливо неравенство $\gamma \leqslant f(x)$, то мы получим, что для всех x из интервала $c < x < c+\delta$ справедливы неравенства

$$\gamma \leqslant f(x) < \gamma + \varepsilon \text{ или } |\gamma - f(x)| < \varepsilon,$$

а это и означает (в силу определения правого предела по Коши), что число γ является правым пределом $f(x)$ в точке c. Лемма доказана.

Замечание к лемме. В предположениях леммы при условии неубывания $f(x)$ для любых c и x, удовлетворяющих соотношениям $a \leqslant c < x \leqslant b$, будут справедливы неравенства

$$f(a) \leqslant f(c) \leqslant f(c+0) \leqslant f(x) \leqslant f(b), \tag{4.1}$$

а для любых c и x, удовлетворяющих соотношениям $a \leqslant x < c \leqslant b$, будут справедливы неравенства

$$f(a) \leqslant f(x) \leqslant f(c-0) \leqslant f(c) \leqslant f(b). \tag{4.2}$$

При условии невозрастания $f(x)$ все знаки в неравенствах (4.1) и (4.2) следует заменить на противоположные.

Пусть, например, $f(x)$ не убывает на $[a, b]$ и $a \leqslant c < x \leqslant b$. Тогда $f(a) \leqslant f(c) \leqslant f(x) \leqslant f(b)$. Из последних неравенств сразу же вытекает, что $f(a) \leqslant f(c) \leqslant f(c+0) \leqslant f(b)$. Для завершения доказательства неравенств (4.1) следует убедиться в том, что $f(c+0) \leqslant f(x) \leqslant f(b)$ для любого x из полуинтервала $c < x \leqslant b$, но это сразу вытекает из того, что число $\gamma = f(c+0)$ является, как доказано в лемме, точной нижней гранью множества значений $f(x)$ на полуинтервале $c < x \leqslant b$. Справедливость неравенств (4.2) проверяется аналогично.

Докажем теперь три теоремы о строго монотонных функциях.

Теорема 4.3. *Пусть функция $y = f(x)$ возрастает (убывает) на сегменте $[a, b]$, и пусть $\alpha = f(a)$, $\beta = f(b)$. Тогда, если множеством всех значений функции $y = f(x)$ является сегмент $[\alpha, \beta]$ (соответственно сегмент $[\beta, \alpha]$), то на этом последнем сегменте определена обратная для $y = f(x)$ функция $x = f^{-1}(y)$, которая также возрастает (убывает) на указанном сегменте.*

Доказательство. Проведем все рассуждения в предположении, что $f(x)$ возрастает на сегменте $[\alpha, \beta]$ (для убывающей функции рассуждения аналогичны).

Убедимся в том, что функция $y = f(x)$ устанавливает взаимно однозначное соответствие между сегментами $a \leqslant x \leqslant b$ и $\alpha \leqslant y \leqslant \beta$. Действительно, то, что каждому x из $[a, b]$ соответствует только одно значение y из $[\alpha, \beta]$, следует из самого понятия функции $y = f(x)$, а то, что каждому y из $[\alpha, \beta]$ соответствует только одно x из $[a, b]$, вытекает из возрастания функции $y = f(x)$.

Убедимся теперь, что если $y = f(x)$ возрастает на $[a, b]$, то и $x = f^{-1}(y)$ также возрастает на $[\alpha, \beta]$. Пусть $y_1 < y_2$, где y_1 и y_2 — любые два числа из $[\alpha, \beta]$. Тогда $x_1 = f^{-1}(y_1) < x_2 = f^{-1}(y_2)$, ибо из неравенства $x_1 \geqslant x_2$ и из возрастания функции $y = f(x)$ следовало бы, что $y_1 \geqslant y_2$, что противоречит неравенству $y_1 < y_2$. Теорема доказана.

Замечание 1. Совершенно аналогично доказывается более общее утверждение: *пусть $y=f(x)$ задана и возрастает (убывает) на некотором множестве $\{x\}$, а $\{y\}$ — множество всех значений этой функции. Тогда на множестве $\{y\}$ определена обратная для $y=f(x)$ функция $x=f^{-1}(y)$, которая также возрастает (убывает) на указанном множестве $\{y\}$.*

Теорема 4.4. *Пусть функция $y=f(x)$ возрастает (или убывает) на сегменте $[a, b]$, и пусть $\alpha=f(a)$, $\beta=f(b)$. Тогда для того, чтобы функция $y=f(x)$ являлась непрерывной на сегменте $[a, b]$, необходимо и достаточно, чтобы любое число γ, заключенное между α и β, было значением этой функции.*

Доказательство. Все рассуждения проведем для возрастающей на сегменте $[a, b]$ функции, ибо для убывающей функции они аналогичны.

1) *Необходимость*. Пусть функция $y=f(x)$ возрастает и непрерывна на сегменте $[a, b]$. Требуется доказать, что любое число γ, удовлетворяющее условиям $\alpha<\gamma<\beta$, является значением функции в некоторой точке c сегмента $[a, b]$.

Пусть $\{x\}$ — множество всех значений x из сегмента $[a, b]$, для которых $f(x)\leqslant\gamma$. Это множество $\{x\}$ непусто (ему принадлежит, например, точка a, ибо $f(a)=\alpha<\gamma$) и ограничено сверху (например, числом b). По основной теореме 2.1 гл. 2 у множества $\{x\}$ существует точная верхняя грань, которую мы обозначим через c: $c=\sup\{x\}$. Остается доказать, что $f(c)=\gamma$.

Сначала убедимся в том, что $f(x)\leqslant\gamma$ для всех x из $[a, b]$, лежащих левее c, и $f(x)>\gamma$ для всех x из $[a, b]$, лежащих правее c.

В самом деле, если $x<c$, то по определению точной верхней грани найдется x' из полуинтервала $x<x'\leqslant c$, принадлежащее множеству $\{x\}$, т. е. такое, что $f(x')\leqslant\gamma$. Но тогда из возрастания $f(x)$ будет вытекать, что и $f(x)\leqslant\gamma$ (ибо $f(x)<f(x')$).

Далее, любое x, лежащее правее c, не принадлежит множеству $\{x\}$, а потому для такого x справедливо неравенство $f(x)>\gamma$.

Теперь убедимся в том, что c является внутренней точкой сегмента $[a, b]$. Докажем, что $c<b$. Предположим, что это не так, т. е. допустим, что $c=b$. Возьмем любую сходящуюся к $c=b$ возрастающую последовательность $\{x_n\}$ точек сегмента $[a, b]$. Так как все ее элементы x_n лежат левее c, то $f(x_n)\leqslant\gamma$ для всех номеров n, а поэтому (в силу теоремы 3.13 гл. 3) и $\lim\limits_{n\to\infty} f(x_n)\leqslant\gamma$. Но так как функция $f(x)$ непрерывна в точке $c=b$, то $\lim\limits_{n\to\infty} f(x_n)=f(b)$. Тем самым мы получаем неравенство $\beta=f(b)\leqslant\gamma$, которое противоречит условию $\gamma<\beta$. Полученное противоречие доказывает, что $c<b$.

Совершенно аналогично доказывается, что $a<c$.

Итак, доказано, что c — внутренняя точка сегмента $[a, b]$.

Теперь для того, чтобы доказать, что $f(c)=\gamma$, рассмотрим две сходящиеся к c с разных сторон последовательности точек сегмента $[a, b]$ — возрастающую последовательность $\{x_n'\}$ и убывающую последовательность $\{x_n''\}$. В силу того, что функция $f(x)$ непрерывна в точке c,

$$\lim_{n\to\infty} f(x_n') = \lim_{n\to\infty} f(x_n'') = f(c).$$

С другой стороны, поскольку $x_n' < c < x_n''$ для любого номера n, то $f(x_n') \leqslant \gamma$, $f(x_n'') > \gamma$ (для любого номера n). Но тогда в силу теоремы 3.13

$$\lim_{n\to\infty} f(x_n') = f(c) \leqslant \gamma, \quad \lim_{n\to\infty} f(x_n'') = f(c) \geqslant \gamma,$$

т. е. $f(c) = \gamma$.

Необходимость доказана.

2) *Достаточность*. Пусть функция $f(x)$ возрастает на сегменте $[a, b]$, и пусть любое число γ из сегмента $[\alpha, \beta]$ является значением этой функции. Докажем, что функция $f(x)$ непрерывна на сегменте $[a, b]$. Достаточно доказать, что $f(x)$ непрерывна справа в любой точке c, удовлетворяющей условиям $a \leqslant c < b$, и непрерывна слева в любой точке c, удовлетворяющей условиям $a < c \leqslant b$.

Мы ограничимся доказательством непрерывности справа в любой точке c, удовлетворяющей условиям $a \leqslant c < b$, ибо вторая часть утверждения доказывается аналогично.

Предположим, что функция $f(x)$ не является непрерывной справа в некоторой точке c, удовлетворяющей условиям $a \leqslant c < b$. Тогда ее правый предел $f(c+0)$, который существует согласно доказанной выше лемме, будет отличен от значения $f(c)$, и поэтому справедливые в силу замечания к указанной лемме неравенства (4.1) примут вид

$$\alpha = f(a) \leqslant f(c) < f(c+0) \leqslant f(x) \leqslant f(b) = \beta \tag{4.2'}$$

(для всех x из полуинтервала $c < x \leqslant b$).

Неравенства (4.2′) означают, что содержащийся в $[\alpha, \beta]$ интервал $(f(c), f(c+0))$ не содержит значений функции $f(x)$*, а это противоречит тому, что любое число γ из сегмента $[\alpha, \beta]$ является значением этой функции. Теорема 4.4 полностью доказана.

Теорема 4.5. *Пусть функция $y = f(x)$ возрастает (убывает) и непрерывна на сегменте $[a, b]$, и пусть $\alpha = f(a)$, $\beta = f(b)$. Тогда на сегменте $[\alpha, \beta]$ (соответственно на сегменте $[\beta, \alpha]$) определена обратная для $y = f(x)$ функция $x = f^{-1}(y)$, которая возрастает (убывает) и непрерывна на указанном сегменте.*

Кратко можно сказать, что из строгой монотонности и непрерывности на сегменте $[a, b]$ данной функции вытекают существо-

* В самом деле, для $x \leqslant c$ значение $f(x)$ удовлетворяет неравенству $f(x) \leqslant f(c)$ (в силу возрастания функции), а для $x > c$ значение $f(x)$ удовлетворяет неравенству $f(c+0) \leqslant f(x)$ (в силу (4.2′)).

вание, строгая монотонность и непрерывность на соответствующем сегменте обратной функции.

Доказательство. Проведем все рассуждения для возрастающей функции, ибо для убывающей функции они проводятся аналогично.

Так как $f(x)$ возрастает и непрерывна на сегменте $[a, b]$, то в силу необходимости теоремы 4.4 множеством всех значений этой функции является сегмент $[\alpha, \beta]$. Но тогда теорема 4.3 обеспечивает существование на этом сегменте возрастающей обратной функции $x=f^{-1}(y)$. Остается доказать непрерывность указанной обратной функции на сегменте $[\alpha, \beta]$. Для этого достаточно учесть, что множеством всех значений обратной функции $x=f^{-1}(y)$ служит сегмент $[a, b]$, где $a=f^{-1}(\alpha)$, $b=f^{-1}(\beta)$, и использовать для этой обратной функции достаточность теоремы 4.4. Доказательство теоремы 4.5 завершено.

Замечание 2. *Можно показать, что из существования обратной функции для функции $f(x)$, непрерывной на сегменте $[a, b]$, следует, что $f(x)$ строго монотонна на этом сегменте* (см. п. 2 § 6 настоящей главы).

§ 3. ПРОСТЕЙШИЕ ЭЛЕМЕНТАРНЫЕ ФУНКЦИИ

Простейшими элементарными функциями, как уже отмечалось, обычно называют следующие функции: $y=x^{\alpha}$, $y=a^x$, $y=\log_a x$, $y=\sin x$, $y=\cos x$, $y=\operatorname{tg} x$, $y=\operatorname{ctg} x$, $y=\arcsin x$, $y=\arccos x$, $y=\operatorname{arctg} x$, $y=\operatorname{arcctg} x$.

Нашей основной целью является изучение вопроса об определении и непрерывности простейших элементарных функций. Следует заметить, что вопрос об определении простейших элементарных функций далеко не прост. Так, например, показательная функция $y=a^x$ легко может быть определена для рациональных значений аргумента x, вместе с тем эту функцию следует определить для произвольных вещественных значений x, т. е. следует определить возведение вещественного числа в любую вещественную степень x. Далее, определение тригонометрических функций $\sin x$ и $\cos x$ с помощью наглядных геометрических соображений имеет логический пробел. Возможность определить эти функции для всех вещественных значений аргумента x сводится к возможности установления взаимно однозначного соответствия между точками единичной окружности и всеми вещественными числами полусегмента $[0, 2\pi)$.

Всеми этими вопросами мы и будем заниматься в настоящем параграфе.

1. Показательная функция. Начнем наше рассмотрение с определения рациональных степеней положительных чисел. Для того чтобы возвести любое вещественное число x в целую положительную степень n, следует умножить это число x само на себя n раз.

Следовательно, при целом n мы можем считать определенной степенную функцию $y=x^n$ для всех вещественных значений x. Установим некоторые простейшие свойства этой функции.

Утверждение 1. *Степенная функция $y=x^n$ при $x\geqslant 0$ и целом положительном n возрастает и непрерывна.*

Доказательство. Покажем, что функция $y=x^n$ возрастает. Пусть $0\leqslant x_1<x_2$. Тогда $x_2^n-x_1^n=(x_2-x_1)(x_2^{n-1}+x_2^{n-2}\cdot x_1+\ldots+x_1^{n-1})$. Оба сомножителя в правой части, в соответствии с выбором значений x_2 и x_1, положительны. Поэтому положительна и левая часть равенства, т. е. $x_2^n>x_1^n$, а это означает возрастание функции $y=x^n$ при $x\geqslant 0$.

Непрерывность функции $y=x^n$ в любой точке a бесконечной прямой $(-\infty, +\infty)$ была установлена в примере 1 п. 1 § 1 настоящей главы. Утверждение 1 доказано.

Рассмотрим степенную функцию $y=x^n$ на сегменте $[0, N]$, где N — любое положительное число. Так как эта функция непрерывна и возрастает на указанном сегменте, то в силу теоремы 4.5 она имеет на сегменте $[0, N^n]$ возрастающую и непрерывную обратную функцию, которую мы обозначим через $x=y^{1/n}$. Поскольку N можно выбрать как угодно большим, то и N^n также можно сделать сколь угодно большим. Следовательно, функция $x=y^{1/n}$ определена для всех неотрицательных значений y. Меняя для этой функции обозначение аргумента y на x, а обозначение функции x на y, мы получим степенную функцию $y=x^{1/n}$, определенную для всех вещественных $x\geqslant 0$.

Теперь мы в состоянии определить любую рациональную степень r положительного числа a. Определим, прежде всего, $a^{1/n}$ как вещественное число b, равное значению функции $x^{1/n}$ в точке a. Далее, если $r=\frac{m}{n}$, где m и n — целые положительные числа, то мы положим

$$a^r=a^{\frac{m}{n}}=(a^{1/n})^m.$$

Кроме того, положим по определению

$$a^0=1,\ a^{-r}=\left(\frac{1}{a}\right)^r \text{ при } r>0.$$

Тем самым, мы определили любую рациональную степень положительного вещественного числа a.

Выполняются следующие свойства рациональной степени положительных вещественных чисел:

$$(a^r)^s=a^{r\cdot s},\ a^r\cdot b^r=(a\cdot b)^r,\ a^r\cdot a^s=a^{r+s}. \qquad (*)$$

Докажем сначала справедливость первого свойства $(*)$.

Заметим, что при целом положительном p равенство $(a^{\frac{m}{n}})^p=$

$=a^{\frac{m\cdot p}{n}}$, в котором под m и n понимаются любые целые положительные числа, заведомо справедливо, ибо как левая, так и правая части этого равенства равны произведению числа $a^{1/n}$ самого на себя $m\cdot p$ раз.

Полагая $r=\frac{m_1}{n_1}$, $s=\frac{m_2}{n_2}$, докажем равенство $(a^r)^s=a^{r\cdot s}$ в ситуации любых положительных рациональных r и s. Положим $c_1=\left(a^{\frac{m_1}{n_1}}\right)^{\frac{m_2}{n_2}}$, $c_2=a^{\frac{m_1\cdot m_2}{n_1\cdot n_2}}$. Если бы c_1 было отлично от c_2, то из возрастания степенной функции $y=x^{n_2}$ следовало бы, что и $c_1^{n_2}\neq c_2^{n_2}$, а последнее соотношение, в силу уже доказанной справедливости равенства $\left(a^{\frac{m}{n}}\right)^p=a^{\frac{m\cdot p}{n}}$ при целом p, означало бы, что $\left(a^{\frac{m_1}{n_1}}\right)^{m_2}\neq a^{\frac{m_1\cdot m_2}{n_1}}$. Полученное соотношение противоречит уже доказанному нами для целых положительных m_1, n_1 и m_2 равенству $\left(a^{\frac{m_1}{n_1}}\right)^{m_2}=a^{\frac{m_1\cdot m_2}{n_1}}$. Тем самым $c_1=c_2$, и первое равенство (*) доказано для любых положительных рациональных r и s.

Распространение этого равенства на неположительные r и s не представляет труда в силу нашей договоренности о том, что

$$a^0=1,\ a^{-r}=\frac{1}{a^r}\ \text{при}\ r>0.$$

Второе равенство (*) также достаточно доказать для положительного рационального r. Полагая это r равным m/n, где m и n — целые положительные числа, заметим, что нам достаточно доказать равенство $a^{1/n}\cdot b^{1/n}=(a\cdot b)^{1/n}$, ибо перемножением m таких равенств будет доказано общее соотношение $a^r\cdot b^r=(a\cdot b)^r$.

Для доказательства равенства $a^{1/n}\cdot b^{1/n}=(a\cdot b)^{1/n}$ заметим, что в силу свойств взаимно обратных функций $y=x^{1/n}$ и $x=y^n$ мы можем утверждать, что $(b^{1/n})^n=b$, $(a^{1/n})^n=a$, $((ab)^{1/n})^n=ab$. Поэтому, положив $c_1=a^{1/n}\cdot b^{1/n}$, $c_2=(ab)^{1/n}$ и предполагая, что $c_1\neq c_2$, мы получили бы, что $c_1^n\neq c_2^n$, что противоречит равенству $a\cdot b=ab$.

Докажем теперь последнее свойство (*), учитывая, что первые два уже доказаны. Пусть $r=\frac{m_1}{n_1}$, $s=\frac{m_2}{n_2}$, тогда $r=\frac{m_1\cdot n_2}{n_1\cdot n_2}$, $s=\frac{m_2\cdot n_1}{n_2\cdot n_1}$, и мы приходим к следующему равенству:

$$a^r\cdot a^s=\left(a^{\frac{1}{n_1\cdot n_2}}\right)^{m_1\cdot n_2}\cdot\left(a^{\frac{1}{n_1\cdot n_2}}\right)^{m_2\cdot n_1}=\left(a^{\frac{1}{n_1\cdot n_2}}\right)^{m_1\cdot n_2+m_2\cdot n_1}.$$

(Последнее равенство справедливо, так как $m_1\cdot n_2$ и $m_2\cdot n_1$ — целые числа.)

Таким образом,

$$a^r \cdot a^s = a^{\frac{m_1 \cdot n_2 + m_2 \cdot n_1}{n_1 \cdot n_2}} = a^{\frac{m_1}{n_1} + \frac{m_2}{n_2}} = a^{r+s},$$

что и требовалось.

При $a>1$ и рациональном $r>0$ справедливо неравенство $a^r>1$.

В самом деле, пусть $r=\frac{m}{n}$ и $a^r=a^{m/n}\leqslant 1$. Перемножая почленно n указанных неравенств, получим $a^m\leqslant 1$. Но это неравенство противоречит неравенству $a^m>1$, полученному почленным перемножением m неравенств вида $a>1$.

Отметим также, что если рациональная дробь $r=\frac{m}{n}$ имеет нечетный знаменатель n, то определение рациональной степени можно распространить и на отрицательные числа, полагая при $a>0$, что

$$(-a)^r=a^r, \text{ если } m \text{ четное},$$

$$(-a)^r=-a^r, \text{ если } m \text{ нечетное}.$$

Убедимся в том, что *функция $y=a^x$ при $a>1$, определенная нами на множестве рациональных чисел, монотонно возрастает на этом множестве.*

Действительно, пусть r_1 и r_2 — два рациональных числа таких, что $r_2>r_1$. Тогда

$$a^{r_2}-a^{r_1}=a^{r_1}(a^{r_2-r_1}-1). \tag{4.3}$$

Поскольку $r_2-r_1>0$ и $a>1$, то (в силу установленного выше) $a^{r_2-r_1}>1$. Таким образом, правая часть равенства (4.3) положительна. Следовательно,

$$a^{r_2}-a^{r_1}>0, \text{ т. е. } a^{r_2}>a^{r_1},$$

что и требовалось.

Определим, наконец, функцию $y=a^x$ не только для рациональных значений x, но и для любых вещественных значений. Пусть x — произвольное вещественное число. Рассмотрим всевозможные рациональные числа α и β, удовлетворяющие неравенствам

$$\alpha<x<\beta. \tag{4.4}$$

Определим a^x при $a>1$ как вещественное число y, удовлетворяющее неравенствам

$$a^\alpha\leqslant y\leqslant a^\beta \tag{4.5}$$

при всевозможных рациональных α и β, удовлетворяющих неравенствам (4.4).

Оказывается, что такое число y существует и притом только одно. Следовательно, таким путем функция $y=a^x$ будет определена на множестве всех вещественных x.

Мы покажем, что эта функция возрастает и непрерывна на всей вещественной прямой. Эти результаты содержатся в доказываемых ниже утверждениях.

Утверждение 2. *Для любых фиксированных вещественных чисел x и $a>1$ и всевозможных рациональных чисел α и β, удовлетворяющих неравенствам* (4.4), *существует и притом единственное вещественное число y, удовлетворяющее неравенствам* (4.5).

Доказательство. Докажем сначала существование такого числа y. Фиксируем произвольное рациональное число β, удовлетворяющее правому неравенству (4.4), и рассмотрим всевозможные рациональные числа α, удовлетворяющие левому неравенству (4.4). Так как $\alpha<\beta$ и показательная функция, определенная на множестве рациональных чисел, возрастает, то $a^\alpha<a^\beta$. Таким образом, множество $\{a^\alpha\}$ ограничено сверху, и число a^β является одной из верхних граней этого множества. Из основной теоремы 2.1 следует, что множество $\{a^\alpha\}$ имеет точную верхнюю грань, которую мы обозначим через y. Покажем, что y удовлетворяет неравенствам (4.5). Из определения верхней грани вытекает справедливость левого неравенства (4.5), а справедливость правого неравенства (4.5) вытекает из того, что a^β — одна из верхних граней, а y — точная верхняя грань множества $\{a^\alpha\}$.

Докажем теперь, что такое число y только одно. Достаточно доказать, что для любого $\varepsilon>0$ найдутся такие рациональные числа α и β, удовлетворяющие неравенствам (4.4), для которых $a^\beta-a^\alpha<\varepsilon$. Тогда любые два числа y_1 и y_2, удовлетворяющие неравенствам (4.5), обязаны совпасть, так как разность между ними по модулю меньше любого наперед заданного числа $\varepsilon>0$.

Фиксируем произвольное положительное число ε и некоторое рациональное число β_0, удовлетворяющее правому неравенству (4.4). Тогда так как $a^\alpha<a^{\beta_0}$, то

$$a^\beta-a^\alpha=a^\alpha(a^{\beta-\alpha}-1)\leqslant a^{\beta_0}(a^{\beta-\alpha}-1).$$

Неравенство $a^\beta-a^\alpha<\varepsilon$ будет доказано, если мы установим возможность выбора в неравенствах (4.4) таких рациональных α и β, что $a^{\beta-\alpha}-1<\dfrac{\varepsilon}{a^{\beta_0}}$.

В гл. 2 было доказано, что для любого натурального n можно выбрать рациональные числа α и β, удовлетворяющие неравенствам (4.4), так, что разность $\beta-\alpha$ будет меньше $1/n$. Таким образом, достаточно доказать, что существует такое натуральное n, что

$$a^{1/n}-1<\frac{\varepsilon}{a^{\beta_0}}.$$

Пусть $a^{1/n}=1+\delta_n$. Так как $a^{1/n}>1$, то δ_n положительно. Используя первые два члена бинома Ньютона, мы получим, что

$$a=(a^{1/n})^n=(1+\delta_n)^n>1+n\cdot\delta_n.$$

Отсюда $a-1>n\cdot\delta_n$, т. е. $0<\delta_n<\frac{a-1}{n}$. Значит $a^{1/n}-1<\frac{a-1}{n}$. Выберем теперь натуральное n удовлетворяющим неравенству $\frac{a-1}{n}<\frac{\varepsilon}{a^{\beta_0}}$ или $n>\frac{(a-1)\cdot a^{\beta_0}}{\varepsilon}$. Тогда $a^{1/n}-1<\frac{a-1}{n}<\frac{\varepsilon}{a^{\beta_0}}$, и доказательство однозначной определенности числа y, удовлетворяющего неравенствам (4.5), завершено. Утверждение 2 доказано.

Заметим, что если x — рациональное число и a^x — значение в точке x показательной функции, первоначально определенной лишь на множестве рациональных чисел, то a^x и является тем единственным числом y, которое удовлетворяет неравенствам (4.5).

Утверждение 3. *Показательная функция $y=a^x$ при $a>1$ возрастает на всей бесконечной прямой.*

Доказательство. Пусть x_1 и x_2 — любые два вещественных числа такие, что $x_1<x_2$. Всегда существуют рациональные числа α и β такие, что $x_1<\alpha<\beta<x_2$ (см. лемму 2 § 3 гл. 2). Так как $x_1<\alpha$ и $\beta<x_2$, то по определению показательной функции выполнены неравенства $a^{x_1}\leqslant a^\alpha$, $a^\beta\leqslant a^{x_2}$. С другой стороны, так как $\alpha<\beta$, то из возрастания показательной функции на множестве рациональных чисел вытекает $a^\alpha<a^\beta$. Сопоставляя неравенства $a^{x_1}\leqslant a^\alpha$, $a^\alpha<a^\beta$ и $a^\beta\leqslant a^{x_2}$ и используя свойство транзитивности знаков $>$ и $=$, получим, что $a^{x_1}<a^{x_2}$, а это и доказывает возрастание функции a^x. Утверждение доказано.

Утверждение 4. *Показательная функция $y=a^x$ при $a>1$ является непрерывной функцией в любой точке бесконечной прямой.*

Доказательство. Пусть x — произвольное вещественное число, а $\{x_n\}$ — любая сходящаяся к x последовательность. В силу определения непрерывности по Гейне достаточно доказать, что для любого $\varepsilon>0$ существует такой номер N, что $|a^{x_n}-a^x|<\varepsilon$ при всех $n\geqslant N$. Фиксируем произвольное $\varepsilon>0$ и по нему рациональные числа α и β такие, что $\alpha<x<\beta$ и $a^\beta-a^\alpha<\varepsilon$. Возможность фиксировать по любому $\varepsilon>0$ такие рациональные числа α и β была установлена в утверждении 2. Поскольку последовательность $\{x_n\}$ сходится к x и $\alpha<x<\beta$, то существует такой номер N, что при всех $n\geqslant N$ справедливы неравенства $\alpha<x_n<\beta$. Так как показательная функция монотонно возрастает, то $a^\alpha<a^x<a^\beta$, $a^\alpha<a^{x_n}<a^\beta$ при всех $n\geqslant N$.

Таким образом, оба числа a^{x_n} и a^x при $n\geqslant N$ заключены между двумя числами a^α и a^β, разность между которыми $a^\beta-a^\alpha$ меньше ε. Отсюда следует, что при $n\geqslant N$ справедливо неравенство

$|a^{x_n} - a^x| < \varepsilon$, которое и доказывает непрерывность показательной функции в произвольной точке x. Утверждение 4 доказано.

Получим теперь некоторые следствия из доказанных свойств показательной функции. Прежде всего заметим, что если $0<a<1$, то $a = \frac{1}{b}$, где $b>1$. Поэтому функцию $y=a^x$ при $0<a<1$ можно определить как функцию $y=b^{-x}$ при $b=\frac{1}{a}>1$.

Следствие 1. *Показательная функция $y=a^x$ при $a>1$ положительна (при всех значениях x).*

Если x — произвольная точка числовой оси, а r — рациональное число такое, что $r<x$, то по определению показательной функции на множестве рациональных чисел $a^r>0$, а по утверждению 3 $a^x>a^r$ при $a>1$. Следовательно, $a^x>0$.

Следствие 2. *Показательная функция $y=a^x$ при $a>1$ удовлетворяет условиям:* $\lim\limits_{n\to-\infty} a^x = 0$, $\lim\limits_{x\to+\infty} a^x = +\infty$.

В самом деле, так как $a>1$, то $a=1+\delta$, где $\delta>0$ и $a^n = (1+\delta)^n > n\delta$. Следовательно, $\lim\limits_{n\to\infty} a^n = +\infty$. В силу монотонности функции $y=a^x$ получаем, что и $\lim\limits_{x\to+\infty} a^x = +\infty$. Так как $a^{-n} = \frac{1}{a^n}$, то $\lim\limits_{n\to\infty} a^{-n} = 0$, и поэтому $\lim\limits_{x\to-\infty} a^x = 0$.

Следствие 3. *Значения функции $y=a^x$ при $a>1$ заполняют всю положительную полупрямую $y>0$.*

Действительно, по следствию 1 функция $y=a^x$ принимает только положительные значения, а по следствию 2 она принимает как сколь угодно малые, так и сколь угодно большие положительные значения. Из непрерывности и строгой монотонности a^x и из теоремы 4.4 вытекает, что любое положительное число является значением функции $y=a^x$.

Следствие 4. *Для любых вещественных чисел x_1 и x_2 справедливы соотношения*

$$(a^{x_1})^{x_2} = a^{x_1 \cdot x_2}, \quad a^{x_1} \cdot b^{x_1} = (a \cdot b)^{x_1}, \quad a^{x_1} \cdot a^{x_2} = a^{x_1 + x_2}.$$

В самом деле, эти соотношения уже были установлены нами для рациональных показателей. Отсюда вытекает справедливость их и для произвольных вещественных показателей. Убедимся, например, в справедливости первого соотношения. Пусть $\{r'_n\}$ и $\{r''_n\}$ — последовательности рациональных чисел, сходящиеся соответственно к x_1 и x_2. Тогда $(a^{r'_n})^{r''_n} = a^{r'_n \cdot r''_n}$. Переходя к пределу при $n\to\infty$ и используя непрерывность показательной функции, получим $(a^{x_1})^{x_2} = a^{x_1 \cdot x_2}$. Аналогично устанавливаются и остальные равенства.

Заметим теперь, что мы фактически изучили и свойства показательной функции $y=a^x$ при $0<a<1$. Действительно, ее непрерывность следует из самого определения. Из определения следует также, что эта функция монотонно убывает на бесконечной прямой. Следствия 1, 3, 4 верны и для функции $y=a^x$ при $0<a<1$, а следствие 2, очевидно, будет выглядеть так:

$$\lim_{x\to-\infty} a^x=+\infty,\quad \lim_{x\to+\infty} a^x=0.$$

На рис. 4.1 и 4.2 изображены графики показательной функции $y=a^x$ для случаев $a>1$ и $0<a<1$.

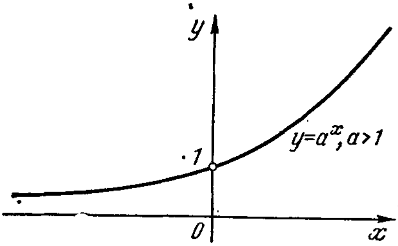


Рис. 4.1

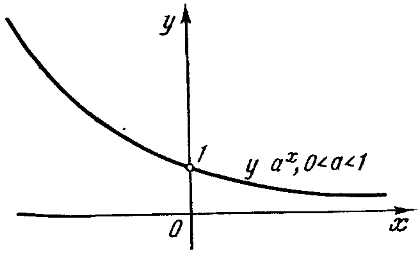


Рис. 4.2

З а м е ч а н и е. Показательную функцию можно было бы определить как решение некоторого функционального уравнения, удовлетворяющее определенным условиям. Можно доказать, что существует, и притом единственная, функция $f(x)$, определенная на всей бесконечной прямой и удовлетворяющая трем требованиям:

1) для любых вещественных x_1 и x_2 выполнено соотношение $f(x_1+x_2)=f(x_1)\cdot f(x_2)$;

2) $f(0)=1$*, $f(1)=a$ при $a>1$;

3) функция $f(x)$ непрерывна при $x=0$.

Такой функцией и является построенная выше функция $f(x)=a^x$ при $a>1$.

2. Логарифмическая функция. Логарифмическую функцию мы определим как обратную к показательной. Пусть $[c, d]$ — произвольный сегмент бесконечной прямой. На этом сегменте функция $y=a^x$ при $a>1$ возрастает и непрерывна. Поэтому в силу теоремы 4.5 функция $y=f(x)=a^x$ имеет на сегменте $[a^c, a^d]$ возрастающую и непрерывную обратную функцию $x=f^{-1}(y)$, которая и называется л о г а р и ф м и ч е с к о й и обозначается так:

$$x=f^{-1}(y)=\log_a y.$$

Заменяя обозначение аргумента y на x, а обозначение функции x на y, запишем ее в более привычном нам виде:

$$y=\log_a x.$$

---

* Можно доказать, что требование $f(0)=1$ является следствием остальных требований (и потому может быть опущено).

Случай $0<a<1$ рассматривается аналогично. Отметим следующие свойства логарифмической функции, вытекающие из ее определения:

1) *Логарифмическая функция определена для всех положительных значений* x. В самом деле, аргумент логарифмической функции представляет собой значения показательной функции, которые, как мы знаем, только положительны и заполняют всю полупрямую $x>0$.

2) *Логарифмическая функция непрерывна и возрастает на всей полупрямой* $x>0$ *при* $a>1$ *и непрерывна и убывает на всей полупрямой* $x>0$ *при* $0<a<1$, *причем*

$$\lim_{x\to 0+0}\log_a x = -\infty,\quad \lim_{x\to +\infty}\log_a x = +\infty \text{ при } a>1,$$

$$\lim_{x\to 0+0}\log_a x = +\infty,\quad \lim_{x\to +\infty}\log_a x = -\infty \text{ при } 0<a<1.$$

Справедливость этих свойств вытекает из свойств показательной функции.

3) *Для любых положительных* x_1 *и* x_2

$$\log_a(x_1\cdot x_2) = \log_a x_1 + \log_a x_2.$$

Это свойство также вытекает из свойств показательной функции.

Замечание. Следует особо выделить логарифмическую функцию $y=\log_e x$, где $e=\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n$. Для этой функции используется обозначение $y=\ln x$. Логарифмы по основанию e называются натуральными.

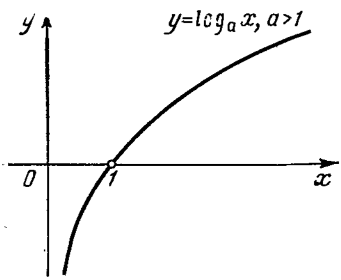


Рис. 4.3

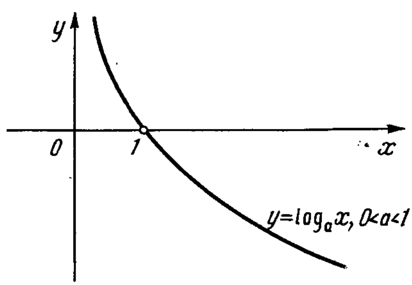


Рис. 4.4

На рис. 4.3 и 4.4 изображены графики логарифмической функции $y=\log_a x$ для случаев $a>1$ и $0<a<1$.

3. Степенная функция. Определим теперь степенную функцию с любым вещественным показателем α через суперпозицию логарифмической функции и показательной. Пусть $x>0$. Тогда общая степенная функция определяется так:

$$y = x^\alpha = (a^{\log_a x})^\alpha = a^{\alpha \cdot \log_a x},$$

где a — любое фиксированное число, ради определенности больше единицы:

Из этого определения и из того, что при $a>1$ логарифмическая функция возрастает на всей полупрямой $x>0$, а показательная функция возрастает на всей бесконечной прямой, вытекает, что *степенная функция* $y=x^\alpha=a^{\alpha\cdot\log_a x}$ *возрастает при* $\alpha>0$ *и убывает при* $\alpha<0$ *на полупрямой* $x>0$.

Справедливы следующие свойства:

1) *Для степенной функции выполнены соотношения:*

$$\lim_{x\to 0+0} x^\alpha = 0 \text{ при } \alpha > 0, \ \lim_{x\to 0+0} x^\alpha = +\infty \text{ при } \alpha < 0.$$

В самом деле, пусть $\{x_n\}$ — любая сходящаяся к нулю справа последовательность значений аргумента x_n. Так как $\lim\limits_{n\to\infty} \log_a x_n = -\infty$, то из свойств показательной функции вытекает, что $\lim\limits_{n\to\infty} a^{\alpha\cdot\log_a x_n} = 0$ при $\alpha>0$ и $\lim\limits_{n\to\infty} a^{\alpha\cdot\log_a x_n} = +\infty$ при $\alpha<0$. По определению положим $0^\alpha=0$ при $\alpha>0$ и будет считать это выражение неопределенным при $\alpha\leqslant 0$.

2) *Степенная функция* $y=x^\alpha=a^{\alpha\cdot\log_a x}$ *непрерывна в каждой точке* x *открытой полупрямой* $x>0$.

Это сразу же вытекает из теоремы 4.2 непрерывности сложной функции с учетом того, что функция $u \doteq \alpha\cdot\log_a x$ непрерывна в любой точке $x>0$, а функция $y=a^u$ непрерывна в любой точке u бесконечной прямой.

Замечание. Если показатель α степенной функции представляет собой рациональное число m/n, где n — нечетное целое число, то степенную функцию $y=x^\alpha$ можно определить на всей числовой оси, полагая при $x<0$:

$$y=|x|^\alpha, \text{ если } \alpha=\frac{m}{n} \text{ и } m \text{ четное},$$

$$y=-|x|^\alpha, \text{ если } \alpha=\frac{m}{n} \text{ и } m \text{ нечетное}.$$

На рис. 4.5—4.7 изображены графики степенной функции $y=x^\alpha$ для различных значений α.

4. Тригонометрические функции. Мы уже имеем представление о тригонометрических функциях $y=\sin x$ и $y=\cos x$ и функциях, которые через них выражаются, $y=\operatorname{tg} x=\frac{\sin x}{\cos x}$, $y=\operatorname{ctg} x=\frac{\cos x}{\sin x}$, $y=\sec x=\frac{1}{\cos x}$, $y=\operatorname{cosec} x=\frac{1}{\sin x}$.

Во введении к этому параграфу мы уже говорили о логических пробелах, возникающих при определении функций $y=\sin x$ и

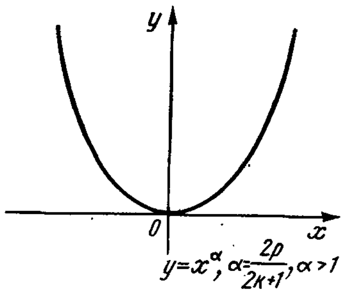


Рис. 4.5

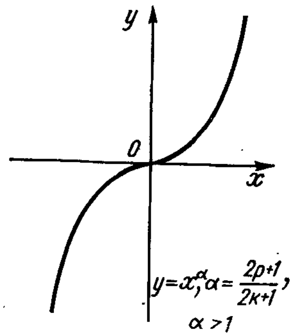


Рис. 4.6

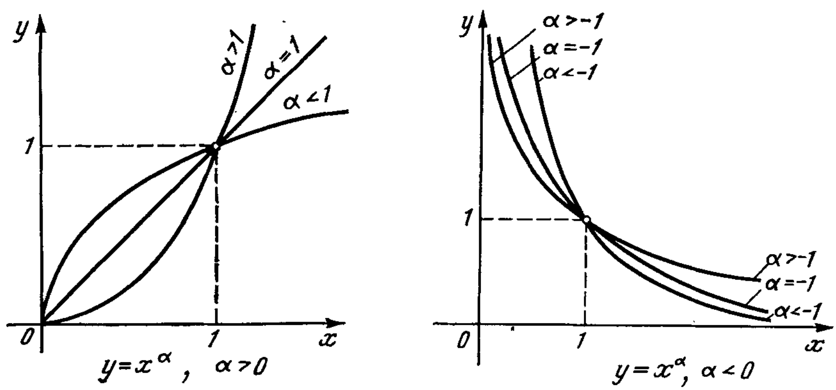


Рис. 4.7

$y=\cos x$ с помощью наглядных геометрических соображений. Логически безупречно эти функции можно определить как решение некоторой системы функциональных уравнений. Точнее, можно доказать следующее утверждение: *существует и притом единственная пара функций $f(x)$ и $g(x)$, определенных для всех вещественных значений аргумента x и удовлетворяющих условиям:*

1) $f(x_1+x_2)=f(x_1)g(x_2)+f(x_2)\cdot g(x_1)$,
$g(x_1+x_2)=g(x_1)\cdot g(x_2)-f(x_1)\cdot g(x_2)$, $\quad f^2(x)+g^2(x)=1$;

2) $f(0)=0$, $g(0)=1$, $f\left(\frac{\pi}{2}\right)=1$, $g\left(\frac{\pi}{2}\right)=0$;

3) *если* $0<x<\frac{\pi}{2}$, *то* $0<f(x)<x<\frac{f(x)}{g(x)}$*. $\qquad$ (4.6)

---

* Справедливость неравенства $x<\frac{f(x)}{g(x)}$ для $0<x<\frac{\pi}{2}$ следует из остальных сформулированных здесь условий (см. по этому поводу указываемое ниже дополнение к книге В. А. Ильина и Э. Г. Позняка).

Первую из этих функций назовем **синусом** и обозначим символом $f(x)=\sin x$, вторую назовем **косинусом** и обозначим символом $g(x)=\cos x$.

Доказательство приведенного утверждения можно найти в дополнении к гл. 4 книги В. А. Ильина и Э. Г. Позняка «Основы математического анализа», 1 (М., Наука, 1982).

Нетрудно доказать, что из свойств 1), 2) и 3) можно извлечь в виде следствий все другие свойства функций $y=\sin x$ и $y=\cos x$, известные читателю из школьных учебников и устанавливаемые в средней школе из наглядных геометрических соображений. Впрочем, этот факт сразу вытекает из того, что свойства 1), 2) и 3) определяют **единственную** пару функций $f(x)$ и $g(x)$ и что введенные в средней школе из наглядных геометрических соображений функции $f(x)=\sin x$ и $g(x)=\cos x$ этими тремя свойствами обладают.

В качестве примера установим с помощью свойств 1), 2), 3) некоторые свойства функций $y=\sin x$ и $y=\cos x$, которые понадобятся нам при доказательстве непрерывности этих функций и для отыскания участков их монотонности.

а) Из третьего соотношения 1), имеющего вид $\sin^2 x+\cos^2 x=1$, сразу же вытекает, что $\sin^2 x\leqslant 1$ и $\cos^2 x\leqslant 1$, т. е.

$$|\sin x|\leqslant 1, \qquad |\cos x|\leqslant 1. \tag{4.7}$$

б) Далее, с помощью первых двух соотношений 1) и первых двух равенств 2) мы получим, что

$$\sin 0=\sin[x+(-x)]=\sin x\cdot\cos(-x)+\cos x\cdot\sin(-x)=0,$$
$$\cos 0=\cos[x+(-x)]=\cos x\cdot\cos(-x)-\sin x\cdot\sin(-x)=1.$$

Полученные два равенства представляют собой систему двух уравнений относительно двух неизвестных $\cos(-x)$ и $\sin(-x)$. Решая эту систему и учитывая, что $\sin^2 x+\cos^2 x=1$, мы получим, что

$$\cos(-x)=\cos x, \qquad \sin(-x)=-\sin x, \tag{4.8}$$

т. е. $\cos x$ представляет собой четную функцию, а $\sin x$ — нечетную функцию*.

в) Из соотношений 1), в свою очередь, вытекают равенства

$$\begin{aligned}\sin(x_1-x_2)&=\sin[x_1+(-x_2)]=\sin x_1\cdot\cos(-x_2)+\\&+\cos x_1\cdot\sin(-x_2)=\sin x_1\cdot\cos x_2-\cos x_1\cdot\sin x_2,\\ \cos(x_1-x_2)&=\cos[x_1+(-x_2)]=\cos x_1\cdot\cos(-x_2)-\\&-\sin x_1\cdot\sin(-x_2)=\cos x_1\cdot\cos x_2+\sin x_1\cdot\sin x_2.\end{aligned} \tag{4.9}$$

* Функция $\varphi(x)$, определенная для всех вещественных значений x, называется четной, если $\varphi(-x)=\varphi(x)$ (для любого значения x), и нечетной, если $\varphi(-x)=-\varphi(x)$ (также для любого значения x).

г) Из первого соотношения 1) и первого соотношения (4.9) мы получим, что

$$\sin x_2 = \sin\left(\frac{x_2+x_1}{2}+\frac{x_2-x_1}{2}\right)=$$

$$=\sin\frac{x_2+x_1}{2}\cdot\cos\frac{x_2-x_1}{2}+\sin\frac{x_2-x_1}{2}\cdot\cos\frac{x_2+x_1}{2},$$

$$\sin x_1 = \sin\left(\frac{x_2+x_1}{2}-\frac{x_2-x_1}{2}\right)=$$

$$=\sin\frac{x_2+x_1}{2}\cdot\cos\frac{x_2-x_1}{2}-\sin\frac{x_2-x_1}{2}\cdot\cos\frac{x_2+x_1}{2}.$$

Складывая и вычитая полученные два равенства, мы придем к соотношениям

$$\begin{aligned}\sin x_2+\sin x_1 &= 2\sin\frac{x_2+x_1}{2}\cdot\cos\frac{x_2-x_1}{2},\\ \sin x_2-\sin x_1 &= 2\cos\frac{x_2+x_1}{2}\cdot\sin\frac{x_2-x_1}{2}.\end{aligned}\qquad(4.10)$$

д) Далее, из первого соотношения (4.9) и из последних двух равенств 2) получим, что

$$\sin\left(\frac{\pi}{2}-x\right)=\sin\frac{\pi}{2}\cdot\cos x-\cos\frac{\pi}{2}\cdot\sin x=\cos x,$$

т. е.

$$\cos x=\sin\left(\frac{\pi}{2}-x\right).\qquad(4.11)$$

е) Убедимся, наконец, в периодичности функций $f(x)=\sin x$ и $g(x)=\cos x$ с периодом 2π. Из первых двух соотношений 1) при $x=x_1=x_2$ получим, что

$$\sin 2x=2\sin x\cdot\cos x,\qquad \cos 2x=\cos^2 x-\sin^2 x.\qquad(4.12)$$

Учитывая, что в силу равенств 2) $\sin\frac{\pi}{2}=1$, $\cos\frac{\pi}{2}=0$, мы получим из соотношений (4.12) при $x=\frac{\pi}{2}$, что $\sin\pi=0$, $\cos\pi=-1$, а из последних двух равенств и из соотношений (4.12) при $x=\pi$ получим, что $\sin 2\pi=0$, $\cos 2\pi=1$.

Используя последние два равенства, мы получим из первых двух соотношений 1), что

$$\sin(x+2\pi)=\sin x\cdot\cos 2\pi+\sin 2\pi\cdot\cos x=\sin x,$$

$$\cos(x+2\pi)=\cos x\cdot\cos 2\pi-\sin x\cdot\sin 2\pi=\cos x,$$

а это и означает периодичность функций $\sin x$ и $\cos x$ с периодом 2π*.

ж) В заключение несколько усилим неравенства, содержащиеся в свойстве 3). Мы установим, что для всех вещественных x справедливо следующее несколько общее неравенство:

$$|\sin x| \leqslant |x|. \qquad (4.13)$$

При $0 < x < \frac{\pi}{2}$ соотношение (4.13) следует из неравенств, содержащихся в свойстве 3).

При $-\frac{\pi}{2} < x < 0$, в силу соотношения $\sin(-x) = -\sin x$, неравенство (4.13) следует из следующих неравенств:

$$0 < \sin(-x) < -x \text{ при } -\frac{\pi}{2} < x < 0,$$

а эти последние неравенства справедливы вследствие того, что $(-x)$ лежит в интервале $\left(0, \frac{\pi}{2}\right)$. При $x=0$ $\sin x = x$.

При $\frac{\pi}{2} \leqslant |x|$ имеем $|\sin x| \leqslant 1 < \frac{\pi}{2} \leqslant |x|$, т. е. $|\sin x| \leqslant \leqslant |x|$.

Перейдем к установлению двух основных свойств функций $f(x) = \sin x$ и $g(x) = \cos x$.

1°. *Функции* $\sin x$ *и* $\cos x$ *непрерывны в каждой точке* x *бесконечной прямой.*

Доказательство. Достаточно установить непрерывность в каждой точке x только функции $f(x) = \sin x$, ибо непрерывность в каждой точке x функции $g(x) = \cos x$ будет при этом вытекать из соотношения (4.11) и теоремы 4.2.

Сначала докажем, что функция $\sin x$ непрерывна в точке $x=0$. Так как в силу первого равенства 2) $\sin 0 = 0$, то в силу определения непрерывности по Гейне достаточно доказать, что для любой бесконечно малой последовательности $\{x_n\}$ последовательность значений функции $\{\sin x_n\}$ также является бесконечно малой.

Из неравенства (4.13) и из условия $|\sin x| \geqslant 0$ вытекает, что для всех x

$$0 \leqslant |\sin x| \leqslant |x|. \qquad (4.13')$$

Следовательно,

$$0 \leqslant |\sin x_n| \leqslant |x_n|.$$

Последние неравенства в силу принципа двустороннего ограничения (см. теорему 3.14 гл. 3) означают, что последовательность

---

* Функция $\varphi(x)$, определенная для всех вещественных x, называется периодической с периодом T, если $\varphi(x+T) = \varphi(x)$ для всех x.

$\{|\sin x_n|\}$, а значит, и последовательность $\{\sin x_n\}$ является бесконечно малой. Непрерывность функции $\sin x$ в точке $x=0$ доказана.

Докажем теперь, что функция $\sin x$ непрерывна в любой точке x бесконечной прямой. Пусть $\{x_n\}$ — произвольная последовательность, сходящаяся к x. Достаточно доказать, что соответствующая последовательность $\{\sin x_n\}$ сходится к $\sin x$.

Воспользуемся вторым соотношением (4.10), положив в нем $x_2=x_n$, $x_1=x$. Получим

$$\sin x_n-\sin x=2\cos\frac{x_n+x}{2}\cdot\sin\frac{x_n-x}{2}. \tag{4.14}$$

Достаточно доказать, что в правой части (4.14) стоит элемент бесконечно малой последовательности, но это сразу вытекает из того, что последовательность $\left\{\sin\frac{x_n-x}{2}\right\}$, в силу уже доказанной непрерывности синуса в нуле, является бесконечно малой, а последовательность $\left\{2\cos\frac{x_n+x}{2}\right\}$, в силу второго неравенства (4.7) является ограниченной.

2°. *Функция* $\sin x$ *возрастает на каждом из сегментов* $\left[2k\pi-\frac{\pi}{2},\ 2k\pi+\frac{\pi}{2}\right]$ *и убывает на каждом из сегментов* $\left[(2k+1)\pi-\frac{\pi}{2},\ (2k+1)\pi+\frac{\pi}{2}\right]$; *функция* $\cos x$ *убывает на каждом из сегментов* $[2k\pi,\ 2k\pi+\pi]$ *и возрастает на каждом из сегментов* $[2k\pi-\pi,\ 2k\pi]$ (*здесь всюду* k — *любое целое число, т. е.* $k=0,\ \pm1,\ \pm2,\dots$).

Доказательство. Все рассуждения достаточно провести для функции $\sin x$, ибо после нахождения всех участков монотонности функции $\sin x$ участки монотонности функции $\cos x$ могут быть получены, исходя из равенства (4.11).

Далее, поскольку $\sin x$ — периодическая функция с периодом 2π, то достаточно найти участки ее монотонности, лежащие в пределах одного периода, т. е., например, для значений аргумента из сегмента $\left[-\frac{\pi}{2},\ 2\pi-\frac{\pi}{2}\right]$.

Докажем сначала, что функция $\sin x$ возрастает на сегменте $\left[0,\ \frac{\pi}{2}\right]$. Пусть x_1 и x_2 — любые два числа из этого сегмента такие, что $x_2>x_1$. Тогда, очевидно, числа $\frac{x_2+x_1}{2}$ и $\frac{x_2-x_1}{2}$ принадлежат интервалу $\left(0,\ \frac{\pi}{2}\right)$, причем в силу второго равенства (4.10)

$$\sin x_2-\sin x_1=2\cos\frac{x_2+x_1}{2}\cdot\sin\frac{x_2-x_1}{2}. \tag{4.15}$$

Достаточно доказать, что в правой части (4.15) стоит положительное число, а для этого достаточно убедиться в том, что для всех значений аргумента из интервала $\left(0, \frac{\pi}{2}\right)$ функции $\sin x$ и $\cos x$ принимают только положительные значения. Для функции $f(x)=\sin x$ это вытекает из свойства 3), а для функции $\cos x$ следует из равенства (4.11).

Итак, доказано, что функция $\sin x$ возрастает на сегменте $\left[0, \frac{\pi}{2}\right]$. Из нечетности функции $\sin x$, т. е. из второго соотношения (4.8), в таком случае следует, что функция $\sin x$ возрастает и на сегменте $\left[-\frac{\pi}{2}, 0\right]$.

Тем самым доказано, что функция $\sin x$ возрастает на сегменте $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Остается исследовать поведение функции $\sin x$ на сегменте $\left[\frac{\pi}{2}, \pi+\frac{\pi}{2}\right]$. В пункте е) мы убедились в том, что $\sin \pi=0$, $\cos \pi=-1$, а из этих равенств и из первого соотношения 1) вытекает, что

$$\sin(x+\pi)=\sin x\cdot\cos\pi+\cos x\cdot\sin\pi=-\sin x.$$

Полученное соотношение позволяет заключить, что из установленного нами возрастания функции $\sin x$ на сегменте $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ вытекает убывание этой функции на сегменте $\left[\frac{\pi}{2}, \pi+\frac{\pi}{2}\right]$.

Изучение участков монотонности функций $\sin x$ и $\cos x$ полностью завершено.

В силу представлений $\operatorname{tg} x=\frac{\sin x}{\cos x}$, $\operatorname{ctg} x=\frac{\cos x}{\sin x}$ и в силу теоремы 4.1 для случая частного функция $\operatorname{tg} x$ непрерывна в любой точке $x\neq\frac{\pi}{2}+k\pi$, а функция $\operatorname{ctg} x$ непрерывна в любой точке $x\neq k\pi$. Пользуясь соотношениями $\sin(x+\pi)=-\sin x$, $\cos(x+\pi)=-\cos x$, мы получим, что $\operatorname{tg}(x+\pi)=\frac{\sin x}{\cos x}=\operatorname{tg} x$ и аналогично $\operatorname{ctg}(x+\pi)=\operatorname{ctg} x$. Это означает, что $\operatorname{tg} x$ и $\operatorname{ctg} x$ *являются периодическими функциями с периодом* π. Значит, достаточно произвести исследование участков монотонности этих функций только в пределах интервала длины π. Из равенства

$$\operatorname{tg} x_2-\operatorname{tg} x_1=\frac{\sin x_2}{\cos x_2}-\frac{\sin x_1}{\cos x_1}=\frac{\sin(x_2-x_1)}{\cos x_2\cdot\cos x_1} \tag{4.16}$$

и из того, что $\sin x$ принимает только положительные значения на интервале $(0, \pi)$, а $\cos x$ принимает только положительные значения на интервале $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, вытекает, что функция $\operatorname{tg} x$ возрастает на интервале $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. (Для любых x_1 и x_2 из интервала $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ таких, что $x_2 > x_1$, в правой части (4.16) будет стоять положительная величина.)

Аналогично устанавливается, что функция $\operatorname{ctg} x$ убывает на интервале $(0, \pi)$.

Мы не останавливаемся на изучении функций $\sec x = \frac{1}{\cos x}$ и $\operatorname{cosec} x = \frac{1}{\sin x}$.

Графики всех тригонометрических функций изображены на рис. 4.8—4.13.

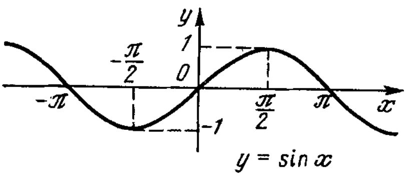


Рис. 4.8

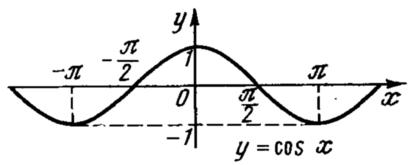


Рис. 4.9

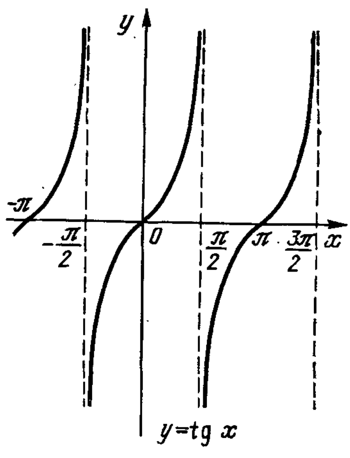


Рис. 4.10

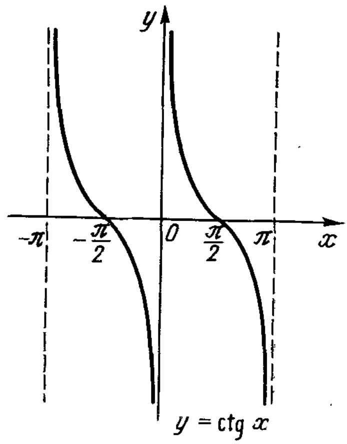


Рис. 4.11

5. Обратные тригонометрические функции. Остановимся на вопросах определения и свойствах непрерывности и монотонности обратных тригонометрических функций.

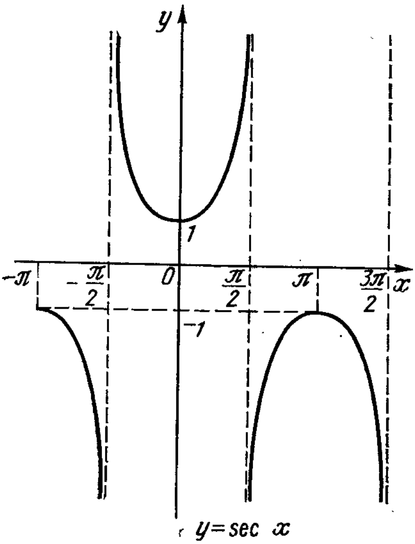


Рис. 4.12

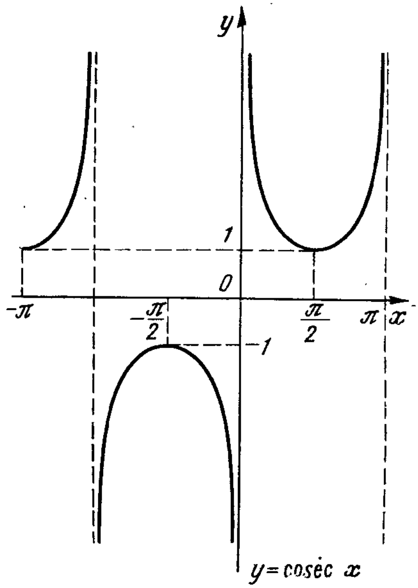


Рис. 4.13

Для определения функции $y=\arcsin x$ рассмотрим функцию $y=\sin x$ на сегменте $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Согласно п. 4 на этом сегменте функция $y=\sin x$ монотонно возрастает и непрерывна. Множество ее значений есть сегмент $[-1, 1]$. В силу теоремы 4.5 на сегменте $[-1, 1]$ существует непрерывная возрастающая обратная функция, принимающая значение $-\frac{\pi}{2}$ в точке -1 и значение $\frac{\pi}{2}$ в точке 1. Эту функцию обозначают символом $x=\arcsin y$ или, меняя обозначение аргумента y на x и обозначение x для функции на y, символом $y=\arcsin x$. Точно так же определяется на сегменте $[-1, 1]$ функция $y=\arccos x$, обратная по отношению к функции $x=\cos y$, убывающей и непрерывной на сегменте $0\leqslant y\leqslant\pi$.

Функция $y=\arccos x$ убывает и непрерывна на сегменте $[-1, +1]$ и принимает в точках $x=-1$ и $x=+1$ значения, соответственно равные π и 0.

Функции $y=\operatorname{arc\,tg} x$ и $y=\operatorname{arc\,ctg} x$ определяются как обратные для тангенса и котангенса, рассматриваемых на интервалах $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ и $(0, \pi)$. Эти функции определены и монотонны

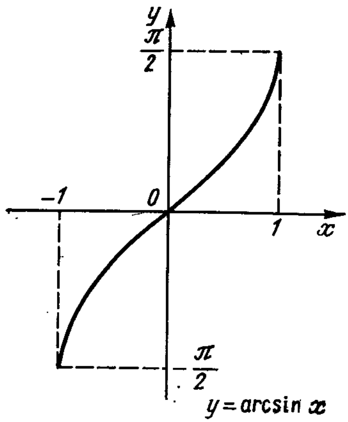


Рис. 4.14

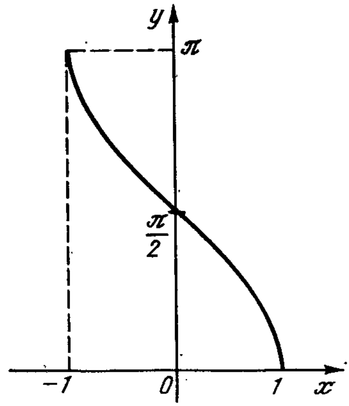


Рис. 4.15

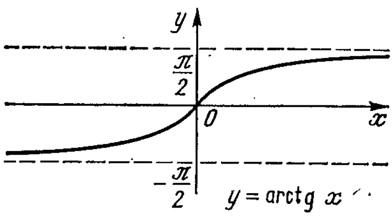


Рис. 4.16

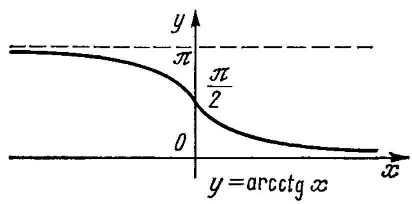


Рис. 4.17

на всей бесконечной прямой. На рис. 4.14—4.17 изображены графики обратных тригонометрических функций.

6. Гиперболические функции. Функции $\dfrac{e^x+e^{-x}}{2}$ и $\dfrac{e^x-e^{-x}}{2}$ называются соответственно *гиперболическим косинусом* и *гиперболическим синусом* и обозначаются символами $\operatorname{ch} x$ и $\operatorname{sh} x$:

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

Гиперболический тангенс и гиперболический котангенс определяются соответственно формулами

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

Из определения гиперболических функций следует, что гиперболический косинус, гиперболический синус и гиперболический тангенс заданы на всей числовой оси, а гиперболический котангенс определен всюду на числовой оси, за исключением точки $x=0$. На рис. 4.18—4.21 изображены графики этих функций.

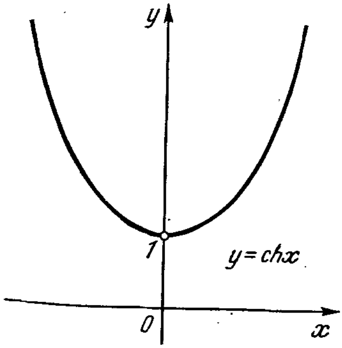


Рис. 4.18

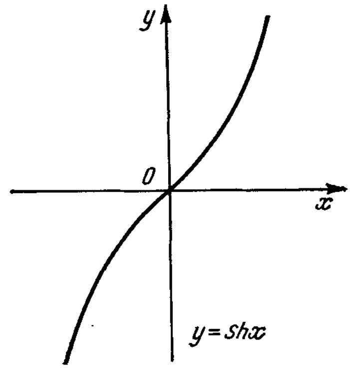


Рис. 4.19

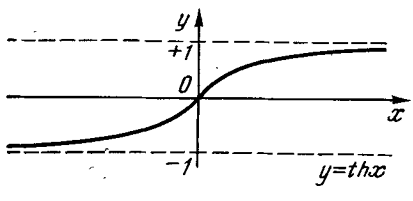


Рис. 4.20

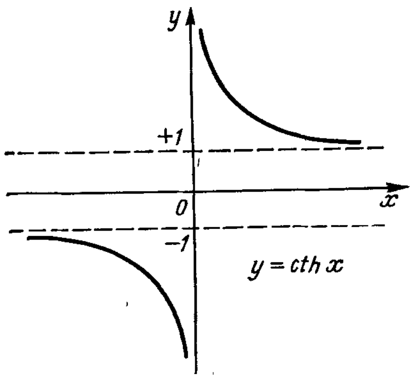


Рис. 4.21

Гиперболические функции непрерывны в каждой точке области их задания (это следует из непрерывности показательной функции и теоремы 4.1).

Гиперболические функции обладают рядом свойств, аналогичных свойствам тригонометрических функций. Например, для гиперболических функций имеют место теоремы сложения, аналогичные теоремам сложения для тригонометрических функций:

$$\operatorname{sh}(x+y)=\operatorname{sh}x\cdot\operatorname{ch}y+\operatorname{ch}x\cdot\operatorname{sh}y,$$
$$\operatorname{ch}(x+y)=\operatorname{ch}x\cdot\operatorname{ch}y+\operatorname{sh}x\cdot\operatorname{sh}y.$$

Непосредственно также проверяются формулы $\operatorname{sh}2x=2\operatorname{sh}x\cdot\operatorname{ch}x$,

$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1$. Эпитет же «гиперболический» связан с тем обстоятельством, что равенства $x = a \cdot \operatorname{ch} t$, $y = a \cdot \operatorname{sh} t$ задают гиперболу, подобно тому, как равенства $x = a \cdot \cos t$, $y = a \cdot \sin t$ задают окружность. Действительно, в первом случае мы, очевидно, имеем $x^2 - y^2 = a^2$, т. е. уравнение гиперболы, а во втором $x^2 + y^2 = a^2$ — уравнение окружности.

§ 4. ДВА ЗАМЕЧАТЕЛЬНЫХ ПРЕДЕЛА

1. Первый замечательный предел. Прежде всего докажем следующую теорему, представляющую собой функциональный аналог теоремы 3.14.

Теорема 4.6 (функциональный аналог принципа двустороннего ограничения). *Пусть в некоторой проколотой δ-окрестности точки a* заданы три функции $f(x)$, $h(x)$ и $g(x)$, две из которых $f(x)$ и $g(x)$ имеют в точке a общий предел, равный b. Тогда если всюду в указанной проколотой δ-окрестности точки a справедливы неравенства*

$$f(x) \leqslant h(x) \leqslant g(x), \tag{4.17}$$

то и функция $h(x)$ имеет в точке a предел, равный b.

Доказательство. Пусть $\{x_n\}$ — произвольная, сходящаяся к a последовательность значений аргумента, все элементы которой отличны от a. Тогда, с одной стороны, в силу определения предела по Гейне, обе последовательности соответствующих значений функций $\{f(x_n)\}$ и $\{g(x_n)\}$ сходятся к b, а с другой стороны, в силу (4.17), для всех номеров n справедливы неравенства

$$f(x_n) \leqslant h(x_n) \leqslant g(x_n).$$

В силу теоремы 3.14 мы можем утверждать, что последовательность $h(x_n)$ также сходится к b, а это и означает, что число b является пределом функции $h(x)$ в точке a.

Теорема доказана.

Теорема 4.7. *Предел функции* $h(x) = \dfrac{\sin x}{x}$ *в точке* $x=0$ *существует и равен единице,* т. е.

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

Доказательство. Будем отправляться от неравенств

$$0 < \sin x < x < \operatorname{tg} x \left(\text{при } 0 < x < \frac{\pi}{2}\right), \tag{4.18}$$

---

* Напомним, что проколотой δ-окрестностью точки a называется интервал $(a-\delta,\ a+\delta)$, из которого выкинута точка a.

указанных в п. 4 § 3. Посредством деления на $\sin x>0$ мы получим из (4.18) следующие неравенства:

$$1<\frac{x}{\sin x}<\frac{1}{\cos x}\left(\text{при } 0<x<\frac{\pi}{2}\right).$$

Для обратных величин, очевидно, справедливы обратные неравенства

$$\cos x<\frac{\sin x}{x}<1\ \left(\text{при } 0<x<\frac{\pi}{2}\right). \tag{4.19}$$

Заметим, что из того, что неравенства (4.19) справедливы при $0<x<\frac{\pi}{2}$, вытекает, что эти неравенства справедливы и при $-\frac{\pi}{2}<x<0$, ибо при замене x на $-x$ все три функции $\cos x$, $\frac{\sin x}{x}$ и 1 не меняют своих значений *.

Таким образом, неравенства (4.19) справедливы для всех значений x из интервала $-\frac{\pi}{2}<x<\frac{\pi}{2}$, за исключением точки $x=0$, т. е. справедливы всюду в проколотой $\frac{\pi}{2}$-окрестности точки $x=0$. Так как, кроме того, обе функции $f(x)=\cos x$ и $g(x)=1$ имеют в точке $x=0$ равный единице предел, то в силу теоремы 4.6 и функция $h(x)=\frac{\sin x}{x}$ имеет в точке $x=0$ предел, равный единице. Теорема доказана.

2. Второй замечательный предел.

Теорема 4.8. *Предел функции $f(x)=(1+x)^{1/x}$ в точке $x=0$ существует и равен числу e **.*

Доказательство. Достаточно доказать, что как правый, так и левый пределы функции $f(x)=(1+x)^{1/x}$ в точке $x=0$ существуют и оба равны e.

1) *Сначала докажем, что правый предел указанной функции в точке $x=0$ существует и равен e.*

В силу определения правого предела по Коши достаточно доказать, что для любого $\varepsilon>0$ найдется отвечающее ему $\delta>0$ такое,

* В самом деле в п. 4 § 3 мы установили, что $\cos(-x)=\cos x$, $\sin(-x)=-\sin x$ и, значит, $\frac{\sin(-x)}{(-x)}=\frac{\sin x}{x}$.

** Число e введено в п. 3 § 2 гл. 3 как предел последовательности $\left\{\left(1+\frac{1}{n}\right)^n\right\}$.

что для любого x из интервала $0<x<\delta$ справедливо неравенство

$$|(1+x)^{1/x}-e|<\varepsilon. \tag{4.20}$$

Фиксируем произвольное $\varepsilon>0$ и рассмотрим две последовательности $\{a_n\}$ и $\{b_n\}$ с элементами $a_n=\left(1+\frac{1}{n+1}\right)^n$, $b_n=\left(1+\frac{1}{n}\right)^{n+1}$. Убедимся в том, что обе эти последовательности сходятся к e. В самом деле, поскольку в силу п. 3 § 2 гл. 3 $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n=e$, то на основании теорем о пределе частного и произведения двух сходящихся последовательностей мы получим, что

$$\lim_{n\to\infty} a_n=\lim_{n\to\infty}\left[\frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{\left(1+\frac{1}{n+1}\right)}\right]=\frac{\lim\limits_{n\to\infty}\left(1+\frac{1}{n+1}\right)^{n+1}}{\lim\limits_{n\to\infty}\left(1+\frac{1}{n+1}\right)}=\frac{e}{1}=e,$$

$$\lim_{n\to\infty} b_n=\lim_{n\to\infty}\left[\left(1+\frac{1}{n}\right)^n\cdot\left(1+\frac{1}{n}\right)\right]=$$

$$=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n\cdot\lim_{n\to\infty}\left(1+\frac{1}{n}\right)=e\cdot 1=e.$$

Так как обе последовательности $\{a_n\}$ и $\{b_n\}$ сходятся к e, то для фиксированного выше $\varepsilon>0$ найдутся номера N_1 и N_2 такие, что $|a_n-e|<\varepsilon$ при $n\geqslant N_1$, $|b_n-e|<\varepsilon$ при $n\geqslant N_2$. Пусть N — это наибольший из двух номеров N_1 и N_2. Тогда, очевидно, при $n\geqslant N$ будут справедливы оба неравенства:

$$|a_n-e|<\varepsilon \quad \text{и} \quad |b_n-e|<\varepsilon. \tag{4.21}$$

Теперь для завершения доказательства существования равного e правого предела функции $f(x)=(1+x)^{1/x}$ в точке $x=0$ убедимся в том, что если взять $\delta=\frac{1}{N}$, то для любого x из интервала $0<x<\delta=\frac{1}{N}$ будет справедливо неравенство (4.20).

В самом деле, пусть x — любое число из интервала $0<x<\delta=\frac{1}{N}$. Тогда $\frac{1}{x}>N$. Обозначив через n целую часть числа $\frac{1}{x}$, т. е. положив $n=\left[\frac{1}{x}\right]$, мы, во-первых, с помощью неравенства $\frac{1}{x}>N$ можем утверждать, что $n\geqslant N$, а во-вторых, можем утверждать, что справедливы неравенства

$$n \leqslant \frac{1}{x} < n+1. \tag{4.22}$$

Из (4.22) вытекают неравенства

$$\frac{1}{n+1} < x \leqslant \frac{1}{n} \text{ и } 1+\frac{1}{n+1} < 1+x \leqslant 1+\frac{1}{n}. \tag{4.23}$$

Из сопоставления неравенства (4.22) со вторым неравенством (4.23) и из свойства возрастания показательной функции с основанием, бо́льшим единицы, вытекает, что

$$\left(1+\frac{1}{n+1}\right)^n < (1+x)^{1/x} < \left(1+\frac{1}{n}\right)^{n+1} \text{ или } a_n < (1+x)^{1/x} < b.$$

Итак, мы доказали, что для любого x из интервала $0 < x < \delta = \frac{1}{N}$ при некотором $n \geqslant N$, зависящем, конечно, от x, будут справедливы неравенства $a_n < (1+x)^{1/x} < b_n$, а значит, и неравенства

$$a_n - e < (1+x)^{1/x} - e < b_n - e. \tag{4.24}$$

Из сопоставления (4.24) с неравенствами (4.21), справедливыми для любого $n \geqslant N$, мы окончательно убедимся в том, что для любого x из интервала $0 < x < \delta = \frac{1}{N}$ будут справедливы неравенства (4.20).

2) *Докажем теперь, что и* л е в ы й *предел функции* $f(x) = (1+x)^{1/x}$ *в точке* $x=0$ *существует и равен* e.

В силу определения левого предела по Гейне достаточно доказать, что для любой бесконечно малой последовательности отрицательных чисел $\{x_n\}$ соответствующая последовательность значений функции $f(x_n) = (1+x_n)^{1/x_n}$ сходится к e.

Пусть $\{x_n\}$ — произвольная бесконечно малая последовательность отрицательных чисел. Эту последовательность мы будем рассматривать, начиная с того номера N, с которого все элементы x_n по модулю меньше единицы.

Положим $y_n = \frac{-x_n}{1+x_n}$, так что $x_n = \frac{-y_n}{1+y_n}$. Тогда, очевидно, $\{y_n\}$ будет являться бесконечно малой последовательностью, состоящей из п о л о ж и т е л ь н ы х чисел, причем

$$f(x_n) = (1+x_n)^{1/x_n} = \left(1 - \frac{y_n}{1+y_n}\right)^{-\frac{1+y_n}{y_n}} =$$

$$= \left(\frac{1}{1+y_n}\right)^{-\left(\frac{1}{y_n}+1\right)} = (1+y_n)^{\frac{1}{y_n}+1}.$$

Таким образом,

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} (1+y_n)^{\frac{1}{y_n}} \cdot \lim_{n\to\infty}(1+y_n) \tag{4.25}$$

при условии, что существуют пределы в правой части (4.25). Но поскольку $\{y_n\}$ сходится к нулю и состоит из положительных чисел, то $\lim\limits_{n\to\infty}(1+y_n)^{\frac{1}{y_n}} = e$ (в силу уже доказанного существования равного e правого предела), а $\lim\limits_{n\to\infty}(1+y_n)=1$. Тем самым доказано, что последовательность $\{f(x_n)\}$ сходится к числу e.

Теорема 4.8 полностью доказана.

Следствие. *Предел функции* $f(t)=\left(1+\dfrac{1}{t}\right)^t$ *при* $t\to\infty$ *существует и равен* e.

В силу определения предела при $t\to\infty$ по Гейне требуется доказать, что для любой бесконечно большой последовательности $\{t_n\}$ соответствующая последовательность значений функции $f(t_n)=\left(1+\dfrac{1}{t_n}\right)^{t_n}$ сходится к числу e. Мы будем рассматривать бесконечно большую последовательность $\{t_n\}$, начиная с того номера N, с которого все ее элементы t_n по модулю превосходят единицу. Положим $x_n=\dfrac{1}{t_n}$, так что $t_n=\dfrac{1}{x_n}$. В силу теоремы 3.6 из гл. 3 последовательность $\{x_n\}$ является бесконечно малой, причем $f(t_n)=\left(1+\dfrac{1}{t_n}\right)^{t_n}=(1+x_n)^{1/x_n}$.

Остается заметить, что в силу теоремы 4.8

$$\lim_{n\to\infty} f(t_n) = \lim_{n\to\infty}(1+x_n)^{1/x_n} = e.$$

§ 5. ТОЧКИ РАЗРЫВА ФУНКЦИИ И ИХ КЛАССИФИКАЦИЯ

1. Классификация точек разрыва функции. В § 1 мы договорились называть точками разрыва функции $f(x)$ те точки, в которых эта функция не обладает свойством непрерывности. При этом подразумевается, что функция $f(x)$ определена в той точке, которую мы апробируем на предмет наличия или отсутствия в ней свойства непрерывности.

Расширяя наше рассмотрение, мы можем подвергнуть изучению и те точки, в которых функция $f(x)$ не определена (при условии, конечно, что эти точки являются предельными для множества задания функции).

Выясним возможные типы точек разрыва.

1°. Устранимый разрыв. *Точка a называется точкой устранимого разрыва функции $y=f(x)$, если предел функции $f(x)$ в точке a существует, но в точке a функция $f(x)$ либо не определена, либо имеет частное значение $f(a)$, отличное от предела $f(x)$ в этой точке.*

Например, функция

$$f(x)=\begin{cases} \dfrac{\sin x}{x} & \text{при } x\neq 0, \\ 0{,}5 & \text{при } x=0 \end{cases}$$

имеет в точке $x=0$ устранимый разрыв.

Действительно, предельное значение этой функции в точке $x=0$, как мы доказали в п. 1 § 4, равно 1. Частное же значение $0{,}5\neq 1$.

Если функция $f(x)$ имеет в точке a устранимый разрыв, то этот разрыв можно устранить, не изменяя при этом значений функции в точках, отличных от a. Для этого достаточно положить значение функции в точке a равным ее предельному значению в этой точке. Так, в рассмотренном выше примере достаточно положить $f(0)=1$ и тогда $\lim\limits_{x\to 0}\dfrac{\sin x}{x}=f(0)=1$, т. е. функция $f(x)$ станет непрерывной в точке $x=0$.

В физических процессах точки устранимого разрыва встречаются при сосредоточенных распределениях физических величин.

2°. Разрыв первого рода. *Точка a называется точкой разрыва первого рода, если в этой точке функция $f(x)$ имеет конечные, но не равные друг другу правый и левый пределы*

$$\lim_{x\to a+0} f(x)\neq \lim_{x\to a-0} f(x).$$

Образно выражаясь, разрыв первого рода можно назвать конечным скачком. Приведем некоторые примеры.

1) Функция

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

в точке $x=0$ имеет разрыв первого рода*. Действительно,

---

* Функция $f(x)=\operatorname{sgn} x$, очевидно, может быть записана так:

$$f(x)=\operatorname{sgn} x=\begin{cases} \dfrac{x}{|x|} & \text{при } x\neq 0, \\ 0 & \text{при } x=0. \end{cases}$$

$\lim\limits_{x\to 0+0} \operatorname{sgn} x = 1$, $\lim\limits_{x\to 0-0} \operatorname{sgn} x = -1$, и, таким образом, эти пределы не равны между собой.

2) Другой пример дает функция $f(x) = \frac{\sin x}{|x|}$ при $x \neq 0$. Для этой функции $\lim\limits_{x\to 0+0} \frac{\sin x}{|x|} = 1$, $\lim\limits_{x\to 0-0} \frac{\sin x}{|x|} = -1$, так что точка $x=0$ является точкой разрыва первого рода.

3) Функция $f(x) = \frac{1}{1+2^{\frac{1}{x-1}}}$, определенная всюду, кроме точки $x=1$, имеет в точке $x=1$ разрыв первого рода.

В самом деле, если $\{x_n\}$ сходится к 1 и состоит из элементов $x_n>1$, то $\left\{\frac{1}{x_n-1}\right\}$ является бесконечно большой последовательностью с положительными членами. Поэтому $\left\{1+2^{\frac{1}{x_n-1}}\right\}$ — бесконечно большая последовательность и, значит, последовательность $f(x_n) = \frac{1}{1+2^{\frac{1}{x_n-1}}}$ является бесконечно малой, т. е. $\lim\limits_{x\to 1+0} f(x) = 0$.

Если же $\{x_n\}$ сходится к 1 и состоит из элементов $x_n<1$, то $\left\{\frac{1}{x_n-1}\right\}$ является бесконечно большой последовательностью с отрицательными членами. Поэтому $\left\{2^{\frac{1}{x_n-1}}\right\}$ сходится к нулю и, значит, последовательность $f(x_n) = \frac{1}{1+2^{\frac{1}{x_n-1}}}$ сходится к единице, т. е. $\lim\limits_{x\to 1-0} f(x) = 1$.

3°. **Разрыв второго рода.** *Точка a называется* ***точкой разрыва второго рода****, если в этой точке функция $f(x)$ не имеет по крайней мере одного из односторонних пределов или если хотя бы один из односторонних пределов бесконечен.*

Примеры. 1) Функция

$$f(x) = \begin{cases} x\cdot\cos\frac{1}{x} & \text{при } x<0, \\ 0 & \text{при } x=0, \\ \cos\frac{1}{x} & \text{при } x>0 \end{cases}$$

обладает левым пределом в точке $x=0$, равным нулю: $\lim\limits_{x\to 0-0} f(x) = 0$. В самом деле, если $\{x_n\}$ — последовательность, сходящаяся к нулю и состоящая из чисел $x_n<0$, то

$$0 \leqslant |f(x_n)| = |x_n| \cdot \left|\cos \frac{1}{x_n}\right| \leqslant |x_n|.$$

Поскольку $|x_n| \to 0$ при $n \to \infty$, то $\lim\limits_{n\to\infty} f(x_n)=0$.

Убедимся теперь в том, что рассматриваемая функция не имеет в точке $x=0$ правого предела. Для этого рассмотрим две сходящиеся к нулю и состоящие из положительных чисел последовательности $x_n = \dfrac{1}{\frac{\pi}{2}+\pi n}$ и $x'_n = \dfrac{1}{2\pi n}$. Если бы функция обладала в точке $x=0$ правым пределом, то обе последовательности $\{f(x_n)\}$ и $\{f(x'_n)\}$ сходились бы к одному и тому же числу. Однако $f(x_n') = \cos 2\pi n = 1$ сходится к единице, а $f(x_n) = \cos\left(\frac{\pi}{2} + \pi n\right) = 0$ сходится к нулю.

Итак, рассматриваемая функция имеет в точке $x=0$ разрыв второго рода.

2) Функция $f(x) = \operatorname{tg} x$, очевидно, имеет разрыв второго рода в каждой из точек $x_k = \frac{\pi}{2} + \pi k$, где $k=0, \pm 1, \pm 2, \ldots$, ибо в каждой такой точке x_k

$$\lim_{x\to x_k - 0} f(x) = +\infty, \quad \lim_{x\to x_k+0} f(x) = -\infty.$$

Можно образно сказать, что в каждой точке x_k функция $\operatorname{tg} x$ имеет бесконечный скачок.

3) Функция

$$f(x) = \begin{cases} \sin\dfrac{1}{x} & \text{при } x \neq 0, \\ 1 & \text{при } x = 0 \end{cases}$$

имеет разрыв второго рода в точке $x=0$, ибо в этой точке у нее не существует ни правого, ни левого пределов. В самом деле, поскольку $\sin\frac{1}{-x} = -\sin\frac{1}{x}$, достаточно удостовериться в отсутствии в точке $x=0$ только правого предела, а для этого достаточно заметить, что двум последовательностям значений аргумента $x_n = \frac{1}{\pi n}$ и $x'_n = \dfrac{1}{\frac{\pi}{2}+2\pi n}$ отвечают последовательности значений функции $f(x_n) = \sin \pi n = 0$ и $f(x'_n) = \sin\left(\frac{\pi}{2} + 2\pi n\right) = 1$, сходящиеся первая к нулю, а вторая — к единице.

Введем понятие кусочно непрерывной функции, часто встречающееся в математике и в ее приложениях.

Функция $f(x)$ называется ***кусочно непрерывной на сегменте*** *$[a, b]$, если эта функция определена всюду на сегменте $[a, b]$, непрерывна во всех внутренних точках этого сегмента, за исключением, быть может, конечного числа точек, в которых она имеет разрыв первого рода и, кроме того, имеет правый предел в точке a и левый предел в точке b.*

Функция $f(x)$ называется ***кусочно непрерывной на интервале (или на бесконечной прямой), если*** *$f(x)$ кусочно непрерывна на любом принадлежащем этому интервалу (или бесконечной прямой) сегменте.*

Например, функция $y=[x]$ кусочно непрерывна как на любом сегменте, так и на бесконечной прямой.

2. О точках разрыва монотонной функции. Следующее утверждение проливает свет на природу точек разрыва монотонной функции.

Теорема 4.9. *Если функция $f(x)$ определена на сегменте $[a, b]$ и является монотонной* на этом сегменте, то она может иметь на этом сегменте только точки разрыва первого рода, причем множество всех ее точек разрыва не более чем счетно.*

Доказательство. В силу леммы, доказанной в п. 1 § 2, монотонная функция имеет конечные правый и левый пределы в любой внутренней точке сегмента $[a, b]$ и, кроме того, конечный правый предел в точке a и конечный левый предел в точке b. Отсюда и вытекает, что точками разрыва монотонной функции могут быть только точки разрыва первого рода.

Чтобы доказать вторую часть теоремы о том, что множество всех точек разрыва не более чем счетно, будем ради определенности считать, что $f(x)$ является неубывающей на сегменте $[a, b]$. Достаточно доказать не более чем счетность множества точек разрыва, расположенных на интервале (a, b), т. е. точек разрыва, являющихся внутренними точками сегмента $[a, b]$. Заметим, что в каждой такой точке разрыва x справедливо неравенство для правого и левого пределов $f(x+0)>f(x-0)$ (см. замечание к указанной выше лемме). В силу леммы 2 гл. 2, каковы бы ни были два вещественных числа, не равных друг другу, всегда найдется рациональное число, заключенное между ними.

Так как в каждой точке разрыва x справедливо неравенство $f(x+0)>f(x-0)$, то каждой точке разрыва x можно поставить в соответствие некоторое рациональное число $r(x)$, удовлетворяющее неравенствам $f(x+0)>r(x)>f(x-0)$.

Заметим, что при этом разным точкам разрыва будут поставлены в соответствие разные рациональные числа. В самом деле, если x_1 и x_2 — две точки разрыва такие, что $x_1<x_2$, то из не-

* Т. е. либо неубывающей, либо невозрастающей.

убывания функции следует, что $f(x_1+0) \leqslant f(x_2-0)$, а поэтому $r(x_1) < r(x_2)$.

Таким образом, множество всех точек разрыва функции $f(x)$, расположенных внутри сегмента $[a, b]$ эквивалентно подмножеству множества рациональных чисел, которое, как доказано в § 7 гл. 2, является не более чем счетным.

Теорема доказана.

§ 6. ЛОКАЛЬНЫЕ И ГЛОБАЛЬНЫЕ СВОЙСТВА НЕПРЕРЫВНЫХ ФУНКЦИЙ

К локальным свойствам относятся те свойства функции, которые справедливы в сколь угодно малой окрестности фиксированной точки области определения функции. Эти свойства характеризуют поведение функции при стремлении аргумента к исследуемой точке. Например, непрерывность функции в некоторой точке области ее определения является локальным свойством этой функции.

Глобальные свойства — это свойства, связанные со всей областью определения функции. Например, монотонность функции на сегменте $[a, b]$ является ее глобальным свойством.

1. Локальные свойства непрерывных функций. Введем новые понятия. Предположим, что функция $f(x)$ задана на множестве $\{x\}$.

Определение 1. *Функция $f(x)$ называется ограниченной сверху (снизу) на множестве $\{x\}$, если существует такое вещественное число M (вещественное число m), что для всех значений аргумента x из множества $\{x\}$ справедливо неравенство $f(x) \leqslant M$ ($f(x) \geqslant m$). При этом число M (число m) называется верхней (нижней) гранью функции $f(x)$ на множестве $\{x\}$.*

Определение 2. *Функция $f(x)$ называется ограниченной с обеих сторон (или просто ограниченной) на множестве $\{x\}$, если она ограничена на этом множестве и сверху, и снизу, т. е. если найдутся такие вещественные числа m и M, что для всех значений аргумента x из множества $\{x\}$ справедливы неравенства $m \leqslant f(x) \leqslant M$.*

Таким образом, ограниченность функции $f(x)$ на множестве $\{x\}$ фактически означает ограниченность множества всех значений этой функции (отвечающих значениям аргумента из множества $\{x\}$).

Примеры. 1) Функция $f(x) = \operatorname{tg} x$, рассматриваемая на интервале $\left(0, \frac{\pi}{2}\right)$, ограничена на этом интервале снизу (в качестве нижней грани можно взять число $m \leqslant 0$), а сверху не ограничена.

2) Функция Дирихле $D(x)$, равная нулю в иррациональных точках и единице в рациональных точках, ограничена (с обеих сторон) на любом множестве $\{x\}$.

Справедлива следующая теорема о локальной ограниченности функции, имеющей конечное предельное значение.

Теорема 4.10 (о локальной ограниченности функции, имеющей конечный предел). *Пусть функция $f(x)$ задана на множестве $\{x\}$ и имеет конечное предельное значение в точке a. Тогда существует такое положительное число δ, что функция $f(x)$ ограничена на множестве B_δ, представляющем собой пересечение множества $\{x\}$ с интервалом $(a-\delta, a+\delta)$, т. е. с δ-окрестностью точки a.*

Замечание 1. Если множество задания функции $f(x)$ сплошь покрывает некоторую δ-окрестность точки a, то в качестве B_δ можно взять сам интервал $(a-\delta, a+\delta)$.

Доказательство. Пусть предел $f(x)$ в точке a равен b. В силу определения предела по Коши для некоторого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех значений аргумента x из проколотой δ-окрестности точки a* справедливо неравенство $|f(x)-b|<\varepsilon$ или

$$b-\varepsilon<f(x)<b+\varepsilon. \tag{4.26}$$

Если множество $\{x\}$ задания функции не содержит точку a, то теорема доказана, ибо в этом случае неравенства (4.26) означают, что для всех точек множества $B_\delta=\{x\}\cap(a-\delta, a+\delta)$ значения функции $f(x)$ заключены между числами $m=b-\varepsilon$ и $M=b+\varepsilon$.

Если же множество $\{x\}$ задания функции $f(x)$ содержит точку a и этой точке отвечает некоторое значение функции $f(a)$, то, обозначив через m наименьшее из двух чисел $b-\varepsilon$ и $f(a)$, а через M наибольшее из двух чисел $b+\varepsilon$ и $f(a)$, мы получим, что для всех точек множества $B_\delta=\{x\}\cap(a-\delta, a+\delta)$ будут справедливы неравенства**

$$m\leqslant f(x)\leqslant M, \tag{4.27}$$

которые и означают ограниченность $f(x)$ на множестве B_δ. Теорема доказана.

Иллюстрацией к доказанной теореме может служить рис. 4.22.

Следствие из теоремы 4.10. *Если функция $f(x)$ непрерывна в точке a, то эта функция ограничена на множестве всех значений ее аргумента, принадлежащих некоторой δ-окрестности точки a.*

---

* Напомним, что проколотой δ-окрестностью точки a называется интервал $(a-\delta, a+\delta)$ без точки a.

** В самом деле, для всех точек множества $B_\delta=\{x\}\cap(a-\delta, a+\delta)$, за исключением точки a, будут справедливы неравенства (4.26), а потому и неравенства (4.27). Для точки $x=a$ неравенства (4.27) справедливы вследствие выбора чисел m и M.

Достаточно заметить, что непрерывная в точке a функция имеет в этой точке конечное предельное значение.

Теорема 4.11 (об устойчивости знака непрерывной в точке функции). *Пусть функция $f(x)$ задана на множестве $\{x\}$, непрерывна в точке a этого множества и ее значение $f(a)$ положительно [отрицательно]. Тогда существует такое положительное число δ, что функция $f(x)$ является положительной [отрицательной] всюду на множестве B_δ, представляющем собой пересечение множества $\{x\}$ с δ-окрестностью точки a.*

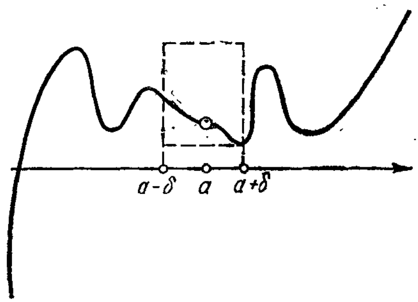


Рис. 4.22

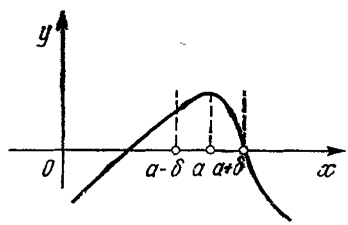


Рис. 4.23

Замечание 2. Если множество задания функции $f(x)$ сплошь покрывает некоторую δ-окрестность точки a, то в качестве B_δ можно взять саму δ-окрестность точки a.

Доказательство. В силу определения непрерывности по Коши для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех значений аргумента x из δ-окрестности точки a справедливо неравенство

$$|f(x)-f(a)|<\varepsilon,$$

или

$$f(a)-\varepsilon<f(x)<f(a)+\varepsilon. \tag{4.28}$$

Если взять в качестве ε положительное число $\frac{|f(a)|}{2}$, то оба числа $f(a)-\varepsilon$ и $f(a)+\varepsilon$ будут положительны при $f(a)>0$ и отрицательны при $f(a)<0$.

Поэтому неравенства (4.28) будут означать, что для всех значений аргумента из δ-окрестности точки a функция $f(x)$ является положительной при $f(a)>0$ и отрицательной при $f(a)<0$. Теорема доказана.

Иллюстрацией к теореме 4.11 может служить рис. 4.23.

Теорему 4.11 легко переформулировать на случай, когда функция $f(x)$ непрерывна в точке a только справа или только слева.

Договоримся называть *полусегмент* $[a, a+\delta)$ ***правой*** *δ-полуокрестностью точки a, а полусегмент* $(a-\delta, a]$ — ***левой*** *δ-полуокрестностью точки a.*

Теорема 4.11*. *Пусть функция $f(x)$ задана на множестве $\{x\}$, непрерывна в точке a этого множества справа [слева] и ее значение $f(a)$ не равно нулю. Тогда найдется такое положительное число δ, что функция $f(x)$ не обращается в нуль и имеет тот же знак, что и в точке a, для всех значений x из множества $\{x\}$, принадлежащих правой [левой] δ-полуокрестности точки a.*

Для доказательства этой теоремы следует дословно повторить доказательство теоремы 4.11 с заменой термина «δ-окрестность точки a» термином «правая [левая] δ-полуокрестность точки a».

Замечание 3. К числу локальных свойств непрерывных в данной точке функций следует отнести доказанные выше теоремы 4.1 и 4.2 о непрерывности суммы, разности, произведения и частного двух непрерывных в данной точке функций и о непрерывности сложной функции.

2. Глобальные свойства непрерывных функций.

Теорема 4.12 (прохождение непрерывной функции через нуль при смене знаков). *Пусть функция $f(x)$ непрерывна на сегменте $[a, b]$, и пусть значения этой функции на концах сегмента $f(a)$ и $f(b)$ суть числа разных знаков. Тогда внутри сегмента $[a, b]$ найдется такая точка ξ, значение функции в которой равно нулю.*

Доказательство. Не ограничивая общности, можно считать, что $f(a)<0$, $f(b)>0$. Пусть $\{x\}$ — множество всех значений x из сегмента $[a, b]$, для которых $f(x)<0$. Это множество непусто (ему, например, принадлежит точка $x=a$) и ограничено сверху (например, числом b).

Согласно теореме 2.1 у множества $\{x\}$ существует точная верхняя грань, которую мы обозначим через ξ.

Заметим, что точка ξ — внутренняя точка сегмента $[a, b]$, так как из непрерывности функции $f(x)$ на $[a, b]$ и из условий $f(a)<0$, $f(b)>0$, в силу теоремы 4.11*, вытекает, что найдется правая δ-полуокрестность точки a, в пределах которой $f(x)<0$, и левая δ-полуокрестность точки b, в пределах которой $f(x)>0$.

Убедимся в том, что $f(\xi)=0$. Если бы это было не так, то по теореме 4.11 нашлась бы δ-окрестность $\xi-\delta<x<\xi+\delta$ точки ξ, в пределах которой функция $f(x)$ имела бы определенный знак. Но это невозможно, так как по определению точной верхней грани найдется хотя бы одно значение x из полусегмента $\xi-\delta<x\leqslant\xi$ такое, что $f(x)<0$, а для любого значения x из интервала $\xi<x<\xi+\delta$ справедливо неравенство $f(x)\geqslant 0$. Полученное противоречие доказывает, что $f(\xi)=0$. Теорема доказана. Иллюстрацией к теореме 4.12 может служить рис. 4.24.

Теорема 4.13 (прохождение непрерывной функции через любое промежуточное значение).

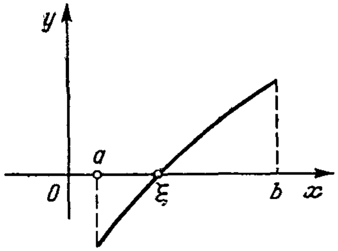


Рис. 4.24

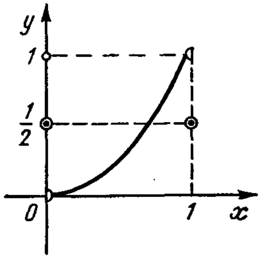


Рис. 4.25

Пусть функция $f(x)$ непрерывна на сегменте $[a, b]$, причем $f(a)=\alpha$, $f(b)=\beta$. Пусть, далее, γ — любое число, заключенное между α и β. Тогда на сегменте $[a, b]$ найдется точка ξ такая, что $f(\xi)=\gamma$.

Доказательство. В доказательстве нуждается, очевидно, лишь случай $\alpha\neq\beta$ (в противном случае $\gamma=\alpha=\beta$ и можно взять $\xi=a$). По этой же причине отпадает случай, когда γ совпадает с одним из чисел α или β.

Не ограничивая общности, будем считать, что $\alpha<\beta$, $\alpha<\gamma<\beta$. Рассмотрим функцию $\varphi(x)=f(x)-\gamma$. Эта функция непрерывна на сегменте $[a, b]$ (как разность непрерывных функций) и принимает на концах этого сегмента значения разных знаков:

$$\varphi(a)=f(a)-\gamma=\alpha-\gamma<0, \qquad \varphi(b)=f(b)-\gamma=\beta-\gamma>0.$$

По теореме 4.12 внутри сегмента $[a, b]$ найдется точка ξ такая, что $\varphi(\xi)=f(\xi)-\gamma=0$, т. е. $f(\xi)=\gamma$. Теорема доказана.

Используя только что доказанную теорему, мы убедимся в справедливости замечания 2, высказанного в п. 2 § 2.

Пусть функция $f(x)$ непрерывна на сегменте $[a, b]$ и существует функция, обратная для $f(x)$, рассматриваемой на этом сегменте. Тогда $f(x)$ строго монотонна на указанном сегменте $[a, b]$.

Доказательство. Из существования обратной функции для $f(x)$ следует, что $f(a)\neq f(b)$. Пусть $f(a)<f(b)$ $[f(a)>f(b)]$. Покажем, что $f(x)$ строго монотонно возрастает [убывает] на сегменте $[a, b]$. Рассмотрим случай $f(a)<f(b)$. (Если $f(a)>f(b)$, то рассуждения аналогичны.) Предварительно установим справедливость неравенства $f(x)<f(b)$ для всех x из (a, b). Действительно, пусть существует такое $x_1\in(a, b)$, что $f(x_1)>f(b)$. (Равенство $f(x_1)=f(b)$ невозможно ввиду существования обратной функции для функции $f(x)$.) Применяя теорему 4.13 для сегментов $[a, x_1]$ и $[x_1, b]$ и используя вытекающие из $f(a)<f(b)<f(x_1)$ неравенства

$$f(a)<\frac{f(x_1)+f(b)}{2}<f(x_1),$$

$$f(x_1) > \frac{f(x_1)+f(b)}{2} > f(b),$$

убедимся в существовании двух таких чисел $\xi_1 \in (a, x_1)$ и $\xi_2 \in (x_1, b)$, что $f(\xi_1) = f(\xi_2) = \frac{f(x_1)+f(b)}{2}$. Итак, $\xi_1 \neq \xi_2$, но $f(\xi_1) = f(\xi_2)$, что противоречит существованию обратной функции для функции $f(x)$ на сегменте $[a, b]$.

Установим теперь строго монотонное возрастание $f(x)$ на сегменте $[a, b]$. Пусть существуют два числа $x_1 < x_2$, принадлежащие полусегменту $[a, b)$, такие, что $f(x_1) > f(x_2)$. Покажем, что это предположение приводит нас к противоречию. Применяя теорему 4.13 для сегментов $[x_1, x_2]$ и $[x_2, b]$ и используя вытекающие из $f(x_1) > f(x_2)$, $f(x_1) < f(b)$ неравенства

$$f(x_1) > \frac{f(x_1)+f(x_2)}{2} > f(x_2),$$

$$f(x_2) < \frac{f(x_1)+f(x_2)}{2} < f(b),$$

убедимся в существовании двух таких чисел $\xi_3 \in (x_1, x_2)$ и $\xi_4 \in (x_2, b)$, что $f(\xi_3) = f(\xi_4) = \frac{f(x_1)+f(x_2)}{2}$. Итак, $\xi_3 \neq \xi_4$, но $f(\xi_3) = f(\xi_4)$, что также противоречит существованию обратной функции для функции $f(x)$ на сегменте $[a, b]$. Замечая, что условие $f(x_1) = f(x_2)$ для $x_1 < x_2$ также невозможно, мы приходим к выводу, что $f(x_1) < f(x_2)$ для любых $x_1 < x_2$ из сегмента $[a, b]$. Замечание доказано.

Теорема 4.14 (первая теорема Вейерштрасса). *Если функция $f(x)$ непрерывна на сегменте $[a, b]$, то она ограничена на этом сегменте.*

Доказательство. Докажем, что функция $f(x)$ ограничена на сегменте $[a, b]$ сверху (ограниченность снизу доказывается аналогично).

Доказательство проведем от противного, т. е. предположим, что $f(x)$ не является ограниченной сверху на сегменте $[a, b]$. Тогда для любого натурального n ($n=1, 2, \ldots$) найдется хотя бы одна точка x_n из $[a, b]$ такая, что $f(x_n) > n$. (В противном случае $f(x)$ была бы ограничена сверху на сегменте $[a, b]$.)

Таким образом, мы указали последовательность значений $\{x_n\}$ из сегмента $[a, b]$ такую, что соответствующая последовательность значений функции $\{f(x_n)\}$ является бесконечно большой. В силу теоремы Больцано—Вейерштрасса (см. следствие 3 из теоремы 3.16 п. 1 § 3 гл. 3) из последовательности $\{x_n\}$ можно выделить подпоследовательность $\{x_{k_n}\}$, $n=1, 2, \ldots$, сходящуюся к некоторой точке ξ. Так как все элементы подпоследовательности $\{x_{k_n}\}$ лежат на сегменте $[a, b]$, то и точка ξ принадлежит сегменту $[a, b]$ (в силу следствия 2 из теоремы 3.13). В силу непрерывности функции $f(x)$ в точке ξ соответствующая подпоследовательность значений функции $\{f(x_{k_n})\}$ обязана сходиться к $f(\xi)$. Но это противоречит тому, что подпоследовательность $\{f(x_{k_n})\}$, будучи выделена из бесконечно большой последовательности $\{f(x_n)\}$, сама является бесконечно большой. Полученное противоречие доказывает теорему.

Замечание 1. Для интервала (или полусегмента) утверждение, высказанное в теореме 4.14, уже несправедливо, т. е. из непрерывности функции на интервале (или полусегменте) не вытекает ее ограниченность на этом множестве. Рассмотрим, например, функцию $f(x)=\frac{1}{x}$ на интервале $(0, 1)$ (или на полусегменте $(0, 1]$). Эта функция непрерывна на указанном множестве, но неограниченна на нем. В самом деле, последовательность $x_n=\frac{1}{n}$ $(n=2, 3, \ldots)$ принадлежит указанному множеству, а последовательность значений функции $\{f(x_n)\}=\{n\}$ является бесконечно большой.

Рассмотрим функцию $f(x)$, ограниченную на данном множестве $\{x\}$ сверху (снизу).

Определение. *Число M (число m) называется точной верхней (точной нижней) гранью функции $f(x)$ на множестве $\{x\}$, если выполнены два требования:* 1) *для каждого значения x из множества $\{x\}$ справедливо неравенство $f(x)\leqslant M$ $(f(x)\geqslant m)$;* 2) *для любого числа $\varepsilon>0$ существует такое значение x из множества $\{x\}$, что для соответствующего значения функции $f(x)$ справедливо неравенство*

$$f(x)>M-\varepsilon \qquad (f(x)<m+\varepsilon).$$

Заметим, что в данном определении условие 1) означает, что число M (число m) является одной из верхних (нижних) граней функции $f(x)$ на множестве $\{x\}$, а условие 2) означает, что эта грань является наименьшей (наибольшей) и уменьшена (увеличена) быть не может.

Точная верхняя грань M функции $f(x)$ на множестве $\{x\}$ обычно обозначается символом

$$M=\sup_{\{x\}}\{f(x)\}=\sup_{\{x\}} f(x).$$

Аналогично точная нижняя грань m функции $f(x)$ на множестве $\{x\}$ обозначается символом

$$m=\inf_{\{x\}}\{f(x)\}=\inf_{\{x\}} f(x).$$

В частности, точная верхняя грань функции $f(x)$ на сегменте $[a, b]$ может обозначаться любым из следующих четырех символов:

$$\sup_{a\leqslant x\leqslant b} f(x)=\sup_{a\leqslant x\leqslant b}\{f(x)\}=\sup_{x\in[a,b]} f(x)=\sup_{x\in[a,b]}\{f(x)\}.$$

Аналогичные четыре символа для точной нижней грани имеют вид

$$\inf_{a\leqslant x\leqslant b} f(x)=\inf_{a\leqslant x\leqslant b}\{f(x)\}=\inf_{x\in[a,b]} f(x)=\inf_{x\in[a,b]}\{f(x)\}.$$

Справедливы следующие утверждения:

1) *если функция $f(x)$ ограничена на множестве $\{x\}$ сверху [снизу], то у нее существует на этом множестве точная верхняя грань [точная нижняя грань]*;

2) *если функция $f(x)$ ограничена на множестве $\{x\}$ (с обеих сторон), то у нее существуют на этом множестве как точная верхняя, так и точная нижняя грани.*

Эти утверждения являются прямым следствием теоремы 2.1 гл. 2, ибо ограниченность функции $f(x)$ на множестве $\{x\}$ сверху [снизу] означает, что множество всех значений этой функции ограничено сверху [снизу].

Возникает вопрос о том, является ли точная верхняя [точная нижняя] грань ограниченной на множестве $\{x\}$ функции $f(x)$ достижимой, т. е. существует ли среди точек множества $\{x\}$ такая точка x_0, значение функции в которой $f(x_0)$ равно точной верхней [соответственно точной нижней] грани $f(x)$ на множестве $\{x\}$.

Следующий пример показывает, что точные грани ограниченной на данном множестве функции, вообще говоря, не являются достижимыми.

Рассмотрим на сегменте $[0, 1]$ функцию $f(x)$ следующего вида (рис. 4.25):

$$f(x)=\begin{cases} x^2 & \text{при } 0<x<1,\\ 1/2 & \text{при } x=0 \text{ и } x=1.\end{cases}$$

Эта функция ограничена на сегменте $[0, 1]$ и имеет на нем точную верхнюю грань $M=1$ и точную нижнюю грань $m=0$. Однако эти грани недостижимы: среди точек сегмента $[0, 1]$ не существует точек, значения функции в которых были бы равны нулю или единице.

Заметим, что рассматриваемая функция $f(x)$ не является непрерывной на сегменте $[0, 1]$ (она имеет разрывы в точках $x=0$ и $x=1$). Оказывается, это обстоятельство не является случайным, ибо справедливо следующее утверждение.

Теорема 4.15 (вторая теорема Вейерштрасса). *Если функция $f(x)$ непрерывна на сегменте $[a, b]$, то она достигает на этом сегменте своих точных верхней и нижней граней, т. е. среди точек сегмента $[a, b]$ найдутся такие точки x_1 и x_2, что значение $f(x_1)$ равно точной верхней грани $f(x)$ на сегменте*

$[a, b]$, а значение $f(x_2)$ равно точной нижней грани $f(x)$ на сегменте $[a, b]$.

Доказательство. В силу первой теоремы Вейерштрасса 4.14 функция $f(x)$ ограничена на сегменте $[a, b]$, а поэтому у нее существует на этом сегменте точная верхняя грань M и точная нижняя грань m.

Остановимся на доказательстве достижимости точной верхней грани M, ибо достижимость точной нижней грани m доказывается аналогично.

Предположим, что точная верхняя грань M не является достижимой, т. е. предположим, что во всех точках сегмента $[a, b]$ функция $f(x)$ принимает значения, строго меньшие M. Тогда мы можем рассмотреть функцию

$$F(x)=\frac{1}{M-f(x)}.$$

Знаменатель $M-f(x)$ представляет собой функцию, непрерывную и строго положительную на сегменте $[a, b]$. Поэтому по теореме 4.1 (для случая частного) функция $F(x)$ будет являться непрерывной на сегменте $[a, b]$. Значит, по первой теореме Вейерштрасса 4.14 функция $F(x)$ ограничена на сегменте $[a, b]$, т. е. найдется положительное число A такое, что $F(x)=\frac{1}{M-f(x)}\leqslant$ $\leqslant A$ для всех x из сегмента $[a, b]$. Так как функция $M-f(x)$ строго положительна на $[a, b]$, то последнее неравенство эквивалентно неравенству $f(x)\leqslant M-\frac{1}{A}$ для всех x из сегмента $[a, b]$, а это противоречит тому, что число M является точной верхней гранью, т. е. наименьшей из всех верхних граней функции $f(x)$ на сегменте $[a, b]$.

Полученное противоречие доказывает, что наше предположение о недостижимости точной верхней грани является неверным. Теорема доказана.

Замечание 2. После того как доказана достижимость непрерывной на сегменте $[a, b]$ функцией своих точных верхней и нижней граней, мы можем назвать точную верхнюю грань M максимальным значением, а точную нижнюю грань m минимальным значением функции $f(x)$ на сегменте $[a, b]$. Теорему 4.15 можно переформулировать в виде: *непрерывная на сегменте $[a, b]$ функция $f(x)$ имеет на этом сегменте максимальное и минимальное значения.*

Максимальное значение функции $f(x)$ на сегменте $[a, b]$ обозначается одним из следующих символов:

$$\max_{a\leqslant x\leqslant b} f(x)=\max_{a\leqslant x\leqslant b}\{f(x)\}=\max_{x\in[a,b]} f(x)=\max_{x\in[a,b]}\{f(x)\}.$$

Аналогичные символы для минимального значения $f(x)$ на сегменте $[a, b]$ имеют вид

$$\min_{a\leqslant x\leqslant b} f(x) = \min_{a\leqslant x\leqslant b} \{f(x)\} = \min_{x\in[a,b]} f(x) = \min_{x\in[a,b]} \{f(x)\}.$$

Замечание 3. Заметим, что и функции, не являющиеся непрерывными на данном сегменте, могут достигать на этом сегменте своих точной верхней и точной нижней граней. Примером может служить функция Дирихле $D(x)$, равная нулю для всех иррациональных x и равная единице для всех рациональных x. Эта функция разрывна в каждой точке сегмента $[0, 1]$, но, очевидно, достигает на этом сегменте своей точной верхней грани, равной единице, и своей точной нижней грани, равной нулю.

Замечание 4. Утверждение теоремы 4.15 окажется неверным, если в ее формулировке термин «сегмент» заменить термином «интервал» или «полусегмент».

Так, функция $f(x)=x$ является непрерывной на интервале $(0, 1)$ или на полусегменте $[0, 1)$, однако точная верхняя грань этой функции на указанном интервале или полусегменте $M=1$ хотя и существует, но не достигается.

К этому следует добавить, что у функции, являющейся непрерывной на интервале или полусегменте, точные грани могут даже не существовать, ибо такая функция может не являться ограниченной на указанном интервале или полусегменте (см. замечание 1).

3. Понятие равномерной непрерывности функции. Предположим, что функция $y=f(x)$ задана на таком множестве $\{x\}$, каждая точка которого является предельной точкой этого множества. Примером такого множества могут служить сегмент, интервал, полусегмент, полупрямая, бесконечная прямая, множество всех рациональных точек, принадлежащих любому из перечисленных множеств.

Определение. *Функция $y=f(x)$ называется равномерно непрерывной на множестве $\{x\}$, если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для любых двух точек x' и x'' множества $\{x\}$, удовлетворяющих условию $|x'-x''|<\delta$, справедливо неравенство*

$$|f(x')-f(x'')|<\varepsilon. \tag{4.29}$$

Замечание 1. Сразу же подчеркнем, что если функция $f(x)$ равномерно непрерывна на множестве $\{x\}$, то она непрерывна в каждой точке x множества $\{x\}$. В самом деле, взяв в сформулированном определении в качестве x'' данную фиксированную точку x_0 множества $\{x\}$, а в качестве x' — любую точку этого множества, мы придем к определению непрерывности функции $f(x)$ в точке x_0 по Коши.

Замечание 2. Основным в сформулированном определении равномерной непрерывности является требование, гарантирующее существование по любому $\varepsilon>0$ такого универсального $\delta>0$, которое обеспечивает справедливость неравенства (4.29) сразу для всех точек x' и x'' множества $\{x\}$, удовлетворяющих условию $|x'-x''|<\delta$.

Если потребовать непрерывности функции $f(x)$ в каждой точке x_0 множества $\{x\}$, то для любого $\varepsilon>0$ и любой точки x_0 множества $\{x\}$ можно гарантировать существование «своего» положительного числа $\delta=\delta(\varepsilon, x_0)$, зависящего не только от ε, но и от x_0 и обеспечивающего справедливость неравенства $|f(x)-f(x_0)|<\varepsilon$ для всех x из множества $\{x\}$, удовлетворяющих условию $|x-x_0|<\delta(\varepsilon, x_0)$. При этом, вообще говоря, может не существовать положительной точной нижней грани указанных $\delta(\varepsilon, x_0)$ по всем точкам x_0 множества $\{x\}$, т. е. равномерная непрерывность функции на множестве $\{x\}$ не вытекает, вообще говоря, из непрерывности этой функции в каждой точке x_0 множества $\{x\}$.

Замечание 3. Из данного нами определения равномерной непрерывности непосредственно вытекает, что если функция $f(x)$ равномерно непрерывна на множестве $\{x\}$, то эта функция равномерно непрерывна и на любом подмножестве множества $\{x\}$.

Рассмотрим примеры функций, как обладающих, так и не обладающих на данном множестве $\{x\}$ свойством равномерной непрерывности.

1. Убедимся в том, что функция $f(x)=\dfrac{1}{x}$ равномерно непрерывна на полупрямой $x\geqslant 1$. В самом деле, для любых двух точек x' и x'' из указанной полупрямой справедливо неравенство

$$|f(x')-f(x'')|=\left|\frac{1}{x'}-\frac{1}{x''}\right|=\left|\frac{x''-x'}{x'\cdot x''}\right|=\frac{|x''-x'|}{x'\cdot x''}\leqslant|x''-x'|.$$

Поэтому, взяв для любого $\varepsilon>0$ положительное число δ равным ε, мы получим, что для любых двух точек x' и x'' полупрямой $[1, +\infty)$, удовлетворяющих условию $|x''-x'|<\delta$, справедливо неравенство $|f(x')-f(x'')|<\delta=\varepsilon$.

2. Функция $f(x)=\sin\dfrac{1}{x}$ не является равномерно непрерывной на интервале $(0, 1)$*. Чтобы убедиться в этом, достаточно доказать, что для некоторого $\varepsilon>0$ и для любого как угодно малого $\delta>0$ найдется хотя бы одна пара точек x' и x'' интервала $(0, 1)$ таких, что

$$|x'-x''|<\delta, \quad \text{но} \quad |f(x')-f(x'')|\geqslant\varepsilon.$$

* Хотя эта функция и является непрерывной в каждой точке интервала $(0, 1)$.

Рассмотрим две последовательности точек, принадлежащих интервалу $(0, 1)$, $\{x'_n\}$ и $\{x''_n\}$ с элементами

$$x'_n = \frac{1}{\pi \cdot n} \text{ и } x''_n = \frac{1}{\frac{\pi}{2} + 2\pi n}, \quad n = 1, 2, \ldots .$$

Обе эти последовательности, а значит, и их разность являются бесконечно малыми. Поэтому для любого как угодно малого $\delta > 0$ найдется номер n такой, что $|x'_n - x''_n| < \delta$. Вместе с тем для любого номера n

$$|f(x'_n) - f(x''_n)| = \left|\sin \pi n - \sin\left(\frac{\pi}{2} + 2\pi n\right)\right| = 1.$$

Поэтому для $\varepsilon = \frac{1}{2} > 0$ и для как угодно малого $\delta > 0$ найдется пара точек x' и x'' из интервала $(0, 1)$ таких, что $|x_n' - x_n''| < \delta$, в то время как $|f(x_n') - f(x_n'')| > \varepsilon$, это и означает, что рассматриваемая функция не является равномерно непрерывной на интервале $(0, 1)$.

Заметим, что если бы мы рассмотрели ту же самую функцию $f(x) = \sin\frac{1}{x}$ не на интервале $(0, 1)$, а на интервале $(\gamma, 1)$, где γ — любое число из интервала $0 < \gamma < 1$, то приведенные выше рассуждения уже не имели бы места. Этот факт не является случайным, ибо ниже мы покажем, что указанная функция является равномерно непрерывной на интервале $(\gamma, 1)$ при $0 < \gamma < 1$.

3. Докажем, что функция $f(x) = x^2$ не является равномерно непрерывной на полупрямой $x \geqslant 1$. Заметим, что для любых двух точек x' и x'' полупрямой $x \geqslant 1$ справедливо неравенство

$$\begin{gathered} |f(x') - f(x'')| = |(x')^2 - (x'')^2| = \\ = |x' + x''| \cdot |x' - x''| > x' \cdot |x' - x''|. \end{gathered} \tag{4.30}$$

Убедимся теперь в том, что не только для **некоторого** $\varepsilon > 0$, а даже для **любого** $\varepsilon > 0$ и для любого как угодно малого $\delta > 0$ найдется пара точек x' и x'' из полупрямой $x \geqslant 1$ таких, что

$$|x' - x''| < \delta, \quad \text{но} \quad |f(x') - f(x'')| \geqslant \varepsilon.$$

(Это и будет означать отсутствие свойства равномерной непрерывности у функции $f(x) = x^2$ на рассматриваемой полупрямой.)

Фиксировав произвольные $\varepsilon > 0$ и $\delta > 0$, возьмем в качестве x' произвольное число, превосходящее единицу и такое, что $x' > \frac{2\varepsilon}{\delta}$, и положим $x'' = x' + \frac{\delta}{2}$. Для таких x' и x'' будет справедливо неравенство $|x'-x''|=\frac{\delta}{2}<\delta$. С другой стороны, в силу (4.30) для этих же x' и x'' будет справедливо неравенство

$$|f(x')-f(x'')|\geqslant\frac{\delta}{2}\cdot\frac{2\varepsilon}{\delta}=\varepsilon.$$

Заметим, что если бы мы рассмотрели ту же самую функцию $f(x)=x^2$ не на полупрямой $x\geqslant 1$, а на любом сегменте $[1, b]$, где b — любое число, то проведенные нами рассуждения уже не имели бы места.

Этот факт становится понятным в силу следующей фундаментальной теоремы.

Основная теорема 4.16. *Если функция $f(x)$ непрерывна на сегменте $[a, b]$, то она и равномерно непрерывна на этом сегменте.*

Доказательство. Предположим, что функция $f(x)$ непрерывна на сегменте $[a, b]$, но не является равномерно непрерывной на этом сегменте.

Тогда для некоторого $\varepsilon>0$ и для любого как угодно малого $\delta>0$ найдутся две точки x' и x'' сегмента $[a, b]$ такие, что

$$|x'-x''|<\delta, \quad \text{но} \quad |f(x')-f(x'')|\geqslant\varepsilon.$$

Выберем бесконечно малую последовательность положительных чисел $\delta_n=\frac{1}{n}$ $(n=1, 2, \ldots)$. Можно утверждать, что для указанного $\varepsilon>0$ и для любого номера n найдутся две точки x_n' и x_n'' сегмента $[a, b]$ такие, что

$$|x_n'-x_n''|<\frac{1}{n}, \quad \text{но} \quad |f(x_n')-f(x_n'')|\geqslant\varepsilon. \tag{4.31}$$

Так как последовательность $\{x_n'\}$ состоит из точек сегмента $[a, b]$, то она ограничена и по теореме Больцано—Вейерштрасса (см. следствие 3 из теоремы 3.16 гл. 3) из нее можно выделить сходящуюся подпоследовательность $\{x'_{k_n}\}$, $n=1, 2, \ldots$. Предел ξ указанной подпоследовательности (в силу следствия 2 из теоремы 3.13 гл. 3) будет также принадлежать сегменту $[a, b]$. В силу левого неравенства (4.31) соответствующая подпоследовательность $\{x''_{k_n}\}$ будет сходиться к той же самой точке ξ.

Поскольку функция $f(x)$ непрерывна в каждой точке сегмента $[a, b]$, она непрерывна и в точке ξ*. Но тогда, в силу определения непрерывности по Гейне, обе подпоследовательности соответствующих значений функции $\{f(x'_{k_n})\}$ и $\{f(x''_{k_n})\}$ обязаны сходиться к $f(\xi)$, т. е. разность указанных подпоследовательностей

* В случае, если ξ совпадает с одним из концов сегмента $[a, b]$, под непрерывностью следует понимать одностороннюю непрерывность.

$\{f(x'_{k_n})-f(x''_{k_n})\}$ обязана быть бесконечно малой. Это противоречит правому неравенству (4.31), справедливому для всех номеров n и потому для всех номеров k_n.

Полученное противоречие доказывает, что наше предположение о том, что непрерывная на сегменте $[a, b]$ функция не является равномерно непрерывной на этом сегменте, является неверным. Теорема доказана.

Возвратимся теперь к рассмотренному выше примеру 2 и покажем, что *функция* $f(x)=\sin\frac{1}{x}$ *является равномерно непрерывной на интервале* $(\gamma, 1)$ *при любом* γ *из интервала* $0<\gamma<1$. В самом деле, при любом таком γ функция $f(x)=\sin\frac{1}{x}$ непрерывна на сегменте $[\gamma, 1]$. Значит, по теореме 4.16 функция $f(x)=\sin\frac{1}{x}$ равномерно непрерывн на сегменте $[\gamma, 1]$. В силу замечания 3 к определению равномерной непрерывности функция $f(x)=\sin\frac{1}{x}$ тем более является равномерно непрерывной на интервале $(\gamma, 1)$, представляющем собой подмножество сегмента $[\gamma, 1]$.

Теорему 4.16 удобно переформулировать в терминах колебания функции на данном сегменте.

Пусть функция $f(x)$ ограничена на данном сегменте $[c, d]$. *Назовем колебанием функции $f(x)$ на сегменте $[c, d]$ разность $\omega=M-m$ между точной верхней и точной нижней гранями функции $f(x)$ на этом сегменте.*

Для непрерывной на сегменте $[c, d]$ функции $f(x)$ колебание равно разности между максимальным и минимальным значениями этой функции на указанном сегменте.

Из теоремы 4.16 непосредственно вытекает следующее утверждение.

Следствие из теоремы 4.16. *Если функция $f(x)$ непрерывна на сегменте $[a, b]$, то для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что колебание функции $f(x)$ на любом содержащемся в сегменте $[a, b]$ сегменте длины, меньшей δ, будет меньше числа ε.*

Замечание 4. Анализируя доказательства теорем 4.14 и 4.15 Вейерштрасса и теоремы 4.16, нетрудно заметить, что в этих трех теоремах вместо сегмента $[a, b]$ можно взять произвольное множество $\{x\}$, для которого выполнены два требования: 1) *это множество $\{x\}$ является ограниченным*; 2) *это множество $\{x\}$ содержит любую свою предельную точку* (такое множество договоримся называть замкнутым).

Множество $\{x\}$, удовлетворяющее указанным двум требованиям, договоримся называть компактным множеством или компактом. Таким образом, указанные три теоремы (т. е. две теоремы Вейерштрасса и теорема 4.16) справедливы не только для функции, непрерывной на сегменте, но и для функции, непрерывной на любом компакте.

В § 7 настоящей главы будут сформулированы более точные определения замкнутого и компактного множеств. Впрочем, для случая числовых множеств эти более точные определения оказываются эквивалентными приведенным нами выше определениям.

4. Понятие модуля непрерывности функции. Предположим, что функция $f(x)$ определена и непрерывна на некотором множестве $\{x\}$, каждая точка которого является предельной точкой этого множества.

Определение. *Для каждого $\delta>0$ назовем модулем непрерывности функции $f(x)$ на множестве $\{x\}$ точную верхнюю грань модуля разности $|f(x')-f(x'')|$ по всем точкам x' и x'', принадлежащим множеству $\{x\}$ и удовлетворяющим неравенству $|x'-x''|<\delta$.*

Для обозначения указанной точной верхней грани обычно употребляют следующий символ:

$$\sup\{|f(x')-f(x'')| : |x'-x''|\leqslant\delta; \quad x', x''\in\{x\}\}.$$

Сам же модуль непрерывности функции $f(x)$ на множестве $\{x\}$ принято обозначать символом $\omega(f, \delta)$.

Таким образом, по определению

$$\omega(f, \delta)=\sup\{|f(x')-f(x'')| : |x'-x''|\leqslant\delta; \quad x', x''\in\{x\}\}. \tag{4.32}$$

Замечание. При определении модуля непрерывности $\omega(f, \delta)$ в правой части (4.32) вместо $|f(x')-f(x'')|$ можно было бы писать разность $[f(x')-f(x'')]$ без знака модуля. Это вытекает из того, что точки x' и x'' можно поменять местами (при этом разность $[f(x')-f(x'')]$ изменит знак на противоположный, в то время как величина $|x'-x''|$ не изменится).

Отметим два свойства модуля непрерывности $\omega(f, \delta)$.

1°. *Модуль непрерывности $\omega(f, \delta)$ всегда неотрицателен: $\omega(f, \delta)\geqslant 0$.*

Это свойство непосредственно вытекает из определения модуля непрерывности (4.32).

2°. *Модуль непрерывности $\omega(f, \delta)$ представляет собой неубывающую функцию δ всюду на полупрямой $\delta>0$.*

В самом деле, при уменьшении δ множество, по которому берется супремум (4.32), сужается, а супремум на части множества не превосходит супремума на всем множестве.

Вычислим модули непрерывности некоторых функций.

1. Вычислим модуль непрерывности функции $f(x)=x^2$ на сегменте $[0, 1]$.

Пусть x' и x'' — любые две точки сегмента $[0, 1]$ такие, что $x''=x'-\delta$, где $0<\delta<1$. Тогда, очевидно,

$$[f(x')-f(x'')]=[(x')^2-(x'')^2]=[(x')^2-(x'-\delta)^2]\leqslant 2\delta-\delta^2.$$

Из последнего неравенства, учитывая замечание к определению модуля непрерывности, мы получим, что

$$\omega(f,\ \delta)=\sup\{[f(x')-f(x'')]:|x'-x''|<\delta;\ x',\ x''\in[0,1]\}\leqslant 2\delta-\delta^2.$$

С другой стороны, взяв $x'=1$, $x''=1-\delta$, так что $|x'-x''|=\delta$, мы получим, что

$$[f(x')-f(x'')]=1-(1-\delta)^2=2\delta-\delta^2.$$

Значит, $\omega(f,\ \delta)=\omega(x^2,\ \delta)=2\delta-\delta^2$ (если $\delta<1$).

2. Вычислим далее модуль непрерывности функции $f(x)=\sin\frac{1}{x}$ на интервале $(0,\ 1)$.

Так как

$$|f(x')-f(x'')|=\left|\sin\frac{1}{x'}-\sin\frac{1}{x''}\right|\leqslant\left|\sin\frac{1}{x'}\right|+\left|\sin\frac{1}{x''}\right|\leqslant 2,$$

то $\omega(f,\ \delta)\leqslant 2$.

С другой стороны, взяв две бесконечно малые последовательности $\{x_n'\}$ и $\{x_n''\}$ тóчек интервала $(0,\ 1)$ вида $x_n'=\dfrac{1}{\frac{\pi}{2}+2\pi n}$, $x_n''=\dfrac{1}{-\frac{\pi}{2}+2\pi n}$, где $n=1,\ 2,\ \ldots$, мы для любого $\delta>0$ сможем указать номер n такой, что $0<x_n'<\delta$ и $0\leqslant x_n''<\delta$, так что $|x_n'-x_n''|\leqslant\delta$, причем

$$[f(x_n')-f(x_n'')]=\sin\frac{1}{x_n'}-\sin\frac{1}{x_n''}=2.$$

Отсюда следует, что $\omega(f,\ \delta)=\omega\left(\sin\frac{1}{x},\ \delta\right)=2$.

3. Вычислим, наконец, модуль непрерывности функции $f(x)=\frac{1}{x}$ на интервале $(0,\ 1)$. Убедимся в том, что этот модуль непрерывности равен $+\infty$.

Фиксировав произвольное $\delta>0$, рассмотрим только такие точки x' и x'', которые удовлетворяют соотношениям $0<x'\leqslant\delta$, $x''=\delta$, так что $|x'-x''|\leqslant\delta$. Очевидно, что *

$$\omega\left(\frac{1}{x},\ \delta\right)\geqslant\sup\left\{\frac{1}{x'}-\frac{1}{\delta}:0<x'<\delta\right\}=+\infty.$$

В заключение докажем теорему, устанавливающую связь между свойством равномерной непрерывности функции $f(x)$ на

* Мы учитываем, что при сужении множества значений x', и x'', по которым берется супремум, этот супремум может только уменьшаться.

множестве $\{x\}$ и величиной модуля непрерывности этой функции на указанном множестве.

Теорема 4.17. *Для того чтобы функция $f(x)$ являлась равномерно непрерывной на множестве $\{x\}$, необходимо и достаточно, чтобы модуль непрерывности $\omega(f, \delta)$ этой функции на указанном множестве удовлетворял соотношению*

$$\lim_{\delta\to 0+0} \omega(f, \delta)=0. \tag{4.33}$$

Доказательство. 1) *Необходимость.* Пусть функция $f(x)$ равномерно непрерывна на множестве $\{x\}$. Требуется доказать, что справедливо соотношение (4.33), т. е. требуется доказать, что для любого $\varepsilon>0$ найдется отвечающее ему $\delta_\varepsilon>0$ такое, что для всех δ, удовлетворяющих условию $0<\delta<\delta_\varepsilon$, справедливо неравенство $\omega(f, \delta)<\varepsilon$*.

По определению равномерной непрерывности для любого $\varepsilon>0$ найдется отвечающее ему $\delta_\varepsilon>0$ такое, что для всех x' и x'' из множества $\{x\}$, удовлетворяющих условию $|x'-x''|<\delta_\varepsilon$, справедливо неравенство $|f(x')-f(x'')|<\frac{\varepsilon}{2}$. Но это и означает, что для любого δ из интервала $0<\delta<\delta_\varepsilon$ справедливо неравенство

$$\omega(f, \delta)=\sup\{|f(x')-f(x'')|:|x'-x''|<\delta;\ x', x''\in\{x\}\}\leqslant\frac{\varepsilon}{2}<\varepsilon.$$

2) *Достаточность.* Пусть выполнено соотношение (4.33), т. е. для любого $\varepsilon>0$ существует отвечающее ему $\delta_\varepsilon>0$ такое, что для всех δ, удовлетворяющих условию $0<\delta<\delta_\varepsilon$, справедливо неравенство $\omega(f, \delta)<\varepsilon$.

Из определения модуля непрерывности следует, что для всех x' и x'' из множества $\{x\}$, удовлетворяющих условию $|x'-x''|\leqslant\delta<\delta_\varepsilon$, справедливо неравенство $|f(x')-f(x'')|<\varepsilon$, а это и означает, что функция $f(x)$ равномерно непрерывна на множестве $\{x\}$. Теорема доказана.

Выше мы вычислили модули непрерывностей трех функций: функции x^2 на сегменте $[0, 1]$ и функций $\sin\frac{1}{x}$ и $\frac{1}{x}$ на интервале $(0, 1)$.

Так как $\omega(x^2, \delta)=2\delta-\delta^2$, $\omega\left(\sin\frac{1}{x}, \delta\right)=2$, $\omega\left(\frac{1}{x}, \delta\right)=+\infty$, то из теоремы 4.17 сразу же вытекает, что функция x^2 равномерно непрерывна на сегменте $[0, 1]$, а функции $\sin\frac{1}{x}$ и $\frac{1}{x}$ не являются равномерно непрерывными на интервале $(0, 1)$.

---

* Мы учитываем при этом, что $\omega(f, \delta)\geqslant 0$.

§ 7. ПОНЯТИЕ КОМПАКТНОСТИ МНОЖЕСТВА

1. Открытые и замкнутые множества. Рассмотрим произвольное множество вещественных чисел $\{x\}$.

Определение 1. *Точка x множества $\{x\}$ называется внутренней точкой этого множества, если существует положительное число δ такое, что δ-окрестность точки x также принадлежит множеству $\{x\}$.*

Определение 2. *Множество $\{x\}$ называется открытым, если любая точка этого множества является внутренней его точкой.*

Примерами открытых множеств могут служить интервал, открытая полупрямая, бесконечная прямая, объединение нескольких непересекающихся интервалов.

Определение 3. *Множество $\{x\}$ называется замкнутым, если дополнение этого множества (т. е. разность $(-\infty, +\infty)\setminus\{x\}$) является открытым множеством.*

В замечании 4 в конце п. 3 § 6 было дано другое определение замкнутого множества. Напомним его формулировку.

Определение 3*. *Множество $\{x\}$ называется замкнутым, если это множество содержит все свои предельные точки.*

Убедимся в том, что для случая произвольных числовых множеств определения 3 и 3* эквивалентны.

1) Пусть сначала множество $\{x\}$ является дополнением открытого множества. Докажем, что в таком случае любая предельная точка x этого множества $\{x\}$ обязательно ему принадлежит.

В самом деле, предположив, что предельная точка x не принадлежит множеству $\{x\}$, мы бы получили, что x принадлежит дополнению множества $\{x\}$, которое является открытым множеством. Но тогда x принадлежала бы этому открытому множеству вместе с некоторой своей δ-окрестностью, т. е. некоторая δ-окрестность точки x не содержала бы точек множества $\{x\}$, а это противоречило бы тому, что x является предельной точкой множества $\{x\}$.

2) Пусть любая предельная точка множества $\{x\}$ принадлежит этому множеству. Докажем, что множество $\{x\}$ является дополнением открытого множества. Пусть x — любая точка дополнения множества $\{x\}$. Требуется доказать, что это дополнение содержит и некоторую δ-окрестность точки x.

Если бы это было не так, то любая δ-окрестность точки x содержала бы точки множества $\{x\}$, т. е. точка x являлась бы предельной точкой множества $\{x\}$ и по условию ему принадлежала, а это противоречило бы тому, что x — точка дополнения множества $\{x\}$.

2. О покрытиях множества системой открытых множеств.

Определение 1. *Будем говорить, что система $\{\Sigma_\alpha\}$ открытых множеств Σ_α образует покрытие множества $\{x\}$, если любая точка x множества $\{x\}$ принадлежит хотя бы одному из множеств системы $\{\Sigma_\alpha\}$.*

Докажем две замечательные леммы о покрытиях множества системой открытых множеств.

Лемма Гейне—Бореля для сегмента. *Из любой системы $\{\Sigma_\alpha\}$ открытых множеств Σ_α, образующей покрытие сегмента $[a, b]$, можно выделить конечную подсистему, также образующую покрытие этого сегмента.*

Доказательство. Пусть бесконечная система $\{\Sigma_\alpha\}$ открытых множеств Σ_α образует покрытие сегмента $[a, b]$*.

Допустим, что сегмент $I=[a, b]$ нельзя покрыть конечным набором открытых множеств из системы $\{\Sigma_\alpha\}$. Тогда, разделив этот сегмент I пополам, мы получим, что хотя бы одну из половин сегмента I нельзя покрыть конечным набором открытых множеств из системы $\{\Sigma_\alpha\}$. Обозначим эту половину через I_1.

Поделив I_1 пополам, мы получим, что хотя бы одну из половин I_1 (обозначим ее через I_2) нельзя покрыть конечным набором множеств из системы $\{\Sigma_\alpha\}$.

Продолжая далее эти рассуждения, мы получим систему вложенных сегментов $\{I_n\}$, каждый из которых нельзя покрыть конечным набором множеств из системы $\{\Sigma_\alpha\}$. Длина n-го сегмента I_n составляет $1/2^n$ часть длины основного сегмента $I=[a, b]$ и стремится к нулю при $n\to\infty$.

В силу следствия из теоремы 3.15 (см. п. 2 § 2 гл. 3) существует единственная точка c, принадлежащая всем сегментам $\{I_n\}$. Так как эта точка c принадлежит основному сегменту $I=[a, b]$, то в системе $\{\Sigma_\alpha\}$ найдется открытое множество Σ_α, которому принадлежит точка c. В силу того, что множество Σ_α является открытым, найдется $\delta>0$ такое, что δ-окрестность точки c, т. е. интервал $(c-\delta, c+\delta)$ также принадлежит множеству Σ_α.

В силу того, что все сегменты I_n содержат точку c и длины этих интервалов стремятся к нулю при $n\to\infty$, можно утверждать, что найдется номер n_0 такой, что при $n\geqslant n_0$ все сегменты I_n содержатся в интервале $(c-\delta, c+\delta)$.

Но это означает, что все сегменты I_n при $n\geqslant n_0$ могут быть покрыты одним множеством Σ_α системы $\{\Sigma_\alpha\}$. Тем самым мы получили противоречие с утверждением о том, что ни один сегмент I_n нельзя покрыть конечным набором множеств из системы $\{\Sigma_\alpha\}$. Полученное противоречие завершает доказательство леммы.

Докажем теперь более общее утверждение.

* Если бы система $\{\Sigma_\alpha\}$, образующая покрытие сегмента $[a, b]$, не являлась бесконечной, то лемма была бы доказана.

Лемма Гейне—Бореля для замкнутого ограниченного множества. *Из любой системы $\{\Sigma_\alpha\}$ открытых множеств Σ_α, образующей покрытие замкнутого ограниченного множества $\{x\}$, можно выделить конечную подсистему, также образующую покрытие множества $\{x\}$.*

Доказательство. Пусть $\{x\}$ — замкнутое ограниченное множество, $\{\Sigma_\alpha\}$ — система открытых множеств, образующая покрытие множества $\{x\}$.

Так как множество $\{x\}$ ограничено, то найдется сегмент $[a, b]$, содержащий это множество. Обозначим через Σ_β открытое множество, являющееся дополнением к замкнутому множеству $\{x\}$ и заметим, что объединение системы $\{\Sigma_\alpha\}$ с открытым множеством Σ_β образует покрытие сегмента $[a, b]$. В силу леммы Гейне—Бореля для сегмента из этого покрытия можно выделить конечную подсистему, также образующую покрытие сегмента $[a, b]$.

Если множество Σ_β входит в эту конечную подсистему, то, удалив его из нее, мы получим конечную подсистему системы $\{\Sigma_\alpha\}$, образующую покрытие множества $\{x\}$*.

Если же множество Σ_β не входит в конечную подсистему, образующую покрытие сегмента $[a, b]$, то эта конечная подсистема состоит исключительно из множеств Σ_α системы $\{\Sigma_\alpha\}$ и образует покрытие множества $\{x\}$, содержащегося в сегменте $[a, b]$. Лемма доказана.

3. Понятие компактности множества. Пусть $\{x\}$ — произвольное множество вещественных чисел.

Определение 1. *Множество $\{x\}$ называется компактным множеством (или компактом), если из любой системы $\{\Sigma_\alpha\}$ открытых множеств, образующей покрытие множества $\{x\}$ можно выделить конечную подсистему, также образующую покрытие множества $\{x\}$.*

В замечании 4 в конце п. 3 § 6 было дано другое определение компактного множества. Напомним его формулировку.

Определение 1*. *Множество $\{x\}$ называется компактным, если оно является замкнутым и ограниченным***.

Докажем, что *для произвольных числовых множеств определения* 1 и 1* *эквивалентны.*

1) Пусть сначала множество $\{x\}$ является замкнутым и ограниченным. Тогда тот факт, что из любой системы открытых множеств $\{\Sigma_\alpha\}$, образующей покрытие $\{x\}$, можно выделить конечную подсистему, также образующую покрытие $\{x\}$, сразу вытекает из леммы Гейне—Бореля (см. п. 2).

---

* Мы учитываем, что множество Σ_β, являясь дополнением к множеству $\{x\}$, не содержит ни одной точки множества $\{x\}$.

** Относительно термина «компактность» см. п. 2 § 3 гл. 12.

2) Пусть множество $\{x\}$ таково, что из любой системы открытых множеств $\{\Sigma_\alpha\}$, образующей покрытие $\{x\}$, можно выделить конечную подсистему, также образующую покрытие $\{x\}$.

Докажем, что множество $\{x\}$ является замкнутым и ограниченным.

Сначала докажем замкнутость множества $\{x\}$.

Достаточно доказать, что дополнение D множества $\{x\}$ является открытым множеством.

Фиксируем произвольную точку y дополнения D. Требуется доказать, что существует некоторая δ-окрестность точки y, также принадлежащая дополнению D.

Пусть x — любая точка множества $\{x\}$. Так как $x \neq y$, то число $\delta(x) = \frac{|x-y|}{2}$ является положительным, причем $\delta(x)$ — окрестности точек x и y

$$\Sigma_x = (x - \delta(x),\ x + \delta(x)),\quad \Psi_x = (y - \delta(x),\ y + \delta(x))$$

не пересекаются.

Поскольку система открытых множеств $\{\Sigma_x\}$, отвечающих всевозможным точкам x множества $\{x\}$ образует покрытие множества $\{x\}$, то из этой системы можно выделить конечную подсистему $\Sigma_{x_1}, \Sigma_{x_2}, \ldots, \Sigma_{x_n}$, также образующую покрытие множества $\{x\}$.

Обозначим через $\Psi_{x_1}, \Psi_{x_2}, \ldots, \Psi_{x_n}$ соответствующую конечную подсистему δ-окрестностей точки y. Наименьшая из этих δ-окрестностей будет содержаться во всех множествах $\Psi_{x_1}, \Psi_{x_2}, \ldots, \Psi_{x_n}$ и потому не будет иметь общих точек ни с одним из множеств $\Sigma_{x_1}, \Sigma_{x_2}, \ldots, \Sigma_{x_n}$. Но тогда, поскольку подсистема $\Sigma_{x_1}, \Sigma_{x_2}, \ldots, \Sigma_{x_n}$ образует покрытие множества $\{x\}$, указанная наименьшая δ-окрестность точки y не будет содержать точек множества $\{x\}$, т. е. будет целиком принадлежать дополнению D множества $\{x\}$.

Тем самым доказано, что множество D является открытым и потому множество $\{x\}$ является замкнутым.

Докажем теперь, что множество $\{x\}$ является ограниченным. Если бы это было не так, то нашлась бы последовательность $\{x_n\}$ не совпадающих друг с другом точек множества $\{x\}$, удовлетворяющая условию

$$|x_n| > n \quad (n = 1, 2, \ldots).$$

Так как эта последовательность не имеет конечных предельных точек, то каждая точка x_n имеет δ-окрестность Σ_{x_n}, свободную от других точек последовательности $\{x_n\}$.

Очевидно, что из бесконечной системы открытых множеств $\{\Sigma_{x_n}\}$, образующих покрытие множества точек $\{x_n\}$, нельзя выбрать конечной подсистемы, образующей покрытие всех точек $\{x_n\}$.

Так как множество $\{x_n\}$ является подмножеством $\{x\}$, то тем более нельзя из всякой системы открытых множеств, образующих покрытие $\{x\}$, выделить конечную подсистему, также образующую покрытие $\{x\}$. Но это противоречит нашему предположению о множестве $\{x\}$. Полученное противоречие доказывает ограниченность множества $\{x\}$.

Заметим в заключение, что все введенные в этом параграфе понятия в более общей ситуации изучаются в дополнении 2 к гл. 12.

Глава 5

ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ

В настоящей главе будут введены фундаментальные понятия производной и дифференциала функции. Мы установим основные правила дифференцирования и вычислим производные всех простейших элементарных функций, уже приведенные нами в гл. 1 и известные из школьного курса. В конце главы будут рассмотрены производные и дифференциалы высших порядков и вопрос о дифференцировании функции, заданной параметрически.

§ 1. ПОНЯТИЕ ПРОИЗВОДНОЙ

1. Приращение функции. Разностная форма условия непрерывности. Рассмотрим функцию $y=f(x)$, заданную на интервале (a, b)*. Пусть x — любая фиксированная точка интервала (a, b), а Δx — произвольное число, настолько малое, что значение $x+\Delta x$ также находится на интервале (a, b). Это число Δx обычно называют приращением аргумента.

Приращением функции $y=f(x)$ в точке x, отвечающим приращению аргумента Δx, будем называть число

$$\Delta y=f(x+\Delta x)-f(x). \tag{5.1}$$

Так, для функции $y=\sin x$ приращение в точке x, отвечающее приращению аргумента Δx, имеет вид

$$\Delta y=\sin(x+\Delta x)-\sin x=2\cos\left(x+\frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}. \tag{5.2}$$

Справедливо следующее *утверждение:*

Для того чтобы функция $y=f(x)$ являлась непрерывной в точке x, необходимо и достаточно, чтобы приращение Δy этой функции в точке x, отвечающее приращению аргумента Δx, являлось бесконечно малым при $\Delta x\to 0$.

В самом деле, по определению функция $y=f(x)$ непрерывна в точке x, если существует предел

$$\lim_{\Delta x\to 0} f(x+\Delta x)=f(x). \tag{5.3}$$

В силу п. 4 § 4 гл. 3 существование предельного значения (5.3)

---

* В качестве множества задания функции вместо интервала (a, b) можно взять любое плотное в себе множество $\{x\}$ (см. гл. 2, § 5, п. 3).

эквивалентно тому, что функция $[f(x+\Delta x)-f(x)]$ аргумента Δx является бесконечно малой при $\Delta x \to 0$.

Доказанное утверждение позволяет выразить условие непрерывности функции $y=f(x)$ в точке x в следующей форме: *функция $y=f(x)$ непрерывна в точке x, если приращение Δy этой функции в точке x, отвечающее приращению аргумента Δx, является бесконечно малым при $\Delta x \to 0$, т. е. если*

$$\lim_{\Delta x \to 0} \Delta y = \lim_{\Delta x \to 0} [f(x+\Delta x) - f(x)] = 0. \tag{5.4}$$

Условие (5.4) мы будем называть *разностной формой условия непрерывности* функции $y=f(x)$ в точке x. Это условие мы будем неоднократно использовать в дальнейшем.

С помощью условия (5.4) еще раз убедимся в том, что функция $y=\sin x$ непрерывна в любой точке x бесконечной прямой. В самом деле, из формулы (5.2), из условия $\left|\cos\left(x+\frac{\Delta x}{2}\right)\right| \leqslant 1$ и из равенства $\lim\limits_{\Delta x \to 0} \sin\frac{\Delta x}{2} = 0$ непосредственно вытекает, что $\lim\limits_{\Delta x \to 0} \Delta y = 0$.

2. Определение производной. Пусть, как и в п. 1, функция $y=f(x)$ определена на интервале (a, b), x — фиксированная точка этого интервала, Δx — любое приращение аргумента, настолько малое, что число $x+\Delta x$ также принадлежит интервалу (a, b).

Считая, что $\Delta x \neq 0$, рассмотрим в данной *фиксированной* точке x отношение приращения Δy функции $y=f(x)$ в этой точке к соответствующему приращению аргумента Δx:

$$\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x) - f(x)}{\Delta x}. \tag{5.5}$$

Отношение (5.5) будем называть *разностным отношением* (в данной точке x). Так как x фиксировано, разностное отношение (5.5) представляет собой функцию аргумента Δx. Эта функция определена для всех значений аргумента Δx, принадлежащих некоторой достаточно малой δ-окрестности точки $\Delta x=0$, за исключением самой точки $\Delta x=0$, т. е. определена всюду в достаточно малой проколотой δ-окрестности точки $\Delta x=0$. Это дает нам право рассматривать вопрос о существовании предела указанной функции при $\Delta x \to 0$.

Определение 1. *Производной функции $y=f(x)$ в данной фиксированной точке x называется предел при $\Delta x \to 0$ разностного отношения* (5.5) *(при условии, что этот предел существует).*

Производную функции $y=f(x)$ в данной фиксированной точке x будем обозначать символом $f'(x)$ или $y'(x)$ или кратким символом y'.

Итак, по определению

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

Если функция имеет производную для всех точек x интервала (a, b), то эта производная будет представлять собой некоторую функцию аргумента x, определенную на интервале (a, b).

Приведем два тривиальных примера вычисления производных.

1°. $f(x) = C = \text{const}$. Совершенно очевидно, что производная $f'(x)$ этой функции тождественно равна нулю, ибо приращение этой функции $\Delta y = f(x+\Delta x) - f(x)$ *равно нулю для всех x и всех Δx.*

2°. $f(x) = x$. Для этой функции разностное отношение (5.5) равно

$$\frac{\Delta y}{\Delta x} = \frac{(x+\Delta x) - x}{\Delta x} = \frac{\Delta x}{\Delta x} = 1.$$

Отсюда следует, что и производная указанной функции равна единице в любой точке x бесконечной прямой.

В полной аналогии с понятиями правого и левого пределов функции в данной точке вводятся понятия п р а в о й и л е в о й производных функции $y = f(x)$ в данной фиксированной точке x.

О п р е д е л е н и е 2. *П р а в о й [л е в о й] производной функции $y = f(x)$ в данной фиксированной точке x называется правый [левый] предел разностного отношения* (5.5) *в точке $\Delta x = 0$ (при условии, что этот предел существует).*

Для обозначения правой [левой] производной функции в точке x используют символ $f'(x+0)$ $[f'(x-0)]$.

Из сопоставления определений 1 и 2 и из свойства правого и левого пределов функции, установленного в п. 2 § 4 гл. 3 *, вытекают следующие утверждения:

1) *если функция $f(x)$ имеет в точке x производную $f'(x)$, то эта функция имеет в точке x как правую, так и левую производные, причем $f'(x+0) = f'(x-0) = f'(x)$;*

2) *если функция $f(x)$ имеет в точке x как правую, так и левую производные, причем эти производные равны друг другу, то функция $f(x)$ имеет в точке x производную $f'(x)$, причем*

$$f'(x) = f'(x+0) = f'(x-0).$$

В дополнение к утверждению 2) следует отметить, что если у функции $f(x)$ существуют правая и левая производные в точке x, но эти производные не равны друг другу, то у этой функции не

---

* Если функция $f(x)$ имеет в точке x и правый, и левый пределы, оба равные одному и тому же числу b, то функция $f(x)$ имеет в этой точке предел, равный числу b.

существует производной в точке x^*. Примером такой функции может служить функция

$$f(x) = |x| = \begin{cases} x \text{ при } x \geqslant 0, \\ -x \text{ при } x < 0. \end{cases}$$

Эта функция имеет в точке $x=0$ правую производную, равную $\lim\limits_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = 1$, и левую производную, равную $\lim\limits_{\Delta x \to 0} \frac{(-\Delta x)}{\Delta x} = -1$, но не имеет в точке $x=0$ производной.

3. Геометрический смысл производной. Рассмотрим график функции $y=f(x)$, определенной и непрерывной на интервале (a, b). Фиксируем произвольную точку x интервала (a, b) и рассмотрим приращение $\Delta x \neq 0$ аргумента x, настолько малое, что число $x+\Delta x$ также принадлежит интервалу (a, b). Пусть M и P — точки графика функции $y=f(x)$, абсциссы которых соответственно равны x и $x+\Delta x$ (рис. 5.1). Координаты точек M и P, очевидно, будут иметь вид

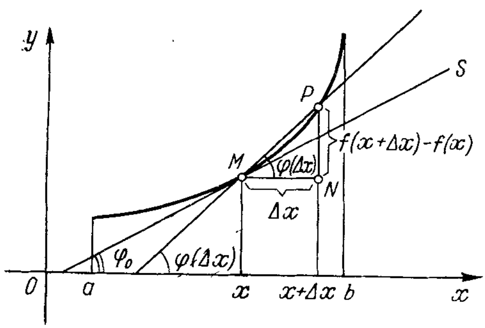


Рис. 5.1

$$M(x, f(x)),$$
$$P(x+\Delta x, f(x+\Delta x)).$$

Прямую, проходящую через точки M и P графика функции $y=f(x)$, будем называть с е к у щ е й. Поскольку точку M мы предполагаем фиксированной, то угол наклона каждой секущей MP к оси Ox будет функцией аргумента Δx (ибо значение Δx однозначно определяет точку P графика функции $y=f(x)$). Обозначим указанный угол наклона секущей MP к оси Ox символом $\varphi(\Delta x)$.

О п р е д е л е н и е. *Если существует предельное положение секущей MP при стремлении точки P графика функции к точке M (или, что то же самое, при стремлении Δx к нулю), то это предельное положение называется* к а с а т е л ь н о й *к графику функции $y=f(x)$ в данной фиксированной точке M этого графика.*

Из этого определения следует, что для существования касательной к графику функции $y=f(x)$ в точке M достаточно, чтобы существовал предел $\lim\limits_{\Delta x \to 0} \varphi(\Delta x) = \varphi_0$, причем указанный предел φ_0 равен углу наклона касательной к оси Ox.

* Иначе мы получили бы противоречие с утверждением 1).

Докажем следующее *у т в е р ж д е н и е*:

Если функция $y=f(x)$ имеет в данной фиксированной точке x производную, то существует касательная к графику функции $y=f(x)$ в точке $M(x, f(x))$, причем угловой коэффициент этой касательной (т. е. тангенс угла наклона ее к оси Ox) равен производной $f'(x)$.

Опустим из точек M и P перпендикуляры на ось абсцисс (см. рис. 5.1). Проведем через точку M прямую, параллельную оси абсцисс, и обозначим через N точку пересечения этой прямой с перпендикуляром, опущенным из P на ось абсцисс. Из треугольника MNP очевидно, что

$$\operatorname{tg}\varphi(\Delta x)=\frac{\Delta y}{\Delta x}=\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

Таким образом,

$$\varphi(\Delta x)=\operatorname{arctg}\frac{\Delta y}{\Delta x}. \tag{5.6}$$

Убедимся в том, что существует предел правой (а значит, и левой) части (5.6) при $\Delta x\to 0$. В самом деле, в силу существования производной $f'(x)$ существует предел $\lim\limits_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=f'(x)$. Отсюда и из непрерывности функции $\operatorname{arctg} u$ для всех значений u следует, что существует предел правой части (5.6), равный $\operatorname{arctg} f'(x)$.

Итак, мы доказали, что существует предел

$$\lim_{\Delta x\to 0}\varphi(\Delta x)=\operatorname{arctg} f'(x).$$

Но это и означает, что существует предельное положение секущей, т. е. существует касательная к графику функции в точке $M(x, f(x))$, причем угол наклона φ_0 этой касательной к оси Ox равен $\varphi_0=\operatorname{arctg} f'(x)$.

Значит, угловой коэффициент указанной касательной $\operatorname{tg}\varphi_0$ равен $f'(x)$.

Сформулированное утверждение доказано.

§ 2. ПОНЯТИЕ ДИФФЕРЕНЦИРУЕМОСТИ ФУНКЦИИ

1. Определение дифференцируемости функции. Пусть, как и в предыдущем параграфе, функция $y=f(x)$ определена на интервале (a, b), x — любое фиксированное число из этого интервала, Δx — произвольное приращение аргумента, настолько малое, что значение аргумента $x+\Delta x$ также принадлежит интервалу (a, b), $\Delta y=f(x+\Delta x)-f(x)$ — приращение функции в точке x, отвечающее приращению аргумента Δx.

О п р е д е л е н и е. *Функция $y=f(x)$ называется дифференцируемой в точке x, если приращение Δy этой функции в точке x, отве-*

чающее приращению аргумента Δx, может быть представлено в виде

$$\Delta y = A\Delta x + \alpha(\Delta x)\Delta x, \tag{5.7}$$

где A — некоторое число, не зависящее от Δx, а $\alpha(\Delta x)$ — функция аргумента Δx, бесконечно малая в точке $\Delta x = 0$.

В самой точке $\Delta x = 0$ эта функция $\alpha(\Delta x)$, вообще говоря, не определена, и ей можно приписать в этой точке любое значение. Для дальнейшего удобно считать это значение $\alpha(0)$ равным нулю. При такой договоренности функция $\alpha(\Delta x)$ будет непрерывна в точке $\Delta x = 0$ и равенство (5.7) можно распространить и на значение $\Delta x = 0$.

Замечание. Второе слагаемое в правой части (5.7) $\alpha(\Delta x)\Delta x$ можно переписать в виде $o(\Delta x)$ *. В самом деле, так как обе функции $\alpha(\Delta x)$ и Δx являются бесконечно малыми в точке $\Delta x = 0$, то произведение этих функций $\alpha(\Delta x)\Delta x$ представляет собой бесконечно малую в точке $\Delta x = 0$ функцию более высокого порядка, чем Δx (см. п. 5 § 4 гл. 3). Таким образом, представление (5.7) можно переписать в виде

$$\Delta y = A\Delta x + o(\Delta x).$$

Докажем следующее утверждение.

Теорема 5.1. *Для того чтобы функция $y = f(x)$ была дифференцируемой в точке x, необходимо и достаточно, чтобы она имела в этой точке конечную производную $f'(x)$.*

Доказательство. 1) *Необходимость.* Пусть функция $y = f(x)$ дифференцируема в точке x, т. е. ее приращение Δy в этой точке, отвечающее приращению аргумента Δx, представимо в виде (5.7). Считая Δx отличным от нуля и поделив (5.7) на Δx, получим

$$\frac{\Delta y}{\Delta x} = A + \alpha(\Delta x). \tag{5.8}$$

Правая (а поэтому и левая) часть (5.8) имеет равный A предел в точке $\Delta x = 0$ **. Остается заметить, что предел при $\Delta x \to 0$ левой части (5.8) (в случае, если он существует) по определению равен производной $f'(x)$.

Итак, мы доказали, что если для функции $f(x)$ справедливо представление (5.7), то эта функция имеет в точке x производную $f'(x)$, причем $f'(x) = A$.

2) *Достаточность.* Пусть существует конечная производная $f'(x)$, т. е. существует конечный предел

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x). \tag{5.9}$$

---

* Напомним, что символ $o(\Delta x)$ обозначает бесконечно малую в точке $\Delta x = 0$ функцию более высокого порядка, чем Δx.

** Ибо $\lim\limits_{\Delta x \to 0} A = A$, $\lim\limits_{\Delta x \to 0} \alpha = 0$.

Обозначим символом $\alpha(\Delta x)$ разность $\frac{\Delta y}{\Delta x} - f'(x)$, т. е. положим

$$\alpha(\Delta x) = \frac{\Delta y}{\Delta x} - f'(x). \tag{5.10}$$

Из существования предела (5.9) вытекает, что функция $\alpha(\Delta x)$, определяемая соотношением (5.10), имеет предел при $\Delta x \to 0$, равный нулю.

Умножая соотношение (5.10) на Δx, мы придем к представлению

$$\Delta y = f'(x)\,\Delta x + \alpha(\Delta x)\,\Delta x,$$

совпадающему с представлением (5.7) при $A = f'(x)$.

Тем самым доказано, что из существования конечной производной $f'(x)$ вытекает дифференцируемость функции $y = f(x)$ в точке x, причем в условии дифференцируемости (5.7) число A совпадает с $f'(x)$. Теорема доказана.

Доказанная теорема позволяет нам в дальнейшем отождествлять понятие дифференцируемости функции в данной точке с понятием существования у этой функции в данной точке конечной производной.

Операцию нахождения производной в дальнейшем договоримся называть дифференцированием.

2. Дифференцируемость и непрерывность. Легко доказывается следующее утверждение.

Теорема 5.2. *Если функция $y = f(x)$ дифференцируема в данной точке x, то она и непрерывна в этой точке.*

Доказательство. Так как функция $f(x)$ дифференцируема в точке x, то для ее приращения Δy в этой точке справедливо представление (5.7), из которого следует, что $\lim\limits_{\Delta x \to 0} \Delta y = 0$, а это и означает непрерывность функции $y = f(x)$ в данной точке (в силу разностной формы условия непрерывности (5.4), введенной в п. 1 § 1). Теорема доказана.

Заметим, что утверждение, обратное к теореме 5.2, несправедливо, т. е. из непрерывности функции $y = f(x)$ в данной точке x, вообще говоря, не вытекает дифференцируемость функции $f(x)$ в этой точке.

Примером может служить функция $y = |x|$, которая, очевидно, непрерывна в точке $x = 0$, но (как мы уже видели в конце п. 2 § 1) не имеет в этой точке производной.

Отметим, что существуют функции, непрерывные в каждой точке некоторого интервала, но не имеющие производной ни в одной точке этого интервала. (Первый пример такой функции был построен Вейерштрассом. Один из примеров такой функции строится в дополнении к гл. 10.)

3. Понятие дифференциала функции. Рассмотрим функцию $y=f(x)$, дифференцируемую в данной точке x. Приращение Δy такой функции в точке x может быть представлено в виде (5.7).

Заметим, что приращение (5.7) представляет собой сумму двух слагаемых, первое из которых $A\Delta x$ линейно относительно Δx, а второе $\alpha(\Delta x)\Delta x$ является в точке $\Delta x=0$ бесконечно малой функцией более высокого порядка, чем Δx. Если число A, равное согласно теореме 5.1 производной $f'(x)$, отлично от нуля, то указанное первое слагаемое $A\Delta x=f'(x)\Delta x$ представляет собой главную часть приращения Δy дифференцируемой функции $y=f(x)$. Эта главная часть приращения является линейной однородной функцией аргумента Δx* и называется дифференциалом функции $y=f(x)$.

В случае, если $A=f'(x)=0$, дифференциал функции по определению считается равным нулю.

Итак, *дифференциалом функции $y=f(x)$ в данной фиксированной точке x, отвечающим приращению аргумента Δx, называется число, обозначаемое символом dy и равное*

$$dy=f'(x)\Delta x. \tag{5.11}$$

В случае $f'(x)\neq 0$ это число представляет собой главную часть приращения Δy функции $y=f(x)$, линейную и однородную относительно приращения аргумента Δx.

Сразу же отметим, что дифференциал dy и приращение Δy функции $y=f(x)$ в данной точке x, оба отвечающие одному и тому же приращению аргумента Δx, вообще говоря, не равны друг другу.

Это легко уяснить из рассмотрения графика функции $y=f(x)$ (рис. 5.2). Пусть M и P — точки графика функции $y=f(x)$, отвечающие значениям аргумента, соответственно равным x и $x+\Delta x$, MS — касательная к графику в точке M, $MN \| Ox$, $NP \| Oy$, Q — точка пересечения касательной MS с прямой PN. Тогда приращение Δy функции $y=f(x)$ в точке x, отвечающее приращению аргумента Δx, очевидно, равно величине отрезка NP, в то время как дифференциал dy этой функции в точке x, отвечающий тому же самому Δx, равен величине отрезка NQ. (Это сразу вытекает из формулы (5.11) и из того, что в прямоугольном

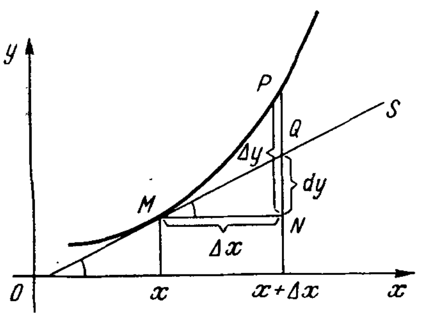


Рис. 5.2

* Напомним, что линейной функцией аргумента t называется функция вида $y=At+B$, где A и B — некоторые постоянные. В случае $B=0$ линейная функция называется однородной.

треугольнике MQN величина отрезка MN равна Δx, а тангенс угла QMN равен $f'(x)$.) Ясно, что величины отрезков NP и NQ являются, вообще говоря, различными.

Весьма удобно ввести в рассмотрение понятие дифференциала аргумента x. При этом следует различать два случая: 1) случай, когда указанный аргумент x представляет собой *независимую переменную*; 2) случай, когда аргумент x сам является *дифференцируемой функцией вида* $x=\varphi(t)$ некоторой новой переменной t, которую мы можем считать независимой.

Договоримся для случая, когда аргумент x является независимой переменной, отождествлять дифференциал этого аргумента с его приращением Δx*, т. е. считать, что $dx=\Delta x$.

В силу этой договоренности равенство (5.11) принимает вид

$$dy=f'(x)dx. \tag{5.12}$$

Таким образом, для случая, когда аргумент x является независимой переменной, для дифференциала функции $y=f(x)$ справедливо представление (5.12).

Ниже в п. 3 § 3 мы докажем, что представление (5.12) носит универсальный характер и справедливо также и в случае, когда аргумент x не является независимой переменной, а является дифференцируемой функцией вида $x=\varphi(t)$ некоторой независимой переменной t. (В этом случае в формуле (5.12) величину dx нельзя считать равной Δx, ибо в силу сказанного выше она равна $dx=\varphi'(t)dt$.)

§ 3. ДИФФЕРЕНЦИРОВАНИЕ СЛОЖНОЙ ФУНКЦИИ И ОБРАТНОЙ ФУНКЦИИ

1. Дифференцирование сложной функции. Установим правило, позволяющее найти производную сложной функции $y=f[\varphi(t)]$ в точке t при условии, что известны производные составляющих ее функций $x=\varphi(t)$ и $y=f(x)$ в точках t и $x=\varphi(t)$ соответственно.

Теорема 5.3. *Пусть функция $x=\varphi(t)$ дифференцируема в точке t, а функция $y=f(x)$ дифференцируема в соответствующей точке $x=\varphi(t)$. Тогда сложная функция $y=f[\varphi(t)]$ дифференцируема в указанной точке t, причем для ее производной в этой точке справедлива формула*

$$\{f[\varphi(t)]\}'=f'(x)\cdot\varphi'(t)=f'[\varphi(t)]\cdot\varphi'(t). \tag{5.13}$$

Доказательство. Придадим аргументу функции $x=\varphi(t)$ в данной точке t произвольное отличное от нуля приращение Δt. Этому приращению отвечает приращение $\Delta x=\varphi(t+\Delta t)-\varphi(t)$

---

* Эта договоренность согласуется с рассмотрением независимой переменной x как функции вида $y=f(x)=x$, для которой $dy=f'(x)\Delta x=\Delta x$, т. е. $dx=\Delta x$.

функции $x=\varphi(t)$, причем указанное приращение Δx может обращаться в нуль.

Приращению Δx, в свою очередь, отвечает приращение $\Delta y=f(x+\Delta x)-f(x)$ функции $y=f(x)$ в соответствующей точке $x=\varphi(t)$. Поскольку функция $y=f(x)$ по условию дифференцируема в указанной точке $x=\varphi(t)$, то ее приращение Δy в этой точке может быть представлено в виде *

$$\Delta y=f'(x)\Delta x+\alpha(\Delta x)\Delta x, \tag{5.14}$$

где $\alpha(\Delta x)$ имеет при $\Delta x\to 0$ предел, равный нулю.

Подчеркнем, что, как указано в п. 1 § 2, представление (5.14) остается справедливым и при $\Delta x=0$.

Поделив (5.14) на $\Delta t\neq 0$, будем иметь

$$\frac{\Delta y}{\Delta t}=f'(x)\frac{\Delta x}{\Delta t}+\alpha(\Delta x)\frac{\Delta x}{\Delta t}. \tag{5.15}$$

Докажем, что правая (а значит, и левая) часть (5.15) имеет предел при $\Delta t\to 0$, причем этот предел равен величине, стоящей в правой части (5.13). Этим будет доказана дифференцируемость сложной функции и формула (5.13) для ее производной.

Из дифференцируемости функции $x=\varphi(t)$ в точке t вытекает, что отношение $\frac{\Delta x}{\Delta t}$ имеет предел при $\Delta t\to 0$, равный $\varphi'(t)$. Остается доказать, что функция $\alpha(\Delta x)$ имеет предел при $\Delta t\to 0$, равный нулю, но это сразу вытекает из того, что $\alpha(\Delta x)\to 0$ при $\Delta x\to 0$ и что $\Delta x\to 0$ при $\Delta t\to 0$ на основании разностной формы условия непрерывности дифференцируемой в точке t функции $x=\varphi(t)$ **. Итак, вся правая часть (5.15) имеет предел при $\Delta t\to 0$, и этот предел равен величине, стоящей в правой части (5.13).

Теорема доказана.

Замечание 1. Теорема 5.3 и содержащееся в ее формулировке правило вычисления производной сложной функции последовательно переносится на сложную функцию, являющуюся суперпозицией трех и большего числа функций. Так, для сложной функции, являющейся суперпозицией трех функций $y=F[f(\varphi(f))]$, правило дифференцирования имеет вид

$$\{F[f(\varphi(t))]\}'=F'[f(\varphi(t))]f'(\varphi(t))\varphi'(t), \tag{5.16}$$

причем формула (5.16) справедлива при условии, что функция $x=\varphi(t)$ дифференцируема в данной точке t, функция $u=f(x)$ дифференцируема в соответствующей точке $x=\varphi(t)$, а функция $y=F(u)$ дифференцируема в соответствующей точке $u=f(x)=f[\varphi(t)]$.

---

* См. п. 1 § 2 определение дифференцируемости и теорему 5.1.

** В силу теоремы 5.2 дифференцируемая в точке t функция $x=\varphi(t)$ является непрерывной в этой точке.

Замечание 2. При доказательстве теоремы 5.3 мы рассматривали сложную функцию вида $y=f(x)$, где $x=\varphi(t)$, т. е. обозначили символом x промежуточный аргумент. Эта символика, конечно, может быть изменена. Чаще удобнее бывает обозначать символом x окончательный аргумент, т. е. рассматривать сложную функцию вида $y=f[\varphi(x)]$. В этих обозначениях правило дифференцирования сложной функции (5.13) принимает вид

$$\{f[\varphi(x)]\}'=f'[\varphi(x)]\cdot\varphi'(x). \qquad (5.13^*)$$

Примеры применения правила дифференцирования сложной функции будут приведены в следующем параграфе.

2. Дифференцирование обратной функции.

Теорема 5.4. *Пусть функция $y=f(x)$ возрастает (или убывает) и непрерывна в некоторой окрестности точки x. Пусть, кроме того, эта функция дифференцируема в указанной точке x и ее производная в этой точке $f'(x)$ отлична от нуля. Тогда в некоторой окрестности соответствующей точки $y=f(x)$ определена обратная для $y=f(x)$ функция $x=f^{-1}(y)$, причем указанная обратная функция дифференцируема в соответствующей точке $y=f(x)$ и для ее производной в этой точке y справедлива формула*

$$\{f^{-1}(y)\}'=\frac{1}{f'(x)}. \qquad (5.17)$$

Доказательство. Так как функция $y=f(x)$ строго монотонна и непрерывна в некоторой окрестности данной точки x, то в силу теоремы 4.5 (см. § 2 гл. 4) обратная функция $x=f^{-1}(y)$ определена, строго монотонна и непрерывна в некоторой окрестности соответствующей точки $y=f(x)$.

Придадим аргументу этой обратной функции в указанной точке произвольное достаточно малое и отличное от нуля приращение Δy. Этому приращению Δy отвечает приращение $\Delta x=f^{-1}(y+\Delta y)-f^{-1}(y)$ обратной функции в соответствующей точке $y=f(x)$, причем в силу строгой монотонности обратной функции указанное приращение Δx отлично от нуля.

Это дает нам право написать следующее тождество *:

$$\frac{\Delta x}{\Delta y}=\frac{1}{\dfrac{\Delta x}{\Delta y}}. \qquad (5.18)$$

Пусть теперь в тождестве (5.18) приращение Δy стремится к нулю. Тогда в силу разностной формы условия непрерывности обратной функции $x=f^{-1}(y)$ в соответствующей точке $y=f(x)$ приращение этой функции Δx также стремится к нулю. Убедимся в том, что в таком случае существует предел правой части (5.18),

* Указанное тождество можно написать для любых двух чисел Δy и Δx, отличных от нуля.

равный величине, стоящей в правой части (5.17). Этим будет доказано, что тот же самый предел имеет и левая часть (5.18), т. е. будет доказано, что обратная функция имеет производную в соответствующей точке $y=f(x)$ и для этой производной справедливо равенство (5.17).

Итак, для завершения доказательства теоремы остается убедиться в том, что правая часть (5.18) имеет предел при $\Delta x \to 0$ равный $1/f'(x)$, где x — данная точка.

Так как $x=f^{-1}(y)$, $\Delta x=f^{-1}(y+\Delta y)-f^{-1}(y)$, то $x+\Delta x=f^{-1}(y+\Delta y)$, т. е. $y+\Delta y=f(x+\Delta x)$ и $\Delta y=f(x+\Delta x)-f(x)$.

Отсюда следует, что правая часть (5.18) может быть переписана в виде

$$\frac{1}{\dfrac{\Delta y}{\Delta x}}=\frac{1}{\dfrac{f(x+\Delta x)-f(x)}{\Delta x}}.$$

Из последнего равенства в силу определения производной $f'(x)$ и предположения $f'(x)\neq 0$ сразу же вытекает, что предел при $\Delta x \to 0$ правой части (5.18) существует и равен $1/f'(x)$.

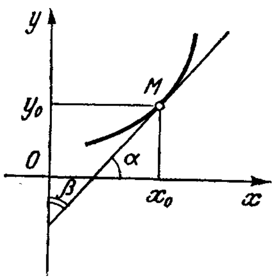


Рис. 5.3

Теорема доказана.

Примеры применения доказанной теоремы будут даны в следующем параграфе.

Доказанная теорема имеет простой геометрический смысл. Пусть M — точка графика функции $y=f(x)$, отвечающая данному значению аргумента x (рис. 5.3). Тогда, очевидно, производная $f'(x)$ равна тангенсу угла наклона α касательной, проходящей через точку M, к оси Ox, а производная обратной функции $\{f^{-1}(y)\}'$ в соответствующей точке $y=f(x)$ равна тангенсу угла наклона β той же самой касательной к оси Oy. Поскольку углы наклона α и β в сумме составляют $\pi/2$, то формула (5.17) выражает очевидный факт: $\operatorname{tg}\beta=\dfrac{1}{\operatorname{tg}\alpha}$ при $\alpha+\beta=\pi/2$.

3. Инвариантность формы первого дифференциала. В п. 3 § 2 мы убедились в том, что для случая, когда аргумент x дифференцируемой функции $y=f(x)$ представляет собой независимую переменную, для дифференциала dy этой функции справедливо представление

$$dy=f'(x)\,dx. \tag{5.12}$$

Сейчас мы докажем, что представление (5.12) является универсальным и справедливо также и в случае, когда аргумент x сам является дифференцируемой функцией вида $x=\varphi(t)$ некоторой независимой переменной t. Это свойство дифференциала функции принято называть инвариантностью его формы*.

Итак, пусть аргумент x дифференцируемой функции $y=f(x)$ сам является дифференцируемой функцией вида $x=\varphi(t)$ некоторой независимой переменной t. В таком случае y можно рассматривать как сложную функцию вида $y=f[\varphi(t)]$ аргумента t. Поскольку этот аргумент t является независимой переменной, то для указанной функции $y=f[\varphi(t)]$ и для функции $x=\varphi(t)$ дифференциалы представимы в форме (5.12), т. е. в виде

$$dy=\{f[\varphi(t)]\}'dt, \qquad dx=\varphi'(t)dt. \tag{5.19}$$

По правилу дифференцирования сложной функции

$$\{f[\varphi(t)]\}'=f'(x)\varphi'(t). \tag{5.13}$$

Подставляя (5.13) в первую из формул (5.19), придадим этой формуле вид

$$dy=f'(x)\varphi'(t)dt. \tag{5.20}$$

Сопоставляя полученное равенство (5.20) со вторым из равенств (5.19), окончательно получим для dy выражение

$$dy=f'(x)dx,$$

совпадающее с представлением (5.12).

Инвариантность формы (5.12) первого дифференциала функции dy установлена.

Замечание. Можно дать и другую эквивалентную формулировку свойства инвариантности формы первого дифференциала, сразу же вытекающую из универсальности представления (5.12): *производная дифференцируемой функции $y=f(x)$ равна отношению дифференциала этой функции dy к дифференциалу ее аргумента dx, т. е. определяется равенством*

$$f'(x)=\frac{dy}{dx}, \tag{5.21}$$

как в случае, когда аргумент x является независимой переменной, так и в случае, когда аргумент x сам является дифференцируемой функцией вида $x=\varphi(t)$ некоторой независимой переменной t.

Универсальность представления для производной (5.21) позволяет использовать отношение $\frac{dy}{dx}$ для обозначения производной функции $y=f(x)$ по аргументу x.

4. Применение дифференциала для установления приближенных формул. Пусть ради простоты аргумент x функции $y=f(x)$ является независимой переменной. В п. 3 § 2 мы показали, что диф-

* Ниже мы введем понятия второго и последующих дифференциалов функции $y=f(x)$ и обнаружим, что эти дифференциалы уже не обладают инвариантностью формы. Вследствие этого доказываемое свойство называют инвариантностью формы первого дифференциала.

ференциал dy функции $y=f(x)$, вообще говоря, не равен приращению Δy этой функции. Тем не менее с точностью до бесконечно малой функции более высокого порядка малости, чем Δx, справедливо приближенное равенство

$$\Delta y \approx dy. \tag{5.22}$$

Отношение $\frac{\Delta y - dy}{\Delta x}$ естественно назвать *относительной погрешностью* равенства (5.22). Так как $\Delta y - dy = o(\Delta x)$ *, то относительная погрешность равенства (5.22) становится как угодно малой при уменьшении $|\Delta x|$.

Соотношение (5.22) позволяет приближенно заменять приращение Δy дифференцируемой функции $y=f(x)$ дифференциалом dy этой функции. Целесообразность такой замены оправдывается тем, что дифференциал dy является линейной функцией Δx, в то время как приращение Δy, вообще говоря, представляет собой более сложную функцию аргумента Δx.

Учитывая, что $\Delta y = f(x+\Delta x) - f(x)$, $dy = f'(x)\Delta x$, мы можем переписать приближенное равенство (5.22) в виде $f(x+\Delta x) - f(x) \approx \approx f'(x)\Delta x$ или, что то же самое, в виде

$$f(x+\Delta x) \approx f(x) + f'(x)\Delta x. \tag{5.23}$$

Приближенное равенство (5.23), так же как и (5.22), справедливо для любой дифференцируемой в данной точке x функции $f(x)$ с точностью до величины $o(\Delta x)$ более высокого порядка малости, чем Δx.

Это приближенное равенство позволяет с ошибкой $o(\Delta x)$ заменить функцию $f(x)$ в малой окрестности точки x (т. е. для малых значений Δx) линейной функцией аргумента Δx, стоящей в правой части (5.23).

Приближенная формула (5.23) часто применяется для различных конкретных видов функции $f(x)$.

§ 4. ДИФФЕРЕНЦИРОВАНИЕ СУММЫ, РАЗНОСТИ, ПРОИЗВЕДЕНИЯ И ЧАСТНОГО ФУНКЦИЙ

Теорема 5.5. *Если каждая из функций $u(x)$ и $v(x)$ дифференцируема в данной точке x, то сумма, разность, произведение и частное этих функций (частное при условии, что значение $v(x)\neq 0$) также дифференцируемы в этой точке, причем имеют место формулы*

$$\begin{cases} [u(x) \pm v(x)]' = u'(x) \pm v'(x), \\ [u(x)\cdot v(x)]' = u'(x)\cdot v(x) + u(x)\cdot v'(x), \\ \left[\dfrac{u(x)}{v(x)}\right]' = \dfrac{u'(x)\cdot v(x) - u(x)\cdot v'(x)}{v^2(x)}. \end{cases} \tag{5.24}$$

* См. п. 3 § 2.

Доказательство. Рассмотрим отдельно случаи суммы (разности), произведения и частного.

1°. Пусть $y(x)=u(x)\pm v(x)$. Обозначим символами Δu, Δv и Δy приращения функций $u(x)$, $v(x)$ и $y(x)$ в данной точке x, отвечающие приращению аргумента Δx. Тогда, очевидно,

$$\Delta y=y(x+\Delta x)-y(x)=[u(x+\Delta x)\pm v(x+\Delta x)]-[u(x)\pm v(x)]=$$
$$=[u(x+\Delta x)-u(x)]\pm[v(x+\Delta x)-v(x)]=\Delta u\pm\Delta v.$$

Таким образом,

$$\frac{\Delta y}{\Delta x}=\frac{\Delta u}{\Delta x}\pm\frac{\Delta v}{\Delta x}. \tag{5.25}$$

Пусть теперь $\Delta x\to 0$. Тогда в силу существования производных функций $u(x)$ и $v(x)$ в точке x существует предел правой части (5.25), равный $u'(x)\pm v'(x)$. Значит, существует предел (при $\Delta x\to 0$) и левой части (5.25). По определению производной указанный предел равен $y'(x)$, и мы приходим к требуемому равенству $y'(x)=u'(x)\pm v'(x)$.

2°. Пусть, далее, $y(x)=u(x)v(x)$. Сохраняя за Δu, Δv и Δy тот же смысл, что и выше, будем иметь

$$\Delta y=y(x+\Delta x)-y(x)=u(x+\Delta x)v(x+\Delta x)-u(x)v(x)=$$
$$=[u(x+\Delta x)v(x+\Delta x)-u(x+\Delta x)v(x)]+[u(x+\Delta x)v(x)-u(x)v(x)].$$

(Мы прибавили и вычли слагаемое $u(x+\Delta x)v(x)$.)

Далее можем записать

$$\Delta y=u(x+\Delta x)[v(x+\Delta x)-v(x)]+v(x)[u(x+\Delta x)-u(x)]=$$
$$=u(x+\Delta x)\Delta v+v(x)\Delta u.$$

Таким образом,

$$\frac{\Delta y}{\Delta x}=u(x+\Delta x)\frac{\Delta v}{\Delta x}+v(x)\frac{\Delta u}{\Delta x}. \tag{5.26}$$

Пусть теперь $\Delta x\to 0$. Тогда в силу дифференцируемости функций $u(x)$ и $v(x)$ в точке x существуют пределы отношений $\frac{\Delta u}{\Delta x}$ и $\frac{\Delta v}{\Delta x}$, соответственно равные $u'(x)$ и $v'(x)$. Далее, из дифференцируемости функции $u(x)$ в точке x в силу теоремы 5.2 следует непрерывность $u(x)$ в этой точке. Значит, существует предел $\lim\limits_{\Delta x\to 0} u(x+\Delta x)$, равный $u(x)$.

Таким образом, существует предел правой части (5.26) при $\Delta x\to 0$, равный $u(x)v'(x)+u(x)v'(x)$.

Значит, существует предел (при $\Delta x\to 0$) и левой части (5.26). По определению производной указанный предел равен $y'(x)$, и мы приходим к требуемой формуле $y'(x)=u'(x)v(x)+u(x)v'(x)$.

3°. Пусть, наконец, $y(x)=\frac{u(x)}{v(x)}$. Тогда, поскольку $v(x)\neq 0$, по теореме 4.11 об устойчивости знака непрерывной в данной точке x функции $v(x+\Delta x)\neq 0$ для всех достаточно малых Δx, и мы можем записать, что

$$\Delta y=y(x+\Delta x)-y(x)=\frac{u(x+\Delta x)}{v(x+\Delta x)}-\frac{u(x)}{v(x)}=$$
$$=\frac{u(x+\Delta x)\,v(x)-v(x+\Delta x)\,v(x)}{v(x)\,v(x+\Delta x)}.$$

Добавляя и вычитая в числителе слагаемое $u(x)v(x)$, будем иметь

$$\Delta y=\frac{[u(x+\Delta x)\,v(x)-u(x)\,v(x)]-[v(x+\Delta x)\,u(x)-u(x)\,v(x)]}{v(x)\,v(x+\Delta x)}=$$
$$=\frac{v(x)\,[u(x+\Delta x)-u(x)]-u(x)\,[v(x+\Delta x)-v(x)]}{v(x)\,v(x+\Delta x)}=\frac{v(x)\,\Delta u-u(x)\,\Delta v}{v(x)\,v(x+\Delta x)}.$$

Таким образом,

$$\frac{\Delta y}{\Delta x}=\frac{v(x)\dfrac{\Delta u}{\Delta x}-u(x)\dfrac{\Delta v}{\Delta x}}{v(x)\,v(x+\Delta x)}. \tag{5.27}$$

Пусть теперь $\Delta x\to 0$. В силу дифференцируемости (и вытекающей из нее непрерывности) функций $u(x)$ и $v(x)$ в точке x существуют пределы

$$\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}=u'(x),\ \lim_{\Delta x\to 0}\frac{\Delta v}{\Delta x}=v'(x),\ \lim_{\Delta x\to 0}v(x+\Delta x)=v(x).$$

Таким образом, поскольку $v(x)\neq 0$, существует предел при $\Delta x\to 0$ правой части (5.27), равный

$$\frac{v(x)\,u'(x)-u(x)\,v'(x)}{v^2(x)}.$$

Значит, существует предел при $\Delta x\to 0$ и левой части (5.27). По определению производной указанный предел равен $y'(x)$, и мы получим требуемую формулу

$$y'(x)=\frac{u'(x)\,v(x)-v'(x)\,u(x)}{v^2(x)}.$$

Теорема 5.5 полностью доказана.

Следствие из теоремы 5.5. *Если для функций $u(x)$ и $v(x)$ выполнены в данной точке x те же предположения, что и в теореме 5.5, то в этой точке x справедливы следующие соотношения для дифференциалов:*

$$\begin{cases} d(u \pm v) = du \pm dv, \\ d(u \cdot v) = vdu + udv, \\ d\left(\dfrac{u}{v}\right) = \dfrac{vdu - udv}{v^2}. \end{cases} \tag{5.28}$$

Для установления соотношений (5.28) достаточно умножить равенство (5.24) на dx и воспользоваться универсальным представлением (5.12) дифференциала произвольной функции $y=f(x)$.

§ 5. ПРОИЗВОДНЫЕ ПРОСТЕЙШИХ ЭЛЕМЕНТАРНЫХ ФУНКЦИЙ

Из вводной главы и из гл. 4 нам уже извество, что **простейшими элементарными функциями** принято называть следующие функции: показательную функцию $y=a^x$ и логарифмическую функцию $y=\log_a x$, рассматриваемые для любого фиксированного значения a такого, что $0<a\neq 1$, степенную функцию $y=x^\alpha$, где α — фиксированное вещественное число, четыре тригонометрические функции $y=\sin x$, $y=\cos x$, $y=\operatorname{tg} x$ и $y=\operatorname{ctg} x$ и четыре обратные тригонометрические функции $y=\arcsin x$, $y=\arccos x$, $y=\operatorname{arctg} x$ и $y=\operatorname{arcctg} x$.

В настоящем параграфе мы вычислим и систематизируем в таблицу производные всех простейших элементарных функций, уже выписанные нами в гл. 1.

1. Производные тригонометрических функций.

1°. *Производная функции* $y=\sin x$. Так как для этой функции

$$\Delta y = \sin(x+\Delta x) - \sin x = 2\cos\left(x + \frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2},$$

то при любом $\Delta x \neq 0$ разностное отношение имеет вид

$$\frac{\Delta y}{\Delta x} = \cos\left(x + \frac{\Delta x}{2}\right)\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}.$$

По определению производной

$$(\sin x)' = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\left\{\cos\left(x+\frac{\Delta x}{2}\right)\frac{\sin\frac{\Delta x}{2}}{\left(\frac{\Delta x}{2}\right)}\right\}. \tag{5.29}$$

В силу непрерывности функции $y=\cos x$ в любой точке x бесконечной прямой

$$\lim_{\Delta x\to 0}\cos\left(x+\frac{\Delta x}{2}\right) = \cos x. \tag{5.30}$$

Далее, в силу первого замечательного предела и элементарной замены переменной $t=\frac{\Delta x}{2}$ *

$$\lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}=\lim_{t\to 0}\frac{\sin t}{t}=1. \tag{5.31}$$

Из существования пределов (5.30) и (5.31) и из теоремы о пределе произведения двух функций вытекает существование предела в правой части (5.29) и равенство

$$(\sin x)'=\lim_{\Delta x\to 0}\left\{\cos\left(x+\frac{\Delta x}{2}\right)\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\right\}=\cos x.$$

Итак,

$$(\sin x)'=\cos x \tag{5.32}$$

(для любой точки x бесконечной прямой).

2°. Производная функции $y=\cos x$. Так как для любой точки x бесконечной прямой

$$\cos x=\sin\left(\frac{\pi}{2}-x\right),$$

то по правилу дифференцирования сложной функции ** и по формуле (5.32)

$$(\cos x)'=\left[\sin\left(\frac{\pi}{2}-x\right)\right]'=\cos\left(\frac{\pi}{2}-x\right)\left(\frac{\pi}{2}-x\right)'=$$
$$=\cos\left(\frac{\pi}{2}-x\right)(-1)=-\sin x.$$

Итак,

$$(\cos x)'=-\sin x \tag{5.33}$$

(для любой точки x бесконечной прямой).

3°. Производная функции $y=\operatorname{tg}x$. Так как $\operatorname{tg}x=\frac{\sin x}{\cos x}$, то в силу правила дифференцирования частного *** и соотношений (5.32) и (5.33) в любой точке x, в которой $\cos x\neq 0$,

---

* Очевидно, что $t=\frac{\Delta x}{2}\to 0$ при $\Delta x\to 0$.

** См. п. 1 § 3, формулу (5.13).

*** См. § 4, третью формулу (5.24).

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

Итак,

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x} = 1 + \operatorname{tg}^2 x \tag{5.34}$$

$\left(\text{в любой точке } x \neq \frac{\pi}{2} + n\pi, \text{ где } n = 0, \pm 1, \pm 2, \ldots\right)$.

4°. Производная функции $y = \operatorname{ctg} x$. Так как $\operatorname{ctg} x = \frac{\cos x}{\sin x}$, то в силу правила дифференцирования частного и соотношений (5.32) и (5.33) в любой точке x, в которой $\sin x \neq 0$,

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

Итак,

$$(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x} = -(1 + \operatorname{ctg}^2 x) \tag{5.35}$$

(в любой точке $x \neq \pi n$, где $n = 0, \pm 1, \pm 2, \pm 3, \ldots$).

2. Производная логарифмической функции. Пусть $y = \log_a x$, где $0 < a \neq 1$, $x > 0$ — фиксированная точка. Тогда для любого достаточно малого $\Delta x \neq 0$

$$\frac{\Delta y}{\Delta x} = \frac{\log_a (x + \Delta x) - \log_a x}{\Delta x} = \frac{1}{\Delta x} \log_a \left(1 + \frac{\Delta x}{x}\right) =$$
$$= \frac{1}{x} \frac{x}{\Delta x} \log_a \left(1 + \frac{\Delta x}{x}\right) = \frac{1}{x} \log_a \left[\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}\right].$$

По определению производной

$$(\log_a x)' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{1}{x} \lim_{\Delta x \to 0} \log_a \left[\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}\right]. \tag{5.36}$$

В силу второго замечательного предела и элементарной замены переменной $t = \frac{\Delta x}{x}$ *

$$\lim_{\Delta x \to 0} \left[\left(1 + \frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}\right] = \lim_{t \to 0} \left[(1 + t)^{\frac{1}{t}}\right] = e. \tag{5.37}$$

* Так как $x > 0$ фиксировано, то $t = \frac{\Delta x}{x} \to 0$ при $\Delta x \to 0$.

Из существования предела (5.37) и из непрерывности функции $y=\log_a x$ в точке $x=e$* вытекает, что предел в правой части (5.36) существует и равен $\frac{1}{x}\log_a e$.

Итак,

$$(\log_a x)' = \frac{1}{x}\log_a e \tag{5.38}$$

(для любых $0<a\neq 1$ и $x>0$).

В частности, при $a=e$

$$(\ln x)' = \frac{1}{x} \text{ (для любого } x>0).$$

3. Производные показательной и обратных тригонометрических функций. Для вычисления производных указанных функций используем теорему 5.4 о дифференцировании обратной функции, доказанную нами в п. 2 § 3 настоящей главы.

1°. Производная функции $y=a^x$ $(0<a\neq 1)$. Так как функция $y=a^x$, определенная на всей бесконечной прямой $-\infty<x<+\infty$, является обратной для функции $x=\log_a y$, определенной на полупрямой $0<y<+\infty$, и для функции $x=\log_a y$ в окрестности любой точки y полупрямой $0<y<+\infty$ выполнены все условия теоремы 5.4, то в силу этой теоремы функция $y=a^x$ дифференцируема в любой точке $x=\log_a y$ и для ее производной в этой точке справедлива формула

$$(a^x)' = \frac{1}{(\log_a y)'} = \frac{1}{\frac{1}{y}\log_a e} = \frac{y}{\log_a e}.$$

(Мы использовали выражение (5.38) для производной логарифмической функции.)

Из полученного равенства в силу элементарного соотношения $\frac{1}{\log_a e}=\ln a$ и соотношения $y=a^x$ окончательно получим

$$(a^x)' = a^x \ln a \tag{5.39}$$

(для любой точки x бесконечной прямой).

В частности, при $a=e$

$$(e^x)' = e^x.$$

2°. Производная функции $y=\arcsin x$. Так как функция $y=\arcsin x$, определенная на интервале $-1<x<1$, является обратной для функции $x=\sin y$, определенной на интервале $-\frac{\pi}{2}<y<$

---

* В п. 2 § 3 гл. 4 доказано, что функция $y=\log_a x$ непрерывна в любой точке $x>0$.

$<+\frac{\pi}{2}$, и для функции $x=\sin y$ в окрестности любой точки y интервала $-\frac{\pi}{2}<y<+\frac{\pi}{2}$ выполнены все условия теоремы 5.4, то по этой теореме функция $y=\arcsin x$ дифференцируема в любой точке $x=\sin y$ и для ее производной в этой точке справедлива формула

$$(\arcsin x)'=\frac{1}{(\sin y)'}=\frac{1}{\cos y}=\frac{1}{\sqrt{1-\sin^2 y}}. \qquad (5.40)$$

Мы использовали равенство (5.32) и взяли перед корнем знак + в силу того, что $\cos y$ положителен всюду на интервале $-\frac{\pi}{2}<$ $<y<+\frac{\pi}{2}$.

Учитывая, что $\sin y=x$, мы окончательно получим из (5.40)

$$(\arcsin x)'=\frac{1}{\sqrt{1-x^2}}$$

(для всех x из интервала $-1<x<+1$).

3°. Производная функции $y=\arccos x$. Так как функция $y=\arccos x$, определенная на интервале $-1<x<+1$, является обратной для функции $x=\cos y$, определенной на интервале $0<y<\pi$, и для функции $x=\cos y$ в окрестности любой точки y интервала $0<y<\pi$ выполнены все условия теоремы 5.4, то по этой теореме функция $y=\arccos x$ дифференцируема в любой точке $x=\cos y$ и для ее производной в этой точке справедлива формула

$$(\arccos x)'=\frac{1}{(\cos y)'}=\frac{1}{(-\sin y)}=\frac{1}{-\sqrt{1-\cos^2 y}} \qquad (5.41)$$

Мы использовали равенство (5.33) и взяли перед корнем знак — в силу того, что $\sin y$ положителен всюду на интервале $0<y<\pi$.

Учитывая, что $\cos y=x$, мы окончательно получим из (5.41)

$$(\arccos x)'=-\frac{1}{\sqrt{1-x^2}}$$

(для всех x из интервала $-1<x<+1$).

4°. Производная функции $y=\operatorname{arctg} x$. Так как функция $y=\operatorname{arctg} x$, определенная на бесконечной прямой $-\infty<x<+\infty$, является обратной для функции $x=\operatorname{tg} y$, определенной на интервале $-\frac{\pi}{2}<y<\frac{\pi}{2}$, и для функции $x=\operatorname{tg} y$ в окрестности каждой точки у интервала $-\frac{\pi}{2}<y<\frac{\pi}{2}$ выполнены все условия теоремы 5.4, то по этой теореме функция $y=\operatorname{arctg} x$ дифференцируема в каждой точке $x=\operatorname{tg} y$ и для ее производной в этой точке справедлива формула

$$(\operatorname{arctg} x)' = \frac{1}{(\operatorname{tg} y)'} = \frac{1}{1+\operatorname{tg}^2 y} = \frac{1}{1+x^2}.$$

(Мы использовали соотношение (5.34).)

Итак,

$$(\operatorname{arctg} x)' = \frac{1}{1+x^2}$$

(для любой точки x бесконечной прямой).

5°. Производная функции $y=\operatorname{arcctg} x$. Так как функция $y=\operatorname{arcctg} x$, определенная на бесконечной прямой $-\infty<x<+\infty$ является обратной для функции $x=\operatorname{ctg} y$, определенной на интервале $0<y<\pi$, и для функции $x=\operatorname{ctg} y$ в окрестности каждой точки интервала $0<y<\pi$ выполнены все условия теоремы 5.4, то по этой теореме функция $y=\operatorname{arcctg} x$ дифференцируема в каждой точке $x=\operatorname{ctg} y$ и для ее производной в этой точке справедлива формула

$$(\operatorname{arcctg} x)' = \frac{1}{(\operatorname{ctg} y)'} = \frac{1}{-(1+\operatorname{ctg}^2 y)} = -\frac{1}{1+x^2}.$$

(Мы использовали соотношение (5.35).)

Итак,

$$(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}$$

(для любой точки x бесконечной прямой).

4. Производная степенной функции. Пусть $y=x^\alpha$, где α — любое вещественное число, x — любая точка полупрямой $0<x<+\infty$. В гл. 4 мы уже рассматривали степенную функцию $y=x^\alpha$ как суперпозицию логарифмической и показательной функций

$$y = x^\alpha = (a^{\log_a x})^\alpha = a^{\alpha \log_a x}$$

(где a — любое фиксированное число $0<a\neq 1$).

По правилу дифференцирования сложной функции $y=a^u$, где $u=\alpha\log_a x$, получим

$$y' = (a^u)' \cdot (\alpha \log_a x)' = a^u \cdot \ln a \cdot \alpha \cdot \frac{1}{x} \log_a e =$$

$$= a^{\alpha \log_a x} \cdot \alpha \cdot \frac{1}{x} = x^\alpha \cdot \alpha \cdot \frac{1}{x} = \alpha x^{\alpha-1}.$$

Итак, окончательно

$$(x^\alpha)' = \alpha x^{\alpha-1}$$

(для любого $x>0$).

5. Таблица производных простейших элементарных функций. Соберем теперь в таблицу все вычисленные нами производные простейших элементарных функций.

1°. $(x^\alpha)' = \alpha x^{\alpha-1}$ $(x>0)$.

В частности, $\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$, $(\sqrt{x})' = \frac{1}{2\sqrt{x}}$.

2°. $(\log_a x)' = \frac{1}{x}\log_a e \quad (0 < a \neq 1,\ x > 0)$.

В частности, $(\ln x)' = \frac{1}{x}$.

3°. $(a^x)' = a^x \cdot \ln a \quad (0 < a \neq 1)$.

В частности, $(e^x)' = e^x$.

4°. $(\sin x)' = \cos x$.

5°. $(\cos x)' = -\sin x$.

6°. $(\operatorname{tg} x)' = \frac{1}{\cos^2 x} = 1 + \operatorname{tg}^2 x \qquad \left(x \neq \frac{\pi}{2} + n\pi;\ n = 0, \pm 1, \ldots\right)$.

7°. $(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x} = -(1 + \operatorname{ctg}^2 x) \quad (x \neq n\pi;\ n = 0, \pm 1, \ldots)$.

8°. $(\arcsin x)' = 1/\sqrt{1-x^2} \quad (-1 < x < +1)$.

9°. $(\arccos x)' = -1/\sqrt{1-x^2} \quad (-1 < x < +1)$.

10°. $(\operatorname{arctg} x)' = \frac{1}{1+x^2}$.

11°. $(\operatorname{arcctg} x)' = -\frac{1}{1+x^2}$.

В гл. 4 мы ввели в рассмотрение четыре гиперболические функции $y = \operatorname{sh} x$, $y = \operatorname{ch} x$, $y = \operatorname{th} x$ и $y = \operatorname{cth} x$, являющиеся простыми комбинациями показательных функций. Из представления этих функций через показательные функции элементарно вытекают следующие выражения для их производных:

12°. $(\operatorname{sh} x)' = \operatorname{ch} x$.

13°. $(\operatorname{ch} x)' = \operatorname{sh} x$.

14°. $(\operatorname{th} x)' = \frac{1}{\operatorname{ch}^2 x}$.

15°. $(\operatorname{cth} x)' = -\frac{1}{\operatorname{sh}^2 x} \quad (x \neq 0)$.

Установленная нами таблица производных вместе с правилом дифференцирования сложной функции (установленным в п. 1 § 3) и правилами дифференцирования суммы, разности, произведения и частного (установленными в § 4) составляет вычислительный аппарат той части математического анализа, которую принято называть **дифференциальным исчислением**.

В гл. 1 и 4 мы уже ввели понятие элементарной функции, как такой функции, которая выражается через простейшие элементарные функции посредством четырех арифметических действий и суперпозиций, последовательно применяемых конечное число раз.

Из установленной нами таблицы производных и правил дифференцирования суммы, разности, произведения, частного и сложной функции вытекает следующий важный вывод: *производная любой элементарной функции представляет собой также элементарную* *функцию, т. е. операция дифференцирования не выводит нас из класса элементарных функций.*

6. Таблица дифференциалов простейших элементарных функций. В п. 3 § 3 мы установили, что дифференциал dy функции $y=f(x)$ всегда равен производной этой функции $f'(x)$, умноженной на дифференциал аргумента dx. Поэтому установленная нами таблица производных сразу же приводит нас к соответствующей таблице дифференциалов простейших элементарных функций.

1°. $d(x^\alpha)=\alpha\cdot x^{\alpha-1}dx$.

В частности, $d\left(\frac{1}{x}\right)=-\frac{1}{x^2}dx$, $d(\sqrt{x})=\frac{1}{2\sqrt{x}}dx$.

2°. $d(\log_a x)=\frac{1}{x}\log_a e\,dx \qquad (0<a\neq 1,\ x>0)$.

В частности, $d(\ln x)=\frac{dx}{x}$.

3°. $d(a^x)=a^x\ln a dx \qquad (0<a\neq 1)$.

В частности, $d(e^x)=e^x dx$.

4°. $d(\sin x)=\cos x dx$.

5°. $d(\cos x)=-\sin x dx$.

6°. $d(\operatorname{tg} x)=\frac{dx}{\cos^2 x}=(1+\operatorname{tg}^2 x)\,dx$
$\left(x\neq\frac{\pi}{2}+\pi n;\ n=0,\ \pm 1,\ \pm 2,\ \ldots\right)$.

7°. $d(\operatorname{ctg} x)=-\frac{dx}{\sin^2 x}=-(1+\operatorname{ctg}^2 x)\,dx \quad (x\neq n\pi;\ n=0,\ \pm 1,\ldots)$

8°. $d(\arcsin x)=dx/\sqrt{1-x^2} \quad (-1<x<+1)$.

9°. $d(\arccos x)=-dx/\sqrt{1-x^2} \quad (-1<x<+1)$.

10°. $d(\operatorname{arctg} x)=\frac{dx}{1+x^2}$.

11°. $d(\operatorname{arctg} x)=-\frac{dx}{1+x^2}$.

7. Логарифмическая производная. Производная степенно-показательной функции. Пусть функция $y=f(x)$ положительна и дифференцируема в данной точке x. Тогда и сложная функция аргумента x вида $w=\ln y$, где $y=f(x)$, в силу теоремы 5.3 будет также дифференцируема в указанной точке x, причем для производной этой сложной функции по аргументу x будет справедлива формула

$$[\ln f(x)]'=(\ln y)'\,y'=\frac{y'}{y}=\frac{f'(x)}{f(x)}. \tag{5.42}$$

Величину (5.42) принято называть логарифмической производной функции $y=f(x)$ в данной точке x.

Логарифмическая производная может быть использована для вычисления производных некоторых функций, не являющихся простейшими элементарными.

В качестве примера вычислим производную так называемой степенно-показательной функции, т. е. функции вида $y=u(x)^{v(x)}$, где $u(x)$ и $v(x)$ — две функции, дифференцируемые в данной точке x, первая из которых $u(x)$ строго положительна в этой точке.

При таких ограничениях функция $w=\ln y=v(x)\ln u(x)$ будет дифференцируема в данной точке x. В самом деле, в силу теоремы 5.3 функция $\ln u(x)$ дифференцируема в точке x. Значит, на основании теоремы о дифференцируемости произведения двух дифференцируемых функций можно утверждать дифференцируемость в данной точке x и функции $w=\ln y=v(x)\cdot\ln u(x)$, причем в силу второй формулы (5.24)

$$(\ln y)'=v'(x)\ln u(x)+v(x)[\ln u(x)]'=$$
$$=v'(x)\ln u(x)+v(x)\frac{u'(x)}{u(x)}. \tag{5.43}$$

Из (5.42) и (5.43) получим, что

$$\frac{y'}{y}=v'(x)\ln u(x)+v(x)\frac{u'(x)}{u(x)}.$$

Учитывая, что $y=u(x)^{v(x)}$, окончательно получим следующее выражение для производной степенно-показательной функции:

$$\{u(x)^{v(x)}\}'=u(x)^{v(x)}\left[v'(x)\ln u(x)+v(x)\frac{u'(x)}{u(x)}\right]. \tag{5.44}$$

Формула (5.44) справедлива в предположении, что $u(x)$ и $v(x)$ дифференцируемы в данной точке x и $u(x)>0$ в этой точке.

§ 6. ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИАЛЫ ВЫСШИХ ПОРЯДКОВ

1. Понятие производной n-го порядка. Как уже отмечалось в п. 2 § 1, производная $f'(x)$ функции $y=f(x)$, определенной и дифференцируемой на интервале (a, b), *представляет собой функцию, также определенную на интервале* (a, b). Может случиться, что функция $f'(x)$ сама является дифференцируемой в некоторой точке x интервала (a, b), т. е. имеет в этой точке производную. Тогда указанную производную называют второй производной (или производной второго порядка) функции $y=f(x)$ в точке x и обозначают символом $f^{(2)}(x)$ или $y^{(2)}$ *.

После того как введено понятие второй производной, можно последовательно ввести понятие третьей производной, затем четвертой производной и т. д. Если предположить, что нами уже введено понятие $(n-1)$-й производной и что $(n-1)$-я производная дифференцируема в некоторой точке x интервала (a, b), т. е.

* Для обозначения второй производной используют также символы f'' или y''.

имеет в этой точке производную, то указанную производную называют n-й производной (или производной n-го порядка) функции $y=f(x)$ в точке x и обозначают символом $f^{(n)}(x)$ или $y^{(n)}$.

Таким образом, мы вводим понятие n-й производной индуктивно, переходя от первой производной к последующим. Соотношение, определяющее n-ю производную, имеет вид

$$y^{(n)}=[y^{(n-1)}]'. \tag{5.45}$$

Функцию, имеющую на данном множестве $\{x\}$ конечную производную порядка n, обычно называют n раз дифференцируемой на данном множестве.

Понятие производных высших порядков находит многочисленные применения. Здесь мы ограничимся тем, что укажем механический смысл второй производной. Если функция $y=f(x)$ описывает закон движения материальной точки вдоль оси Oy, то, как мы уже знаем из гл. 1, первая производная $f'(x)$ дает мгновенную скорость движущейся точки в момент времени x. В таком случае вторая производная $f^{(2)}(x)$ равна скорости изменения скорости, т. е. равна ускорению движущейся точки в момент времени x.

Методика вычисления производных высшего порядка предполагает умение вычислять только производные первого порядка, ибо последовательное применение формулы (5.45) есть не что иное, как многократное вычисление первых производных. В качестве примеров вычислим производные n-го порядка некоторых простейших элементарных функций.

2. n-ые производные некоторых функций.

1°. Вычислим n-ю производную степенной функции $y=x^\alpha$ ($x>0$, α — любое вещественное число). Последовательно дифференцируя, будем иметь

$$y'=\alpha x^{\alpha-1},\ y^{(2)}=\alpha(\alpha-1)x^{\alpha-2},\ y^{(3)}=\alpha(\alpha-1)(\alpha-2)x^{\alpha-3},\ \ldots .$$

Отсюда легко уяснить общий закон

$$\boxed{(x^\alpha)^{(n)}=\alpha(\alpha-1)\ldots(\alpha-n+1)x^{\alpha-n}.}$$

Строгое доказательство этого закона легко проводится методом математической индукции.

В частном случае $\alpha=m$, где m — натуральное число, получим

$$(x^m)^{(m)}=m!,$$

$$(x^m)^{(n)}=0 \text{ при } n>m.$$

Таким образом, n-я производная многочлена m-го порядка при $n>m$ равна нулю*.

* При этом мы используем следующую очевидную формулу: $[Au(x)+Bv(x)]^{n}=Au^{(n)}(x)+Bv^{(n)}(x)$, где A и B — постоянные.

2°. Далее вычислим n-ю производную показательной функции $y=a^x$. Последовательно дифференцируя, будем иметь

$$y'=a^x \ln x,\ y^{(2)}=a^x \ln^2 x,\ y^{(3)}=a^x \ln^3 a, \dots .$$

Общая формула, легко устанавливаемая по методу индукции, имеет вид

$$\boxed{(a^x)^{(n)}=a^x \ln^n a.}$$

В частности,

$$\boxed{(e^x)^{(n)}=e^x.}$$

3°. Вычислим n-ю производную функции $y=\sin x$.

Первую производную этой функции можно записать в виде $y'=\cos x=\sin\left(x+\frac{\pi}{2}\right)$. Таким образом, дифференцирование функции $y=\sin x$ прибавляет к аргументу этой функции величину $\frac{\pi}{2}$. Отсюда получаем формулу

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

4°. Совершенно аналогично устанавливается формула

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

5°. Вычислим n-ю производную функции $y=\operatorname{arctg} x$. Докажем с помощью метода математической индукции, что справедлива следующая формула:

$$(\operatorname{arctg} x)^{(n)}=\frac{(n-1)!}{(1+x^2)^{n/2}}\sin\left[n\left(\operatorname{arctg} x+\frac{\pi}{2}\right)\right]. \tag{5.46}$$

Учитывая, что $y=\operatorname{arctg} x$, $x=\operatorname{tg} y$, $\frac{1}{1+x^2}=\frac{1}{1+\operatorname{tg}^2 y}=\cos^2 y$, мы можем переписать устанавливаемую формулу в виде

$$y^{(n)}=(n-1)!\cos^n y \sin\left[n\left(y+\frac{\pi}{2}\right)\right]. \tag{5.46*}$$

Убедимся методом индукции в справедливости формулы (5.46*). При $n=1$, в силу п. 3 § 5, $y'=(\operatorname{arctg} x)'=\frac{1}{1+x^2}=\cos^2 y$. То же самое выражение получается при $n=1$ из (5.46*) (достаточно учесть, что $\sin\left(y+\frac{\pi}{2}\right)=\cos y$).

Таким образом, при $n=1$ формула (5.46*) справедлива.

Предположим теперь, что формула (5.46*) справедлива для некоторого n, и убедимся, что в таком случае эта формула справедлива и для следующего номера $n+1$.

В самом деле, производя дифференцирование (5.46*), получим

$$y^{(n+1)} = (n-1)!\,\frac{d}{dx}\left\{\cos^n y \sin\left[n\left(y+\frac{\pi}{2}\right)\right]\right\} =$$
$$= (n-1)!\,\frac{d}{dy}\left\{\cos^n y \sin\left[n\left(y+\frac{\pi}{2}\right)\right]\right\}\frac{dy}{dx} =$$
$$= (n-1)!\left\{\frac{d}{dy}(\cos^n y)\sin\left[n\left(y+\frac{\pi}{2}\right)\right] +\right.$$
$$\left. + \cos^n y\,\frac{d}{dy}\left[\sin n\left(y+\frac{\pi}{2}\right)\right]\right\} y'.$$

Учитывая, что $y' = \cos^2 y$, $\frac{d}{dy}(\cos^n y) = n\cos^{n-1} y(-\sin y)$, мы получим, что

$$y^{(n+1)} = n!\cos^{n+1} y\left\{-\sin y \sin\left[n\left(y+\frac{\pi}{2}\right)\right] + \cos y \cos\left[n\left(y+\frac{\pi}{2}\right)\right]\right\} =$$
$$= n!\cos^{n+1} y \cos\left[(n+1)y + n\frac{\pi}{2}\right] = n!\cos^{n+1} y \sin\left[(n+1)\left(y+\frac{\pi}{2}\right)\right].$$

Мы получаем для $y^{(n+1)}$ формулу вида (5.46*), взятую для номера $n+1$. Тем самым индукция завершена и формула (5.46) доказана.

6°. В заключение вычислим n-ю производную так называемой *дробно-линейной функции* $y = \frac{ax+b}{cx+d}$, где a, b, c и d — некоторые постоянные.

Последовательно дифференцируя эту функцию, будем иметь

$$y' = \frac{a(cx+d) - c(ax+b)}{(cx+d)^2} = (ad-bc)(cx+d)^{-2},$$
$$y^{(2)} = (ad-bc)(-2)(cx+d)^{-3}c,$$
$$y^{(3)} = (ad-bc)(-2)(-3)(cx+d)^{-4}c^2, \ \ldots\ .$$

Легко усмотреть и общий закон

$$\boxed{y^{(n)} = \left(\frac{ax+b}{cx+d}\right)^{(n)} = (ad-bc)(-1)^{n-1}\,n!\,(cx+d)^{-(n+1)}\,c^{n-1},}$$

который доказывается по методу индукции.

3. **Формула Лейбница для n-й производной произведения двух функций.** В то время как установленное выше правило вычисления

первой производной от суммы или разности двух функций $(u \pm v)' = u' \pm v'$ легко переносится (например, по методу индукции) на случай n-й производной $(u \pm v)^{(n)} = u^{(n)} \pm v^{(n)}$, возникают затруднения при вычислении n-й производной от произведения двух функций u и v.

Соответствующее правило носит название формулы Лейбница и имеет следующий вид:

$$\boxed{(uv)^{(n)} = u^{(n)}v + C_n^1 u^{(n-1)}v' + C_n^2 u^{(n-2)}v^{(2)} + C_n^3 u^{(n-3)}v^{(3)} + \ldots + uv^{(n)}.} \tag{5.47}$$

Легко подметить закон, по которому построена правая часть формулы Лейбница (5.47): она совпадает с формулой разложения бинома $(u+v)^n$, лишь вместо степеней u и v стоят производные соответствующих порядков.

Это сходство становится еще более полным, если вместо самих функций u и v писать соответственно $u^{(0)}$ и $v^{(0)}$ (т. е. если рассматривать саму функцию как производную нулевого порядка).

Докажем формулу Лейбница по методу индукции. При $n=1$ эта формула принимает вид $(uv)' = u'v + uv'$, что совпадает с установленным выше (в § 4) правилом дифференцирования произведения двух функций. Поэтому достаточно, предположив справедливость формулы (5.47) для некоторого номера n, доказать ее справедливость для следующего номера $n+1$.

Итак, пусть для некоторого номера n формула (5.47) верна. Продифференцируем эту формулу и объединим слагаемые, стоящие в правой части, так, как это указано ниже:

$$(uv)^{(n+1)} = u^{(n+1)}v + [C_n^0 u^{(n)}v' + C_n^1 u^{(n)}v'] + [C_n^1 u^{(n-1)}v^{(2)} + C_n^2 u^{(n-1)}v^{(2)}] +$$
$$+ [C_n^2 u^{(n-2)}v^{(3)} + C_n^3 u^{(n-2)}v^{(3)}] + \ldots + uv^{(n+1)}. \tag{5.48}$$

(При этом мы воспользовались тем, что $1 = C_n^0$.)

Легко проверить, что для любого номера k, не превосходящего n, справедлива формула

$$C_n^k + C_n^{k-1} = C_{n+1}^k. \tag{5.49}$$

Для того чтобы убедиться в справедливости формулы (5.49), достаточно заметить, что

$$C_n^k = \frac{n(n-1)\ldots(n-k+1)}{k!}, \quad C_n^{k-1} = \frac{n(n-1)\ldots(n-k+2)}{(k-1)!},$$
$$C_{n+1}^k = \frac{(n+1)n\ldots(n-k+2)}{k!}.$$

Из написанных соотношений вытекает, что

$$C_n^k + C_n^{k-1} = \frac{n(n-1)\ldots(n-k+1)}{k!} + \frac{n(n-1)\ldots(n-k+2)}{(k-1)!} =$$

$$=\frac{n(n-1)\ldots(n-k+2)(n-k+1)+n(n-1)\ldots(n-k+2)k}{k!}=$$

$$=\frac{n(n-1)\ldots(n-k+2)[(n-k+1)+k]}{k!}=$$

$$=\frac{(n+1)n(n-1)\ldots(n-k+2)}{k!}=C_{n+1}^{k}.$$

Используя формулу (5.49), мы можем следующим образом переписать соотношение (5.48):

$$(uv)^{(n+1)}=u^{(n+1)}v+C_{n+1}^{1}u^{(n)}v'+C_{n+1}^{2}u^{(n-1)}v^{(2)}+C_{n+1}^{3}u^{(n-2)}v^{(3)}+\ldots+uv^{(n+1)}.$$

Тем самым мы убедились в справедливости формулы (5.47) для номера $n+1$. Индукция проведена, и вывод формулы Лейбница (5.47) завершен.

Пример. Вычислим n-ю производную функции $y=x^2e^x$. Полагая в формуле Лейбница (5.47) $u=e^x$, $v=x^2$ и учитывая, что $u^{(k)}=e^x$ (для любого номера k), $v'=2x$, $v^{(2)}=2$, $v^{(3)}=v^{(4)}=\ldots=0$, мы получим, что

$$(x^2e^x)^{(n)}=e^xx^2+C_n^1e^x\,2x+C_2^n\,e^x\cdot 2=e^x[x^2+2nx+n(n-1)].$$

Подчеркнем, что формула Лейбница особенно эффективна в том случае, когда одна из перемножаемых функций имеет лишь конечное число отличных от нуля производных и не представляет затруднения вычисление всех производных другой из перемножаемых функций.

4. Дифференциалы высших порядков. Выше для обозначения дифференциала аргумента и соответствующего ему дифференциала функции мы использовали символы dx и соответственно dy. В рассуждениях настоящего пункта нам придется использовать для обозначения дифференциала аргумента и соответствующего ему дифференциала функции и другие символы. В частности, мы будем обозначать дифференциал аргумента и соответствующий ему дифференциал функции символами δx и δy соответственно. В этих обозначениях инвариантное по форме выражение для первого дифференциала функции $y=f(x)$ будет иметь вид $\delta y=f'(x)\,\delta x$.

Рассмотрим выражение для первого дифференциала дифференцируемой в данной точке x функции $y=f(x)$:

$$dy=f'(x)\,dx. \tag{5.50}$$

Предположим, что величина, стоящая в правой части (5.50), представляет собой функцию аргумента x, дифференцируемую в данной точке x. Для этого достаточно потребовать, чтобы функция $y=f(x)$ была два раза дифференцируема в данной точке x, а аргумент либо являлся независимой переменной, либо представлял собой дважды дифференцируемую функцию некоторой независимой переменной t.

При этих предположениях мы можем рассмотреть дифференциал

$$\delta(dy)=\delta[f'(x)dx]$$

от величины (5.50).

Определение 1. *Значение $\delta(dy)$ дифференциала от первого дифференциала* (5.50), *взятое при $\delta x=dx$, называется* **вторым** *дифференциалом функции $y=f(x)$ (в данной точке x) и обозначается символом d^2y.*

Итак, по определению*

$$d^2y=\delta(dy)|_{\delta x=dx}=\{\delta[f'(x)dx]\}|_{\delta x=dx}.$$

Дифференциал d^ny любого порядка n вводится по индукции. Предположим, что уже введен дифференциал $d^{n-1}y$ порядка $n-1$ и что функция $y=f(x)$ n раз дифференцируема в данной точке x, а ее аргумент x либо является независимой переменной, либо представляет собой n раз дифференцируемую функцию некоторой независимой переменной t.

Определение 2. *Значение $\delta(d^{n-1}y)$ дифференциала от $(n-1)$-го дифференциала $d^{n-1}y$, взятое при $\delta x=dx$, называется* n-м **дифференциалом** *функции $y=f(x)$ (в данной точке x) и обозначается символом d^ny.*

Итак, по определению $d^ny=\delta(d^{n-1}y)|_{\delta x=dx}$.

При вычислении второго и последующих дифференциалов приходится существенно различать два случая: 1) случай, когда аргумент x является независимой переменной; 2) случай, когда аргумент x представляет собой соответствующее число раз дифференцируемую функцию некоторой независимой переменной t.

В первом случае, когда x является независимой переменной, мы имеем право считать, что dx не зависит от x и равен одному и тому же приращению аргумента Δx (для всех точек x). При этом мы получим, что $\delta(dx)=(dx)'\delta x=0$.

Последнее соотношение и второе соотношение (5.28) позволяют нам записать следующую цепочку равенств:

$$\begin{aligned}d^2y&=\delta(dy)|_{\delta x=dx}=\{\delta[f'(x)dx]\}|_{\delta x=dx}=\\&=\{\delta[f'(x)]dx+f'(x)\delta(dx)\}|_{\delta x=dx}=\{\delta[f'(x)]dx\}|_{\delta x=dx}=\\&=\{f^{(2)}(x)\delta x dx\}|_{\delta x=dx}=f^{(2)}(x)(dx)^2.\end{aligned}\tag{5.51}$$

Итак, в случае, когда аргумент x является независимой переменной, для второго дифференциала функции $y=f(x)$ справедливо представление

$$d^2y=f^{(2)}(x)(dx)^2.\tag{5.52}$$

---

* Символ $\{...\}|_{\delta x=dx}$ означает, что в выражении, заключенном в фигурные скобки, следует положить $\delta x=dx$.

Совершенно аналогично, по индукции легко убедиться в том, что в случае, когда аргумент x является независимой переменной, для n-го дифференциала n раз дифференцируемой функции $y=f(x)$ справедливо представление

$$d^n y = f^{(n)}(x)\,(dx)^n.$$

Таким образом, в случае, когда аргумент x является независимой переменной, производная порядка n функции $y=f(x)$ равна отношению n-го дифференциала этой функции $d^n y$ к n-й степени дифференциала аргумента dx.

Совсем другой вид имеют представления для второго и последующих дифференциалов в случае, когда аргумент x является соответствующее число раз дифференцируемой функцией некоторой независимой переменной t.

Установим выражение для второго дифференциала, считая, что функция $y=f(x)$ два раза дифференцируема в данной точке x, а ее аргумент x является два раза дифференцируемой функцией некоторой независимой переменной t.

Повторяя рассуждения из цепочки (5.51), мы на этот раз получим

$$\begin{aligned} d^2 y &= \delta(dy)\,|_{\delta x = dx} = \{\delta[f'(x)\,dx]\}\,|_{\delta x = dx} = \\ &= \{\delta[f'(x)]\,dx + f'(x)\,\delta(dx)\}\,|_{\delta x = dx} = \\ &= \{f^{(2)}(x)\,\delta x dx\}\,|_{\delta x = dx} + \{f'(x)\,\delta(dx)\}\,|_{\delta x = dx}. \end{aligned}$$

Заметим, что в силу определения второго дифференциала функции $y=x$

$$\delta(dx)\,|_{\delta x = dx} = d^2 x.$$

Учитывая это соотношение, мы приходим к следующему представлению:

$$d^2 y = f^{(2)}(x)\,(dx)^2 + f'(x)\,d^2 x. \qquad (5.53)$$

Сравнивая полученное представление (5.53) с представлением (5.52), мы убедимся в том, что (в отличие от первого дифференциала) второй дифференциал уже не обладает свойством инвариантности формы.

Тем более не обладают свойством инвариантности формы последующие дифференциалы.

§ 7. ДИФФЕРЕНЦИРОВАНИЕ ФУНКЦИИ, ЗАДАННОЙ ПАРАМЕТРИЧЕСКИ

Мы будем говорить, что переменная y как функция аргумента x з а д а н а н а м п а р а м е т р и ч е с к и, если обе переменные x и y заданы как функции некоторой третьей переменной t: $x=\varphi(t)$, $y=\psi(t)$.

При этом указанную переменную t обычно называют **параметром**.

Мы, естественно, будем предполагать, что функции $\varphi(t)$ и $\psi(t)$ имеют нужное число производных по параметру t в рассматриваемой области изменения этого параметра.

Кроме того, мы всегда будем считать, что функция $x=\varphi(t)$ в окрестности рассматриваемой точки $t=t_0$ имеет обратную функцию $t=\varphi^{-1}(x)$*, ибо это позволяет рассматривать y как функцию x.

Рассмотрим вопрос о вычислении производных функции $y=y(x)$ по аргументу x.

В силу свойства инвариантности формы первого дифференциала (см. п. 3 § 3) мы можем записать равенства:

$$y'(x)=\frac{dy}{dx},\quad dy=\psi'(t)\,dt,\quad dx=\varphi'(t)\,dt. \tag{5.54}$$

Из этих равенств сразу же вытекает, что

$$y'(x)=\frac{\psi'(t)}{\varphi'(t)}. \tag{5.55}$$

Для вычисления второй производной $y^{(2)}(x)$ достаточно заметить, что в силу свойства инвариантности формы первого дифференциала

$$y^{(2)}(x)=\frac{d\,[y'(x)]}{dx}. \tag{5.56}$$

Используя в правой части (5.56) соотношение (5.55) и третье равенство (5.54), получим

$$y^{(2)}(x)=\frac{\left[\dfrac{\psi'(t)}{\varphi'(t)}\right]' dt}{\varphi'(t)dt}=\frac{\psi^{(2)}(t)\,\varphi'(t)-\varphi^{(2)}(t)\psi'(t)}{[\varphi'(t)]^2}. \tag{5.57}$$

По такому же принципу вычисляются производные третьего и последующих порядков. Так, для вычисления третьей производной $y^{(3)}(x)$ достаточно представить эту производную в виде

$$y^{(3)}(x)=\frac{d\,[y^{(2)}(y)]}{dx}$$

и воспользоваться соотношением (5.57) и третьим равенством (5.54).

В качестве примера вычислим первую и вторую производные следующей функции, заданной параметрически:

$$\begin{cases} x=a(t-\sin t), \\ y=a(1-\cos t), \end{cases} \quad -\infty<t<+\infty.$$

* В гл. 14 будет доказано, что для существования у функции $x=\varphi(t)$ в окрестности точки $t=t_0$ обратной функции достаточно, чтобы производная $\varphi'(t)$ была непрерывна и отлична от нуля в точке t_0.

Кривая, являющаяся графиком этой функции, называется циклоидой. Эта кривая представляет собой траекторию некоторой фиксированной точки окружности радиуса a, которая катится без скольжения по прямой линии (параметр t равен углу поворота радиуса этой окружности).

Пользуясь формулами (5.55) и (5.57), получим

$$y'(x) = \frac{a \sin t}{a(1-\cos t)} = \operatorname{ctg}\frac{t}{2},$$

$$y^{(2)}(x) = \frac{\left(\operatorname{ctg}\frac{t}{2}\right)'}{a(1-\cos t)} = \frac{-1}{4a\sin^4\frac{t}{2}}.$$

§ 8. ПРОИЗВОДНАЯ ВЕКТОРНОЙ ФУНКЦИИ

В приложениях математики часто встречаются понятия векторной функции и ее производной.

Если каждому значению переменной t из некоторого множества $\{t\}$ ставится в соответствие по известному закону определенный вектор a, то говорят, что на множестве $\{t\}$ задана **векторная функция** $\mathbf{a}=\mathbf{a}(t)$.

Так как каждый вектор $\mathbf{a}$ в заданной декартовой прямоугольной системе координат однозначно определяется тремя координатами x, y и z, то задание векторной функции $\mathbf{a}=\mathbf{a}(t)$ эквивалентно заданию трех скалярных функций $x=x(t)$, $y=y(t)$ и $z=z(t)$.

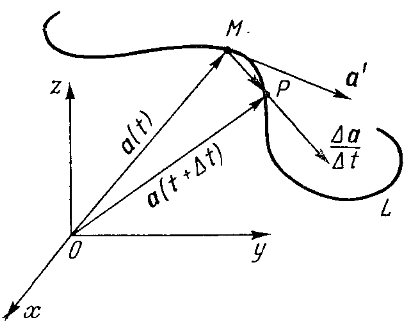


Рис. 5.4

Понятие векторной функции приобретает особую наглядность при обращении к так называемому **годографу** этой функции (т. е. геометрическому месту концов всех векторов $\mathbf{a}=\mathbf{a}(t)$, приложенных к началу координат O). На рис. 5.4 кривая L представляет собой годограф векторной функции $\mathbf{a}=\mathbf{a}(t)$.

Понятие годографа векторной функции является естественным обобщением понятия графика скалярной функции.

Введем понятие производной векторной функции $\mathbf{a}=\mathbf{a}(t)$ в данной фиксированной точке t.

Зададим аргументу в точке t произвольное приращение $\Delta t \neq 0$ и рассмотрим соответствующий вектор приращения $\Delta\mathbf{a}=\mathbf{a}(t+\Delta t)-\mathbf{a}(t)$. (На рис. 5.4 указанный вектор совпадает с MP.)

Умножив указанный вектор на число $\frac{1}{\Delta t}$, мы получим новый вектор

$$\frac{\Delta \mathbf{a}}{\Delta t} = \frac{1}{\Delta t}\left[\mathbf{a}(t+\Delta t) - \mathbf{a}(t)\right], \tag{5.58}$$

коллинеарный прежнему. Этот вектор (5.58) является аналогом разностного отношения (5.5).

Вектор (5.58), очевидно, представляет собой среднюю скорость изменения векторной функции на сегменте $[t, t+\Delta t]$.

Производной векторной функции $\mathbf{a}=\mathbf{a}(t)$ в данной фиксированной точке t называется предел при $\Delta t \to 0$ вектора (5.58) (при условии, что этот предел существует).

Для обозначения производной векторной функции $\mathbf{a}=\mathbf{a}(t)$ используются символы $\mathbf{a}'(t)$ или $\frac{d\mathbf{a}}{dt}$.

Из геометрических соображений очевидно, что производная векторной функции $\mathbf{a}=\mathbf{a}(t)$ представляет собой вектор, касательный к годографу этой функции.

Так как координаты разностного отношения (5.58) соответственно равны

$$\frac{x(t+\Delta t)-x(t)}{\Delta t}, \quad \frac{y(t+\Delta t)-y(t)}{\Delta t}, \quad \frac{z(t+\Delta t)-z(t)}{\Delta t},$$

то ясно, что координаты производной $\mathbf{a}'(t)$ равны производным функций $x'(t)$, $y'(t)$, $z'(t)$. Таким образом, вычисление производной векторной функции сводится к вычислению производных ее координат.

Замечание 1. Так как векторная функция $\mathbf{a}=\mathbf{a}(t)$ определяет закон движения материальной точки по кривой L, представляющей собой годограф этой функции, то производная $\mathbf{a}'(t)$ равна скорости движения по указанной кривой.

Замечание 2. Из курса аналитической геометрии известны различные типы произведений векторов (скалярное произведение, векторное произведение и смешанное произведение). Выражение всех этих произведений в координатах дает возможность указать правила, по которым вычисляются производные соответствующих произведений векторных функций. В качестве примера приведем правило вычисления производной скалярного произведения двух векторных функций $\mathbf{a}(t)=(a_1(t), a_2(t), a_3(t))$ и $\mathbf{b}(t)=(b_1(t), b_2(t), b_3(t))$: $\{\mathbf{a}(t)\cdot\mathbf{b}(t)\}'=\mathbf{a}'(t)\cdot\mathbf{b}(t)+\mathbf{a}(t)\cdot\mathbf{b}'(t) = \{a_1'(t)\cdot b_1(t) + a_2'(t)\cdot b_2(t) + a_3'(t)\cdot b_3(t)\} + \{a_1(t)\cdot b_1'(t) + a_2(t)\cdot b_2'(t) + a_3(t)\cdot b_3'(t)\}$. Аналогичное правило справедливо и для векторного произведения двух векторных функций:

$$[\mathbf{a}(t)\mathbf{b}(t)]' = [\mathbf{a}'(t)\mathbf{b}(t)] + [\mathbf{a}(t), \mathbf{b}'(t)].$$

Глава 6

ОСНОВНЫЕ ТЕОРЕМЫ О ДИФФЕРЕНЦИРУЕМЫХ ФУНКЦИЯХ

В настоящей главе будет установлен ряд важных теорем, относящихся к произвольным дифференцируемым функциям. Эти теоремы чрезвычайно эффективны при исследовании поведения функции (как в окрестности отдельных точек области ее задания, так и на целых участках области ее задания).

§ 1. ВОЗРАСТАНИЕ (УБЫВАНИЕ) ФУНКЦИИ В ТОЧКЕ. ЛОКАЛЬНЫЙ ЭКСТРЕМУМ

Рассмотрим функцию $y=f(x)$, определенную всюду в некоторой окрестности фиксированной точки c.

Определение 1. *Будем говорить, что функция $y=f(x)$ возрастает в точке c, если найдется такая δ-окрестность точки c, в пределах которой*

$$f(x)<f(c) \text{ при } x<c$$

и

$$f(x)>f(c) \text{ при } x>c.$$

Определение 2. *Будем говорить, что функция $y=f(x)$ убывает в точке c, если найдется такая δ-окрестность точки c, в пределах которой*

$$f(x)>f(c) \text{ при } x<c$$

и

$$f(x)<f(c) \text{ при } x>c.$$

Определение 3. *Будем говорить, что функция $y=f(x)$ имеет в точке c локальный максимум [локальный минимум], если найдется такая δ-окрестность точки c, в пределах которой значение $f(c)$ является наибольшим [наименьшим] среди всех значений $f(x)$ этой функции.*

Определение 4. *Будем говорить, что функция $y=f(x)$ имеет в точке c локальный экстремум, если эта функция имеет в указанной точке либо локальный максимум, либо локальный минимум.*

На рис. 6.1 изображена функция, возрастающая в точке c_1, убывающая в точке c_2, имеющая локальный максимум в точке c_3 и локальный минимум в точке c_4.

Докажем следующие две теоремы.

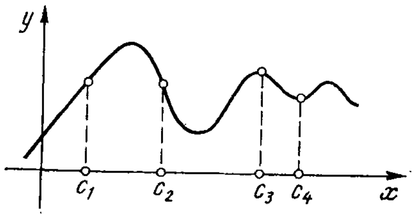


Рис. 6.1

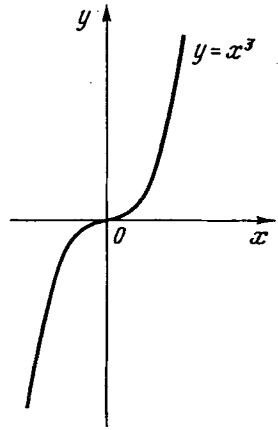


Рис. 6.2

Теорема 6.1 (достаточное условие возрастания или убывания функции в точке). *Если функция $y=f(x)$ дифференцируема в точке c и ее производная в этой точке $f'(c)$ положительна [отрицательна], то функция $y=f(x)$ возрастает [убывает] в точке c.*

Доказательство. Проведем все рассуждения для случая $f'(c)>0$ (для случая $f'(c)<0$ они аналогичны).

Так как (по определению производной)

$$f'(c)=\lim_{x\to c}\frac{f(x)-f(c)}{x-c},$$

то на основании определения предела функции по Коши для положительного числа $\varepsilon=f'(c)$ найдется $\delta>0$ такое, что

$$\left|\frac{f(x)-f(c)}{x-c}-f'(c)\right|<f'(c) \text{ при } 0<|x-c|<\delta$$

или, что то же самое,

$$0<\frac{f(x)-f(c)}{x-c}<2f'(c) \text{ при } c-\delta<x<c+\delta,\ x\neq c.$$

Таким образом, всюду в проколотой δ-окрестности точки c

$$\frac{f(x)-f(c)}{x-c}>0.$$

Но это и означает, что всюду в пределах δ-окрестности точки c

$$f(x)>f(c) \text{ при } x>c,$$

и

$$f(x)<f(c) \text{ при } x<c,$$

т. е. функция $y=f(x)$ возрастает в точке c.

Случай $f'(c)<0$ рассматривается аналогично. Теорема доказана.

Замечание 1. *Подчеркнем, что положительность [отрицательность] производной $f'(c)$ не является необходимым условием возрастания [убывания] дифференцируемой в точке c функции $y=f(x)$.*

Так, функция $y=x^3$ возрастает в точке $c=0$, в то время как производная этой функции $f'(x)=3x^2$ обращается в нуль в точке $c=0$ (график функции — на рис. 6.2).

Теорема 6.2 (необходимое условие локального экстремума дифференцируемой в данной точке функции). *Если функция $y=f(x)$ дифференцируема в точке c и имеет в этой точке локальный экстремум, то $f'(c)=0$.*

Доказательство. По условию теоремы существует конечная производная $f'(c)$. Так как функция $y=f(x)$ имеет в точке c локальный экстремум, то она не может в этой точке c ни возрастать, ни убывать. Значит, в силу теоремы 6.1 производная $f'(c)$ не может быть ни положительной, ни отрицательной. Тем самым доказано, что $f'(c)=0$.

Теорема 6.2 имеет очень простой геометрический смысл: она утверждает, что *если в той точке кривой $y=f(x)$, в которой достигается локальный экстремум, существует касательная к этой кривой, то эта касательная обязательно параллельна оси Ox* (рис. 6.3).

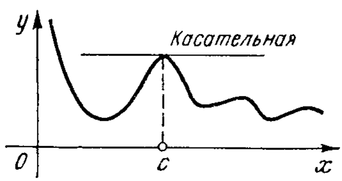


Рис. 6.3

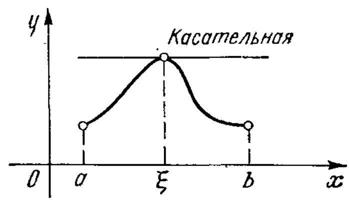


Рис. 6.4

Замечание 2. Пример той же самой функции $y=x^3$ (см. рис. 6.2) показывает, что *обращение в нуль производной является лишь необходимым и не является достаточным условием локального экстремума.* (Производная $f'(x)=3x^2$ этой функции обращается в нуль в точке $c=0$, но никакого экстремума в этой точке нет.)

§ 2. ТЕОРЕМА О НУЛЕ ПРОИЗВОДНОЙ

Теорема 6.3 (теорема Ролля*). *Пусть функция $f(x)$ непрерывна на сегменте $[a, b]$ и дифференцируема во всех внут-*

* Мишель Ролль — французский математик (1652—1719).

ренних точках этого сегмента. Пусть, кроме того, $f(a)=f(b)$. Тогда внутри сегмента $[a, b]$ найдется точка ξ такая, что значение производной в этой точке $f'(\xi)$ равно нулю.

Кратко можно сказать, что между двумя равными значениями дифференцируемой функции обязательно лежит нуль производной этой функции.

Доказательство. Так как функция $f(x)$ непрерывна на сегменте $[a, b]$, то согласно теореме 4.15 эта функция достигает на этом сегменте максимального значения M и своего минимального значения m. Могут представиться два случая: 1) $M=m$; 2) $M>m$. В случае 1) $f(x)=M=m=\text{const}$. Поэтому производная $f'(x)$ равна нулю в любой внутренней точке сегмента $[a, b]$. В случае $M>m$, поскольку $f(a)=f(b)$, можно утверждать, что хотя бы одно из двух значений M или m достигается функцией в некоторой внутренней точке ξ сегмента $[a, b]$. Но тогда функция $f(x)$ имеет в этой точке ξ локальный экстремум. Поскольку функция $f(x)$ дифференцируема в точке ξ, то по теореме 6.2 $f'(\xi)=0$. Теорема полностью доказана.

Теорема Ролля имеет простой геометрический смысл: *если крайние ординаты кривой $y=f(x)$ равны, то согласно теореме Ролля на кривой $y=f(x)$ найдется точка, в которой касательная к кривой параллельна оси Ox* (рис. 6.4).

Как мы увидим ниже, теорема Ролля лежит в основе многих формул и теорем математического анализа.

Замечание. В теореме Ролля требуется, чтобы функция $y=f(x)$ была непрерывна на сегменте $[a, b]$ и дифференцируема во всех внутренних точках этого сегмента. Так как из дифференцируемости $f(x)$ во всех внутренних точках вытекает непрерывность $f(x)$ во всех внутренних точках, то по существу вместо непрерывности $f(x)$ на сегменте $[a, b]$ можно было бы потребовать непрерывность $f(x)$ в точке a справа и в точке b слева.

§ 3. ФОРМУЛА КОНЕЧНЫХ ПРИРАЩЕНИЙ (ФОРМУЛА ЛАГРАНЖА)

Большое значение в анализе и его приложениях имеет следующая теорема, принадлежащая Лагранжу*.

Теорема 6.4 (теорема Лагранжа). *Если функция $f(x)$ непрерывна на сегменте $[a, b]$ и дифференцируема во всех внутренних точках этого сегмента, то внутри сегмента $[a, b]$ найдется точка ξ такая, что справедлива формула*

$$f(b)-f(a)=f'(\xi)(b-a). \tag{6.1}$$

Формулу (6.1) называют формулой Лагранжа или формулой конечных приращений.

* Жозеф Луи Лагранж — великий французский математик и механик (1736—1813).

Доказательство. Рассмотрим на сегменте $[a, b]$ следующую вспомогательную функцию

$$F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x-a). \tag{6.2}$$

Проверим, что для функции $F(x)$ выполнены все условия теоремы Ролля. В самом деле, $F(x)$ непрерывна на сегменте $[a, b]$ (как разность функции $f(x)$ и линейной функции) и во всех внутренних точках сегмента $[a, b]$ имеет производную, равную

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

Из формулы (6.2) очевидно, что $F(a)=F(b)=0$.

Согласно теореме Ролля внутри сегмента $[a, b]$ найдется точка ξ такая, что

$$F'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a} = 0. \tag{6.3}$$

Из равенства (6.3) вытекает формула Лагранжа (6.1). Подчеркнем, что в формуле (6.1) вовсе не обязательно считать, что $b>a$, эта формула верна и при $b<a$.

Замечание. Мы получили теорему Лагранжа как следствие теоремы Ролля. Заметим вместе с тем, что сама теорема Ролля является частным случаем теоремы Лагранжа (при $f(a)=f(b)$).

Для выяснения геометрического смысла теоремы Лагранжа заметим, что величина $\frac{f(b)-f(a)}{b-a}$ есть угловой коэффициент секущей, проходящей через точки $A(a, f(a))$ и $B(b, f(b))$ кривой $y=f(x)$, а $f'(\xi)$ есть угловой коэффициент касательной к кривой $y=f(x)$, проходящей через точку $C(\xi, f(\xi))$. Формула Лагранжа (6.1) означает, что на кривой $y=f(x)$ между точками A и B найдется такая точка C, касательная в которой параллельна секущей AB (рис. 6.5).

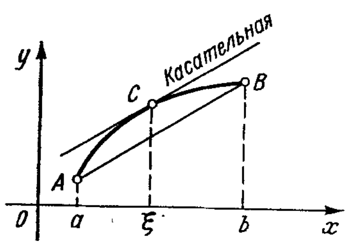


Рис. 6.5

Часто бывает удобно записывать формулу Лагранжа в виде, несколько отличном от (6.1). Пусть $f(x)$ удовлетворяет условиям теоремы 6.4. Зафиксируем любое x_0 из сегмента $[a, b]$ и зададим ему приращение Δx произвольное, но такое, чтобы значение $x_0+\Delta x$ также лежало на сегменте $[a, b]$. Тогда, записывая формулу Лагранжа для сегмента, ограниченного точками x_0 и $x_0+\Delta x$, будем иметь

$$f(x_0+\Delta x)-f(x_0)=\Delta x f'(\xi), \tag{6.4}$$

где ξ — некоторая точка, лежащая между x_0 и $x_0+\Delta x$. Можно утверждать, что найдется такое (зависящее от Δx) число θ из интервала $0<\theta<1$, что

$$\xi = x_0 + \theta\Delta x.$$

Таким образом, формуле (6.4) можно придать вид

$$f(x_0+\Delta x) - f(x_0) = \Delta x f'(x_0+\theta\Delta x), \tag{6.5}$$

где θ — некоторое число из интервала $0<\theta<1$. Формула Лагранжа в виде (6.5) дает точное выражение для приращения функции через вызвавшее его произвольное конечное приращение Δx аргумента. Этот вид формулы Лагранжа оправдывает термин «формула конечных приращений».

§ 4. НЕКОТОРЫЕ СЛЕДСТВИЯ ИЗ ФОРМУЛЫ ЛАГРАНЖА

1. Постоянство функции, имеющей на интервале равную нулю производную.

Теорема 6.5. *Если функция $f(x)$ дифференцируема всюду на интервале (a, b) и если всюду на этом интервале $f'(x)=0$, то функция $f(x)$ является постоянной на интервале (a, b).*

Доказательство. Пусть x_0 — некоторая фиксированная точка интервала (a, b), а x — любая точка этого интервала.

Сегмент $[x_0, x]$ или соответственно $[x, x_0]$ целиком принадлежит интервалу (a, b). Поэтому функция $f(x)$ дифференцируема (а значит, и непрерывна) всюду на этом сегменте. Это дает право применить к функции $f(x)$ на этом сегменте теорему Лагранжа. Согласно этой теореме внутри указанного сегмента найдется точка ξ такая, что

$$f(x) - f(x_0) = (x - x_0) f'(\xi). \tag{6.6}$$

По условию производная функции $f(x)$ равна нулю всюду в интервале (a, b). Значит, $f'(\xi)=0$ и из формулы (6.6) мы получим

$$f(x) = f(x_0). \tag{6.7}$$

Равенство (6.7) утверждает, что значение функции $f(x)$ в любой точке x интервала (a, b) равно ее значению в фиксированной точке x_0. Это и означает, что функция $f(x)$ постоянна всюду на интервале (a, b). Теорема доказана.

Теорема 6.5 имеет простой геометрический смысл: *если касательная в каждой точке некоторого участка кривой $y=f(x)$ параллельна оси Ox, то указанный участок кривой $y=f(x)$ представляет собой отрезок прямой, параллельной оси Ox.*

Замечание. Теорема 6.5 будет использована нами в гл. 8 «Первообразная функция и неопределенный интеграл». Во всем остальном гл. 8 является независимой от гл. 6 и 7 и может читаться сразу после гл. 5.

2. Условия монотонности функции на интервале. В качестве второго следствия формулы Лагранжа рассмотрим вопрос об условиях, обеспечивающих неубывание (невозрастание) функции на данном интервале.

Прежде всего напомним определения неубывания, невозрастания, возрастания и убывания функции на данном интервале.

1°. *Говорят, что функция $f(x)$ не убывает (не возрастает) на интервале (a, b), если для любых точек x_1 и x_2 из интервала (a, b), удовлетворяющих условию $x_1<x_2$, справедливо неравенство*

$$f(x_1)\leqslant f(x_2)\ (f(x_1)\geqslant f(x_2)).$$

2°. *Говорят, что функция $f(x)$ возрастает (убывает) на интервале (a, b), если для любых точек x_1, x_2 из интервала (a, b), связанных условием $x_1<x_2$, справедливо неравенство*

$$f(x_1)<f(x_2)\quad (f(x_1)>f(x_2)).$$

Теорема 6.6. *Для того чтобы дифференцируемая на интервале (a, b) функция $f(x)$ не убывала (не возрастала) на этом интервале, необходимо и достаточно, чтобы производная этой функции была неотрицательной (неположительной) всюду на этом интервале.*

Доказательство. 1) *Достаточность.* Пусть $f'(x)\geqslant 0$ $(\leqslant 0)$ всюду на интервале (a, b). Требуется доказать, что $f(x)$ не убывает (не возрастает) на интервале (a, b). Пусть x_1 и x_2 — любые две точки из интервала (a, b), удовлетворяющие условию $x_1<x_2$. Функция $f(x)$ дифференцируема (а значит, и непрерывна) всюду на сегменте $[x_1, x_2]$. Поэтому к $f(x)$ можно применить на сегменте $[x_1, x_2]$ теорему Лагранжа, в результате чего получим

$$f(x_2)-f(x_1)=(x_2-x_1)f'(\xi), \tag{6.8}$$

где $x_1<\xi<x_2$.

По условию $f'(\xi)\geqslant 0$ $(\leqslant 0)$, $x_2-x_1>0$. Поэтому правая, а значит, и левая части (6.8) неотрицательны (неположительны), что и доказывает неубывание (невозрастание) $f(x)$ на (a, b).

2) *Необходимость.* Пусть функция $f(x)$ дифференцируема на интервале (a, b) и не убывает (не возрастает) на этом интервале. Требуется доказать, что $f'(x)\geqslant 0$ $(\leqslant 0)$ всюду на этом интервале. Так как $f(x)$ не убывает (не возрастает) на интервале (a, b), то эта функция не может убывать (возрастать) ни в одной точке интервала (a, b). Значит, в силу теоремы 6.1 производная $f'(x)$ ни в одной точке интервала (a, b) не может быть отрицательной (положительной), что и требовалось доказать.

Теорема 6.7. *Для того, чтобы функция $f(x)$ возрастала (убывала) на интервале (a, b), достаточно, чтобы производная $f'(x)$ была положительной (отрицательной) всюду на этом интервале.*

Доказательство проводится по той же схеме, что и доказательство достаточности в теореме 6.6. Пусть x_1, x_2 — любые две точки из интервала (a, b), удовлетворяющие условию $x_1 < x_2$. Записывая для сегмента $[x_1, x_2]$ формулу Лагранжа, получим равенство (6.8), но на этот раз в этом равенстве $f'(\xi) > 0$ (< 0).

Вследствие этого левая часть (6.8) положительна (отрицательна), что и доказывает возрастание (убывание) $f(x)$ на интервале (a, b).

Замечание. Подчеркнем, что *положительность (отрицательность) производной $f'(x)$ на интервале (a, b) не является необходимым условием возрастания (убывания) функции $f(x)$ на интервале* (a, b). Так, функция $y = x^3$ возрастает на интервале $(-1, +1)$, но производная этой функции $f'(x) = 3x^2$ не является всюду положительной на этом интервале (она обращается в нуль в точке $x = 0$). Вообще, легко доказать, что *функция $f(x)$ возрастает (убывает) на интервале (a, b), если производная этой функции $f'(x)$ положительна (отрицательна) всюду на этом интервале, за исключением конечного числа точек, в которых эта производная равна нулю.* (Для доказательства достаточно применить теорему 6.7 к каждому из конечного числа интервалов, на которых $f'(x)$ строго положительна (отрицательна), и учесть непрерывность $f(x)$ в тех точках, в которых производная равна нулю.)

Установленную теоремой 6.7 связь между знаком производной и направлением изменения функции легко понять из геометрических соображений. Поскольку производная равна угловому коэффициенту касательной к графику функции $y = f(x)$, знак производной указывает, острый или тупой угол с положительным направлением оси Ox составляет луч касательной, лежащей в верхней полуплоскости. Если $f'(x) > 0$ всюду на интервале (a, b), то всюду на этом интервале луч касательной, лежащей в верхней полуплоскости, составляет с положительным направлением оси Ox острый угол, значит, и кривая $y = f(x)$ идет вверх всюду на этом интервале (рис. 6.6).

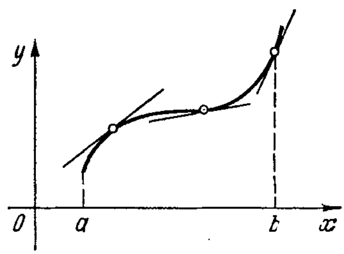


Рис. 6.6

3. **Отсутствие разрывов первого рода и устранимых разрывов у производной.** Теорема Лагранжа позволяет установить одно замечательное свойство производной. Начнем с доказательства следующей леммы.

Лемма. *Пусть функция $y = f(x)$ имеет конечную производную $f'(x)$ всюду на интервале $(c, c+\delta)$ [на интервале $(c-\delta, c)$], где δ — некоторое положительное число и, кроме того, имеет правую производную $f'(c+0)$ [левую производную $f'(c-0)$]. Тогда, если производная $f'(x)$ имеет в точке c правый предел [левый предел],* *то этот предел совпадает с правой производной $f'(c+0)$ [с левой производной $f'(c-0)$].*

Доказательство. Из существования правой производной $f'(c+0)$ [левой производной $f'(c-0)$] вытекает существование конечного предела

$$\lim_{x\to c+0}\frac{f(x)-f(c)}{x-c}\quad\left[\lim_{x\to c-0}\frac{f(x)-f(c)}{x-c}\right].$$

Но это означает, что существует равный нулю предел

$$\lim_{x\to c+0}\{f(x)-f(c)\}=0\quad[\lim_{x\to c-0}\{f(x)-f(c)\}=0],$$

т. е. функция $y=f(x)$ является непрерывной в точке c справа [слева].

Фиксируем любое x из интервала $(c, c+\delta)$ [$(c-\delta, c)$]. Так как функция $y=f(x)$ дифференцируема (а значит, и непрерывна) всюду на указанном интервале и, кроме того, непрерывна в точке c справа [слева], то для этой функции выполнены на сегменте $[c, x]$ [на сегменте $[x, c]$] все условия теоремы Лагранжа 6.4.

В силу этой теоремы между x и c найдется точка ξ такая, что справедливо равенство

$$\frac{f(x)-f(c)}{x-c}=f'(\xi). \tag{6.9}$$

Перейдем теперь в равенстве (6.9) к пределу при $x\to c+0$ [при $x\to c-0$]. Если производная $f'(x)$ имеет в точке c конечный правый предел $\lim\limits_{x\to c+0}f'(x)$ [конечный левый предел $\lim\limits_{x\to c-0}f'(x)$], то правая часть (6.9) обязана стремиться к этому пределу (ибо $\xi\to c+0$ [$\xi\to c-0$] при $x\to c+0$ [$x\to c-0$]).

Тот же самый предел при $x\to c+0$ [$x\to c-0$] обязана иметь и левая часть (6.9). Но предел левой части (6.9) при $x\to c+0$ [$x\to c-0$] по определению равен правой производной $f'(c+0)$ [левой производной $f'(c-0)$]. Лемма доказана.

Применяя только что доказанную лемму в каждой точке c некоторого интервала (a, b), мы придем к следующему утверждению: *если функция $f(x)$ имеет конечную производную всюду на интервале (a, b), то эта производная $f'(x)$ не может иметь на этом интервале ни точек устранимого разрыва, ни точек разрыва первого рода.*

В самом деле, если в некоторой точке c интервала (a, b) существуют конечные правый и левый пределы функции $f'(x)$, то $f'(x)$ непрерывна в точке c (в силу доказанной нами леммы*).

---

* В силу этой леммы $\lim\limits_{x\to c+0}f'(x)=f'(c+0)$, $\lim\limits_{x\to c-0}f'(x)=f'(c-0)$, а поскольку $f'(c+0)=f'(c-0)=f'(c)$, то $\lim\limits_{x\to c+0}f'(x)=\lim\limits_{x\to c-0}f'(x)=f'(c)$. Это и означает, что $f'(x)$ непрерывна в точке c.

Если же не существует хотя бы одного из пределов $\lim\limits_{x\to c+0} f'(x)$ и $\lim\limits_{x\to c-0} f'(x)$, то функция $f'(x)$ по определению имеет в точке c разрыв второго рода. Итак, производная $f'(x)$ в каждой точке c интервала (a, b) либо непрерывна, либо имеет разрыв второго рода. Сформулированное нами утверждение доказано.

Приведем пример функции $f(x)$, производная $f'(x)$ которой существует и конечна всюду на некотором интервале и имеет в некоторой точке этого интервала разрыв второго рода.

Рассмотрим на интервале $(-1, +1)$ функцию

$$f(x) = \begin{cases} x^2 \cos\dfrac{1}{x} & \text{при } x \neq 0, \\ 0 & \text{при } x = 0. \end{cases}$$

Очевидно, что для любого $x\neq 0$ производная $f'(x)$ этой функции существует и равна $f'(x) = 2x\cos\frac{1}{x} + \sin\frac{1}{x}$. Существование производной $f'(x)$ в точке $x=0$ непосредственно вытекает из существования предела

$$\lim_{\Delta x\to 0} \frac{f(0+\Delta x)-f(0)}{\Delta x} = \lim_{\Delta x\to 0} \Delta x \cos\frac{1}{\Delta x} = 0.$$

Производная $f'(x)$ не имеет в точке $x=0$ ни правого, ни левого пределов, ибо первое слагаемое $2x\cos\frac{1}{x}$ имеет в точке $x=0$ равный нулю предел, а второе слагаемое $\sin\frac{1}{x}$ не имеет в точке $x=0$ ни правого, ни левого пределов. (Это было доказано в примере, рассмотренном в п. 3 § 5 гл. 4.)

4. Вывод некоторых неравенств. В заключение покажем, как с помощью теоремы Лагранжа могут быть получены некоторые весьма полезные неравенства. В качестве примера установим следующие два неравенства:

$$|\sin x_1 - \sin x_2| \leqslant |x_1 - x_2|, \tag{6.10}$$

$$|\operatorname{arctg} x_1 - \operatorname{arctg} x_2| \leqslant |x_1 - x_2|. \tag{6.11}$$

(Здесь под x_1 и x_2 можно понимать любые значения аргумента.) Для установления неравенства (6.10) применим теорему Лагранжа к функции $f(x) = \sin x$ по сегменту $[x_1, x_2]$. Получим

$$\sin x_1 - \sin_2 = (x_1 - x_2) f'(\xi). \tag{6.12}$$

Учитывая, что $f'(\xi) = \cos\xi$ и что $|\cos\xi| \leqslant 1$ для любого ξ, получим, переходя в (6.12) к модулям, неравенство (6.10).

Для установления неравенства (6.11) следует применить теорему Лагранжа по сегменту $[x_1, x_2]$ к функции $f(x) = \operatorname{arctg} x$ и учесть, что $f'(\xi) = \dfrac{1}{1+\xi^2} \leqslant 1$.

§ 5. ОБОБЩЕННАЯ ФОРМУЛА КОНЕЧНЫХ ПРИРАЩЕНИЙ (ФОРМУЛА КОШИ)

В этом параграфе мы докажем теорему, принадлежащую Коши и обобщающую установленную выше теорему Лагранжа.

Теорема 6.8 (теорема Коши). *Если каждая из двух функций $f(x)$ и $g(x)$ непрерывна на сегменте $[a, b]$ и дифференцируема во всех внутренних точках этого сегмента и если, кроме того, производная $g'(x)$ отлична от нуля всюду внутри сегмента $[a, b]$, то внутри этого сегмента найдется точка ξ такая, что справедлива формула*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)}. \tag{6.13}$$

Формулу (6.13) называют **обобщенной формулой конечных приращений** или **формулой Коши**.

Доказательство. Прежде всего докажем, что $g(a) \neq g(b)$. В самом деле, если бы это было не так, то для функции $g(x)$ были бы выполнены на сегменте $[a, b]$ все условия теоремы 6.3 (Ролля) и по этой теореме внутри сегмента $[a, b]$ нашлась бы точка ξ такая, что $g'(\xi) = 0$. Последнее противоречит условию теоремы. Итак, $g(a) \neq g(b)$, и мы имеем право рассмотреть следующую вспомогательную функцию:

$$F(x) = f(x) - f(a) - \frac{f(b)-f(a)}{g(b)-g(a)} [g(x) - g(a)]. \tag{6.14}$$

В силу требований, наложенных на функции $f(x)$ и $g(x)$, функция $F(x)$ непрерывна на сегменте $[a, b]$ и дифференцируема во всех внутренних точках этого сегмента. Кроме того, очевидно, что $F(a) = F(b)$. Таким образом, для $F(x)$ выполнены все условия теоремы 6.3 (Ролля). Согласно этой теореме внутри сегмента $[a, b]$ найдется точка ξ такая, что

$$F'(\xi) = 0. \tag{6.15}$$

Имея в виду, что $F'(x) = f'(x) - \dfrac{f(b)-f(a)}{g(b)-g(a)} g'(x)$, и используя равенство (6.15), будем иметь

$$f'(\xi) - \frac{f(b)-f(a)}{g(a)-g(b)} g'(\xi) = 0. \tag{6.16}$$

Учитывая, что $g'(\xi) \neq 0$, из равенства (6.16) получим формулу Коши (6.13). Теорема доказана.

Замечание 1. Формула Лагранжа (6.1) является частным случаем формулы Коши (6.13) при $g(x)=x$.

Замечание 2. В формуле (6.13) вовсе не обязательно считать, что $b>a$. Эта формула верна и при $b<a$.

§ 6. РАСКРЫТИЕ НЕОПРЕДЕЛЕННОСТЕЙ (ПРАВИЛО ЛОПИТАЛЯ)

1. Раскрытие неопределенности вида $\frac{0}{0}$. Будем говорить, что *отношение двух функций $\frac{f(x)}{g(x)}$ представляет собой при $x \to a$ неопределенность вида $\frac{0}{0}$, если*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0.$$

Раскрыть эту неопределенность — это значит вычислить предел $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ (при условии, что этот предел существует).

Аналогично вводится понятие неопределенности вида $\frac{0}{0}$ при $x\to a+0$ $[x\to a-0]$, при $x\to\infty$, а также при $x\to+\infty$ $[x\to-\infty]$.

Следующая теорема дает правило раскрытия неопределенности вида $\frac{0}{0}$ для предела в точке a.

Теорема 6.9 (первое правило Лопиталя*). *Пусть множество C_δ представляет собой проколотую δ-окрестность точки a**, функции $f(x)$ и $g(x)$ определены и дифференцируемы на C_δ и, кроме того, производная $g'(x)$ не обращается на C_δ в нуль. Пусть, далее,*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0. \tag{6.17}$$

Тогда если существует (конечный или бесконечный) предел

$$\lim_{x\to a} \frac{f'(x)}{g'(x)}, \tag{6.18}$$

то существует и предел

$$\lim_{x\to a} \frac{f(x)}{g(x)}, \tag{6.19}$$

причем справедливо соотношение

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}. \tag{6.20}$$

---

* Гильом Франсуа де Лопиталь — французский математик (1661—1704).

** Напомним, что проколотой δ-окрестностью точки a называется интервал $(a-\delta, a+\delta)$ с выкинутой точкой a (при условии, конечно, что $\delta>0$).

Теорема 6.9 дает правило раскрытия неопределенности вида $\frac{0}{0}$, сводящее вычисление предела в точке a отношения двух функций к вычислению предела в этой точке отношения производных этих функций.

Доказательство. Пусть $\{x_n\}$ — произвольная последовательность значений аргумента, сходящаяся к a и состоящая из чисел отличных от a. Доопределим функции $f(x)$ и $g(x)$ в точке a, положив их равными нулю в этой точке. При таком доопределении функции $f(x)$ и $g(x)$ окажутся непрерывными всюду на множестве C_δ, дополненном точкой a, т. е. всюду в δ-окрестности точки a. В самом деле, непрерывность $f(x)$ и $g(x)$ во всех точках δ-окрестности точки a, за исключением самой точки a, вытекает из их дифференцируемости в этих точках, а непрерывность $f(x)$ и $g(x)$ в точке a вытекает из того, что в силу нашего доопределения этих функций их пределы в точке a равны частным значениям в этой точке.

Учитывая, что все элементы последовательности $\{x_n\}$ принадлежат множеству C_δ, рассмотрим произвольный сегмент, ограниченный точками a и x_n.

В силу сказанного выше обе функции $f(x)$ и $g(x)$ будут непрерывными на таком сегменте. Кроме того, функции $f(x)$ и $g(x)$ дифференцируемы во всех внутренних точках указанного сегмента, и производная $g'(x)$ не обращается в этих внутренних точках в нуль.

Это дает нам право применить к функциям $f(x)$ и $g(x)$ по указанному сегменту, ограниченному точками a и x_n, теорему Коши 6.8.

В силу этой теоремы между точками a и x_n найдется точка ξ_n такая, что справедливо равенство

$$\frac{f(x_n)-f(a)}{g(x_n)-g(a)}=\frac{f'(\xi_n)}{g'(\xi_n)}. \tag{6.21}$$

Учитывая, что по нашему доопределению функций $f(x)$ и $g(x)$ справедливы равенства $f(a)=g(a)=0$, мы можем переписать соотношение (6.21) в виде

$$\frac{f(x_n)}{g(x_n)}=\frac{f'(\xi_n)}{g'(\xi_n)}. \tag{6.22}$$

Пусть теперь в соотношении (6.22) номер n стремится к бесконечности, тогда $x_n \to a$. Поскольку ξ_n заключено между a и x_n, то и $\xi_n \to a$ при $n \to \infty$. В силу существования предела (6.18) и определения предела функции по Гейне правая часть (6.22) имеет предел при $n \to \infty$, равный пределу (6.18). Значит, тот же самый предел при $n \to \infty$ имеет и левая часть (6.22). В силу произвольности последовательности значений аргумента $\{x_n\}$, сходящейся к a, и в силу определения предела функции по Гейне существование равного (6.18) предела при $n\to\infty$ левой части (6.22) означает существование предела функции (6.19), также равного (6.18).

Итак, в пределе из (6.22) при $n\to\infty$ мы получаем соотношение (6.20). Теорема доказана.

Замечание 1. Правило Лопиталя «действует» не всегда, т. е. *предел отношения функций* (6.19) *может существовать и в случае, когда предела отношения производных* (6.18) *не существует.*

Например, при $a=0$, $f(x)=x^2\cos\frac{1}{x}$, $g(x)=\sin x$ существует предел $\lim\limits_{x\to a}\frac{f(x)}{g(x)}=\lim\limits_{x\to 0}\frac{x^2\cos\frac{1}{x}}{\sin x}=\lim\limits_{x\to 0}\frac{x}{\sin x}\lim\limits_{x\to 0}x\cos\frac{1}{x}=0$, в то время как предел $\lim\limits_{x\to 0}\frac{f'(x)}{g'(x)}=\lim\limits_{x\to 0}\frac{2x\cos\frac{1}{x}+\sin\frac{1}{x}}{\cos x}$ не существует (в силу того, что не существует предел $\lim\limits_{x\to 0}\sin\frac{1}{x}$, а предел $\lim\limits_{x\to 0}\left(2x\cos\frac{1}{x}\right)$ существует и равен нулю).

Замечание 2. Если к условиям теоремы 6.9 добавить требование непрерывности производных $f'(x)$ и $g'(x)$ в точке a, то при условии $g'(a)\neq 0$ соотношение (6.20) можно переписать в виде

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

Замечание 3. Если производные $f'(x)$ и $g'(x)$ удовлетворяют тем же требованиям, что и сами функции $f(x)$ и $g(x)$, то правило Лопиталя можно применять повторно, т. е. предел отношения первых производных функций $f(x)$ и $g(x)$ можно заменить пределом отношения вторых производных этих функций. Мы получим при этом, что

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

Примеры. 1)

$$\lim_{x\to 0}\frac{1-\cos x}{x^2}=\lim_{x\to 0}\frac{\sin x}{2x}=\frac{1}{2}.$$

2) Следующий предел вычисляется двукратным применением правила Лопиталя:

$$\lim_{x\to 0}\frac{x-\sin x}{x^3}=\lim_{x\to 0}\frac{1-\cos x}{3x^2}=\lim_{x\to 0}\frac{\sin x}{6x}=\frac{1}{6}.$$

3) Трехкратным применением правила Лопиталя вычисляется предел

$$\lim_{x\to 0}\frac{x^4}{x^2+2\cos x-2}=\lim_{x\to 0}\frac{4x^3}{2x-2\sin x}=\lim_{x\to 0}\frac{12x^2}{2-2\cos x}=$$

$$=\lim_{x\to 0}\frac{24x}{2\sin x}=12.$$

Мы рассмотрели вопрос о раскрытии неопределенности вида $\frac{0}{0}$ для случая предела в точке a. Совершенно аналогичные результаты справедливы и для случаев предела в точке a справа [слева], предела при $x\to\infty$, а также предела при $x\to+\infty$ [при $x\to-\infty$].

Мы сейчас убедимся в том, что теорема 6.9 остается справедливой в каждом из следующих трех случаев:

1) в случае, если в этой теореме в качестве множества C_δ взять интервал $(a, a+\delta)$ [соответственно $(a-\delta, a)$], а все пределы (6.17)—(6.20) взять при $x\to a+0$ [соответственно при $x\to a-0$];

2) в случае, если в теореме 6.9 в качестве C_δ взять множество всех точек, лежащих вне сегмента $[-\delta, \delta]$, а все пределы (6.17)—(6.20) взять при $x\to\infty$;

3) в случае, если в теореме 6.9 в качестве множества C_δ взять полупрямую $(\delta, +\infty)$ [соответственно $(-\infty, -\delta)$], а все пределы (6.17)—(6.20) взять при $x\to+\infty$ [соответственно при $x\to-\infty$].

Случай 1. В этом случае остается справедливой вся схема доказательства теоремы 6.9, только вместо последовательности $\{x_n\}$, сходящейся к a и состоящей из чисел x_n, отличных от a, следует взять последовательность $\{x_n\}$ точек интервала $(a, a+\delta)$ [соответственно $(a-\delta, a)$], сходящуюся к a. Детали рассуждений предоставляем читателю.

Случай 2. Пусть функции $f(x)$ и $g(x)$ определены и дифференцируемы всюду вне сегмента $[-\delta, \delta]$ при некотором $\delta>0$ и производная $g'(x)$ не обращается в нуль вне указанного сегмента. Путь, кроме того, существует предел

$$\lim_{x\to\infty}\frac{f'(x)}{g'(x)}. \tag{6.18'}$$

Сделаем замену переменной $t=\frac{1}{x}$ и положим $G(t)=g\left(\frac{1}{t}\right)=g(x)$, $F(t)=f\left(\frac{1}{t}\right)=f(x)$. Тогда, очевидно, функции $F(t)$ и $G(t)$ будут определены и дифференцируемы в проколотой $\frac{1}{\delta}$ -окрестности точки $t=0$ * и производная

$$G'(t)=g'\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right)=g'(x)\ (-x^2)$$

не будет обращаться в нуль в указанной проколотой $\frac{1}{\delta}$ -окрестности.

Кроме того, в силу существования предела (6.18′) существует предел

$$\lim_{t\to 0}\frac{F'(t)}{G'(t)}=\lim_{t\to 0}\frac{f'\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right)}{g'\left(\frac{1}{t}\right)\left(-\frac{1}{t^2}\right)}=\lim_{t\to 0}\frac{f'\left(\frac{1}{t}\right)}{g'\left(\frac{1}{t}\right)}=\lim_{x\to\infty}\frac{f'(x)}{g'(x)}. \tag{6.23}$$

Но тогда в силу теоремы 6.9 существует и предел

$$\lim_{t\to 0}\frac{F(t)}{G(t)}=\lim_{t\to 0}\frac{f\left(\frac{1}{t}\right)}{g\left(\frac{1}{t}\right)}=\lim_{x\to\infty}\frac{f(x)}{g(x)}, \tag{6.24}$$

причем справедливо соотношение (6.20), принимающее (в силу (6.23) и (6.24)) вид

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=\lim_{t\to 0}\frac{F(t)}{G(t)}=\lim_{t\to 0}\frac{F'(t)}{G'(t)}=\lim_{x\to\infty}\frac{f'(x)}{g'(x)}.$$

Рассмотрение случая 2 завершено.

Случай 3. В этом случае действует та же замена $t=\frac{1}{x}$, что и в случае 2, однако на этот раз эта замена сводит рассмотрение предела при $x\to+\infty\,[x\to-\infty]$ к пределу при $t\to 0+0\,[t\to 0-0]$, рассмотренному в случае 1. Детали рассуждений предоставляем читателю.

Примеры. 1) Вычислить $\lim\limits_{x\to 0+0}\frac{x^\delta}{\ln(1+x)}$ (для любого $\delta>1$). Этот пример относится к случаю 1. Применяя правило Лопиталя, получим

$$\lim_{x\to 0+0}\frac{x^\delta}{\ln(1+x)}=\lim_{x\to 0+0}\frac{\delta x^{\delta-1}}{\left(\frac{1}{1+x}\right)}=\lim_{x\to 0+0}\delta x^{\delta-1}(1+x)=0.$$

---

* Т. е. всюду в интервале $-\frac{1}{\delta}<t<\frac{1}{\delta}$, за исключением точки $t=0$.

2) Вычислить $\lim\limits_{x\to\infty} \dfrac{\frac{\pi}{4} - \operatorname{arctg}\left(1 - \frac{1}{x}\right)}{\sin\frac{1}{x}}$ (этот пример относится к случаю 2).

Применяя правило Лопиталя, получим

$$\lim_{x\to\infty} \frac{\frac{\pi}{4} - \operatorname{arctg}\left(1-\frac{1}{x}\right)}{\sin\frac{1}{x}} = \lim_{x\to\infty} \frac{\dfrac{-1}{1+\left(1-\frac{1}{x}\right)^2}\left(\frac{1}{x^2}\right)}{\left(\cos\frac{1}{x}\right)\left(-\frac{1}{x^2}\right)} =$$

$$= \lim_{x\to\infty} \frac{\dfrac{1}{1+\left(1-\frac{1}{x}\right)^2}}{\cos\frac{1}{x}} = \frac{1}{2}.$$

2. Раскрытие неопределенности вида $\frac{\infty}{\infty}$. Будем говорить, что *отношение двух определенных в окрестности точки a* функций $f(x)$ *и $g(x)$ представляет собой при $x\to a$ неопределенность вида $\frac{\infty}{\infty}$, если*

$$\lim_{x\to a} f(x) = \infty, \quad \lim_{x\to a} g(x) = \infty^*. \tag{6.25}$$

Для раскрытия этой неопределенности, т. е. для вычисления предела $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$, справедливо утверждение, полностью аналогичное теореме 6.9.

Теорема 6.9* (второе правило Лопиталя). *Пусть множество C_δ представляет собой проколотую δ-окрестность точки a, функции $f(x)$ и $g(x)$ определены и дифференцируемы на C_δ и, кроме того, производная $g'(x)$ не обращается на C_δ в нуль. Пусть, далее,*

$$\lim_{x\to a} f(x) = \infty, \quad \lim_{x\to a} g(x) = \infty. \tag{6.17*}$$

Тогда, если существует (конечный или бесконечный) предел

$$\lim_{x\to a} \frac{f'(x)}{g'(x)}, \tag{6.18*}$$

то существует и предел

$$\lim_{x\to a} \frac{f(x)}{g(x)}, \tag{6.19*}$$

---

* Вместо ∞ в (6.25) может стоять $+\infty$ или $-\infty$.

причем справедливо соотношение

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{x\to a}\frac{f'(x)}{g'(x)}. \qquad (6.20^*)$$

Доказательство. 1) Предположим сначала, что существует конечный предел (6.18 *), равный числу b. Докажем, что в этом случае существует предел (6.19 *), также равный числу b.

Пусть $\{x_n\}$ — произвольная последовательность значений аргумента, сходящаяся к a либо справа, либо слева *. Так как все элементы этой последовательности принадлежат множеству C_δ, то, каковы бы ни были два элемента этой последовательности x_m и x_n, для функций $f(x)$ и $g(x)$ выполнены на сегменте $[x_m, x_n]$ все условия теоремы Коши 6.8. По этой теореме между x_m и x_n найдется точка ξ_{mn} такая, что справедливо равенство

$$\frac{f(x_n)-f(x_m)}{g(x_n)-g(x_m)}=\frac{f(x_n)}{g(x_n)}\,\frac{1-\dfrac{f(x_m)}{f(x_n)}}{1-\dfrac{g(x_m)}{g(x_n)}}=\frac{f'(\xi_{mn})}{g'(\xi_{mn})}.$$

Из этого равенства заключаем, что

$$\frac{f(x_n)}{g(x_n)}=\frac{f'(\xi_{mn})}{g'(\xi_{mn})}\,\frac{1-\dfrac{g(x_m)}{g(x_n)}}{1-\dfrac{f(x_m)}{f(x_n)}}. \qquad (6.26)$$

Фиксируем теперь произвольное положительное число ε. Так как по условию $\lim\limits_{x\to a}\dfrac{f'(x)}{g'(x)}=b$, а последовательность $\{x_n\}$ сходится к a, то для положительного числа $\dfrac{\varepsilon}{2}$ можно фиксировать такой номер m, что для любого номера n, превосходящего m, будет выполняться условие **

$$\frac{f'(\xi_{mn})}{g'(\xi_{mn})}=b+\alpha_{mn},\ \text{где } |\alpha_{mn}|<\frac{\varepsilon}{2}. \qquad (6.27)$$

Далее заметим, что в силу условий (6.17*) $\lim\limits_{n\to\infty} f(x_n)=\infty$, $\lim\limits_{n\to\infty} g(x_n)=\infty$, а поэтому, поскольку номер m фиксирован, существует равный единице предел

$$\lim_{n\to\infty}\frac{1-\dfrac{g(x_m)}{g(x_n)}}{1-\dfrac{f(x_m)}{f(x_n)}}=1.$$

* Т. е. $\{x_n\}$ — такая сходящаяся к a последовательность, у которой либо все $x_n>a$, либо все $x_n<a$.

** Мы учитываем, что ξ_{mn} заключено между x_m и x_n.

Это означает, что для положительного числа $\frac{\varepsilon/2}{|b|+\varepsilon/2}$ и для фиксированного нами номера m найдется превосходящий его номер n_0 такой, что при всех $n \geqslant n_0$

$$\frac{1-\dfrac{g(x_m)}{g(x_n)}}{1-\dfrac{f(x_m)}{f(x_n)}} = 1+\beta_{mn}, \text{ где } |\beta_{mn}| < \frac{\varepsilon/2}{|b|+\varepsilon/2}. \qquad (6.28)$$

Из соотношений (6.26), (6.27) и (6.28) вытекает, что

$$\frac{f(x_n)}{g(x_n)} = (b+\alpha_{mn})(1+\beta_{mn}) = b + (b+\alpha_{mn})\beta_{mn} + \alpha_{mn}.$$

Значит, справедливо неравенство

$$\left|\frac{f(x_n)}{g(x_n)} - b\right| \leqslant (|b|+|\alpha_{mn}|)|\beta_{mn}| + |\alpha_{mn}|.$$

Учитывая неравенства, стоящие в условиях (6.27) и (6.28), мы получим, что при всех $n \geqslant n_0$

$$(|b|+|\alpha_{mn}|)\cdot|\beta_{mn}| + |\alpha_{mn}| < \left(|b|+\frac{\varepsilon}{2}\right)\cdot\frac{\varepsilon/2}{|b|+\varepsilon/2} + \frac{\varepsilon}{2} = \varepsilon.$$

Итак, для произвольного фиксированного нами $\varepsilon > 0$ мы нашли номер $n_0 \geqslant m$ такой, что при всех $n \geqslant n_0$

$$\left|\frac{f(x_n)}{g(x_n)} - b\right| < \varepsilon.$$

Это и означает, что предел (6.19*) равен числу b и справедливо соотношение (6.20*). Таким образом, для случая конечного предела (6.18*) теорема доказана.

2) Пусть теперь предел (6.18*) равен ∞. Тогда, очевидно, предел обратного отношения $\lim\limits_{x\to a}\frac{g'(x)}{f'(x)}$ будет равен нулю и по только что рассмотренному нами случаю конечного предела (6.18*) мы получим, что *

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

Последнее соотношение с учетом (6.17*) эквивалентно тому, что

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \infty.$$

Теорема 6.9* полностью доказана.

* Мы учитываем, что для обратного отношения выполнены все условия теоремы 6.9*. В частности, производная $f'(x)$ не обращается в нуль в достаточно малой проколотой δ-окрестности точки a. (Это вытекает из существования равного ∞ предела (6.28*) и из необращения в указанной проколотой δ-окрестности производной $g'(x)$.)

Так же как и теорема 6.9, теорема 6.9* остается справедливой в каждом из следующих трех случаев:

1) в случае, если в этой теореме в качестве множества C_δ взять интервал $(a, a+\delta)$ [соответственно $(a-\delta, a)$], а все пределы (6.17*)—(6.20*) взять при $x \to a+0$ [соответственно при $x \to a-0$];

2) в случае, если в качестве множества C_δ взять совокупность всех x, лежащих вне сегмента $[-\delta, \delta]$, а все пределы (6.17*)—(6.20*) взять при $x \to \infty$;

3) в случае, если в качестве множества C_δ взять полупрямую $(\delta, +\infty)$ [соответственно $(-\infty, -\delta)$], а все пределы (6.17*)—(6.20*) взять при $x \to +\infty$ [соответственно при $x \to -\infty$].

Обоснование справедливости теоремы 6.9* в этих трех случаях может быть заимствовано из предыдущего пункта.

Примеры. 1)

$$\lim_{x \to 0+0} \sqrt{x} \ln x = \lim_{x \to 0+0} \frac{\ln x}{x^{+1/2}} =$$

$$= \lim_{x \to 0+0} \frac{\frac{1}{x}}{\left(-\frac{1}{2}\right) x^{-3/2}} = -2 \lim_{x \to 0+0} \sqrt{x} = 0.$$

2) n-кратным применением правила Лопиталя вычисляется предел

$$\lim_{x \to +\infty} \frac{x^n}{e^x} = \lim_{x \to +\infty} \frac{nx^{n-1}}{e^x} = \lim_{x \to +\infty} \frac{n(n-1)x^{n-2}}{e^x} = \ldots = \lim_{x \to +\infty} \frac{n!}{e^x} = 0.$$

3. Раскрытие неопределенностей других видов. Кроме изученных выше неопределенностей видов $\frac{0}{0}$ и $\frac{\infty}{\infty}$, часто встречаются неопределенности следующих видов: $0 \cdot \infty$, $\infty - \infty$, 1^∞, 0^∞, ∞^0.

Все эти неопределенности сводятся к изученным выше двум неопределенностям путем алгебраических преобразований. Покажем это, например, по отношению к последним трем из указанных выше неопределенностей. Каждая из этих неопределенностей имеет вид

$$y = f(x)^{g(x)}, \tag{6.29}$$

где при $x \to a$ функция $f(x)$ стремится соответственно к 1, 0 или ∞, а $g(x)$ стремится соответственно к ∞ или 0. Логарифмируя выражение (6.29), получим (считая, что $f(x) > 0$)

$$\ln y = g(x) \ln f(x). \tag{6.30}$$

Для нахождения предела выражения (6.29) достаточно найти предел выражения (6.30).

Заметим, что в любом из трех рассматриваемых случаев выражение (6.30) представляет собой при $x \to a$ неопределенность вида $0\cdot\infty$. Значит, достаточно научиться сводить неопределенность вида $0\cdot\infty$ к неопределенности вида $\frac{0}{0}$ или $\frac{\infty}{\infty}$. Покажем, как это делается. Итак, пусть

$$z=\varphi(x)\ \psi(x), \tag{6.31}$$

причем

$$\lim_{x\to a}\varphi(x)=0, \quad \lim_{x\to a}\psi(x)=\infty.$$

Перепишем (6.31) в виде

$$z=\varphi(x)\,\psi(x)=\frac{\varphi(x)}{\dfrac{1}{\psi(x)}}. \tag{6.32}$$

Очевидно, выражение (6.32) представляет собой при $x\to a$ неопределенность вида $\frac{0}{0}$. Наша цель достигнута.

Примеры. 1) Вычислить $\lim\limits_{x\to 0+0} x^x$. Обозначим $y=x^x$. Тогда $\ln y=x\ln x=\dfrac{\ln x}{\dfrac{1}{x}}$. Применяя правило Лопиталя, будем иметь

$$\lim_{x\to 0+0}(\ln y)=\lim_{x\to 0+0}\frac{\ln x}{\dfrac{1}{x}}=\lim_{x\to 0+0}\frac{\dfrac{1}{x}}{-\dfrac{1}{x^2}}=-\lim_{x\to 0+0}x=0.$$

Отсюда ясно, что

$$\lim_{x\to 0+0}y=1.$$

2) $\lim\limits_{x\to 0}(1+x^2)^{\frac{1}{e^x-1-x}}$. Пусть $y=(1+x^2)^{\frac{1}{e^x-1-x}}$.

Тогда

$$\ln y=\frac{1}{e^x-1-x}\ln(1+x^2).$$

Пользуясь правилом Лопиталя, получим

$$\lim_{x\to 0}(\ln y)=\lim_{x\to 0}\frac{\ln(1+x^2)}{e^x-1-x}=\lim_{x\to 0}\frac{\dfrac{2x}{1+x^2}}{e^x-1}=$$

$$=\lim_{x\to 0}\frac{2x}{(e^x-1)(1+x^2)}=\lim_{x\to 0}\frac{2}{e^x(1+x^2)+(e^x-1)2x}=2.$$

Отсюда ясно, что $\lim\limits_{x\to 0}y=e^2$.

§ 7. ФОРМУЛА ТЕЙЛОРА

В этом параграфе мы установим одну из важнейших формул математического анализа, имеющую многочисленные приложения как в математике, так и в смежных дисциплинах.

Теорема 6.10 (теорема Тейлора*). *Пусть функция $f(x)$ имеет в некоторой окрестности точки a производную порядка*** $n+1$ (*n — любой фиксированный номер*). *Пусть, далее, x — любое значение аргумента из указанной окрестности, p — произвольное положительное число. Тогда между точками a и x найдется точка ξ такая, что справедлива следующая формула*:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \ldots$$

$$\ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x), \tag{6.33}$$

где

$$R_{n+1}(x) = \left(\frac{x-a}{x-\xi}\right)^p \frac{(x-\xi)^{n+1}}{n!\,p} f^{(n+1)}(\xi). \tag{6.34}$$

Замечание. Так как точка ξ лежит между x и a, то дробь $\frac{x-a}{x-\xi}$ всегда положительна, а поэтому для любого $p>0$ определена степень $\left(\frac{x-a}{x-\xi}\right)^p$.

Формула (6.33) называется формулой Тейлора (с центром в точке a), а выражение $R_{n+1}(x)$ называется остаточным членом. Как мы увидим ниже, остаточный член может быть записан не только в виде (6.34), но и в других видах. Принято называть остаточный член, записанный в виде (6.34), остаточным членом в общей форме***.

Доказательство. Обозначим символом $\varphi(x, a)$ многочлен относительно x порядка n, фигурирующий в правой части (6.33), т. е. положим

$$\varphi(x, a) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \ldots +$$

$$+ \frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{6.35}$$

Далее обозначим символом $R_{n+1}(x)$ разность

$$R_{n+1}(x) = f(x) - \varphi(x, a). \tag{6.36}$$

* Брук Тейлор — английский математик (1685—1731).

** Отсюда вытекает, что сама функция $f(x)$ и ее производные до порядка n непрерывны в указанной окрестности точки a.

*** Эту форму остаточного члена называют также формой Шлемильха—Роша.

Теорема будет доказана, если мы установим, что $R_{n+1}(x)$ определяется формулой (6.34).

Фиксируем любое значение x из окрестности, указанной в формулировке теоремы. Ради определенности будем считать, что $x>a$. Обозначим через t переменную величину, имеющую областью своего изменения сегмент $[a, x]$, и рассмотрим вспомогательную функцию $\psi(t)$ следующего вида:

$$\psi(t)=f(x)-\varphi(x, t)-(x-t)^p Q(x), \tag{6.37}$$

где

$$Q(x)=\frac{R_{n+1}(x)}{(x-a)^p}. \tag{6.38}$$

Подробнее $\psi(t)$ можно записать так:

$$\psi(t)=f(x)-f(t)-\frac{f'(t)}{1!}(x-t)-\ldots$$
$$\ldots-\frac{f^{(n)}(t)}{n!}(x-t)^n-(x-t)^p Q(x). \tag{6.39}$$

Наша цель — выразить $Q(x)$, исходя из свойств введенной нами функции $\psi(t)$.

Покажем, что функция $\psi(t)$ удовлетворяет на сегменте $[a, x]$ всем условиям теоремы 6.3 (Ролля).

Из формулы (6.39) и из условий, наложенных на функцию $f(x)$, очевидно, что функция $\psi(t)$ непрерывна на сегменте $[a, x]$ и дифференцируема во всех внутренних точках этого сегмента*. Убедимся в том, что $\psi(a)=\psi(x)=0$. Полагая в (6.37) $t=a$ и принимая во внимание равенство (6.38), будем иметь

$$\psi(a)=f(x)-\varphi(x, a)-R_{n+1}(x).$$

Отсюда на основании (6.36) получим $\psi(a)=0$. Равенство $\psi(x)=0$ сразу вытекает из формулы (6.39).

Итак, для функции $\psi(t)$ на сегменте $[a, x]$ выполнены все условия теоремы 6.3 (Ролля). На основании этой теоремы внутри сегмента $[a, x]$ найдется точка ξ такая, что

$$\psi'(\xi)=0. \tag{6.40}$$

Подсчитаем производную $\psi'(t)$. Дифференцируя равенство (6.39), будем иметь

---

* Это вытекает из того, что из условия существования у функции $f(t)$ в окрестности точки a производной порядка $n+1$ вытекает непрерывность самой функции $f(t)$ и всех ее производных до порядка n в указанной окрестности точки a, а потому и на сегменте $[a, x]$. Далее можно утверждать, что сама функция $f(t)$ и все ее производные до порядка n один раз дифференцируемы всюду в указанной окрестности точки a, а потому и всюду внутри сегмента $[a, x]$.

$$\psi'(t) = -f'(t) + \frac{f'(t)}{1!} - \frac{f^{(2)}(t)}{1!}(x-t) + \frac{f^{(2)}(t)}{2!}\cdot 2(x-t) - \ldots$$

$$\ldots + \frac{f^{(n)}(t)}{n!} n(x-t)^{n-1} - \frac{f^{(n+1)}(t)}{n!}(x-t)^n + p(x-t)^{p-1} Q(x). \tag{6.41}$$

Легко видеть, что все члены в правой части (6.41), за исключением последних двух, взаимно уничтожаются. Таким образом,

$$\psi'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + p\cdot(x-t)^{p-1}\cdot Q(x). \tag{6.42}$$

Полагая в формуле (6.42) $t=\xi$ и используя равенство (6.40), получим

$$Q(x) = \frac{(x-\xi)^{n-p+1}}{n!\,p} f^{(n+1)}(\xi). \tag{6.43}$$

Сопоставляя (6.43) и (6.38), окончательно будем иметь

$$R_{n+1}(x) = (x-a)^p Q(x) = \left(\frac{x-a}{x-\xi}\right)^p \frac{(x-\xi)^{n+1}}{n!\,p} f^{(n+1)}(\xi).$$

Случай, когда $x<a$, рассматривается совершенно аналогично. Теорема доказана.

Найдем разложение по формуле Тейлора простейшей функции — алгебраического многочлена n-го порядка. Пусть

$$f(x) = C_0x^n + C_1x^{n-1} + \ldots + C_{n-1}x + C_n.$$

Тогда, поскольку $f^{(n+1)}(x)\equiv 0$, остаточный член $R_{n+1}(x)\equiv 0$ и формула Тейлора (6.33) принимает вид

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

(Здесь в качестве a можно взять любую точку бесконечной прямой.) Таким образом, формула Тейлора позволяет представить любой многочлен $f(x)$ в виде многочлена по степеням $(x—a)$, где a — любое вещественное число.

Пусть теперь $f(x)$ — произвольная функция, удовлетворяющая условиям теоремы 6.10. Постараемся выяснить, какими свойствами обладает многочлен (6.35), фигурирующий в формуле Тейлора для этой функции. Как и выше, будем обозначать этот многочлен символом $\varphi(x, a)$. Символом $\varphi^{(n)}(x, a)$ обозначим n-ю производную $\varphi(x, a)$ по x. Дифференцируя формулу (6.35) по x и затем полагая $x=a$, мы получим следующие равенства:

$$\left\{\begin{array}{l} \varphi(a, a) = f(a), \\ \varphi'(a, a) = f'(a), \\ \varphi^{(2)}(a, a) = f^{(2)}(a), \\ \ldots\ldots\ldots\ldots \\ \varphi^{(n)}(a, a) = f^{(n)}(a). \end{array}\right. \tag{6.44}$$

Таким образом, фигурирующий в формуле Тейлора для произвольной функции $f(x)$ *многочлен* $\varphi(x, a)$ *обладает следующим свойством: он сам и его производные до порядка n включительно равны в точке $x=a$ соответственно $f(x)$ и ее производным до порядка n.*

§ 8. РАЗЛИЧНЫЕ ФОРМЫ ОСТАТОЧНОГО ЧЛЕНА. ФОРМУЛА МАКЛОРЕНА

1. Остаточный член в форме Лагранжа, Коши и Пеано. Выше мы установили формулу Тейлора с остаточным членом в общей форме. Здесь мы установим другие возможные представления для остаточного члена. Два из них могут быть получены в качестве частных случаев из общей формы.

Прежде всего несколько преобразуем формулу для остаточного члена (6.34). Поскольку точка ξ лежит между точками a и x, найдется такое число θ * из интервала $0<\theta<1$; что $\xi-a=\theta(x-a)$. При этом $\xi=a+\theta(x-a)$, $x-\xi=(x-a)(1-\theta)$. Таким образом, формула (6.34) может быть переписана в виде

$$R_{n+1}(x)=\frac{(x-a)^{n+1}(1-\theta)^{n-p+1}}{n!\,p}f^{(n+1)}[a+\theta(x-a)]. \qquad (6.45)$$

Рассмотрим теперь два важных частных случая формулы (6.45): 1) $p=n+1$; 2) $p=1$ (напомним, что в формулах (6.34) и (6.45) в качестве p может быть взято любое положительное число). Первый из этих частных случаев ($p=n+1$) приводит нас к остаточному члену в форме Лагранжа:

$$R_{n+1}(x)=\frac{(x-a)^{n+1}}{(n+1)!}f^{(n+1)}[a+\theta(x-a)]. \qquad (6.46)$$

Эта форма остаточного члена наиболее употребительна в приложениях. Остаточный член в форме Лагранжа напоминает следующий, очередной член формулы Тейлора, лишь только $(n+1)$-я производная функции $f(t)$ вычисляется не в точке a, а в некоторой промежуточной между a и x точке $\xi=a+\theta(x-a)$.

Второй из указанных выше частных случаев ($p=1$) приводит нас к остаточному члену в форме Коши:

$$R_{n+1}(x)=\frac{(x-a)^{n+1}(1-\theta)^{n}}{n!}f^{(n+1)}[a+\theta(x-a)]. \qquad (6.47)$$

Так как формы Лагранжа и Коши отвечают разным значениям p, а θ зависит от p, то значения θ в формулах (6.46) и (6.47) являются, вообще говоря, различными. Для оценки некоторых функций форма Коши является более предпочтительной, чем форма Лагранжа. Обе формы остаточного члена (Лагранжа и

* Следует подчеркнуть, что ξ, а значит, и θ зависит от p.

Коши) обычно используются в тех случаях, когда требуется при тех или иных фиксированных значениях x, отличных от a, приближенно вычислить функцию $f(x)$.

Естественно приближенно заменить $f(x)$ многочленом $\varphi(x, a)$ и численно оценить сделанную при этом ошибку. Наряду с этим встречаются задачи, в которых нас интересует не численная величина указанной ошибки, а лишь порядок ее относительно малой величины $(x-a)$. Для этой цели удобна другая форма записи остаточного члена (так называемая форма Пеано*), к установлению которой мы и переходим.

Предположим, что функция $f(x)$ имеет производную порядка $n-1$ в некоторой окрестности точки a и производную порядка n в самой точке a **.

При этих предположениях мы можем рассмотреть многочлен $\varphi(x, a)$, определяемый соотношением (6.35). Разность между $f(x)$ и этим многочленом, как и при доказательстве теоремы 6.10, обозначим символом $R_{n+1}(x)$, т. е. положим

$$R_{n+1}(x) = f(x) - \varphi(x, a).$$

Докажем, что при сделанных нами предположениях для остаточного члена $R_{n+1}(x)$ справедливо следующее представление ***:

$$R_{n+1}(x) = o[(x-a)^n]. \tag{6.48}$$

Представление (6.48) принято называть остаточным членом в форме Пеано.

Используя установленное в конце предыдущего параграфа свойство многочлена $\varphi(x, a)$, выражающееся соотношениями (6.44), мы получим следующие равенства:

$$R_{n+1}(a) = 0, \; R'_{n+1}(a) = 0, \; R^{(2)}_{n+1}(a) = 0, \; \ldots, \; R^{(n)}_{n+1}(a) = 0. \tag{6.49}$$

Нам остается доказать, что из равенств (6.49) вытекает представление (6.48). Доказательство того, что из равенств (6.49) вытекает представление (6.48), проведем методом математической индукции.

Сначала убедимся в том, что равенства (6.49) влекут представление (6.48) при $n=1$.

При $n=1$ (6.49) превращаются в два равенства:

$$R_2(a) = 0, \quad R'_2(a) = 0,$$

из которых сразу же вытекает, что

$$\lim_{x \to a} \frac{R_2(x)}{x-a} = \lim_{x \to a} \frac{R_2(x) - R_2(a)}{x-a} = R'_2(a) = 0,$$

* Джузеппе Пеано — итальянский математик (1858—1932).

** Таким образом, мы подчиняем функцию $f(x)$ требованиям более слабым, чем в теореме 6.10.

*** Напомним, что символ $o[(x-a)^n]$ обозначает бесконечно малую при $x \to a$ функцию более высокого порядка малости, чем $(x-a)^n$ (см. п. 5 § 4 гл. 3).

а это и означает, что $R_2(x)=o(x-a)$, т. е. при $n=1$ представление (6.48) вытекает из (6.49).

Теперь для завершения индукции предположим, что представление (6.48) вытекает из (6.49) для некоторого номера $n\geqslant 1$, и убедимся, что в таком случае представление (6.48) вытекает из (6.49) и для следующего номера $n+1$. Для номера $n+1$ равенства (6.49) имеют вид

$$R_{n+2}(a)=0,\ R'_{n+2}(a)=0,\ R^{(2)}_{n+2}(a)=0,\ \ldots,\ R^{(n+1)}_{n+2}(a)=0. \quad (6.49^*)$$

Поскольку мы предположили, что для номера n равенства вида (6.49) влекут представление вида (6.48), то, отбросив первое из равенств (6.49*), мы можем утверждать, что остальные равенства

$$R'_{n+2}(a)=0,\ R^{(2)}_{n+2}(a)=0,\ \ldots,\ R^{(n+1)}_{n+2}(a)=0$$

влекут представление

$$R'_{n+2}(x)=o\left[(x-a)^n\right]. \quad (6.50)$$

С другой стороны, применяя к функции $R_{n+2}(x)$ по сегменту $[a,\ x]$ теорему 6.4 Лагранжа* и учитывая первое равенство (6.49*), мы получаем, что между a и x найдется точка ξ такая, что

$$R_{n+2}(x)=R_{n+2}(x)-R_{n+2}(a)=R'_{n+2}(\xi)\ (x-a). \quad (6.51)$$

Сопоставляя соотношения (6.50) и (6.51) и учитывая, что $o[(\xi-a)^n]=o[(x-a)^n]$ (в силу того, что ξ заключено между a и x), мы окончательно получим, что $R_{n+2}(x)=o[(x-a)^n]$, а это совпадает с представлением (6.48), взятым для номера $n+1$.

Тем самым индукция проведена и вывод представления (6.48) остаточного члена в форме Пеано завершен.

В заключение запишем полностью формулу Тейлора с остаточным членом в форме Пеано:

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+o\left[(x-a)^n\right]. \quad (6.52)$$

2. Другая запись формулы Тейлора. Часто записывают формулу Тейлора (6.33) в несколько ином виде. Положим в (6.33) $a=x_0$, $x-a=\Delta x$ и возьмем остаточный член в форме Лагранжа (6.46). При этом $x=x_0+\Delta x$, и мы получим

$$f(x_0+\Delta x)-f(x_0)=\frac{f'(x_0)}{1!}\Delta x+\frac{f^{(2)}(x_0)}{2!}(\Delta x)^2+\ldots$$

* При любом $n\geqslant 1$ для $R_{n+2}(x)$ выполнены все условия применимости теоремы 6.4 Лагранжа.

$$\ldots + \frac{f^{(n)}(x_0)}{n!}(\Delta x)^n + \frac{f^{(n-1)}(x_0+\theta\Delta x)}{(n+1)!}(\Delta x)^{n+1}. \qquad (6.53)$$

(Здесь θ — некоторое число из интервала $0<\theta<1$.) Формула Тейлора (6.53) является естественным обобщением формулы Лагранжа (6.5) (см. § 3). Формула Лагранжа (6.5) получается из формулы (6.53) в частном случае $n=0$.

3. Формула Маклорена. Принято называть формулой Маклорена* формулу Тейлора (6.33) с центром в точке $a=0$. Таким образом, формула Маклорена дает представление функции в окрестности точки $x=0$. Запишем формулу Маклорена для произвольной функции с остаточным членом в форме Лагранжа, Коши и Пеано**:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f^{(2)}(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + R_{n+1}(x), \qquad (6.54)$$

где остаточный член имеет вид:

1) в форме Лагранжа

$$R_{n+1}(x) = \frac{x^{n+1}}{(n+1)!}f^{(n+1)}(\theta x) \quad (0<\theta<1); \qquad (6.55)$$

2) в форме Коши

$$R_{n+1}(x) = \frac{x^{n+1}(1-\theta)^n}{n!}f^{(n+1)}(\theta x); \qquad (6.56)$$

3) в форме Пеано

$$R_{n+1}(x) = o(x^n). \qquad (6.57)$$

(Мы использовали формулы (6.46), (6.47) и (6.48).)

Перейдем к оценке остаточного члена в формуле Тейлора — Маклорена, к отысканию разложения по формуле Маклорена важнейших элементарных функций и к рассмотрению различных приложений этой формулы.

§ 9. ОЦЕНКА ОСТАТОЧНОГО ЧЛЕНА. РАЗЛОЖЕНИЕ НЕКОТОРЫХ ЭЛЕМЕНТАРНЫХ ФУНКЦИЙ

1. Оценка остаточного члена для произвольной функции. Оценим для произвольной функции $f(x)$ остаточный член в формуле Маклорена (6.54), взятый в форме Лагранжа (6.55).

Предположим, что рассматриваемая нами функция $f(x)$ обладает следующим свойством: *существует такое вещественное число M, что для всех номеров n и для всех значений аргумента x из рас-*

* Колин Маклорен — английский математик (1698—1746).

** При этом предполагается, что $f(x)$ имеет $(n+1)$-ю производную в окрестности точки $x=0$, а для остаточного члена в форме Пеано $(n-1)$-ю производную в окрестности точки $x=0$ и n-ю производную в самой точке $x=0$.

сматриваемой окрестности точки $x=0$ справедливо неравенство

$$|f^{(n)}(x)| \leqslant M. \tag{6.58}$$

Функцию, обладающую указанным свойством, будем называть ф у н к ц и е й, с о в о к у п н о с т ь в с е х п р о и з в о д н ы х к о т о р о й о г р а н и ч е н а в о к р е с т н о с т и т о ч к и $x=0$.

Из неравенства (6.58) и из того, что $0<\theta<1$ вытекает, что

$$|f^{(n)}(\theta x)| \leqslant M, \tag{6.59}$$

и поэтому из формулы (6.55) получим

$$|R_{n+1}(x)| = \frac{|x|^{n+1}}{(n+1)!}\,|f^{(n+1)}(\theta x)| \leqslant M\,\frac{|x|^{n+1}}{(n+1)!}.$$

Итак, мы получаем следующую универсальную оценку остаточного члена для функции, совокупность всех производных которой ограничена числом M в окрестности точки $x=0$:

$$|R_{n+1}(x)| \leqslant M\,\frac{|x|^{n+1}}{(n+1)!}. \tag{6.60}$$

Напомним, что при любом фиксированном x

$$\lim_{n\to\infty} \frac{|x|^{n+1}}{(n+1)!} = 0$$

(см. пример из п. 4 § 2 гл. 3). Отсюда вытекает, что, выбирая достаточно большой номер n, мы можем сделать правую часть (6.60) как угодно малой. Это дает нам возможность применять формулу Маклорена для приближенного вычисления функций, обладающих указанным свойством, с любой наперед заданной точностью. Приведем примеры функций, совокупность всех производных которых ограничена в окрестности точки $x=0$.

1) $f(x)=e^x$, $f^{(n)}(x)=e^x$. Совокупность всех производных этой функции ограничена на любом сегменте $[-r, r]$ $(r>0)$ числом $M=e^r$.

2) $f(x)=\cos x$ или $f(x)=\sin x$. Совокупность всех производных каждой из этих функций ограничена всюду на бесконечной прямой числом $M=1$.

2. Разложение по формуле Маклорена некоторых элементарных функций. А. $f(x)=e^x$. Поскольку $f^{(n)}(x)=e^x$, $f^{(n)}(0)=1$ для любого n, формула Маклорена (6.54) имеет вид

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + R_{n+1}(x), \tag{6.61}$$

где остаточный член в форме Лагранжа равен

$$R_{n+1}(x) = \frac{x^{n+1}}{(n+1)!}\,e^{\theta x} \quad (0<\theta<1). \tag{6.62}$$

На любом сегменте $[-r, r]$ $(r>0)$ в силу того, что $|e^{\theta x}|<e^r$, получим следующую оценку для остаточного члена:

$$|R_{n+1}(x)| < \frac{r^{n+1}}{(n+1)!}e^r. \tag{6.62*}$$

Б. $f(x)=\sin x$. Поскольку $\quad f^{(n)}(x) = \sin\left(x + n\frac{\pi}{2}\right)$,

$$f^{(n)}(0) = \sin\left(n\frac{\pi}{2}\right) = \begin{cases} 0 & \text{при четном } n, \\ (-1)^{\frac{n-1}{2}} & \text{при нечетном } n, \end{cases}$$

формула Маклорена (6.54) имеет вид

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^{\frac{n-1}{2}} \cdot \frac{x^n}{n!} + R_{n+1}(x), \tag{6.63}$$

где n — нечетное число, а остаточный член в форме Лагранжа равен

$$R_{n+1}(x) = \frac{x^{n+2}}{(n+2)!}\sin\left(\theta x + n\frac{\pi}{2} + \pi\right) \quad (0<\theta<1).$$

Очевидно, что на любом сегменте $[-r, r]$ $(r>0)$ для остаточного члена справедлива следующая оценка:

$$|R_{n+1}(x)| \leqslant \frac{r^{n+2}}{(n+2)!}. \tag{6.64}$$

В. $f(x)=\cos x$. Поскольку $\quad f^{(n)}(x) = \cos\left(x + n\frac{\pi}{2}\right)$,

$$f^{(n)}(0) = \cos\left(n\frac{\pi}{2}\right) = \begin{cases} 0 & \text{при нечетном } n, \\ (-1)^{\frac{n}{2}} & \text{при четном } n, \end{cases}$$

формула Маклорена (6.54) имеет вид

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^{\frac{n}{2}} \frac{x^n}{n!} + R_{n+2}(x), \tag{6.65}$$

где n — четное число, а остаточный член в форме Лагранжа равен

$$R_{n+2}(x) = \frac{x^{n+2}}{(n+2)!}\cos\left(\theta x + n\frac{\pi}{2} + \pi\right) \quad (0<\theta<1).$$

На любом сегменте $[-r, r]$ $(r>0)$ получаем для остаточного члена оценку (6.64).

Г. $f(x)=\ln(1+x)$. Поскольку

$$f^{(n)}(x) = (-1)^{n-1}\frac{(n-1)!}{(1+x)^n}, \quad f(0)=0, \; f^{(n)}(0) = (-1)^{n-1}(n-1)!,$$

формула Маклорена (6.54) имеет вид

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^{n-1}\frac{x^n}{n} + R_{n+1}(x). \quad (6.66)$$

Остаточный член на этот раз запишем и оценим и в форме Лагранжа, и в форме Коши:

$$R_{n+1}(x) = \frac{(-1)^n x^{n+1}}{(n+1)(1+\theta x)^{n+1}} \text{ (в форме Лагранжа).} \quad (6.67)$$

$$R_{n+1}(x) = (-1)^n x^{n+1}\frac{(1-\theta)^n}{(1+\theta x)^{n+1}} \text{ *(в форме Коши).} \quad (6.68)$$

Для оценки функции $\ln(1+x)$ для значений x, принадлежащих сегменту $0 \leqslant x \leqslant 1$, удобнее исходить из остаточного члена в форме Лагранжа (6.67). Переходя в формуле (6.67) к модулям, получим для всех x из сегмента $0 \leqslant x \leqslant 1$

$$|R_{n+1}(x)| < \frac{1}{n+1}. \quad (6.69)$$

Из оценки (6.69) очевидно, что для всех x из сегмента $0 \leqslant x \leqslant 1$ $R_{n+1}(x) \to 0$ при $n \to \infty$.

Оценим теперь функцию $\ln(1+x)$ для отрицательных значений x из сегмента $-r \leqslant x \leqslant 0$, где $0 < r < 1$. Для этого будем исходить из остаточного члена в форме Коши (6.68).

Перепишем этот остаточный член в виде

$$R_{n+1}(x) = (-1)^n \left(\frac{1-\theta}{1+\theta x}\right)^n \frac{x^{n+1}}{1+\theta x}. \quad (6.70)$$

Принимая во внимание, что для рассматриваемых значений x выражение $\frac{1-\theta}{1+\theta x} < 1$, и переходя в формуле (6.70) к модулям, будем иметь

$$|R_{n+1}(x)| < \frac{r^{n+1}}{1-r}. \quad (6.71)$$

Так как $r < 1$, то оценка (6.71) позволяет утверждать, что $\lim\limits_{n\to\infty} R_{n+1} = 0$.

Д. $f(x) = (1+x)^\alpha$, где α — вещественное число. Поскольку

$$f^{(n)}(x) = \alpha(\alpha-1)\ \ldots\ (\alpha-n+1)(1+x)^{\alpha-n},$$

$$f^{(n)}(0) = \alpha(\alpha-1)\ \ldots\ (\alpha-n+1),$$

---

* Еще раз отметим, что в формулах (6.67) и (6.68) значения θ являются, вообще говоря, различными.

формула Маклорена (6.54) имеет вид

$$(1+x)^{\alpha}=1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\ldots+$$
$$+\frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}x^n+R_{n+1}(x), \qquad (6.72)$$

где остаточный член в форме Лагранжа равен

$$R_{n+1}(x)=\frac{\alpha(\alpha-1)\ldots(\alpha-n)}{(n+1)!}(1+\theta x)^{\alpha-(n+1)}x^{n+1}\ (0<\theta<1). \qquad (6.73)$$

В частном случае, когда $\alpha=n$ — натуральное число, $R_{n+1}(x)=0$, и мы получим известную из элементарного курса формулу бинома Ньютона

$$(1+x)^n=1+\frac{n}{1!}x+\frac{n(n-1)}{2!}x^2+\ldots+x^n. \qquad (6.74)$$

Если нужно получить разложение не двучлена $(1+x)^n$, а двучлена $(a+x)^n$, то можно вынести a^n за скобку и воспользоваться формулой (6.74). При этом получим

$$(a+x)^n=a^n\left(1+\frac{x}{a}\right)^n=$$
$$=a^n\left[1+\frac{n}{1!}\left(\frac{x}{a}\right)+\frac{n(n-1)}{2!}\left(\frac{x}{a}\right)^2+\ldots+\left(\frac{x}{a}\right)^n\right].$$

Таким образом, общий случай бинома Ньютона является частным случаем формулы Маклорена.

Е. $f(x)=\operatorname{arctg}x$. Поскольку

$$f^{(n)}(x)=\frac{(n-1)!}{(1+x^2)^{\frac{n}{2}}}\cdot\sin\left[n\left(\operatorname{arctg}x+\frac{\pi}{2}\right)\right]$$

(см. пример 5 из п. 2 § 6 гл. 5), то

$$f^{(n)}(0)=\begin{cases}0 & \text{при четном } n,\\ (-1)^{\frac{n-1}{2}}(n-1)! & \text{при нечетном } n,\end{cases}$$

и формула Маклорена (6.54) принимает вид

$$\operatorname{arctg}x=x-\frac{x^3}{3}+\frac{x^5}{5}-\frac{x^7}{7}+\ldots+(-1)^{\frac{n-1}{2}}\frac{x^n}{n}+R_{n+2}(x), \qquad (6.75)$$

где n — нечетное число, а остаточный член в форме Лагранжа равен

$$R_{n+2}(x) = \frac{x^{n+2}}{(n+2)} \frac{\sin\left[(n+2)\left(\operatorname{arctg}\theta x + \frac{\pi}{2}\right)\right]}{[1+(\theta x)^2]^{\frac{n+2}{2}}} \quad (0<\theta<1).$$

Для остаточного члена на любом сегменте $[-r, r]$ (где $r>0$) справедлива оценка

$$|R_{n+2}(x)| < \frac{r^{n+2}}{(n+2)}. \qquad (6.76)$$

Из оценки (6.76) очевидно, что при любом $r \leqslant 1$ остаточный член $R_{n+2}(x)$ стремится к нулю при $n \to \infty$.

§ 10. ПРИМЕРЫ ПРИЛОЖЕНИЙ ФОРМУЛЫ МАКЛОРЕНА

1. Вычисление числа e на ЭВМ. В п. 3 § 2 гл. 3 мы ввели число e как предел последовательности $\left\{\left(1+\frac{1}{n}\right)^n\right\}$ и получили для e грубую оценку вида $2 \leqslant e \leqslant 3$.

В этом пункте мы укажем, как вычислить число e с любой интересующей нас точностью.

Воспользуемся формулой Маклорена (6.61) и оценкой остаточного члена (6.62*), положив в этих формулах $x=r=1$. Мы получим при этом, что

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} + R_{n+1}(1). \qquad (6.77)$$

где

$$|R_{n+1}(1)| \leqslant \frac{e}{(n+1)!} < \frac{3}{(n+1)!}. \qquad (6.78)$$

Выбирая номер n в (6.77) и (6.78) достаточно большим, мы можем оценить с помощью этих формул число e с любой наперед заданной интересующей нас точностью.

Алгоритм вычисления числа e, основанный на формулах (6.77) и (6.78), легко реализуется на электронно-вычислительных машинах.

Мы приведем пример вычисления числа e по этим формулам при $n=400$. Вычисления велись в МГУ на электронно-вычислительной машине БЭСМ-6 с 600 знаками после запятой. Учитывая возможные ошибки округления, мы отбросили последние 10 знаков и приводим результат вычисления с 590 знаками после запятой *:

2,718281 828459 045235 360287 471352 662497 757247 093699 959574 966967
627724 076630 353547 594571 382178 525166 427427 466391 932003 059921

* Для читателей, знакомых со стандартным алгоритмическим языком АЛГОЛ, на с. 261 приведена записанная на этом языке программа вычислений.

817413 596629 043572 900334 295260 595630 738132 328627 943490 763233
829880 753195 251019 011573 834187 930702 154089 149934 884167 509244
761460 668082 264800 168477 411853 742345 442437 107539 077744 992069
551702 761838 606261 331384 583000 752044 933826 560297 606737 113200
709328 709127 443747 047230 696977 209310 141692 836819 025515 108657
463772 111252 389784 425056 953696 770785 449969 967946 864454 905987
931636 889230 098793 127736 178215 424999 229576 351482 208269 895193
668033 182528 869398 496465 105820 939239 829488 793320 36...

Таким образом, использование формулы Маклорена дает возможность вычислить e с любой интересующей нас точностью.

2. Доказательство иррациональности числа e. Докажем теперь с помощью той же формулы Маклорена (6.77), что число e является иррациональным.

Используем для $R_{n+1}(1)$ формулу (6.62), положив в ней $x=1$. Мы получим, что

$$R_{n+1}(1)=\frac{e^{\theta}}{(n+1)!}, \tag{6.79}$$

где θ — некоторое число, удовлетворяющее неравенствам $0<\theta<1$. Из этих неравенств и из (6.79) вытекает, что $R_{n+1}(1)$ удовлетворяет неравенствам

$$\frac{1}{(n+1)!}<R_{n+1}(1)<\frac{3}{(n+1)!}. \tag{6.80}$$

Итак, для e справедливо представление (6.77) с неравенствами для $R_{n+1}(1)$ вида (6.80).

Предполжим теперь, что число e является рациональным, т. е. его можно представить в виде $e=\frac{m}{n}$, где m и n — некоторые целые положительные числа, второе из которых (n), мы, не ограничивая общности, можем считать не меньшим двух*.

Выбирая в формуле Маклорена (6.77) номер n равным знаменателю рациональной дроби $e=\frac{m}{n}$, мы получим, умножая формулу Маклорена (6.77) на $n!$, что каждое из чисел $n!\,e$ и $n!\left(1+\frac{1}{1!}+\ldots+\frac{1}{n!}\right)$ является целым, в то время как число $n!\,R_{n+1}$ в силу неравенств (6.80) удовлетворяет условиям $\frac{1}{n+1}<n!\;R_{n+1}(1)<$

* Если бы в представлении $e=\frac{m}{n}$ число n оказалось равным единице, то мы взяли бы эквивалентное ему представление $e=\frac{2m}{2n}$, в знаменателе которого стоит число $2n=2$.

$< \frac{3}{n+1}$ и заведомо не является целым *. Таким образом, при умножении формулы Маклорена (6.77) на число $n!$ мы получаем соотношение

$$n!\,e-n!\left(1+\frac{1}{1!}+\frac{1}{2!}+\ldots+\frac{1}{n!}\right)=n!\,R_{n+1}(1),$$

в левой части которого стоит целое число, а в правой части — число, не являющееся целым.

Мы получаем противоречие, которое доказывает, что наше предположение о том, что e — рациональное число, является ошибочным.

Иррациональность e доказана.

3. Вычисление значений тригонометрических функций. Легко убедиться в том, что значения тригонометрических функций $\sin x$ и $\cos x$ для x, принадлежащих сегменту $\left[0, \frac{\pi}{4}\right]$, полностью определяют значения этих функций для всех x. Поэтому мы можем ограничиться вычислением $\sin x$ и $\cos x$ для значений x из указанного сегмента. Желая обеспечить точность 10^{-4}, положим в формуле (6.63) и в оценке (6.64) $n=5$, $r=\frac{\pi}{4}$. Тогда

$$|R_{n+2}(x)| = |R_7(x)| \leqslant \frac{\left(\frac{\pi}{4}\right)^7}{7!} < 10^{-4},$$

и поэтому для любого x, удовлетворяющего условию $|x| \leqslant \frac{\pi}{4}$, с точностью до 10^{-4}

$$\sin x \approx x - \frac{x^3}{6} + \frac{x^5}{120}.$$

Аналогично, полагая в формуле (6.65) и в оценке (6.64) $n=6$, $r=\frac{\pi}{4}$, получим

$$|R_{n+2}(x)| = |R_8| \leqslant \frac{\left(\frac{\pi}{4}\right)^8}{8!} < 10^{-5},$$

* Достаточно учесть, что в представлении $e=\frac{m}{n}$, как уже сказано выше, можно считать $n \geqslant 2$, а при $n \geqslant 2$ из неравенств $\frac{1}{n+1} < n!\,R_{n+1}(1) < \frac{3}{n+1}$ вытекает, что $0 < n!R_{n+1} < 1$, т. е. $n!R_{n+1}(1)$ не является целым.

и поэтому для любого x, удовлетворяющего условию $|x| \leqslant \frac{\pi}{4}$, с точностью до 10^{-5}

$$\cos x \approx 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720}.$$

4. Асимптотическая оценка элементарных функций и вычисление пределов. Формула Тейлора — Маклорена является мощным средством для вычисления тонких пределов.

Из установленного нами в п. 2 § 9 разложения ряда элементарных функций вытекают асимптотические * оценки этих функций, характеризующие их поведение в окрестности точки $x=0$, т. е. при малых значениях $|x|$ с точностью до членов любой степени n малой величины x.

Беря в формулах Маклорена для функций e^x, $\sin x$, $\cos x$, $\ln(1+x)$, $(1+x)^\alpha$ и $\operatorname{arctg} x$ остаточный член в форме Пеано, мы получим, что для любого номера n справедливы следующие оценки:

$$
\begin{gathered}
e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + o(x^n),\\
\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}),\\
\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1}), \qquad (6.81)\\
\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^n \frac{x^n}{n} + o(x^n),\\
(1+x)^\alpha = 1 + \frac{\alpha}{1!} x + \frac{\alpha(\alpha-1)}{2!} x^2 + \ldots\\
\ldots + \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} x^n + o(x^n),\\
\operatorname{arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+2}).
\end{gathered}
$$

Приведем примеры использования асимптотических оценок (6.81).

1°. Привлекая вторую из оценок (6.81), взятую при $n=1$, вычислим предел

$$\lim_{x\to 0} \frac{\sin x - x}{x^3} = \lim_{x\to 0} \frac{x - \frac{x^3}{3!} + o(x^4) - x}{x^3} = \lim_{x\to 0} \left[-\frac{1}{3!} + o(x) \right] = -\frac{1}{3!}.$$

---

* Формулу или оценку, характеризующую поведение функции $f(x)$ при $x \to a$ (в данном случае при $x \to 0$), называют асимптотической.

$$2°.\ I=\lim_{x\to 0}\frac{e^{-\frac{x^2}{2}}-\cos x}{x^3\sin x}.$$

Исходя из вида знаменателя, можно заключить, что определяющую роль должны играть члены четвертого порядка относительно x (ибо $\sin x=x+o(x)$). Пользуясь (6.81), можно записать:

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}+o(x^5), \tag{6.82}$$

$$\sin x=x+o(x), \tag{6.83}$$

$$e^z=1+z+\frac{z^2}{2}+o(z^2). \tag{6.84}$$

Значит, при $z=-\frac{x^2}{2}$ получим

$$e^{-\frac{x^2}{2}}=1-\frac{x^2}{2}+\frac{x^4}{8}+o(x^4). \tag{6.85}$$

Из формул (6.82), (6.83) и (6.85) следует, что

$$I=\lim_{x\to 0}\frac{1-\frac{x^2}{2}+\frac{x^4}{8}+o(x^4)-1+\frac{x^2}{2}-\frac{x^4}{24}}{x^4+o(x^4)}=\lim_{x\to 0}\frac{\frac{1}{12}+\alpha(x)}{1+\alpha(x)}=\frac{1}{12}.$$

(Здесь символом $\alpha(x)$ мы обозначили величину $\frac{o(x^4)}{x^4}$, являющуюся бесконечно малой при $x\to 0$.)

$3°.\ I=\lim\limits_{x\to 0}\left(\cos x+\frac{x^3}{2}\right)^{\frac{1}{x(\sin x-x)}}$. Обозначим через y величину $y=\left(\cos x+\frac{x^2}{2}\right)^{\frac{1}{x(\sin x-x)}}$. Тогда $I=\lim\limits_{x\to 0}y$. Прологарифмируем y (так как при малых x это выражение положительно):

$$\ln y=\frac{1}{x(\sin x-x)}\ln\left(\cos x+\frac{x^2}{2}\right).$$

Вычислим
$$\lim_{x\to 0}\ln y=\lim_{x\to 0}\frac{\ln\left(\cos x+\frac{x^2}{2}\right)}{x(\sin x-x)}.$$

Поскольку $\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^5)$, $\sin x = x - \frac{x^3}{6} + o(x^4)$,

получим $$\lim_{x\to 0} \ln y = \lim_{x\to 0} \frac{\ln\left(1 + \frac{x^4}{24} + o(x^5)\right)}{-\frac{x^4}{6} + o(x^5)}.$$

Учтем теперь, что $\ln(1+z) = z + o(z)$. Из этой формулы $\ln\left(1 + \frac{x^4}{24} + o(x^5)\right) = \frac{x^4}{24} + o(x^4)$. Таким образом,

$$\lim_{x\to 0} \ln y = \lim_{x\to 0} \frac{\frac{x^4}{24} + o(x^4)}{-\frac{x^4}{6} + o(x^5)} = \lim_{x\to 0} \frac{\frac{1}{24} + \frac{o(x^4)}{x^4}}{-\frac{1}{6} + o(x)} = -\frac{1}{4}.$$

Отсюда $$l = \lim_{x\to 0} y = e^{-\frac{1}{4}}.$$

Программа вычисления числа *e*

Система Алгол—БЭСМ6, вариант 10—12—69

```
begin integer i, c, p, n, m; integer array a, b, e [0 : 601];
  m:=400; marg (39, 50, 39, 10, 0, 0);
  e [0]:=1; b [0]:=1;
  for i :=1 step 1 until 601 do
    a [i] : =b [i] : =e [i] : =0;
  for n : =1 step 1 until m do
  begin for i: =0 step 1 until 600 do
    a [i] : =b [i]; c : =a [0];
    for i =0 step 1 until 600 do
    begin b [i] : =c÷n;
      c : = (c—n×b [i])×10+a [i+1] end
    p : =0
    for i:=600 step—1 until 0 do
    begin c : =e [i]+b [i]+p;
      p : =0
      if c<10 then e [i] : =c else
      begin e [i] : =c—10; p : =1 end
    end
end
for n : =1 step 1 until 6 do
begin output ('10/', 'zd.', e [0]);
      for i : =1 step 1 until 590 do
       output ('zd', e [i])
  end
end
```

Глава 7

ИССЛЕДОВАНИЕ ГРАФИКА ФУНКЦИИ И ОТЫСКАНИЕ ЭКСТРЕМАЛЬНЫХ ЗНАЧЕНИЙ

В настоящей главе мы применим разработанный в предыдущих двух главах аппарат дифференциального исчисления для *исследования графика функции* и для *отыскания как локальных, так и глобальных экстремумов функции.*

§ 1. ОТЫСКАНИЕ СТАЦИОНАРНЫХ ТОЧЕК

1. Признаки монотонности функции. Из предыдущей главы мы уже знаем, что изучение вопроса об участках монотонности дифференцируемой функции $f(x)$ сводится к исследованию знака первой производной этой функции.

Для удобства сформулируем еще раз найденные нами в предыдущей главе условия монотонности функции.

1°. *Для того чтобы дифференцируемая на интервале* (a, b) *функция* $f(x)$ *не убывала (не возрастала) на этом интервале, необходимо и достаточно, чтобы производная* $f'(x)$ *этой функции была неотрицательна (неположительна) всюду на этом интервале.*

2°. *Для того чтобы дифференцируемая на интервале* (a, b) *функция* $f(x)$ *возрастала (убывала) на интервале* (a, b), *достаточно, чтобы производная* $f'(x)$ *была положительна (отрицательна) всюду на этом интервале.*

Найдем участки монотонности функции $f(x)=x^3-3x^2-4$. Производная $f'(x)=3x(x-2)$ этой функции положительна при $-\infty<x<0$, отрицательна при $0<x<2$ и положительна при $2<x<+\infty$. Поэтому, согласно сказанному, данная функция $f(x)$ возрастает на полупрямой $(-\infty, 0)$ убывает на интервале $(0, 2)$ и возрастает на полупрямой $(2, +\infty)$.

График этой функции изображен на рис. 7.1.

2. Отыскание стационарных точек. Напомним определения локального максимума и локального минимума функции.

Пусть функция $f(x)$ *определена всюду в некоторой окрестности точки* c. *Тогда эта функция имеет в точке* c *локальный максимум* [*или соответственно локальный минимум*], *если существует такая окрестность точки* c, *что для всех точек этой окрестности значение* $f(c)$ *является наибольшим* [*или соответственно наименьшим*] *среди всех значений* $f(x)$ *этой функции.*

Локальный максимум и локальный минимум объединяются общим названием **локальный экстремум**.

В § 1 предыдущей главы нами было установлено необходимое условие экстремума дифференцируемой в данной точке функции. Это условие имеет следующий вид: *если функция $y = f(x)$ дифференцируема в данной точке c и имеет в этой точке локальный экстремум, то $f'(c) = 0$.*

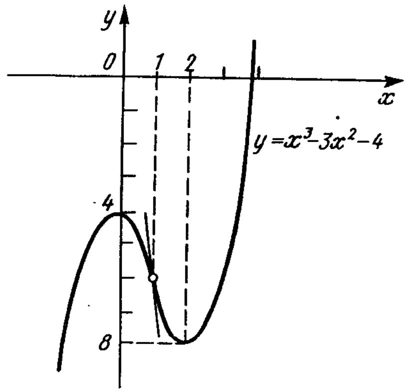


Рис. 7.1

Вместе с тем в § 1 гл. 6 было указано, что обращение в нуль производной является только необходимым и не является достаточным условием локального экстремума дифференцируемой в данной точке функции.

Так, функция $f(x) = x^3$ имеет производную $f'(x) = 3x^2$, обращающуюся в нуль в точке $x = 0$, но никакого экстремума в этой точке $x = 0$ функция $f(x) = x^3$ не имеет (график этой функции см. на рис. 6.2).

Точки, в которых производная $f'(x)$ функции $f(x)$ обращается в нуль, будем называть **стационарными** *точками функции $f(x)$.*

Каждая стационарная точка — это точка возможного экстремума функции.

Однако сделать заключение о том, что в данной стационарной точке на самом деле имеется экстремум, можно лишь на основании дополнительного исследования, для проведения которого мы должны установить **достаточные условия экстремума**.

Такие условия будут установлены в ближайших трех пунктах.

3. Первое достаточное условие экстремума.

Теорема 7.1. *Пусть функция $f(x)$ дифференцируема всюду в некоторой окрестности точки c, и пусть точка c является стационарной точкой функции $f(x)$. Тогда если в пределах указанной окрестности производная $f'(x)$ положительна (отрицательна) слева от точки c и отрицательна (положительна) справа от точки c, то функция $f(x)$ имеет в точке c локальный максимум (минимум). Если же в пределах указанной окрестности точки c производная $f'(x)$ имеет один и тот же знак слева и справа от точки c, то экстремума в точке c нет.*

Доказательство. 1) Пусть сначала производная $f'(x)$ в пределах рассматриваемой окрестности положительна (отрицательна) слева от c и отрицательна (положительна) справа от c. Требуется доказать, что значение $f(c)$ является наибольшим (наименьшим) среди всех значений $f(x)$ в рассматриваемой окрестности. Обозначим через x_0 *любое* значение аргумента из рассматриваемой окрестности, отличное от c. Достаточно доказать, что

$$f(c)-f(x_0)>0 \quad (<0).$$

Так как функция $f(x)$ дифференцируема всюду в рассматриваемой окрестности точки c, то на сегменте, ограниченном точками c и x_0, для функции $f(x)$ выполнены все условия теоремы 6.4 Лагранжа. В силу этой теоремы

$$f(c)-f(x_0)=f'(\xi)(c-x_0), \tag{7.1}$$

где ξ — некоторое значение аргумента между c и x_0. Поскольку производная $f'(\xi)$ положительна (отрицательна) при $x_0<c$ и отрицательна (положительна) при $x_0>c$, правая часть (7.1) положительна (отрицательна).

2) Пусть теперь производная $f'(x)$ имеет один и тот же знак слева и справа от c. Обозначая, как и выше, через x_0 любое значение аргумента, отличное от c, и повторяя проведенные выше рассуждения, мы теперь докажем, что правая часть (7.1) имеет *разные* знаки при $x_0<c$ и при $x_0>c$. Это доказывает отсутствие экстремума в точке c.

Вытекающее из теоремы 7.1 правило можно кратко сформулировать так: 1) *если при переходе через данную стационарную точку c производная $f'(x)$ меняет знак с плюса на минус (с минуса на плюс), то функция $f(x)$ имеет в точке c локальный максимум (минимум);* 2) *если же при переходе через данную стационарную точку c производная $f'(x)$ не меняет знака, то экстремума в точке c нет.*

Примеры. 1) *Найти точки экстремума функции* $f(x)=x^3-3x^2-4$. Поскольку $f'(x)=3x(x-2)$, то функция $f(x)$ имеет две стационарные точки: $x=0$ и $x=2$. При переходе через точку $x=0$ производная меняет знак с плюса на минус, а при переходе через точку $x=2$ — с минуса на плюс. Следовательно, $x=0$ — точка локального максимума, а $x=2$ — точка локального минимума (см. рис. 7.1).

2) *Найти точки экстремума функции* $f(x)=(x-2)^5$. Производная $f'(x)=5(x-2)^4$ обращается в нуль в единственной точке $x=2$. Так как $f'(x)$ положительна как слева, так и справа от этой точки, то функция $f(x)=(x-2)^5$ не имеет точек экстремума. График функции $f(x)=(x-2)^5$ изображен на рис. 7.2.

Иногда вызывает затруднение исследование знака первой производной $f'(x)$ слева и справа от стационарной точки. На этот случай мы укажем другое достаточное условие экстремума в данной стационарной точке c, не требующее исследования знака

$f'(x)$ в окрестности c, но зато *предполагающее существование в точке c отличной от нуля конечной второй производной* $f^{(2)}(x)$.

4. Второе достаточное условие экстремума.

Теорема 7.2. *Пусть функция $f(x)$ имеет в данной стационарной точке c конечную вторую производную. Тогда функция $f(x)$ имеет в точке c локальный максимум, если $f^{(2)}(c)<0$, и локальный минимум, если $f^{(2)}(c)>0$.*

Доказательство. Из условия $f^{(2)}(c)<0$ (>0) и из доказанной в гл. 6 теоремы 6.1 вытекает, что функция $f'(x)$ убывает (возрастает) в точке c. Поскольку по условию $f'(c)=0$, то най-

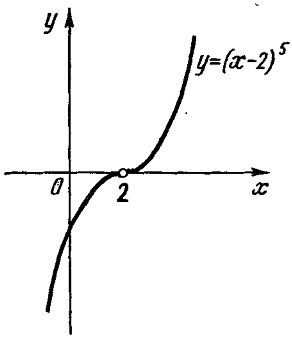


Рис. 7.2

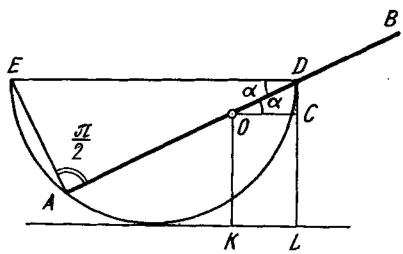


Рис. 7.3

дется такая окрестность точки c, в пределах которой $f'(x)$ положительна (отрицательна) слева от c и отрицательна (положительна) справа от c. Но тогда по предыдущей теореме $f(x)$ имеет в точке c локальный максимум (минимум)

Замечание. Теорема 7.2 имеет, вообще говоря, более узкую сферу действия, чем теорема 7.1. Так, теорема 7.2 не решает вопроса об экстремуме для случая, когда вторая производная $f^{(2)}(x)$ не существует в точке c, а также для случая, когда $f^{(2)}(c)=0$. В последнем случае для решения вопроса о наличии экстремума нужно изучить поведение в точке c производных высших порядков, что будет сделано нами ниже в п. 5.

Примеры. 1) В неподвижную чашку, имеющую форму полушара радиуса r, опущен однородный стержень длины l (рис. 7.3). Предполагая, что $2r<l<4r$, найти положение равновесия стержня.

Положению равновесия стержня соответствует минимальное значение его потенциальной энергии, т. е. наинизшее положение центра его тяжести O (поскольку стержень является однородным, центр тяжести его совпадает с его серединой). Обозначая через OK перпендикуляр к плоскости, на которой стоит чашка, мы сведем задачу к отысканию того положения стержня AB, при котором отрезок OK имеет минимальную длину. Прежде всего вычислим длину отрезка OK как функцию угла α наклона стержня к плоскости, на которой стоит чашка. Пусть прямая DL параллельна прямой OK, а прямая OC перпендикулярна OK (D — точка, в которой стержень опирается на край чашки).

Из рассмотрения прямоугольного треугольника EAD следует, что $AD = ED\cos\alpha = 2r\cos\alpha$. По условию $AO = \frac{l}{2}$. Таким образом,

$$OD = AD - AO = 2r\cos\alpha - \frac{l}{2}.$$

С другой стороны, $DC = DL - OK = r - OK$. Поэтому из рассмотрения прямоугольного треугольника ODC следует, что

$$\sin\alpha = \frac{DC}{OD} = \frac{r - OK}{2r\cos\alpha - \frac{l}{2}}.$$

Таким образом, длина отрезка OK, которую мы обозначим $f(\alpha)$, равна

$$f(\alpha) = r + \frac{l}{2}\sin\alpha - r\sin 2\alpha.$$

Переходим к отысканию того значения угла α, которое доставляет минимум $f(\alpha)$. (Понятно, что мы можем ограничиться значениями угла α из первой четверти.) Так как

$$f'(\alpha) = \frac{l}{2}\cos\alpha - 2r\cos 2\alpha = \frac{l}{2}\cos\alpha + 2r - 4r\cos^2\alpha,$$

то стационарные точки находятся как решения квадратного уравнения

$$4r\cos^2\alpha - \frac{l}{2}\cos\alpha - 2r = 0.$$

Поскольку $\cos\alpha$ в первой четверти положителен, то нам пригоден только положительный корень этого уравнения

$$\cos\alpha_0 = \frac{l + \sqrt{l^2 + 128r^2}}{16r}. \tag{7.2}$$

Хотя по смыслу задачи и ясно, что единственная стационарная точка α_0 является точкой минимума функции $f(\alpha)$, мы установим это строго при помощи теоремы 7.2. Достаточно убедиться в том, что $f^{(2)}(\alpha_0) > 0$. Поскольку

$$f^{(2)}(\alpha) = -\frac{l}{2}\sin\alpha + 4r\sin 2\alpha = 8r\sin\alpha\cdot\left(\cos\alpha - \frac{l}{16r}\right),$$

то в силу (7.2)

$$f^{(2)}(\alpha)=8r\sin\alpha_0\left(\cos\alpha_0-\frac{l}{16r}\right)=\frac{\sin\alpha_0}{2}\sqrt{l^2+128r^2}>0.$$

Тем самым установлено, что положению равновесия стержня отвечает угол наклона его к плоскости, на которой стоит чашка, определяемый формулой (7.2).

2) Еще раз найдем точки экcртемума функции $f(x)=$ $=x^3-3x^2-4$. Стационарными точками этой функции, как мы уже видели выше, являются точки $x=0$ и $x=2$. Так как

$$f^{(2)}(x)=6x-6,\quad f^{(2)}(0)=-6<0,\quad f^{(2)}(2)=6>0,$$

то в силу теоремы 7.2 функция $f(x)$ имеет максимум в точке 0 и минимум в точке 2. Экстремальные значения этой функции равны

$$f_{\max}=f(0)=-4,\quad f_{\min}=f(2)=-8.$$

5. Третье достаточное условие экстремума. Установим еще одно *достаточное* условие локального экстремума, пригодное в случае, когда вторая производная функции в данной стационарной точке обращается в нуль.

Теорема 7.3. *Пусть $n\geqslant 1$ — некоторое нечетное число, и пусть функция $y=f(x)$ имеет производную порядка n в некоторой окрестности точки c и производную порядка $n+1$ в самой точке c. Тогда, если выполнены соотношения*

$$f'(c)=f^{(2)}(c)=\ldots=f^{(n)}(c)=0,\quad f^{(n+1)}(c)\neq 0, \tag{7.3}$$

то функция $y=f(x)$ имеет в точке c локальный экстремум, точнее, локальный максимум при $f^{(n+1)}(c)<0$ и локальный минимум при $f^{(n+1)}(c)>0$.

Доказательство. При $n=1$ теорема 7.3 совпадает с уже доказанной выше теоремой 7.2, так что нужно вести доказательство лишь при нечетном $n\geqslant 3$.

Пусть нечетное число n удовлетворяет условию $n\geqslant 3$, и пусть ради определенности $f^{(n+1)}(c)>0$. Докажем, что функция $y=f(x)$ имеет в точке c локальный минимум.

Так как $f^{(n+1)}(c)>0$, то, в силу теоремы 6.1 о достаточном условии возрастания функции в точке, функция $f^{(n)}(x)$ возрастает в точке c. Но тогда, поскольку $f^{(n)}(c)=0$, можно утверждать, что всюду в достаточно малой окрестности точки c функция $f^{(n)}(x)$ отрицательна слева от c и положительна справа от c.

Заметив это, разложим функцию $f'(x)$ в окрестности точки c по формуле Тейлора, записав остаточный член в форме Лагранжа (см. § 7 и 8 гл. 6). Мы получим, что для всех x из достаточно малой окрестности точки c между x и c найдется точка ξ такая, что

$$f'(x)=f'(c)+\frac{f^{2}(c)}{1!}(x-c)+\ldots+\frac{f^{(n-1)}(c)}{(n-2)!}(x-c)^{n-2}+$$

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

В силу соотношений (7.3) написанное разложение принимает вид

$$f'(x)=\frac{f^{(n)}(\xi)}{(n-1)!}(x-c)^{n-1}. \tag{7.4}$$

Выше мы установили, что для всех x из достаточно малой окрестности точки c производная $f^{(n)}(x)$ отрицательна слева от c и положительна справа от c. Так как ξ лежит между x и c, то для всех x из достаточно малой окрестности точки c величина $f^{(n)}(\xi)$ (а значит, в силу нечетности n и вся правая часть (7.4)) отрицательна слева от c и положительна справа от c.

Итак, с помощью равенства (7.4) мы доказали, что *производная $f'(x)$ для всех x из достаточно малой окрестности точки c отрицательна слева от c и положительна справа от c.*

В этой ситуации, в силу первого достаточного условия экстремума (т. е. теоремы 7.1), функция $y=f(x)$ имеет в точке c локальный минимум.

Случай $f^{(n+1)}(c)<0$ рассматривается совершенно аналогично. Те же рассуждения и формула (7.4) в этом случае дают возможность заключить, что функция $y=f(x)$ имеет в точке c локальный максимум. Теорема полностью доказана.

Замечание. Очень важным является требование нечетности числа n в теореме 7.3. При четном n и при сохранении всех остальных условий теоремы 7.3 никакого экстремума у функции $y=f(x)$ в точке c не будет (см. по этому поводу теорему 7.10 из п. 5 § 3 этой главы).

6. Экстремум функции, недифференцируемой в данной точке. Выше мы рассмотрели вопрос о наличии у функции $f(x)$ экстремума *в такой точке c, в которой функция $f(x)$ дифференцируема.* В этом пункте мы изучим вопрос о наличии экстремума в точке c у такой функции, которая недифференцируема в точке c, но дифференцируема всюду в некоторой окрестности справа и слева от c и, кроме того, непрерывна в точке c.

Оказывается, теорема 7.1 может быть обобщена на случай такой функции. Именно, имеет место следующее утверждение.

Теорема 7.4. *Пусть функция $f(x)$ дифференцируема всюду в некоторой окрестности точки c, за исключением, быть может, самой точки c, и непрерывна в точке c.*

Тогда, если в пределах указанной окрестности производная $f'(x)$ положительна (отрицательна) слева от точки c и отрицательна (положительна) справа от точки c, то функция $f(x)$ имеет в точке c локальный максимум (минимум). Если же производная $f'(x)$ имеет один и тот же знак слева и справа от точки c, то экстремума в точке c нет.

Доказательство в точности совпадает с доказательством теоремы 7.1.

Достаточно заметить, что условия теоремы 7.4 и на этот раз обеспечивают применимость к $f(x)$ теоремы 6.4 Лагранжа по сегменту, ограниченному точками c и x_0, где x_0 — любое число из достаточно малой окрестности точки c.

Примеры. 1) *Найти точки экстремума функции* $f(x) = |x|$. Эта функция дифференцируема всюду на бесконечной прямой, кроме точки $x=0$, и непрерывна в точке $x=0$, причем производная $f'(x)$ равна 1 при $x>0$ и равна -1 при $x<0$.

Теорема 7.1 к этой функции неприменима, а согласно теореме 7.4 она имеет минимум при $x=0$ (рис. 7.4).

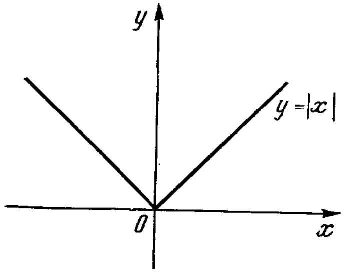


Рис. 7.4

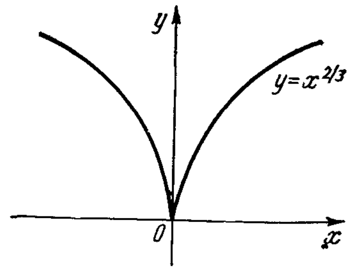


Рис. 7.5

2) *Найти точки экстремума функции* $y=x^{2/3}$. Эта функция непрерывна на всей бесконечной прямой и дифференцируема всюду на этой прямой, за исключением точки $x=0$. Производная при $x \neq 0$ равна

$$y' = \frac{2}{3} \frac{1}{\sqrt[3]{x}}.$$

В предыдущем примере производная имела в точке $x=0$ конечный скачок *, на этот раз производная имеет в точке $x=0$ разрыв 2-го рода («бесконечный скачок»). Из выражения для производной заключаем, что эта производная отрицательна слева от точки $x=0$ и положительна справа от этой точки. Значит, теорема 7.4 позволяет утверждать, что рассматриваемая функция имеет минимум в точке $x=0$ (график рассматриваемой функции изображен на рис. 7.5).

* В том смысле, что эта производная хотя и не существовала в точке $x=0$, но имела в этой точке правое и левое предельные значения, не совпадающие между собой.

3) *Найти точки экстремума функции*

$$y=f(x)=\begin{cases}\dfrac{x}{1+e^{\frac{1}{x}}} & \text{при } x\neq 0,\\ 0 & \text{при } x=0.\end{cases}$$

Легко видеть, что функция непрерывна на всей бесконечной прямой. В самом деле, единственной «сомнительной» точкой является точка $x=0$, но и в этой точке функция непрерывна, ибо

$$\lim_{x\to 0+0} y=\lim_{x\to 0-0} y=0.$$

Далее, очевидно, что рассматриваемая функция дифференцируема на всей бесконечной прямой, кроме точки $x=0$. Всюду, кроме этой точки, производная определяется формулой

$$y'=\frac{1+e^{\frac{1}{x}}+\frac{1}{x}e^{\frac{1}{x}}}{\left(1+e^{\frac{1}{x}}\right)^2}.$$

Легко видеть, что предел $\lim\limits_{x\to 0}\dfrac{f(x)-f(0)}{x}=\lim\limits_{x\to 0}\dfrac{1}{1+e^{\frac{1}{x}}}$ не существует, так что функция $y=f(x)$ недифференцируема в точке $x=0$. Поскольку производная y' положительна и слева, и справа от точки $x=0$, рассматриваемая функция согласно теореме 7.4 не имеет экстремума в точке $x=0$, а значит, и вообще не имеет экстремумов. (График рассматриваемой функции изображен на рис. 7.6.)

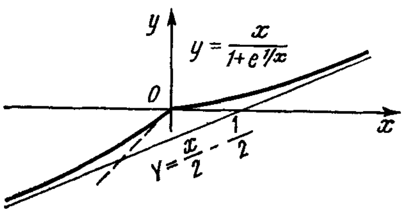


Рис. 7.6

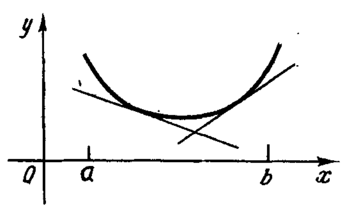


Рис. 7.7

7. Общая схема отыскания экстремумов. Переходим к общей схеме отыскания точек локального экстремума. Предположим, что *функция $f(x)$ непрерывна на интервале* (a, b) и ее производная*

* Вместо интервала (a, b) можно рассматривать бесконечную прямую или открытую полупрямую.

$f'(x)$ существует и непрерывна на этом интервале всюду, кроме конечного числа точек.

Кроме того, предположим, что *производная $f'(x)$ обращается в нуль на интервале (a, b) не более чем в конечном числе точек.* Иными словами, мы предполагаем, что на интервале (a, b) имеется лишь конечное число точек, в которых производная $f'(x)$ не существует или обращается в нуль. Обозначим эти точки символами $x_1, x_2, \ldots, x_n$ $(a<x_1<x_2<\ldots<x_n<b)$. В силу сделанных предположений *производная $f'(x)$ сохраняет постоянный знак на каждом из интервалов (a, x_1), $(x_1, x_2), \ldots, (x_n, b)$.* Значит, *вопрос о наличии экстремума в каждой из точек $x_1, x_2, \ldots, x_n$ может быть решен (в утвердительном или отрицательном смысле) при помощи теоремы 7.4.*

§ 2. ВЫПУКЛОСТЬ ГРАФИКА ФУНКЦИИ

Предположим, что функция $f(x)$ дифференцируема в любой точке интервала (a, b). Тогда, как установлено в п. 3 § 1 гл. 5, *существует касательная к графику функции $y=f(x)$, проходящая через любую точку $M(x, f(x))$ этого графика $(a<x<b)$, причем эта касательная не параллельна* оси Oy.*

Определение. *Будем говорить, что график функции $y=f(x)$ имеет на интервале (a, b) выпуклость, направленную вниз (вверх), если график этой функции в пределах указанного интервала лежит не ниже (не выше) любой своей касательной.*

Замечание 1. Термин «график лежит *не ниже* (или *не выше*) своей касательной» имеет смысл, ибо касательная не параллельна оси Oy.

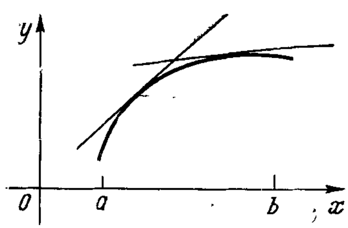


Рис. 7.8

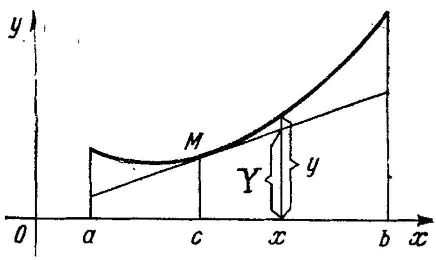


Рис. 7.9

На рис. 7.7 изображен график функции, имеющий на интервале (a, b) выпуклость, направленную вниз, а на рис. 7.8 изображен график функции, имеющий выпуклость, направленную вверх.

* Ибо угловой коэффициент ее, равный производной $f'(x)$, конечен.

Теорема 7.5. *Если функция $y=f(x)$ имеет на интервале (a, b) конечную вторую производную и если эта производная неотрицательна (неположительна) всюду на этом интервале, то график функции $y=f(x)$ имеет на интервале (a, b) выпуклость, направленную вниз (вверх).*

Доказательство. Ради определенности рассмотрим случай, когда вторая производная $f^{(2)}(x)\geqslant 0$ всюду на (a, b). Обозначим через c *любую* точку интервала (a, b) (рис. 7.9). Требуется доказать, что график функции $y=f(x)$ в пределах интервала (a, b) лежит не ниже касательной, проходящей через точку $M(c, f(c))$. Запишем уравнение указанной касательной, обозначая ее текущую ординату через Y. Поскольку угловой коэффициент указанной касательной равен $f'(c)$, то ее уравнение имеет вид*

$$Y-f(c)=f'(c)(x-c). \tag{7.5}$$

Разложим функцию $f(x)$ в окрестности точки c по формуле Тейлора, беря в этой формуле $n=1$. Получим

$$y=f(x)=f(c)+\frac{f'(c)}{1!}(x-c)+\frac{f^{(2)}(\xi)}{2!}(x-c)^2, \tag{7.6}$$

где остаточный член взят в форме Лагранжа, ξ заключено между c и x. (Поскольку по условию $f(x)$ имеет вторую производную на интервале (a, b), формула (7.6) справедлива для любого x из интервала (a, b); см. §§ 7 и 8 гл. 6.)

Сопоставляя (7.6) и (7.5), будем иметь

$$y-Y=\frac{f^{(2)}(\xi)}{2}(x-c)^2. \tag{7.7}$$

Поскольку вторая производная по условию $\geqslant 0$ всюду на (a, b), то правая часть (7.7) *неотрицательна*, т. е. для всех x из (a, b) справедливо $y-Y\geqslant 0$, или $y\geqslant Y$.

Последнее неравенство доказывает, что график функции $y=f(x)$ всюду в пределах интервала (a, b) лежит не ниже касательной (7.5).

Аналогично доказывается теорема для случая $f^{(2)}(x)\leqslant 0$.

Замечание 2. Если $f^{(2)}(x)=0$ всюду на интервале (a, b), то, как легко убедиться, $y=f(x)$ — линейная функция, т. е. график ее есть прямая линия. В этом случае направление выпуклости можно считать произвольным.

Теорема 7.6. *Пусть вторая производная функции $y=f(x)$ непрерывна и положительна (отрицательна) в точке c. Тогда существует такая окрестность*** *точки c, в пределах которой график*

* Уравнение прямой, проходящей через точку $M(a, b)$ и имеющей угловой коэффициент k, имеет вид $Y-b=k(x-a)$ (см. курс аналитической геометрии).

** Напомним, что окрестностью точки c называется интервал, содержащий точку c.

функции $y=f(x)$ имеет выпуклость, направленную вниз (вверх).

Доказательство. По теореме об устойчивости знака непрерывной функции найдется такая окрестность точки c, в пределах которой вторая производная $f^{(2)}(x)$ положительна (отрицательна). По предыдущей теореме график функции $y=f(x)$ имеет в пределах этой окрестности выпуклость, направленную вниз (вверх).

Таким образом, направление выпуклости графика функции полностью характеризуется знаком второй производной этой функции.

Пример. Исследовать направление выпуклости графика функции $y=x^3-3x^2-4$. Из вида второй производной $f^{(2)}(x)=6(x-1)$ вытекает, что эта производная отрицательна при $x<1$ и положительна при $x>1$. Таким образом, выпуклость графика функции $y=x^3-3x^2-4$ направлена вверх на участке $(-\infty, 1)$ и вниз на участке $(1, \infty)$ (см. рис. 7.1).

§ 3. ТОЧКИ ПЕРЕГИБА

1. Определение точки перегиба. Необходимое условие перегиба. Пусть a, b и c — некоторые три числа, связанные неравенствами $a<c<b$. Предположим, что функция $y=f(x)$ дифференцируема на интервале (a, b), т. е. существует касательная к графику этой функции во всех точках, абсциссы которых принадлежат интервалу (a, b). Предположим, кроме того, что график функции $y=f(x)$ имеет определенное направление выпуклости на каждом из интервалов (a, c) и (c, b).

Определение. *Точка $M(c, f(c))$ графика функции $y=f(x)$ называется* ***точкой перегиба*** *этого графика, если существует такая окрестность точки c оси абсцисс, в пределах которой график функции $y=f(x)$ слева и справа от c имеет разные направления выпуклости.*

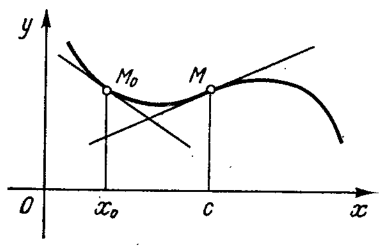


Рис. 7.10

На рис. 7.10 изображен график функции, имеющий перегиб в точке $M(c, f(c))$.

Иногда при определении точки перегиба графика функции $y=f(x)$ дополнительно требуют, *чтобы указанный график всюду в пределах достаточно малой окрестности точки c оси абсцисс слева и справа от c лежал по разные стороны от касательной к этому графику в точке $M(c, f(c))$.* Ниже мы докажем, что это свойство будет вытекать из данного нами определения в предположении, что производная $f'(x)$ является непрерывной в c.

Докажем следующие две леммы.

Лемма 1. *Пусть функция $y=f(x)$ имеет производную $f'(x)$ всюду в δ-окрестности точки c, причем эта производная непрерывна в точке c. Тогда, если график функции $y=f(x)$ имеет на интервале $(c, c+\delta)$ выпуклость, направленную вниз [вверх], то всюду в пределах интервала $(c, c+\delta)$ этот график лежит не ниже [не выше] касательной к графику, проведенной в точке $M(c, f(c))$.*

Доказательство. Рассмотрим последовательность $\{x_n\}$ точек интервала $(c, c+\delta)$, сходящуюся к точке c. Через каждую точку $M_n(x_n, f(x_n))$ графика функции $y=f(x)$ проведем касательную к этому графику, т. е. прямую *

$$Y_n=f(x_n)+f'(x_n)(x-x_n).$$

Так как по условию график функции $y=f(x)$ имеет на интервале $(c, c+\delta)$ выпуклость, направленную вниз [вверх], то для любого номера n и любой *фиксированной* точки x интервала $(c, c+\delta)$

$$f(x)-Y_n=f(x)-f(x_n)-f'(x_n)(x-x_n)\geqslant 0 \quad [\leqslant 0]. \qquad (*)$$

Из условия непрерывности $f'(x)$ (и тем более $f(x)$) в точке c и из определения непрерывности по Гейне вытекает, что существует предел

$$\lim_{n\to\infty}(f(x)-Y_n)=\lim_{n\to\infty}\{f(x)-f(x_n)-f'(x_n)(x-x_n)\}=$$
$$=f(x)-f(c)-f'(c)(x-c).$$

Из существования последнего предела в силу неравенства $(*)$ и теоремы 3.13 из § 1 гл. 3 мы получим, что

$$f(x)-f(c)-f'(c)(x-c)\geqslant 0 \quad [\leqslant 0].$$

Если обозначить через Y текущую ординату касательной (7.5), проходящей через точку $M(c, f(c))$, то последнее неравенство можно переписать в виде

$$f(x)-Y\geqslant 0 \quad [\leqslant 0].$$

Итак, переходя в $(*)$ к пределу при $n\to\infty$ и используя теорему 3.13, мы получим, что $f(x)-Y\geqslant 0$ $[\leqslant 0]$ для любой фиксированной точки x из интервала $(c, c+\delta)$, причем Y обозначает текущую ординату касательной, проведенной через точку $M(c, f(c))$. Лемма доказана.

Замечание. Аналогично формулируется и доказывается лемма 1 и для случая, когда график функции имеет определенное направление выпуклости не на интервале $(c, c+\delta)$, а на интервале $(c-\delta, c)$.

---

* Мы используем уравнение прямой, проходящей через данную точку $M_n(x_n, f(x_n))$ и имеющей угловой коэффициент, равный $f'(x_n)$. Текущую ординату этой прямой обозначаем через Y_n.

Лемма 2. *Пусть функция* $y=f(x)$ *имеет производную* $f'(x)$ *в некоторой окрестности точки* c, *причем эта производная непрерывна в точке* c. *Тогда, если график функции* $y=f(x)$ *имеет перегиб в точке* $M(c, f(c))$, *то в пределах достаточно малой* δ-*окрестности точки* c *этот график слева и справа от* c *лежит по разные стороны от касательной, проведенной через точку* $M(c, f(c))$.

Для доказательства этой леммы следует выбрать $\delta>0$ настолько малым, чтобы на каждом из интервалов $(c-\delta, c)$ и $(c, c+\delta)$ график функции $y=f(x)$ имел определенное направление выпуклости (это направление будет *различным* на интервалах $(c-\delta, c)$ и $(c, c+\delta)$. После этого для доказательства леммы 2 остается применить лемму 1 к функции $y=f(x)$ по каждому из интервалов $(c-\delta, c)$ и $(c, c+\delta)$.

Лемма 2 позволяет нам установить необходимое условие перегиба графика дважды дифференцируемой в данной точке функции $y=f(x)$.

Теорема 7.7 (необходимое условие перегиба графика дважды дифференцируемой функции). *Если функция* $y=f(x)$ *имеет в точке* c *вторую производную и график этой функции имеет перегиб в точке* $M(c, f(c))$, *то* $f^{(2)}(c)=0$.

Доказательство. Пусть, как выше, Y — текущая ордината касательной $Y=f(c)+f'(c)(x-c)$, проходящей через точку графика $M(c, f(c))$.

Рассмотрим функцию

$$F(x)=f(x)-Y=f(x)-f(c)-f'(c)(x-c),$$

равную разности $f(x)$ и линейной функции $f(c)+f'(c)(x-c)$.

Эта функция $F(x)$, как и функция $f(x)$, имеет в точке c вторую производную (а потому имеет первую производную в некоторой окрестности c, причем эта первая производная непрерывна в точке c). В силу леммы 2 в малой окрестности точки c график функции $y=f(x)$ лежит слева и справа от c по разные стороны от касательной, проходящей через точку $M(c, f(c))$, а потому функция $F(x)$ в малой окрестности точки c имеет слева и справа от c *разные знаки*.

Значит, *функция* $F(x)$ *не может иметь в точке* c *локального экстремума*.

Предположим теперь, что $f^{(2)}(c)\neq 0$. Тогда, поскольку $F'(x)=f'(x)-f'(c)$, $F^{(2)}(x)=f^{(2)}(x)$, выполняются условия $F'(c)=0$, $F^{(2)}(c)\neq 0$ и функция $F(x)$ в силу теоремы 7.2 имеет в точке c локальный экстремум. Полученное противоречие доказывает, что предположение $f^{(2)}(c)\neq 0$ является неверным, т. е. $f^{(2)}(c)=0$. Теорема доказана.

Тот факт, что обращение в нуль второй производной является *лишь необходимым* условием перегиба графика дважды дифференцируемой функции, вытекает, например, из рассмотрения графика функции $y=x^4$. Для этой функции вторая производная $y''=12x^2$ обращается в нуль в точке $x=0$, но ее график не имеет перегиба в точке $M(0, 0)$.

В силу теоремы 7.7 для отыскания всех точек перегиба графика дважды дифференцируемой функции $y=f(x)$ нужно рассмотреть все корни уравнения $f^{(2)}(x)=0$.

Поскольку равенство нулю второй производной является лишь необходимым условием перегиба, то нужно дополнительно исследовать вопрос о наличии перегиба в каждой точке, для которой $f^{(2)}(x)=0$. Для проведения такого исследования следует установить *достаточные условия перегиба*, к чему мы и переходим.

2. Первое достаточное условие перегиба.

Теорема 7.8. *Пусть функция $y=f(x)$ имеет вторую производную в некоторой окрестности точки c и $f^{(2)}(c)=0$. Тогда, если в пределах указанной окрестности вторая производная $f^{(2)}(x)$ имеет разные знаки слева и справа от c, то график этой функции имеет перегиб в точке M $(c, f(c))$.*

Доказательство. Заметим, во-первых, что график функции $y=f(x)$ имеет касательную в точке $M(c, f(c))$, ибо из условий теоремы вытекает существование конечной производной $f'(c)$. Далее, из того, что $f^{(2)}(x)$ слева и справа от c имеет разные знаки, и из теоремы 7.5 заключаем, что направление выпуклости слева и справа от c является различным. Теорема доказана.

Пример. Найти точки перегиба графика функции $f(x)=$ $=x^3-3x^2-4$. Эту функцию мы неоднократно рассматривали выше (график ее изображен на рис. 7.1). Поскольку $f^{(2)}(x)=$ $=6x-6=6(x-1)$, то единственное значение аргумента, для которого возможен перегиб, есть $x=1$. Этому значению аргумента соответствует точка графика $M(1, -6)$. Так как $f^{(2)}(x)$ имеет разные знаки при $x>1$ и при $x<1$, то точка $M(1, -6)$ является точкой перегиба графика рассматриваемой функции.

3. Некоторые обобщения первого достаточного условия перегиба. Прежде всего заметим, что *в условиях теоремы 7.8 можно отказаться от требования двукратной дифференцируемости функции $y=f(x)$ в самой точке c, сохраняя это требование лишь для точек, лежащих в некоторой окрестности слева и справа от c. При этом следует дополнительно предположить существование конечной производной $f'(c)$.*

Доказательство теоремы 7.8 с указанными изменениями дословно совпадает с доказательством, приведенным выше.

Далее можно договориться при определении точки перегиба не исключать случая, когда касательная к графику в рассматриваемой точке параллельна оси Oy*. При такой договоренности в теореме 7.8 можно отказаться даже от требования однократной

* В этом случае первая производная $f'(x)$ в точке c принимает бесконечное значение.

дифференцируемости функции $f(x)$ в самой точке c и сформулировать эту теорему следующим образом:

Пусть функция $y=f(x)$ имеет конечную вторую производную всюду в некоторой окрестности точки c, за исключением, быть может, самой точки c. Пусть, далее, функция $y=f(x)$ непрерывна в точке c и график этой функции имеет касательную в точке $M(c, f(c))$. Тогда, если в пределах указанной окрестности вторая производная $f^{(2)}(x)$ имеет разные знаки слева и справа от точки c, то график функции $y=f(x)$ имеет перегиб в точке $M(c, f(c))$.*

Доказательство сформулированного утверждения полностью аналогично доказательству теоремы 7.8.

Пример. Найти точки перегиба графика функции $y=x^{1/3}$. Эта функция имеет вторую производную всюду на бесконечной прямой, за исключением точки $x=0$. В точке $x=0$ рассматриваемая функция непрерывна, но уже первая производная обращается в бесконечность. Однако график функции $y=x^{1/3}$ имеет в точке $(0, 0)$ касательную, параллельную оси Oy** (рис. 7.11).

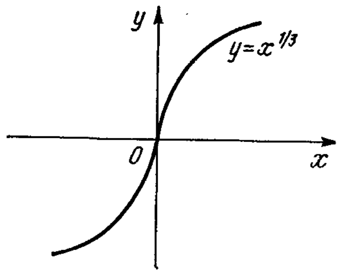


Рис. 7.11

Так как вторая производная $y^{(2)} = -\frac{2}{9}\frac{1}{x^{5/3}}$ имеет слева и справа от точки $x=0$ разные знаки, то график функции $y=x^{1/3}$ имеет перегиб в точке $(0,0)$.

4. Второе достаточное условие перегиба. На случай, когда нежелательно исследование знака второй производной в окрестности точки c, мы сформулируем второе достаточное условие перегиба, предполагающее существование у функции $y=f(x)$ в точке c конечной третьей производной.

Теорема 7.9. *Если функция $y=f(x)$ имеет в точке c конечную третью производную и удовлетворяет в этой точке условиям $f^{(2)}(c)=0$, $f^{(3)}(c)\neq 0$, то график этой функции имеет перегиб в точке $M(c, f(c))$.*

Доказательство. Из условия $f^{(3)}(c)\neq 0$ и из теоремы 6.1 вытекает, что функция $f^{(2)}(x)$ либо возрастает, либо убывает в точке c. Так как $f^{(2)}(c)=0$, то и в том, и в другом случае найдется такая окрестность точки c, в пределах которой $f^{(2)}(x)$ имеет *разные знаки слева и справа от c*. Но тогда по предыдущей теореме график функции $y=f(x)$ имеет перегиб в точке $M(c, f(c))$.

* Быть может, параллельную оси Oy.

** Это вытекает, например, из того, что график обратной функции $x=y^3$ имеет в этой точке касательную $x=0$.

Замечание. Конечно, теорема 7.9 имеет более узкую сферу действия, чем теорема 7.8. Так, теорема 7.9 не решает вопроса о наличии перегиба для случая, когда у функции $y=f(x)$ не существует конечной третьей производной, а также для случая, *когда* $f^{(3)}(c)=0$. В последнем случае для решения вопроса о наличии перегиба нужно изучить поведение в точке c производных высших порядков, что будет сделано нами ниже (см. п. 5).

Возвратимся к примеру, рассмотренному в п. 2, и покажем, что вопрос о наличии перегиба у графика функции $y=x^3-3x^2-4$ может быть решен и при помощи теоремы 7.9. В самом деле, $f^{(3)}(x)=6\neq 0$, значит, точка $M(1, -6)$ является точкой перегиба согласно теореме 7.9.

5. Третье достаточное условие перегиба. Установим еще одно достаточное условие перегиба, пригодное для случая, когда в данной точке c обращаются в нуль как вторая, так и третья производные рассматриваемой функции.

Аналогом теоремы 7.3 является следующее утверждение.

Теорема 7.10. *Пусть $n\geqslant 2$ некоторое четное число, и пусть функция $y=f(x)$ имеет производную порядка n в некоторой окрестности точки c и производную порядка $n+1$ в самой точке c. Тогда, если выполнены соотношения*

$$f^{(2)}(c)=f^{(3)}(c)=\ldots=f^{(n)}(c)=0, \qquad f^{(n+1)}(c)\neq 0, \qquad (7.3^*)$$

то график функции $y=f(x)$ имеет перегиб в точке $M(c, f(c))$.

Доказательство. При $n=2$ теорема 7.10 совпадает с уже доказанной выше теоремой 7.9, так что нужно вести доказательство *лишь для четного $n\geqslant 4$.*

Пусть четное число n удовлетворяет условию $n\geqslant 4$, и пусть $f^{(n+1)}(c)\neq 0$. Тогда, в силу теоремы 6.1 о достаточном условии возрастания или убывания функции в точке, функция $f^{(n)}(x)$ либо убывает в точке c (при $f^{(n+1)}(c)<0$), либо возрастает в этой точке (при $f^{(n+1)}(c)>0$). Поскольку, кроме того, $f^{(n)}(c)=0$, то и в том, и в другом случае всюду в достаточно малой окрестности точки c функция $f^{(n)}(x)$ *имеет разные знаки справа и слева от c.*

Заметив это, разложим функцию $f^{(2)}(x)$ в окрестности точки c по формуле Тейлора, записав остаточный член в форме Лагранжа (см. § 7 и 8 гл. 6). Мы получим, что для всех x из достаточно малой окрестности точки c между x и c найдется точка ξ такая, что

$$f^{(2)}(x)=f^{(2)}(c)+\frac{f^{(3)}(c)}{1!}(x-c)+\ldots+\frac{f^{(n-1)}(c)}{(n-3)!}(x-c)^{n-3}+$$

$$+\frac{f^{(n)}(\xi)}{(n-2)!}(x-c)^{n-2}.$$

В силу соотношений (7.3*) написанное разложение принимает вид

$$f^{(2)}(x) = \frac{f^{(n)}(\xi)}{(n-2)!}(x-c)^{n-2}. \tag{7.4*}$$

Выше мы установили, что для всех x из достаточно малой окрестности точки c производная $f^{(n)}(x)$ имеет разные знаки справа и слева от c. Так как ξ лежит между x и c, то для всех x из достаточно малой окрестности точки c величина $f^{(n)}(\xi)$ (а значит, в силу четности n и вся правая часть (7.4*)) имеет разные знаки справа и слева от c. Итак, в силу равенства (7.4*) для всех x из достаточно малой окрестности точки c производная $f^{(2)}(x)$ имеет разные знаки справа и слева от c. В силу теоремы 7.8 график функции $y=f(x)$ имеет перегиб в точке $M(c, f(c))$. Теорема доказана.

Замечание. Очень важным является требование *четности* n в теореме 7.10 (сравните эту теорему с теоремой 7.3). Рис. 7.12 и 7.13 иллюстрируют исследование на экстремум и перегиб графи-

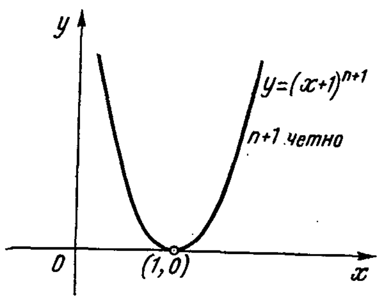


Рис. 7.12

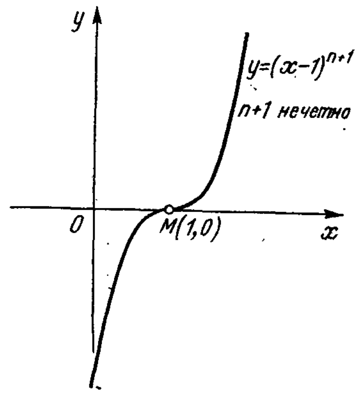


Рис. 7.13

ка функции $y=(x-1)^{n+1}$. В силу теорем 7.3 и 7.10 эта функция имеет минимум в точке $x=1$ при нечетном n, а ее график имеет перегиб в точке $M(1, 0)$ при четном n (проверьте это сами).

§ 4. АСИМПТОТЫ ГРАФИКА ФУНКЦИИ

Определение 1. *Говорят, что прямая $x=a$ является* ***вертикальной асимптотой*** *графика функции $y=f(x)$, если хотя бы один из пределов*

$$\lim_{x\to a+0} f(x) \quad \textit{или} \quad \lim_{x\to a-0} f(x)$$

равен $+\infty$ или $-\infty$.

Пример. График функции $y=\frac{1}{x}$ имеет вертикальную асимптоту $x=0$, ибо $\lim\limits_{x\to 0+0}\frac{1}{x}=+\infty$, $\lim\limits_{x\to 0-0}\frac{1}{x}=-\infty$ (рис. 7.14).

Предположим, далее, что функция $y=f(x)$ определена для сколь угодно больших значений аргумента. Ради определенности будем рассматривать сколь угодно большие значения положительного знака.

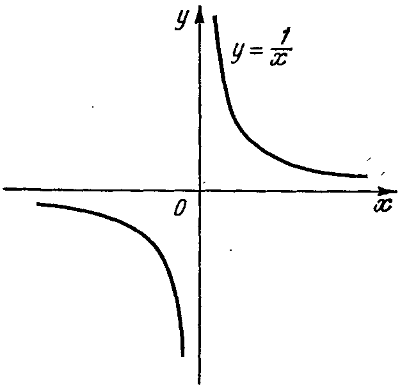


Рис. 7.14

Определение 2. *Говорят, что прямая*

$$Y=kx+b \tag{7.8}$$

является ***наклонной асимптотой*** *графика функции $y=f(x)$ при $x\to+\infty$, если $f(x)$ представима в виде*

$$f(x)=kx+b+\alpha(x), \tag{7.9}$$

где

$$\lim_{x\to+\infty}\alpha(x)=0.$$

Теорема 7.11. *Для того чтобы график функции $y=f(x)$ имел $x\to+\infty$ наклонную асимптоту* (7.8), *необходимо и достаточно, чтобы существовали два предела*

$$\lim_{x\to+\infty}\frac{f(x)}{x}=k \text{ и } \lim_{x\to+\infty}[f(x)-kx]=b. \tag{7.10}$$

Доказательство. 1) *Необходимость.* Пусть график функции $y=f(x)$ имеет при $x\to+\infty$ асимптоту (7.8), т. е. для $f(x)$ справедливо представление (7.9). Тогда

$$\lim_{x\to+\infty}\frac{f(x)}{x}=\lim_{x\to+\infty}\frac{kx+b+\alpha(x)}{x}=\lim_{x\to+\infty}\left[k+\frac{b}{x}+\frac{\alpha(x)}{x}\right]=k,$$

$$\lim_{x\to+\infty}[f(x)-kx]=\lim_{x\to+\infty}[b+\alpha(x)]=b.$$

2) *Достаточность.* Пусть существуют пределы (7.10). Второй из этих пределов дает право утверждать, что разность $f(x)-kx-b$ является бесконечно малой при $x\to+\infty$. Обозначив эту бесконечно малую через $\alpha(x)$, получим для $f(x)$ представление (7.9). Теорема доказана.

Замечание. Аналогично определяется наклонная асимптота и доказывается теорема 7.11 и для случая $x\to-\infty$.

Пример. График функции $y=\frac{2x^2+x}{x+1}$ имеет наклонную асимптоту $Y=2x-1$ и при $x\to+\infty$, и при $x\to-\infty$ и, кроме того, имеет вертикальную асимптоту $x=-1$ (рис. 7.15). В самом деле,

$$\lim_{x\to\pm\infty}\frac{f(x)}{x}=\lim_{x\to\pm\infty}\frac{2x^2+x}{x(x+1)}=2,$$

$$\lim_{x\to\pm\infty}[f(x)-2x]=\lim_{x\to\pm\infty}\left[-1+\frac{1}{x+1}\right]=-1,$$

$$\lim_{x\to-1+0}f(x)=+\infty,\quad \lim_{x\to-1-0}f(x)=-\infty.$$

Наряду с линейной асимптотой (7.8) рассматривают также и асимптоты более сложного вида.

Говорят, что парабола n-го порядка, определяемая многочленом

$$Y=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0, \tag{7.8*}$$

является ***асимптотой*** *графика функции $y=f(x)$ при $x\to+\infty$, если функция $f(x)$ представима в виде*

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0+\alpha(x),$$

$$\lim_{x\to+\infty}\alpha(x)=0.$$

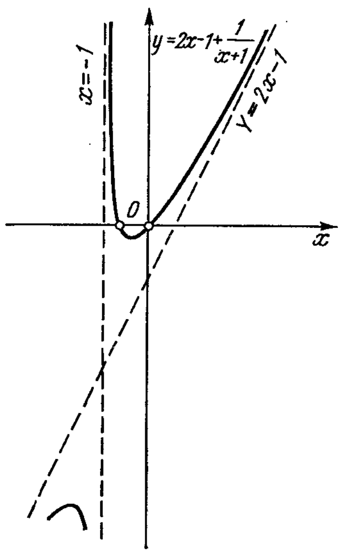


Рис. 7.15

Легко доказать следующее утверждение.

Для того чтобы график функции $y=f(x)$ имел при $x\to+\infty$ асимптоту (7.8), необходимо и достаточно, чтобы существовали следующие $n+1$ пределов:*

$$\lim_{x\to+\infty}\frac{f(x)}{x^n}=a_n,\quad \lim_{x\to+\infty}\frac{f(x)-a_nx^n}{x^{n-1}}=a_{n-1},\ \ldots,$$

$$\lim_{x\to+\infty}\frac{f(x)-(a_nx^n+a_{n-1}x^{n-1}+\ldots+a_2x^2)}{x}=a_1,$$

$$\lim_{x\to+\infty}[f(x)-(a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x)]=a_0.$$

§ 5. ПОСТРОЕНИЕ ГРАФИКА ФУНКЦИИ

В этом параграфе мы изложим схему, по которой целесообразно проводить исследование графика функции, и приведем пример, иллюстрирующий эту схему.

Для исследования графика функции $y=f(x)$ целесообразно прежде всего провести следующие исследования:

1°. Уточнить область задания функции.

2°. Выяснить вопрос о существовании асимптот (вертикальных и наклонных).

3°. Найти области возрастания и убывания функции и точки экстремума.

4°. Найти области сохранения направления выпуклости и точки перегиба.

5°. Найти точки пересечения графика функции с осью Ox.

По полученным данным легко строится эскиз графика функции. В качестве примера построим график функции

$$y=\frac{x^3-5x^2+19x-15}{x^2}. \tag{7.11}$$

Будем следовать изложенной выше схеме.

1°. Поскольку функция (7.11) представляет собой рациональную дробь, то она определена и непрерывна всюду на бесконечной прямой, кроме точки $x=0$, в которой обращается в нуль знаменатель.

2°. Выясним вопрос о существовании асимптот. Очевидно, что

$$\lim_{x\to 0\pm 0}\frac{x^3-5x^2+19x-15}{x^2}=-\infty,$$

поэтому график функции имеет *вертикальную асимптоту* $x=0$.

Далее, из существования пределов

$$\lim_{x\to\pm\infty}\frac{f(x)}{x}=\lim_{x\to\pm\infty}\frac{x^3-5x^2+19x-15}{x^3}=$$

$$=\lim_{x\to\pm\infty}\left(1-\frac{5}{x}+\frac{19}{x^2}-\frac{15}{x^3}\right)=1,$$

$$\lim_{x\to\pm\infty}[f(x)-x]=\lim_{x\to\pm\infty}\frac{x^3-5x^2+19x-15-x^3}{x^2}=$$

$$=\lim_{x\to\pm\infty}\left(-5+\frac{19}{x}-\frac{15}{x^2}\right)=-5$$

вытекает, что при $x\to+\infty$ и при $x\to-\infty$ график функции имеет *наклонную асимптоту* $Y=x-5$.

3°. Для нахождения областей возрастания и убывания вычислим первую производную функции (7.11):

$$y'=\frac{x^3-19x+30}{x^3}=\frac{(x+5)(x-2)(x-3)}{x^3}.$$

Имея в виду, кроме того, что сама функция и первая производная не существуют при $x=0$, мы получим следующие области сохранения знака y':

Область значения x	$-\infty<x<-5$	$-5<x<0$	$0<x<2$	$2<x<3$	$3<x<+\infty$
Знак y'	+	—	+	—	+
Поведение функции	возрастает	убывает	возрастает	убывает	возрастает

Из приведенной таблицы очевидно, что функция имеет следующие точки экстремума:

1) максимум при $x=-5$, причем $f(-5)=-14, 4$;
2) максимум при $x=2$, причем $f(2)=2{,}75$;
3) минимум при $x=3$, причем $f(3)=2{,}666\ldots$.

4°. Для нахождения областей сохранения направления выпуклости вычислим вторую производную

$$y^{(2)}=\frac{38x-90}{x^4}=\frac{38\left(x-\frac{45}{19}\right)}{x^4}.$$

Имея в виду, что сама функция и ее производные не существуют в точке $x=0$, мы получим следующие области сохранения знака $y^{(2)}$:

Область значений x	$-\infty<x<0$	$0<x<\frac{45}{19}$	$\frac{45}{19}<x<+\infty$
Знак $y^{(2)}$	—	—	+
Направление выпуклости графика	вверх	вверх	вниз

Из приведенной таблицы очевидно, что график функции имеет перегиб в точке $\left(\frac{45}{19},\ f\left(\frac{45}{19}\right)\right)$. Легко подсчитать, что

$$f\left(\frac{45}{19}\right)=\frac{6968}{2565}\approx 2{,}72.$$

5°. Остается найти точки пересечения графика с осью Ox. Эти точки соответствуют вещественным корням уравнения

$$x^3-5x^2+19x-15=0.$$

Легко видеть, что

$$x^3-5x^2+19x-15=(x-1)(x^2-4x+15).$$

Поскольку квадратный трехчлен $x^2-4x+15$ не имеет вещественных корней, то рассматриваемое уравнение имеет только один ве-

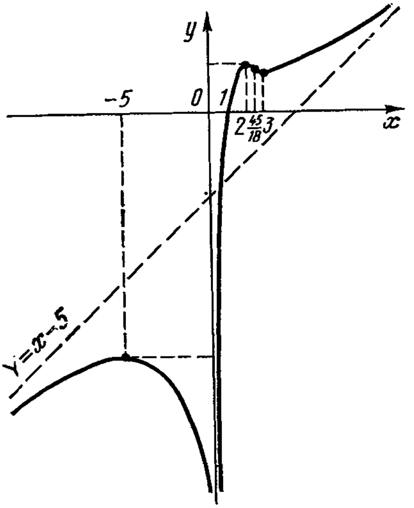


Рис. 7.16

щественный корень $x=1$, так что график функции пересекает ось Ox в точке (1,0). По полученным данным строим эскиз графика рассматриваемой функции (рис. 7.16).

§ 6. ГЛОБАЛЬНЫЕ МАКСИМУМ И МИНИМУМ ФУНКЦИИ НА СЕГМЕНТЕ. КРАЕВОЙ ЭКСТРЕМУМ

1. Отыскание максимального и минимального значений функции, определенной на сегменте. Рассмотрим функцию $y=f(x)$, определенную на сегменте $[a, b]$ и *непрерывную* на нем. До сих пор мы занимались лишь отысканием локальных максимумов и минимумов функции. Теперь поставим задачу об отыскании глобальных максимумов и минимумов или, по-другому, об отыскании максимального и минимального значений $f(x)$ на сегменте $[a, b]$. Подчеркнем, что в силу теоремы Вейерштрасса (см. теорему 4.15 из гл. 4) непрерывная функция $f(x)$ обязательно достигает в некоторой точке сегмента $[a, b]$ своего максимального (минимального) значения. Ради определенности остановимся на отыскании максимального значения $f(x)$ на сегменте $[a, b]$.

Максимальное значение функции $f(x)$ может достигаться либо во внутренней точке x_0 сегмента $[a, b]$ (тогда оно совпадает с

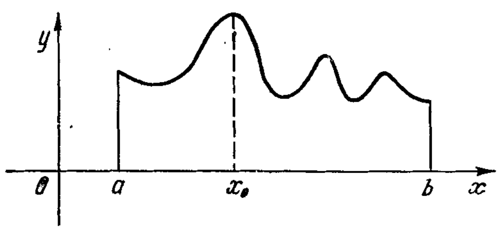


Рис. 7.17

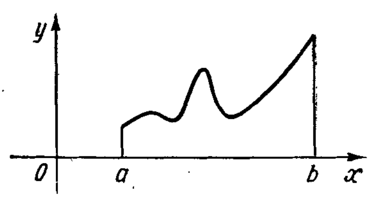


Рис. 7.18

одним из локальных максимумов функции $f(x)$) (рис. 7.17), либо на одном из концов сегмента $[a, b]$ (рис. 7.18). Отсюда ясно, что *для нахождения максимального значения функции $f(x)$ на сегменте $[a, b]$ нужно сравнить между собой значения $f(x)$ во всех* *точках локального максимума и в граничных точках сегмента a и b. Наибольшее из этих значений и будет максимальным значением $f(x)$ на сегменте $[a, b]$. Аналогично находится и минимальное значение $f(x)$ на сегменте $[a, b]$.*

Если желательно избежать исследования стационарных точек, то можно просто сравнить между собой значения $f(x)$ во всех стационарных точках и в граничных точках a и b. Наибольшее (наименьшее) из этих значений, очевидно, и будет максимальным (минимальным) значением функции $f(x)$ на сегменте $[a, b]$.

Отметим далее, что если $f(x)$ имеет на сегменте $[a, b]$ *лишь одну* точку локального максимума (или *лишь одну* точку локального минимума), то без сравнения значения $f(x)$ в этой точке с $f(a)$ и $f(b)$ можно утверждать, что это значение является максимальным (минимальным) значением $f(x)$ на сегменте $[a, b]$ (рис. 7.19).

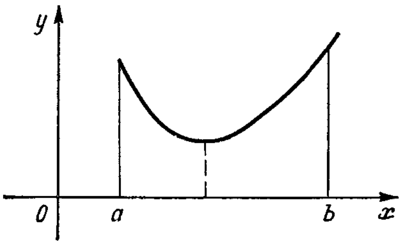


Рис. 7.19

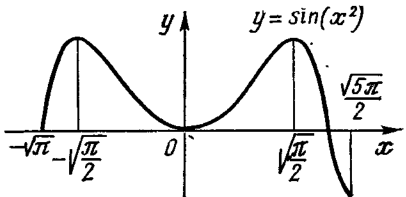


Рис. 7.20

Аналогичными средствами решается вопрос об отыскании максимального (минимального) значения функции $y=f(x)$ на интервале, полупрямой и бесконечной прямой (при условии, что это значение существует).

Может случиться так, что дифференцируемая функция вовсе не имеет на сегменте $[a, b]$ (или полупрямой $a \leqslant x < +\infty$) стационарных точек.

В таком случае $f(x)$ является монотонной на этом сегменте (полупрямой) и ее максимальное и минимальное значения достигаются на концах этого сегмента (на конце этой полупрямой).

В качестве примера *рассмотрим задачу об отыскании максимального и минимального значений функции* $y=\sin(x^2)$ на сегменте $-\sqrt{\pi} \leqslant x \leqslant \frac{\sqrt{5\pi}}{2}$.

Поскольку $y'=2x\cos(x^2)$, указанная функция имеет на рассматриваемом сегменте три стационарные точки: $x=0$ и $x = \pm\sqrt{\frac{\pi}{2}}$. Сравнивая значения функции в указанных точках и на концах сегмента

$$f(0)=0,\ f\left(\pm\sqrt{\frac{\pi}{2}}\right)=1,\ f(-\sqrt{\pi})=0,$$

$$f\left(\frac{\sqrt{5\pi}}{2}\right)=\sin\frac{5\pi}{4}=-\frac{\sqrt{2}}{2},$$

убедимся в том, что максимальное значение рассматриваемой функции равно $+1$ и достигается в двух внутренних точках сегмента: $x_1=-\sqrt{\frac{\pi}{2}}$ и $x_2=\sqrt{\frac{\pi}{2}}$, а минимальное значение рассматриваемой функции равно $-\frac{\sqrt{2}}{2}$ и достигается на правом конце сегмента $\frac{\sqrt{5\pi}}{2}$.

График рассматриваемой функции изображен на рис. 7.20.

2. Краевой экстремум. Пусть функция $y=f(x)$ определена на некотором сегменте $[a, b]$. Будем говорить, что эта *функция имеет в граничной точке b этого сегмента краевой максимум (краевой минимум), если найдется левая полуокрестность точки b, в пределах которой значение $f(b)$ является наибольшим (наименьшим) среди всех других значений этой функции.*

Аналогично определяются краевой максимум и краевой минимум в граничной точке a сегмента $[a, b]$.

Краевой максимум и краевой минимум объединяются общим названием: краевой экстремум.

Имеет место следующее достаточное условие краевого экстремума: *для того чтобы функция $y=f(x)$ имела в точке b сегмента $[a, b]$ краевой максимум (краевой минимум), достаточно, чтобы эта функция имела в точке b положительную (отрицательную) левую производную**. (Доказательство совершенно аналогично доказательству теоремы 7.1.) Из указанного достаточного условия краевого экстремума непосредственно вытекает следующее необходимое условие краевого экстремума функции, имеющей в точке b левую производную: *для того чтобы функция $y=f(x)$, обладающая в точке b левой производной, имела в этой точке краевой максимум (краевой минимум), необходимо, чтобы указанная производная была неотрицательной (неположительной).*

Аналогично, *для того чтобы функция $y=f(x)$, обладающая в точке a правой производной, имела в этой точке краевой макси-*

---

* Для граничной точки a достаточным условием краевого максимума (краевого минимума) является отрицательность (положительность) правой производной в точке a.

мум (краевой минимум), необходимо, чтобы указанная производная была неположительной (неотрицательной).

3. Теорема Дарбу*.

Определение. *Будем говорить, что функция $f(x)$ имеет производную на сегменте $[a, b]$, если $f(x)$ имеет конечную производную в каждой внутренней точке $[a, b]$ и, кроме того, имеет конечные односторонние производные $f'(a+0)$ и $f'(b-0)$.*

Очевидно, функция, имеющая производную на сегменте, будет непрерывной на этом сегменте**.

Докажем теперь следующую теорему.

Теорема 7.12 (теорема Дарбу). *Пусть функция $f(x)$ имеет производную на сегменте $[a, b]$. Тогда, каково бы ни было число C, заключенное между $A=f'(a+0)$ и $B=f'(b-0)$, на этом сегменте найдется точка ξ такая, что $f'(\xi)=C$.*

Итак, производная при одном только условии существования на сегменте $[a, b]$ принимает любое промежуточное значение.

Доказательство. Сначала докажем следующее утверждение: *если $F(x)$ имеет конечную производную на $[a, b]$ и если $F'(a+0)$ и $F'(b-0)$ — числа разных знаков, то на сегменте $[a, b]$ найдется точка ξ такая, что $F'(\xi)=0$.*

Пусть ради определенности $F'(a+0)<0$, $F'(b-0)>0$. Тогда функция $F(x)$ имеет краевой максимум на обоих концах сегмента $[a, b]$. Но это означает, что минимальное значение $F(x)$ на сегменте $[a, b]$ достигается в некоторой внутренней точке ξ этого сегмента (функция $F(x)$ имеет производную, а значит, и непрерывна на сегменте $[a, b]$ и поэтому достигает на этом сегменте своего минимального значения). В указанной точке ξ функция $F(x)$ имеет локальный минимум, и поэтому $F'(\xi)=0$.

Для доказательства теоремы 7.12 остается положить $F(x)=f(x)-Cx$*** и применить к $F(x)$ только что доказанное утверждение.

Заметим, что непрерывность производной $f'(x)$ мы не предполагали.

Из теоремы Дарбу сразу же следует доказанное в п. 3 § 4 гл. 6 утверждение об отсутствии у прозводной точек разрыва первого рода.

---

* Гастон Дарбу — французский математик (1842—1917).

** В самом деле, из существования производной $f'(x)$ во внутренних точках $[a, b]$, вытекает непрерывность $f(x)$ во внутренних точках $[a, b]$, а из существования односторонних производных $f'(a+0)$ и $f'(b-0)$ вытекает непрерывность справа в точке a и слева в точке b.

*** При этом мы, не ограничивая общности, предполагаем, что $f'(a+0)=A<C<B=f'(b-0)$.

ДОПОЛНЕНИЕ

Алгоритм отыскания экстремальных значений функции, использующий только значения этой функции

Предположим, что функция $f(x)$ задана на сегменте $[a, b]$, мы располагаем значениями этой функции в узлах сетки, получающейся при делении сегмента $[a, b]$ на 2^n равных частей ($n=1, 2, 3, \ldots$).

Ради определенности остановимся на отыскании точки минимума функции $f(x)$.

При этом мы будем предполагать, что выполнены следующие два условия: 1) функция $f(x)$ имеет на сегменте $[a, b]$ единственную точку минимума c; 2) при $a<c$ функция $f(x)$ убывает на сегменте $[a, c]$ (т. е. убывает слева от точки минимума), а при $c<b$ функция $f(x)$ возрастает на сегменте $[c, b]$ (т. е. возрастает справа от точки минимума).

Эти условия будут выполнены, например, в случае, если функция $f(x)$ два раза дифференцируема на сегменте $[a, b]$, причем $f'(c)=0$, а $f''(x)$ строго положительна на $[a, b]$. Однако для выполнения указанных двух условий дифференцируемость $f(x)$, вообще говоря, не требуется.

Мы сейчас укажем алгоритм построения стягивающейся системы сегментов*, содержащих точку c минимума функции $f(x)$.

Остановимся на построении первого сегмента стягивающейся системы, имея в виду, что все последующие сегменты этой системы строятся по тому же принципу, что и первый ее сегмент.

Разделим сегмент $[a, b]$ при помощи точек $a=x_0$, x_1, x_2, x_3 и $x_4=b$ на четыре **равных** частичных сегмента $[x_{i-1}, x_i]$ ($i=1, 2, 3, 4$).

Данный частичный сегмент $[x_{i-1}, x_i]$ договоримся называть **сегментом убывания**, если $f(x_{i-1})>f(x_i)$, т. е. если значение функции $f(x)$ на левом его конце строго больше, чем на правом, и соответственно **сегментом возрастания**, если $f(x_{i-1})<f(x_i)$, т. е. если значение функции $f(x)$ на левом конце строго меньше, чем на правом.

Так как функция $f(x)$ имеет на сегменте $[a, b]$ единственную точку минимума c, то указанная точка минимума c принадлежит одному из четырех частичных сегментов $[x_{i-1}, x_i]$.

Тот частичный сегмент $[x_{i-1}, x_i]$, которому принадлежит точка минимума c, можем являться либо сегментом возрастания, либо сегментом убывания, либо, наконец, сегментом, на концах которого функция $f(x)$ имеет равные значения.

---

* Определение и свойства стягивающейся системы сегментов см. в п. 2 § 2 гл. 3.

Кроме того, из того, что по условию функция $f(x)$ убывает слева от точки минимума c и возрастает справа от этой точки, вытекает, что *если данный частичный сегмент содержит точку минимума c, то любой частичный сегмент, лежащий налево от данного, является сегментом убывания, а любой частичный сегмент, лежащий направо от данного, является сегментом возрастания.*

Но тогда можно утверждать, что тот частичный сегмент, который содержит точку минимума c, является либо самым правым сегментом убывания, либо самым левым сегментом возрастания, либо, наконец, частичным сегментом, на концах которого $f(x)$ имеет равные значения.

Сформулированное утверждение позволяет указать алгоритм построения первого сегмента $[a_1, b_1]$ стягивающейся системы сегментов $[a_n, b_n]$, каждый из которых содержит точку минимума c.

Рассмотрим четыре возможных случая.

1) Среди частичных сегментов $[x_{i-1}, x_i]$ есть сегмент, на концах которого $f(x)$ имеет равные значения. В этом случае этот сегмент содержит точку минимума c*, и мы примем его за первый сегмент $[a_1, b_1]$ стягивающейся системы.

2) Все частичные сегменты $[x_{i-1}, x_i]$ $(i=1, 2, 3, 4)$ являются сегментами убывания. В этом случае точка минимума c лежит на самом правом из частичных сегментов, т. е. на сегменте $[x_3, x_4]$, и мы примем этот сегмент за $[a_1, b_1]$.

3) Все частичные сегменты $[x_{i-1}, x_i]$ $(i=1, 2, 3, 4)$ являются сегментами возрастания. В этом случае точка минимума лежит на самом левом из частичных сегментов, т. е. на сегменте $[x_0, x_1]$, и мы примем этот сегмент за $[a_1, b_1]$.

4) Среди частичных сегментов имеются как сегменты убывания, так и лежащие правее их сегменты возрастания. В этом случае можно утверждать, что точка минимума c лежит на объединении самого правого сегмента убывания и самого левого сегмента возрастания. Указанное объединение двух частичных сегментов мы и примем за $[a_1, b_1]$.

Тем самым мы указали однозначный алгоритм построения первого сегмента $[a_1, b_1]$ из стягивающейся системы сегментов $\{[a_n, b_n]\}$.

Второй сегмент этой системы $[a_2, b_2]$ строится, отправляясь от $[a_1, b_1]$, точно так же, как сегмент $[a_1, b_1]$ строился, отправляясь от $[a, b]$. По такому же принципу, отправляясь от n-го сегмента $[a_n, b_n]$, строится $(n+1)$-й сегмент стягивающейся системы.

---

* Точка минимума c не может лежать налево от частичного сегмента $[x_{i-1}, x_i]$, на концах которого $f(x)$ имеет равные значения, ибо при этом указанный сегмент является сегментом возрастания. Аналогично предположение о том, что c лежит направо от частичного сегмента, на концах которого $f(x)$ имеет равные значения, привело бы к тому, что этот частичный сегмент является сегментом убывания.

Ясно, что построенная система сегментов $\{[a_n, b_n]\}$ является стягивающейся и, поскольку все эти сегменты содержат точку минимума c, обе последовательности правых концов $\{b_n\}$ этих сегментов и левых их концов $\{a_n\}$ сходятся к точке минимума c.

Аналогично строится алгоритм отыскания точки максимума функции $f(x)$, имеющей на сегменте $[a, b]$ единственную точку максимума c при условии, что эта функция возрастает слева от c при $c>a$ и убывает справа от c при $c<b$.

Глава 8

ПЕРВООБРАЗНАЯ ФУНКЦИЯ И НЕОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ

В этой главе мы изучим операцию, обратную по отношению к операции дифференцирования, т. е. займемся вопросом о восстановлении функции по известной производной этой функции.

Изучение этого вопроса естественно приведет нас к понятиям первообразной и неопределенного интеграла (уже упоминавшимся в гл. 1).

Откладывая до гл. 9 вопрос о существовании первообразной и неопределенного интеграла, мы изучим в настоящей главе важнейшие методы интегрирования и выделим классы функций, неопределенные интегралы от которых выражаются через элементарные функции.

§ 1. ПОНЯТИЕ ПЕРВООБРАЗНОЙ ФУНКЦИИ И НЕОПРЕДЕЛЕННОГО ИНТЕГРАЛА

1. Понятие первообразной функции. К числу важных задач механики относятся задача об определении закона движения материальной точки по заданной ее скорости, а также задача об определении закона движения и скорости материальной точки по заданному ее ускорению *.

Эти задачи приводят к математической проблеме *отыскания функции по заданной производной этой функции.*

Переходим к рассмотрению этой проблемы.

Определение. *Функция $F(x)$ называется первообразной функцией (или просто первообразной) для функции $f(x)$ на интервале (a, b), если в любой точке x интервала (a, b) функция $F(x)$ дифференцируема и имеет производную $F'(x)$, равную $f(x)$.*

Замечание. Аналогично определяется первообразная для функции $f(x)$ *на бесконечной прямой и на полупрямой* **.

* Вместо ускорения материальной точки можно задать действующую на эту точку силу (ибо согласно второму закону Ньютона сила определяет ускорение этой точки).

** Можно ввести первообразную для функции $f(x)$ и на сегменте $[a, b]$, понимая под такой первообразной функцию $F(x)$, имеющую производную $F'(x)$ в любой внутренней точке сегмента $[a, b]$, равную $f(x)$, и, кроме того, имеющую правую производную $F'(a+0)$, равную $f(a+0)$, и левую производную $F'(b-0)$, равную $f(b-0)$.

Примеры. 1) Функция $F(x)=\sqrt{1-x^2}$ является первообразной для функции $f(x)=-\frac{x}{\sqrt{1-x^2}}$ на интервале $(-1, +1)$, ибо в любой точке x этого интервала $(\sqrt{1-x^2})'=\frac{-x}{\sqrt{1-x^2}}$.

2) Функция $F(x)=\sin x$ является первообразной для функции $f(x)=\cos x$ на бесконечной прямой $(-\infty, +\infty)$, ибо в каждой точке x бесконечной прямой $(\sin x)'=\cos x$.

3) Функция $F(x)=\ln x$ является первообразной для функции $f(x)=\frac{1}{x}$ на открытой полупрямой $x>0$, ибо в каждой точке x этой полупрямой $(\ln x)'=\frac{1}{x}$.

Если $F(x)$ является первообразной для функции $f(x)$ на интервале (a, b), то, очевидно, и функция $F(x)+C$, где C — любая постоянная, является первообразной для функции $f(x)$ на интервале (a, b).

Естественно, возникает вопрос, как связаны между собой различные первообразные для одной и той же функции $f(x)$.

Справедлива следующая основная теорема.

Теорема 8.1. *Если $F_1(x)$ и $F_2(x)$ — любые первообразные для функции $f(x)$ на интервале (a, b), то всюду на этом интервале $F_1(x)-F_2(x)=C$, где C — некоторая постоянная.*

Другими словами, две любые первообразные для одной и той же функции могут отличаться лишь на постоянную.

Доказательство. Положим $\Phi(x)=F_1(x)-F_2(x)$. Так как каждая из функций $F_1(x)$ и $F_2(x)$ дифференцируема на интервале (a, b), то в силу теоремы 5.5 и функция $\Phi(x)$ дифференцируема на интервале (a, b), причем всюду на этом интервале $\Phi'(x)=F_1'(x)-F_2'(x)=f(x)-f(x)=0$.

В п. 1 § 4 гл. 6 была доказана теорема 6.5 следующего содержания: если функция $\Phi(x)$ дифференцируема всюду на интервале (a, b) и если всюду на этом интервале $\Phi'(x)=0$, то функция $\Phi(x)$ является постоянной на интервале (a, b).

Из этой теоремы получим, что $\Phi(x)=F_1(x)-F_2(x)=C=\text{const}$, что и требовалось доказать.

Следствие. Если $F(x)$ — *одна* из первообразных функций для функции $f(x)$ на интервале (a, b), то *любая* первообразная $\Phi(x)$ для функции $f(x)$ на интервале (a, b) имеет вид $\Phi(x)=F(x)+C$, где C — некоторая постоянная.

2. Неопределенный интеграл.

Определение. *Совокупность всех первообразных функций для данной функции $f(x)$ на интервале (a, b) называется неопределенным интегралом от функции $f(x)$ (на этом интервале) и обозначается символом*

$$\int f(x)\,dx. \tag{8.1}$$

В этом обозначении знак $\int$ называется **знаком интеграла**, выражение $f(x)dx$ — **подынтегральным выражением** а сама функция $f(x)$ — **подынтегральной функцией**.

Если $F(x)$ — одна из первообразных функций для функции $f(x)$ на интервале (a, b), то в силу следствия из теоремы 8.1

$$\int f(x)\,dx = F(x) + C, \tag{8.2}$$

где C — любая постоянная *.

Подчеркнем, что если *первообразная (а значит, и неопределенный интеграл) для функции $f(x)$ на интервале (a, b) существует, то подынтегральное выражение в формуле (8.1) представляет собой дифференциал любой из этих первообразных.*

В самом деле, пусть $F(x)$ — любая из первообразных для функции $f(x)$ на интервале (a, b), т. е. для всех x из интервала (a, b) $F'(x) = f(x)$. Тогда $f(x)dx = F'(x)dx = dF$.

Примеры. 1) $\int \frac{-x}{\sqrt{1-x^2}}\,dx = \sqrt{1-x^2} + C$ на интервале $-1 < x < 1$, ибо функция $F(x) = \sqrt{1-x^2}$ является одной из первообразных для функции $f(x) = \frac{-x}{\sqrt{1-x^2}}$ на указанном интервале.

2) $\int \cos x\,dx = \sin x + C$ на всей бесконечной прямой $-\infty < x < +\infty$, ибо функция $F(x) = \sin x$ является одной из первообразных для функции $f(x) = \cos x$ на бесконечной прямой.

В этой главе мы не будем заниматься вопросом о *существовании* первообразных (или неопределенных интегралов) для широких классов функций. Здесь мы лишь отметим, что в § 4 гл. 9 будет доказано, что *для всякой функции $f(x)$, непрерывной на интервале (a, b), существует на этом интервале первообразная функция (и неопределенный интеграл).*

3. Основные свойства неопределенного интеграла. Прежде всего отметим два свойства, непосредственно вытекающие из определения неопределенного интеграла:

1°. $d\int f(x)\,dx = f(x)\,dx$.

2°. $\int dF(x) = F(x) + C$.

Свойство 1° означает, что знаки d и $\int$ взаимно сокращаются в случае, если знак дифференциала стоит перед знаком интеграла.

Свойство 2° означает, что знаки $\int$ и d взаимно сокращаются и в случае, если знак интеграла стоит перед знаком дифференциала, но в этом случае к $F(x)$ следует добавить произвольную постоянную C.

---

* Равенство (8.2) следует понимать как равенство двух множеств.

Для установления свойства 1° достаточно взять дифференциал от обоих частей формулы (8.2) и учесть, что $dF(x) = F'(x)\,dx = f(x)\,dx$.

Для установления свойства 2° достаточно в левой части (8.2) воспользоваться равенством $dF(x) = f(x)\,dx$.

Следующие два свойства обычно называют *линейными свойствами интеграла:*

3°. $\int [f(x) \pm g(x)]\,dx = \int f(x)\,dx \pm \int g(x)\,dx$.

4°. $\int [Af(x)]\,dx = A\int f(x)\,dx$ $(A = \text{const})$.

Подчеркнем, что равенство в формулах 3° и 4° имеет условный характер: его следует понимать как равенство правой и левой частей с точностью до произвольного постоянного слагаемого (это понятно, поскольку каждый из интегралов, фигурирующих в формулах 3° и 4°, определен с точностью до произвольного постоянного слагаемого).

Поскольку две первообразные для одной и той же функции могут отличаться лишь на постоянную, то для доказательства свойства 3° достаточно доказать, что если $F(x)$ — первообразная для $f(x)$, а $G(x)$ — первообразная для $g(x)$, то функция $F(x) \pm G(x)$ является первообразной для функции $f(x) \pm g(x)$. Это последнее непосредственно вытекает из того, что производная (алгебраической) суммы функций равна сумме производных этих функций, т. е.

$$[F(x) \pm G(x)]' = F'(x) \pm G'(x) = f(x) \pm g(x).$$

Аналогично доказывается свойство 4°. В этом случае используется равенство $[AF(x)]' = AF'(x) = Af(x)$.

4. Таблица основных неопределенных интегралов. В гл. 5 мы получили таблицу производных простейших элементарных функций (см. § 5 гл. 5), представляющую собой вычислительный аппарат дифференциального исчисления. Каждая формула этой таблицы, устанавливающая, что та или иная функция $F(x)$ имеет производную, равную $f(x)$, приводит нас, в силу определения неопределенного интеграла, к соответствующей формуле интегрального исчисления

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

Таким путем мы приходим к следующей таблице основных неопределенных интегралов:

1°. $\int 0 \cdot dx = C$.

2°. $\int 1 \cdot dx = x + C$.

3°. $\displaystyle\int x^{\alpha}\,dx = \frac{x^{\alpha+1}}{\alpha+1} + C \quad (\alpha \neq -1)$.

4°. $\displaystyle\int \frac{dx}{x} = \ln|x| + C \quad (x \neq 0)$.

5°. $\int a^x\,dx = a^x/\ln a + C \quad (0 < a \neq 1), \qquad \int e^x\,dx = e^x + C$.

6°. $\int \sin x\,dx = -\cos x + C.$

7°. $\int \cos x\,dx = \sin x + C.$

8°. $\int \frac{dx}{\cos^2 x} = \int (1+\operatorname{tg}^2 x)\,dx =$
$= \operatorname{tg} x + C\left(x \neq \frac{\pi}{2} + \pi n,\ n = 0,\ \pm 1,\ \ldots\right).$

9°. $\int \frac{dx}{\sin^2 x} = \int (1+\operatorname{ctg}^2 x)\,dx = -\operatorname{ctg} x + C$ ($x \neq \pi n$, где $n = 0,\ \pm 1,\ \ldots$).

10°. $\int \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \arcsin x + C, \\ -\arccos x + C \end{cases} \quad (-1 < x < 1).$

11°. $\int \frac{dx}{1+x^2} = \begin{cases} \operatorname{arcctg} x + C, \\ -\operatorname{arc\,tg} x + C. \end{cases}$

12°. $\int \frac{dx}{\sqrt{x^2 \pm 1}} = \ln|x + \sqrt{x^2 \pm 1}| + C$ (в случае знака «минус» либо $x > 1$, либо $x < -1$).

13°. $\int \frac{dx}{1-x^2} = \frac{1}{2} \ln\left|\frac{1+x}{1-x}\right| + C \quad (|x| \neq 1).$

К этим формулам можно присоединить и соответствующие формулы для гиперболических функций:

14°. $\int \operatorname{sh} x\,dx = \operatorname{ch} x + C.$

15°. $\int \operatorname{ch} x\,dx = \operatorname{sh} x + C.$

16°. $\int \frac{dx}{\operatorname{ch}^2 x} = \operatorname{th} x + C.$

17°. $\int \frac{dx}{\operatorname{sh}^2 x} = -\operatorname{cth} x + C \quad (x \neq 0).$

Сделаем замечания в отношении формул 4°, 12° и 13°. Формула 4° справедлива для любого интервала, не содержащего значения $x=0$. В самом деле, если $x>0$, то из формулы $(\ln x)' = \frac{1}{x}$ заключаем, что $\int \frac{dx}{x} = \ln x + C$, , а если $x<0$, то из формулы $[\ln(-x)]' = \frac{1}{x}$ заключаем, что $\int \frac{dx}{x} = \ln(-x) + C$. Тем самым формула 4° оправдана для любого $x \neq 0$.

Формулы 12° и 13° занимают исключительное положение в нашей таблице, ибо эти формулы не имеют аналогов среди формул таблицы производных.

Однако для проверки формул 12° и 13° достаточно убедиться в том, что производные выражений, стоящих в правых частях этих формул, совпадают с соответствующими подынтегральными функциями.

Наша ближайшая цель — дополнить таблицу неопределенных интегралов основными приемами и методами интегрирования. Но прежде чем приступить к реализации этой цели, сделаем одно важное замечание.

В гл. 1 и 4 мы ввели понятие элементарной функции, а в п. 5 § 5 гл. 5 установили, что производная любой элементарной функции представляет собой также элементарную функцию. Иными словами, мы установили, что *операция дифференцирования не выводит нас из класса элементарных функций.*

Отметим сразу же, что с операцией интегрирования дело обстоит иначе. Можно доказать, что *интегралы от некоторых элементарных функций уже не являются элементарными функциями.* Примерами таких интегралов могут служить следующие:

1°. $\int e^{-x^2}\,dx.$

2°. $\int \cos(x^2)\,dx.$

3°. $\int \sin(x^2)\,dx.$

4°. $\int \frac{dx}{\ln x}\quad (0 < x \neq 1).$

5°. $\int \frac{\cos x}{x}\,dx\quad (x \neq 0).$

6°. $\int \frac{\sin x}{x}\,dx.$

Каждый из указанных интегралов *представляет собой функцию, не являющуюся элементарной.* Указанные функции не только реально существуют*, но и играют большую роль в различных вопросах физики. Так, например, интеграл 1°, называемый *интегралом Пуассона* или *интегралом ошибок,* широко используется в статистической физике, в теории теплопроводности и диффузии, интегралы 2° и 3°, называемые *интегралами Френеля,* широко применяются в оптике. Часто встречаются в приложениях и интегралы 4°—6°, первый из которых называется *интегральным логарифмом,* а последние два — *интегральными косинусом* и *синусом.*

Для всех перечисленных новых функций (интеграла Пуассона, интегралов Френеля, интегрального логарифма, синуса и косинуса) составлены таблицы и графики.

Ввиду важности для приложений эти функции изучены с такой же полнотой, как и простейшие элементарные функции. Вообще, следует подчеркнуть *условность* понятия *простейшей элементарной функции.*

---

* Мы уже отмечали, что в § 4 гл. 9 будет доказано существование неопределенного интеграла от любой непрерывной функции. Существование интегралов 1°—6° обеспечивается непрерывностью подынтегральных функций.

§ 2. ОСНОВНЫЕ МЕТОДЫ ИНТЕГРИРОВАНИЯ

1. Интегрирование заменой переменной (подстановкой). Замена переменной — один из самых эффективных приемов интегрирования. Этот прием базируется на следующем элементарном утверждении.

Пусть функция $t=\varphi(x)$ определена и дифференцируема на множестве $\{x\}$, представляющем собой либо интервал, либо открытую полупрямую, либо бесконечную прямую, и пусть символ $\{t\}$ обозначает множество всех значений этой функции. Пусть, далее, для функции $g(t)$ существует на множестве $\{t\}$ первообразная функция $G(t)$, т. е.

$$\int g(t)\,dt = G(t) + C. \tag{8.3}$$

Тогда всюду на множестве $\{x\}$ для функции $g[\varphi(x)]\varphi'(x)$ существует первообразная функция, равная $G[\varphi(x)]$, т. е.

$$\int g[\varphi(x)]\,\varphi'(x)\,dx = G[\varphi(x)] + C. \tag{8.4}$$

Для доказательства этого утверждения достаточно воспользоваться правилом дифференцирования сложной функции*

$$\frac{d}{dx}\{G[\varphi(x)]\} = G'[\varphi(x)]\,\varphi'(x)$$

и учесть, что по определению первообразной $G'(t)=g(t)$. Предположим теперь, что нам требуется вычислить интеграл

$$\int f(x)\,dx. \tag{8.5}$$

В ряде случаев удается выбрать в качестве новой переменной такую дифференцируемую функцию $t=\varphi(x)$, что имеет место равенство

$$f(x) = g[\varphi(x)]\varphi'(x), \tag{8.6}$$

причем функция $g(t)$ легко интегрируется, т. е. интеграл

$$\int g(t)\,dx = G(t) + C$$

просто вычисляется. Доказанное выше утверждение позволяет нам написать следующую формулу для интеграла (8.5):

$$\int f(x)\,dx = G[\varphi(x)] + C.$$

Этот прием вычисления интеграла (8.5) и называется *интегрированием путем замены переменной.*

Конечно, такой прием применим не ко всякому интегралу. Кроме того, следует подчеркнуть, что выбор правильной подста-

---

* См. п. 1 § 3 гл. 5.

новки в значительной мере определяется искусством вычислителя. Приведем ряд примеров, иллюстрирующих только что изложенный метод.

1°. Вычислить $\int \sin 3x\,dx$. Для вычисления этого интеграла следует сделать простейшую подстановку $t=3x$, $dt=3dx$. В результате этой замены получим

$$\int \sin 3x\,dx = \int \frac{1}{3}\sin t\,dt = -\frac{1}{3}\cos t + C = -\frac{1}{3}\cos 3x + C.$$

2°. Вычислить $\int \frac{dx}{x+a}$. Этот интеграл вычисляется посредством замены $t=x+a$, $dt=dx$.

При этом получим

$$\int \frac{dx}{x+a} = \int \frac{dt}{t} = \ln|t| + C = \ln|x+a| + C \quad (x \neq -a).$$

3°. Вычислить $\int e^{\cos x}\sin x\,dx$. Легко видеть, что этот интеграл вычисляется путем замены $t=\cos x$.

В самом деле, при этом $dt=-\sin x\,dx$ и

$$\int e^{\cos x}\sin x\,dx = -\int e^t\,dt = -e^t + C = -e^{\cos x} + C.$$

4°. Вычислить $\int \frac{(\operatorname{arc\,tg} x)^{100}}{1+x^2}\,dx$. Для вычисления этого интеграла удобна замена $t=\operatorname{arc\,tg} x$. В самом деле, при такой замене

$$dt = \frac{dx}{1+x^2} \text{ и } \int \frac{(\operatorname{arc\,tg} x)^{100}}{1+x^2}\,dx = \int t^{100}\,dt = \frac{t^{101}}{101} + C = \frac{(\operatorname{arc\,tg} x)^{101}}{101} + C.$$

5°. Вычислить интеграл $I=\int (5x-6)^{1984}dx$. Конечно, раскрывая подынтегральную функцию по формуле бинома Ньютона, мы можем свести этот интеграл к сумме тысячи девятисот восьмидесяти пяти табличных интегралов. Но гораздо проще сделать замену переменной $t=5x-6$, $dt=5dx$, в результате которой мы получим, что

$$I = \frac{1}{5}\int t^{1984}\,dt = \frac{t^{1985}}{9925} + C = \frac{(5x-6)^{1985}}{9925} + C.$$

6°. Вычислить $\int \frac{dx}{\cos x}$. Чтобы усмотреть ту замену, посредством которой может быть взят этот интеграл, перепишем его в виде

$$\int \frac{dx}{\cos x} = \int \frac{\cos x\,dx}{\cos^2 x} = \int \frac{\cos x\,dx}{1-\sin^2 x}.$$

После этого понятно, что следует положить $t=\sin x$, $dt=\cos x\,dx$.

В результате получим

$$\int \frac{dx}{\cos x} = \int \frac{dt}{1-t^2} = \frac{1}{2} \ln \left| \frac{1+t}{1-t} \right| + C = \ln \left| \operatorname{tg} \left(\frac{x}{2} + \frac{\pi}{4} \right) \right| + C.$$

7°. Вычислить интеграл $\int \frac{x^5\,dx}{(3x)^{12}+1}$. Для вычисления этого интеграла удобна замена $t=(3x)^6$, $dt=4374\,x^5 dx$. В результате указанной замены получим

$$\int \frac{x^5\,dx}{(3x)^{12}+1} = \frac{1}{4374} \int \frac{dt}{t^2+1} = \frac{\operatorname{arc\,tg} t}{4374} + C = \frac{\operatorname{arc\,tg}(3x)^6}{4374} + C.$$

8°. Вычислить $\int \frac{dx}{(x^2+a^2)^{3/2}}$. Для вычисления этого интеграла оказывается удобной тригонометрическая подстановка $t = \operatorname{arc\,tg}\frac{x}{a}$, $x = a \operatorname{tg} t$, $dx = a\frac{dt}{\cos^2 t}$.

В результате этой подстановки интеграл принимает вид

$$\int \frac{dx}{(x^2+a^2)^{3/2}} = \frac{1}{a^2} \int \cos t\,dt = \frac{\sin t}{a^2} + C =$$

$$= \frac{\operatorname{tg} t}{a^2\sqrt{1+\operatorname{tg}^2 t}} + C = \frac{x}{a^2\sqrt{x^2+a^2}} + C.$$

9°. Вычислить $\int \frac{dx}{(a^2-x^2)^{3/2}}$. Здесь оказывается удобной подстановка $t = \arcsin\frac{x}{a}$, $x = a \sin t$, $dx = a\cos t\,dt$. При этом

$$\int \frac{dx}{(a^2-x^2)^{3/2}} = \frac{1}{a^2} \int \frac{dt}{\cos^2 t} = \frac{\operatorname{tg} t}{a^2} + C =$$

$$= \frac{1}{a^2}\,\frac{\sin t}{\sqrt{1-\sin^2 t}} + C = \frac{x}{a^2\sqrt{a^2-x^2}} + C.$$

10. Вычислить $\int \sqrt{\frac{a+x}{a-x}}\,dx$. Для вычисления этого интеграла оказывается удобной замена $2t = \arccos\frac{x}{a}$, $x = a\cos 2t$, $dx = -2a\sin 2t\,dt$. Мы получим

$$\int \sqrt{\frac{a+x}{a-x}}\,dx = -4a \int \cos^2 t\,dt = -4a \int \left(\frac{1}{2} + \frac{1}{2}\cos 2t \right) dt =$$

$$= -2at - 2a \int \cos 2t\,dt = -2at - a\sin 2t + C =$$

$$= -a \left[\arccos\frac{x}{a} + \sqrt{1-\left(\frac{x}{a}\right)^2} \right] + C.$$

2. Интегрирование по частям. К числу весьма эффективных методов интегрирования относится *метод интегрирования по частям*. Этот метод основывается на следующем утверждении.

Пусть каждая из функций $u(x)$ и $v(x)$ дифференцируема на множестве $\{x\}$ и, кроме того, на этом множестве существует первообразная для функции $v(x)u'(x)$. Тогда на множестве $\{x\}$ существует первообразная и для функции $u(x)v'(x)$, причем справедлива формула

$$\int u(x)\,v'(x)\,dx = u(x)\,v(x) - \int v(x)\,u'(x)\,dx. \tag{8.8}$$

Замечание. Определение дифференциала и свойство инвариантности его формы позволяют записать формулу (8.8) в виде

$$\int u\,dv = u(x)\,v(x) - \int v\,du. \tag{8.9}$$

Для доказательства сформулированного утверждения запишем формулу для производной произведения функций $u(x)$ и $v(x)$

$$[u(x)v(x)]' = u(x)v'(x) + u'(x)v(x). \tag{8.10}$$

Умножим равенство (8.10) на dx и возьмем интеграл от обеих частей полученного таким путем равенства. Так как по условию для всех x из множества $\{x\}$ существует $\int v(x)u'(x)dx$ и $\int [u(x)v(x)]'dx = u(x)v(x) + C$ (см. свойство 2° из п. 3 § 1), то для всех x из множества $\{x\}$ существует и интеграл $\int u(x)v'(x)dx$, причем справедлива формула (8.8) (или (8.9)).

Формула (8.9) *сводит вопрос о вычислении интеграла $\int udv$ к вычислению интеграла $\int vdu$.* В ряде конкретных случаев этот последний интеграл без труда вычисляется.

Вычисление интеграла $\int udv$ посредством применения формулы (8.9) и называют *интегрированием по частям*. Заметим, что при конкретном применении формулы интегрирования по частям (8.9) очень удобно пользоваться таблицей дифференциалов, выписанной нами в п. 6 § 4 гл. 5.

Переходим к рассмотрению примеров.

1°. Вычислим интеграл $I = \int x^n \ln x\,dx$ $(n \neq -1)$. Полагая $u = \ln x$, $dv = x^n dx$ и используя формулу (8.9), получим $du = \dfrac{dx}{x}$, $v = \dfrac{x^{n+1}}{n+1}$,

$$I = \frac{x^{n+1}}{n+1}\ln x - \frac{1}{n+1}\int x^n\,dx = \frac{x^{n+1}}{n+1}\left(\ln x - \frac{1}{n+1}\right) + C.$$

2°. Вычислим, далее, интеграл $I = \int x \operatorname{arc\,tg} x\,dx$. Полагая $u = \operatorname{arc\,tg} x$, $dv = xdx$ и используя формулу (8.9), будем иметь

$$du = \frac{dx}{1+x^2}, \quad v = \frac{x^2}{2},$$

$$I = \frac{x^2}{2}\operatorname{arc\,tg} x - \frac{1}{2}\int \frac{x^2}{1+x^2}dx = \frac{x^2}{2}\operatorname{arc\,tg} x - \frac{1}{2}\int \frac{[(1+x^2)-1]\,dx}{1+x^2} =$$

$$= \frac{x^2}{2}\operatorname{arc\,tg} x - \frac{1}{2}\int dx + \frac{1}{2}\int \frac{dx}{1+x^2} = \frac{x^2+1}{2}\operatorname{arc\,tg} x - \frac{x}{2} + C.$$

3°. Вычислим интеграл $I=\int x^2 \cos x\,dx$. Сначала применим формулу (8.9), полагая $u=x^2$, $dv=\cos x dx$. Получим $du=2xdx$, $v=\sin x$, $I=x^2 \sin x - 2\int x \sin x\,dx$. Для вычисления последнего интеграла еще раз применим формулу (8.9), полагая на этот раз $u=x$, $dv=\sin x\,dx$. Получим $du=dx$, $v=-\cos x$, $I=x^2\sin x + 2x\cos x - 2\int \cos x dx = (x^2-2)\sin x + 2x\cos x + C$.

Таким образом, интеграл $\int x^2 \cos x dx$ вычислен нами посредством двукратного интегрирования по частям. Легко понять, что интеграл $\int x^n \cos x\,dx$ (где n — любое целое положительное число) может быть вычислен по аналогичной схеме посредством n-кратного интегрирования по частям.

4°. Вычислим теперь интеграл $I=\int e^{ax}\cos bx dx$ ($a=\text{const}$, $b=\text{const}$). Сначала применим формулу (8.9), полагая $u=e^{ax}$, $dv=\cos bx dx$. Получим $du=ae^{ax}dx$, $v=\frac{\sin bx}{b}$,

$$I = \frac{e^{ax}\sin bx}{b} - \frac{a}{b}\int e^{ax}\sin bx\,dx.$$

Для вычисления последнего интеграла еще раз применим формулу (8.9), полагая на этот раз $u=e^{ax}$, $dv=\sin bx dx$. Получим

$$du = ae^{ax}\,dx, \quad v = -\frac{\cos bx}{b},$$

$$I = \frac{e^{ax}\sin bx}{b} + \frac{a}{b^2}e^{ax}\cos bx - \frac{a^2}{b^2}I. \tag{8.11}$$

Таким образом, посредством двукратного интегрирования I по частям мы получим для интеграла I уравнение первого порядка (8.11). Из этого уравнения находим

$$I = \frac{a\cos bx + b\sin bx}{a^2+b^2}e^{ax}.$$

Практика показывает, что большая часть интегралов, берущихся посредством интегрирования по частям, может быть разбита на следующие *три группы:*

1) К *первой* группе относятся интегралы, подынтегральная функция которых содержит в качестве множителя одну из следующих функций: $\ln x$, $\arcsin x$, $\arccos x$, $\operatorname{arc\,tg} x$, $(\operatorname{arc\,tg} x)^2$, $(\arccos x)^2$, $\ln \varphi(x)$, ... — при условии, что оставшаяся часть по-

дынтегральной функции представляет собой производную известной функции (см. рассмотренные выше примеры 1° и 2°). Для вычисления интегралов первой группы следует применить формулу (8.9), полагая в ней $u(x)$ равной одной из указанных выше функций *.

2) Ко *второй* группе относятся интегралы вида $\int(ax+b)^n\times\times\cos(cx)dx$, $\int(ax+b)^n\sin(cx)dx$, $\int(ax+b)^n e^{cx}dx$, где a, b, c — некоторые постоянные, n — любое целое положительное число (см. рассмотренный выше пример 3°). Интегралы второй группы берутся путем n-кратного применения формулы интегрирования по частям (8.9), причем в качестве $u(x)$ всякий раз следует брать $(ax+b)$ в соответствующей степени. После каждого интегрирования по частям эта степень будет понижаться на единицу.

3) К *третьей* группе относятся интегралы вида $\int e^{ax}\sin bxdx$, $\int e^{ax}\sin bxdx$, $\int\sin(\ln x)dx$, $\int\cos(\ln x)dx$, ... (см. рассмотренный выше пример 4°). Обозначая любой из интегралов этой группы через I и производя двукратное интегрирование по частям, мы составим для I уравнение первого порядка.

Конечно, указанные три группы не исчерпывают всех без исключения интегралов, берущихся посредством интегрирования по частям. Приведем примеры интегралов, не входящих ни в одну из перечисленных трех групп, но вычисляемых при помощи формулы (8.9).

5°. Вычислим интеграл $I=\int\frac{x\,dx}{\sin^2 x}$. Этот интеграл не входит ни в одну из упомянутых трех групп. Тем не менее, применяя формулу (8.9) и полагая в ней $u=x$, $dv=\frac{dx}{\sin^2 x}$, получим $du=dx$, $v=-\operatorname{ctg} x$,

$$I=-x\operatorname{ctg}x+\int\operatorname{ctg}x\,dx=-x\operatorname{ctg}x+\int\frac{\cos x\,dx}{\sin x}=$$

$$=-x\operatorname{ctg}x+\int\frac{d(\sin x)}{\sin x}=-x\operatorname{ctg}x+\ln|\sin x|+C$$

(в проведенных рассуждениях $x\neq\pi n$, где $n=0,\ \pm1\ldots$).

Аналогично вычисляется интеграл $\int\frac{x\,dx}{\cos^2 x}$.

6°. Вычислим, наконец, весьма важный для дальнейшего интеграл $K_\lambda=\int\frac{dt}{(t^2+a^2)^\lambda}$, где $a=\text{const}$, $\lambda=1,\ 2,\ \ldots$ **. Этот интеграл также не входит ни в одну из упомянутых выше трех групп.

---

* В случае, если подынтегральная функция содержит в качестве множителя $(\operatorname{arctg}x)^2$, $(\arccos x)^2,\ldots$, формулу интегрирования по частям (8.9) придется применить дважды.

** Для обозначения переменной под знаком этого интеграла нам удобнее писать букву t.

Для вычисления этого интеграла установим для него рекуррентную формулу, сводящую вопрос о вычислении K_λ к вычислению $K_{\lambda-1}$. Можно записать (при $\lambda\neq 1$):

$$K_\lambda = \frac{1}{a^2}\int \frac{a^2\,dt}{(t^2+a^2)^\lambda} = \frac{1}{a^2}\int \frac{[(t^2+a^2)-t^2]\,dt}{(t^2+a^2)^\lambda} =$$

$$= \frac{1}{a^2}\int \frac{dt}{(t^2-a^2)^{\lambda-1}} - \frac{1}{2a^2}\int t\,\frac{2t\,dt}{(t^2+a^2)^\lambda} = \frac{1}{a^2}K_{\lambda-1} - \frac{1}{2a^2}\int t\,\frac{d(t^2+a^2)}{(t^2+a^2)^\lambda}.$$

Для вычисления последнего интеграла применим формулу интегрирования по частям (8.9), полагая в ней $u=t$, $dv=\dfrac{d(t^2+a^2)}{(t^2+a^2)^\lambda}$.

Получим $du=dt$, $v=-\dfrac{1}{(\lambda-1)(t^2+a^2)^{\lambda-1}}$,

$$K_\lambda = \frac{1}{a^2}K_{\lambda-1} + \frac{t}{2a^2(\lambda-1)(t^2+a^2)^{\lambda-1}} - \frac{1}{2a^2(\lambda-1)}K_{\lambda-1}.$$

Из последнего равенства получим рекуррентную формулу

$$K_\lambda = \frac{t}{2a^2(\lambda-1)(t^2+a^2)^{\lambda-1}} + \frac{2\lambda-3}{a^2(2\lambda-2)}K_{\lambda-1}. \qquad (8.12)$$

Убедимся в том, что рекуррентная формула (8.12) позволяет вычислить интеграл K_λ для любого $\lambda=2, 3, \ldots$. В самом деле, интеграл K_1 вычисляется элементарно:

$$K_1 = \int \frac{dt}{t^2+a^2} = \frac{1}{a}\int \frac{d\left(\frac{t}{a}\right)}{\left(\frac{t}{a}\right)^2+1} = \frac{1}{a}\operatorname{arc\,tg}\frac{t}{a} + C.$$

После того как вычислен интеграл K_1, полагая в формуле (8.12) $\lambda=2$, мы без труда вычислим K_2. В свою очередь, зная K_2 и полагая в формуле (8.12) $\lambda=3$, мы без труда вычислим K_3. Продолжая действовать таким образом дальше, мы вычислим интеграл K_λ для любого натурального λ, выразив его через элементарные функции.

§ 3. КЛАССЫ ФУНКЦИЙ, ИНТЕГРИРУЕМЫХ В ЭЛЕМЕНТАРНЫХ ФУНКЦИЯХ

Хотя, как уже отмечалось выше, неопределенный интеграл от элементарной функции может не выражаться через элементарные функции, все же существуют широкие классы функций, неопределенные интегралы от которых выражаются через элементарные функции. Изучению таких классов функций и будет посвящен настоящий параграф.

Наиболее важным среди указанных классов функций является класс **рациональных дробей**, представляющих собой отношение двух алгебраических многочленов. Изучению класса рациональных дробей мы предпошлем краткие сведения о комплексных числах и об алгебраических многочленах.

1. Краткие сведения о комплексных числах. *Два вещественных числа x и y мы будем называть упорядоченной парой, если указано, какое из этих чисел является первым, какое вторым.* Упорядоченную пару вещественных чисел x и y будем обозначать символом (x, y), записывая на первом месте первый элемент этой пары.

Комплексным числом называется упорядоченная пара (x, y) вещественных чисел, первое из которых x называется действительной частью, а второе y — мнимой частью этого комплексного числа.

В случае, когда мнимая часть y равна нулю, соответствующую пару $(x, 0)$ договариваются отождествлять с вещественным числом x. Это позволяет рассматривать множество всех вещественных чисел как подмножество комплексных чисел.

Два комплексных числа $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$ называются равными, если $x_1=x_2$, $y_1=y_2$. Говорят, что комплексное число $z=(x, y)$ равно нулю, если $x=0$ и $y=0$.

Определим операции сложения и умножения комплексных чисел. Поскольку вещественные числа являются подмножеством комплексных чисел, эти операции должны быть определены так, чтобы в применении к двум вещественным числам они приводили к уже известным нам из § 4 гл. 2 определениям суммы и произведения вещественных чисел.

Суммой двух комплексных чисел $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$ назовем комплексное число z вида

$$z=(x_1+x_2,\ y_1+y_2). \tag{8.13}$$

Произведением двух комплексных чисел $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$ назовем комплексное число z вида

$$z=(x_1x_2-y_1y_2,\quad x_1y_2+x_2y_1). \tag{8.14}$$

Легко проверить, что сумма и произведение комплексных чисел обладают теми же самыми свойствами, что и сумма и произведение вещественных чисел. Именно справедливы следующие свойства:

1°. $z_1+z_2=z_2+z_1$ (переместительное свойство суммы).

2°. $(z_1+z_2)+z_3=z_1+(z_2+z_3)$ (сочетательное свойство суммы).

3°. $z+(0, 0)=z$ (особая роль числа $(0, 0)$).

4°. Для каждого числа $z=(x, y)$ существует противоположное ему число $z'=(-x, -y)$ такое, что $z+z'=(0, 0)$.

5°. $z_1\cdot z_2=z_2\cdot z_1$ (переместительное свойство произведения).

6°. $(z_1\cdot z_2)\cdot z_3=z_1\cdot(z_2\cdot z_3)$ (сочетательное свойство произведения).

7°. $z\cdot(1, 0)=z$ (особая роль числа $(1, 0)$).

8°. Для любого комплексного числа $z=(x, y)$, не равного нулю, существует обратное ему число $z'=\left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right)$ такое, что $z\cdot z'=(1, 0)$.

9°. $(z_1+z_2)\cdot z_3=z_1\cdot z_3+z_2\cdot z_3$ (распределительное свойство произведения относительно суммы).

Свойства 1°—9° позволяют утверждать, что для комплексных чисел полностью сохраняются все правила элементарной алгебры, относящиеся к арифметическим действиям и к сочетанию равенств.

Кроме того, эти свойства полностью решают вопрос о *вычитании комплексных чисел* как о действии, обратном сложению, и о *делении комплексных чисел* как о действии, обратном умножению.

Разностью двух комплексных чисел $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$ *называется такое комплексное число z, которое в сумме с z_2 дает z_1*. С помощью свойств 1°—4° элементарно устанавливается существование и единственность разности двух любых комплексных чисел *.

Легко проверить, что разностью двух комплексных чисел $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$ является комплексное число z вида

$$z=(x_1-x_2;\quad y_1-y_2). \tag{8.15}$$

Частным двух комплексных чисел $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$, второе из которых не равно нулю, называется такое комплексное число z, которое при умножении на z_2 дает z_1. С помощью свойств 5°—8° легко установить, что единственным частным двух указанных комплексных чисел является комплексное число z вида

$$z=\left(\frac{x_1x_2+y_1y_2}{x_2^2+y_2^2}, \frac{x_2y_1-x_1y_2}{x_2^2+y_2^2}\right).$$

В операциях с комплексными числами особую роль играет число, представимое парой $(0, 1)$ и обозначаемое буквой i. Умножая эту пару самое на себя (т. е. возводя ее в квадрат), получим в силу определения произведения комплексных чисел

$$(0, 1)\cdot(0, 1)=(-1,0)=-1, \text{ т. е. } i^2=-1.$$

Заметив это, мы можем любое комплексное число $z=(x, y)$ представить в виде

$$z=(x, y)=(x,0)+(0, y)=(x, 0)+(y, 0)\cdot(0, 1)=x+iy.$$

В дальнейшем мы будем широко использовать для комплексного числа $z=(x, y)$ представление $z=x+iy$. Это представление

* Для этого следует дословно повторить рассуждения, проведенные в гл. 2 для случая разности двух вещественных чисел (см. п. 1 § 5 гл. 2).

и рассмотрение i в качестве множителя, квадрат которого равен —1, позволяет производить операции с комплексными числами так же, как они производятся с алгебраическими многочленами.

Комплексное число $\bar{z}=(x, -y)=x-iy$ принято называть сопряженным *по отношению к комплексному числу* $z=(x, y)=x+iy$.

Очевидно, что *комплексное число равно нулю тогда и только тогда, когда равно нулю сопряженное ему число* *.

Для геометрического изображения комплексных чисел удобно пользоваться декартовой прямоугольной системой координат. При этом комплексное число $z=(x, y)$ изображается или точкой M с координатами (x, y), или вектором $\overrightarrow{OM}$, идущим из начала координат в точку M.

При таком способе изображения сложение и вычитание комплексных чисел сводится к сложению и вычитанию соответствующих им векторов (это понятно из формул (8.13) и (8.15)).

Непосредственно из определения (8.14) произведения двух комплексных чисел вытекает следующее утверждение: *произведение двух (а значит, и нескольких) комплексных чисел равно нулю тогда и только тогда, когда обращается в нуль хотя бы один из сомножителей.*

В самом деле, если хотя бы одно из чисел $z_1=(x_1, y_1)$ и $z_2=(x_2, y_2)$ равно $(0, 0)$, то из (8.14) очевидно, что $z=z_1z_2=(0, 0)$.

Если, наоборот, $z=z_1z_2$ равно $(0, 0)$, то из (8.14) следует, что

$$\begin{cases} x_1x_2-y_1y_2=0, \\ y_1x_2+x_1y_2=0, \end{cases} \tag{8.14'}$$

и если $z_1\neq(0, 0)$, т. е. $x_1^2+y_1^2\neq 0$, то (8.14′) представляет собой однородную систему двух уравнений относительно двух неизвестных x_2 и y_2 с определителем $x_1^2+y_1^2$, отличным от нуля. Такая система имеет только тривиальное решение, т. е. $z_2=(x_2, y_2)=(0, 0)$.

Непосредственно из определения (8.14) произведения двух комплексных чисел вытекает и еще одно утверждение: *комплексное число, сопряженное к произведению двух (а значит, и нескольких) комплексных чисел, равно произведению комплексных чисел, являющихся сопряженными к каждому из сомножителей, т. е.*

$$\overline{(z_1z_2)}=\bar{z}_1\bar{z}_2. \tag{8.14''}$$

---

* Ибо система $\begin{cases} x=0, \\ y=0 \end{cases}$ эквивалентна системе $\begin{cases} x=0, \\ -y=0. \end{cases}$

С помощью правила перемножения комплексных чисел (8.14) легко проверяется, что как правая, так и левая часть (8.14″) равна одному и тому же комплексному числу $(x_1x_2-y_1y_2,$ $-x_1y_2-x_2y_1)$.

2. Краткие сведения о корнях алгебраических многочленов.

1°. Алгебраическим многочленом n-й степени называется выражение вида

$$f(z)=c_0z^n+c_1z^{n-1}+\ldots+c_{n-1}z+c_n, \tag{8.16}$$

где $z=(x,\ y)=x+iy$ — переменное комплексное число, а $c_0,$ $c_1,\ldots,c_n$ — некоторые постоянные комплексные числа, первое из которых отлично от нуля. Как известно, любой алгебраический многочлен степени n можно поделить «столбиком» на другой алгебраический многочлен степени не выше чем n. Таким путем мы приходим к следующему утверждению: *каковы бы ни были два многочлена $f(z)$ и $\varphi(z)$ такие, что степень $\varphi(z)$ не выше, чем степень $f(z)$, справедливо равенство*

$$f(z)=\varphi(z)\cdot q(z)+r(z), \tag{8.17}$$

в котором $q(z)$ и $r(z)$ — некоторые многочлены, причем степень $q(z)$ равна разности степеней многочленов $f(z)$ и $\varphi(z)$, а степень $r(z)$ ниже степени $\varphi(z)$.

По отношению к фигурирующим в равенстве (8.17) многочленам $f(z)$, $\varphi(z)$, $q(z)$ и $r(z)$ обычно применяют вполне понятные термины «делимое», «делитель», «частное» и «остаток».

Говорят, что многочлен $f(z)$ *делится* на многочлен $\varphi(z)$, если в полученной посредством деления столбиком формуле (8.17) остаток $r(z)=0$.

Договоримся называть многочленом нулевой степени любую комплексную постоянную. Тогда совершенно ясно, что *любой многочлен делится на отличный от нуля многочлен нулевой степени.* Изучим вопрос о делимости многочлена $f(z)$ на многочлен первой степени.

Определение. *Назовем комплексное число b корнем многочлена $f(z)$, если $f(b)$ равно нулю.*

Теорема 8.2. *Многочлен ненулевой степени $f(z)$ делится на двучлен $(z-b)$ тогда и только тогда, когда b является корнем многочлена.*

Доказательство. Запишем для многочленов $f(z)$ и $\varphi(z)=(z-b)$ формулу (8.17). Поскольку степень остатка $r(z)$ в этой формуле *обязана быть ниже* степени делителя $\varphi(z)=z-b$, то $r(z)$ — *многочлен нулевой степени,* т. е. $r(z)=c=\text{const}$. Таким образом, формула (8.17) принимает вид

$$f(z)=(z-b)q(z)+c.$$

Полагая в формуле (8.18) $z=b$, найдем, что $c=f(b)$. По определению $f(z)$ делится на $(z-b)$ тогда и только тогда, когда остаток в формуле (8.18) $c=f(b)$ равен нулю, т. е. тогда и только тогда, когда b является корнем $f(z)$. Теорема доказана.

2°. Естественно, возникает вопрос, всякий ли алгебраический многочлен имеет корни? Ответ на этот вопрос дает основная теорема алгебры*: *всякий многочлен ненулевой степени имеет хотя бы один корень.*

Опираясь на эту теорему, докажем, что *алгебраический многочлен n-й степени имеет точно n корней***. В самом деле, пусть $f(z)$ — многочлен n-й степени. Согласно основной теореме алгебры $f(z)$ имеет хотя бы один корень b_1, т. е. для $f(z)$ справедливо представление

$$f(z)=(z-b_1)f_1(z), \tag{8.19^1}$$

в котором через $f_1(z)$ обозначен некоторый многочлен степени $(n-1)$. Если $n\neq 1$, то, согласно основной теореме алгебры, $f_1(z)$ имеет хотя бы один корень b_2, т. е. для $f_1(z)$ справедливо представление

$$f_1(z)=(z-b_2)f_2(z), \tag{8.19^2}$$

в котором через $f_2(z)$ обозначен некоторый многочлен степени $(n-2)$. Повторяя указанные рассуждения далее, мы получим представления

$$f_2(z)=(z-b_3)f_3(z), \tag{8.19^3}$$

$$\cdots\cdots\cdots\cdots$$

$$f_{n-1}(z)=(z-b_n)f_n(z). \tag{8.19^n}$$

В последнем из этих представлений через $f_n(z)$ обозначен некоторый многочлен нулевой степени, т. е. $f_n(z)=c=\text{const}$. Сопоставляя между собой равенства (8.19^1)—(8.19^n) и учитывая, что $f_n(z)=c$, будем иметь

$$f(z)=(z-b_1)(z-b_2)\dots(z-b_n)c. \tag{8.20}$$

Отметим, что комплексная постоянная c не равна нулю, ибо в противном случае многочлен $f(z)$ был бы тождественно равен нулю и не являлся бы многочленом n-й степени***.

Из равенства (8.20) очевидно, что $f(b_1)=f(b_2)=\dots=f(b_n)=0$, т. е. каждое из чисел $b_1, b_2, \dots, b_n$ является корнем многочлена $f(z)$. Кроме того, из (8.20) очевидно, что, каково бы ни было комплексное число b, отличное от $b_1, b_2, \dots, b_n$, комплексное число

---

* Доказательство основной теоремы алгебры обычно приводится в курсах алгебры и в курсе теории функций комплексной переменной.

** При этом, конечно, предполагается, что $n>0$.

*** Здесь мы используем следующее утверждение: если многочлен $f(z)=a_0z^n+a_1z^{n-1}+\dots+a_{n-1}z+a_n$ *тождественно равен нулю, то все его коэффициенты равны нулю.* В самом деле, если $f(z)\equiv 0$, то при $z=0$ получим $a_n=0$. Но тогда $f(z)\equiv z[a_0z^{n-1}+a_1z^{n-2}+\dots+a_{n-1}]\equiv 0$. Так как $z\neq 0$, то выражение в квадратных скобках тождественно равно нулю, откуда при $z=0$ получим $a_{n-1}=0$. Продолжая аналогичные рассуждения далее, докажем, что все коэффициенты равны нулю.

$f(b)$ не равно нулю*. Таким образом, многочлен $f(z)$ имеет ровно n корней: $b_1, b_2, \ldots, b_n$.

Равенство (8.20) дает разложение многочлена $f(z)$ на множители. Если известен вид многочлена $f(z)$ (8.16), то мы можем определить постоянную c в равенстве (8.20). Сравнивая в равенствах (8.20) и (8.16) коэффициенты при z^n, получим $c=c_0$**.

Многочлен (8.16), у которого $c_0=1$, называется **приведенным**. Для приведенного многочлена формула разложения (8.20) принимает вид

$$f(z)=(z-b_1)(z-b_2)\ldots(z-b_n). \tag{8.21}$$

В дальнейшем, если не оговорено противное, мы будем рассматривать *приведенные многочлены*.

Среди корней многочлена $f(z)$ могут быть *совпадающие корни*. Пусть $a, b, \ldots, c$ — **различные** корни приведенного многочлена $f(z)$.

Тогда для этого многочлена представление вида (8.21) принимает форму следующего равенства:

$$f(z)=(z-a)^{\alpha}(z-b)^{\beta}\ldots(z-c)^{\gamma}. \tag{8.22}$$

В этом разложении $\alpha, \beta, \ldots, \gamma$ — некоторые целые числа, каждое из которых не меньше единицы, причем $\alpha+\beta+\ldots+\gamma=n$, где n — степень многочлена $f(z)$.

Если для многочлена $f(z)$ справедливо разложение (8.22), то говорят, что комплексное число a является **корнем** *$f(z)$* **кратности** *α, комплексное число b является* **корнем** *$f(z)$* **кратности** *$\beta, \ldots$, комплексное число c является* **корнем** *$f(z)$* **кратности** *γ.*

Корень, кратность которого равна единице, принято называть **однократным**, а корень, кратность которого больше единицы, принято называть **кратным**.

Можно дать и другое эквивалентное определение корня данной кратности: *комплексное число a называется* **корнем** *многочлена $f(z)$* **кратности** *α, если для $f(z)$ справедливо представление*

$$f(z)=(z-a)^{\alpha}\varphi(z), \quad \text{где} \quad \varphi(a)\neq 0. \tag{8.23}$$

3°. Пусть теперь

$$f(z)=z^n+c_1z^{n-1}+c_2z^{n-2}+\ldots+c_n \tag{8.24}$$

приведенный алгебраический многочлен с **вещественными** коэффициентами $c_1, c_2, \ldots, c_n$.

---

* Ибо произведение нескольких комплексных чисел равно нулю лишь в том случае, когда равен нулю хотя бы один из сомножителей (см. п. 1).

** Здесь мы используем утверждение: *если два многочлена $a_0z^n+a_1z^{n-1}+\ldots+a_n$ и $b_0z^n+b_1z^{n-1}+\ldots+b_n$ тождественно равны друг другу, то $a_0=b_0$, $a_1=b_1, \ldots, a_n=b_n$.* Для доказательства достаточно к разности указанных многочленов применить утверждение, отмеченное в сноске *** на с. 308.

Докажем, что этот многочлен обладает следующим важным свойством.

Теорема 8.3. *Если комплексное число a является корнем многочлена с вещественными коэффициентами (8.24) кратности λ, то и сопряженное ему комплексное число $\bar{a}$ также является корнем этого многочлена той же самой кратности λ.*

Доказательство. Начнем с того, что докажем следующий вспомогательный факт: *если $f(z)$ — многочлен с вещественными коэффициентами, то комплексная величина $f(\bar{z})$ является сопряженной по отношению к величине $f(z)$.*

Так как коэффициенты многочлена (8.24) являются вещественными числами, то для доказательства указанного факта достаточно убедиться в том, что для любого номера n комплексная величина $(\bar{z})^n$ является сопряженной по отношению z^n. Но это последнее сразу вытекает из утверждения, доказанного в самом конце п. 1, точнее, из соотношения (8.14″). Положив в этом соотношении $z_1=z_2=z$, мы получим, что $\overline{(z^2)}=(\bar{z})^2$. Далее, положив в том же соотношении (8.14″), $z_1=z^2$, $z_2=z$, мы получим $\overline{(z^3)}=\overline{(z^2)}\,\bar{z}=(\bar{z})^3$.

Продолжая аналогичные рассуждения, мы убедимся в том, что $\overline{(z^n)}=(\bar{z})^n$ для любого номера n.

Итак, доказано, что величина $f(\bar{z})$ является сопряженной по отношению к величине $f(z)$, т. е.

$$\overline{f(z)}=f(\bar{z})$$

или, что то же самое,

$$f(z)=\overline{f(\bar{z})}. \tag{8.25}$$

Пусть теперь комплексное число a является корнем многочлена с вещественными коэффициентами $f(z)$ кратности λ, т. е. справедливо представление

$$f(z)=(z-a)^\lambda\varphi(z), \tag{8.26}$$

в котором

$$\varphi(a)\neq 0. \tag{8.27}$$

Из сопоставления (8.26) и (8.25) вытекает, что

$$f(z)=\overline{(\bar{z}-a)^\lambda\varphi(\bar{z})},$$

а последнее равенство в силу (8.14″) можно переписать в виде

$$f(z)=\overline{(\bar{z}-a)^\lambda}\;\overline{\varphi(\bar{z})}. \tag{8.28}$$

Заметим теперь, что в силу установленного выше соотношения $\overline{(z^n)}=(\bar{z}^n)$ справедливо равенство

$$\overline{(\bar{z}-a)^\lambda}=\left(\overline{\bar{z}-a}\right)^\lambda=(z-\bar{a})^\lambda. \tag{8.29}$$

Подставляя (8.29) и (8.28), мы получим представление

$$f(z)=(z-\bar{a})^{\lambda}\psi(z), \tag{8.30}$$

в котором через $\psi(z)$ обозначена величина

$$\psi(z)=\overline{\varphi(\bar{z})}. \tag{8.31}$$

Для завершения доказательства теоремы 8.3 остается убедиться в том, что $\psi(\bar{a})\neq 0$*, но это сразу вытекает из того, что в силу того, что согласно (8.27) $\varphi(a)$ не обращается нуль **. Теорема 8.3 доказана.

3. Разложение алгебраического многочлена с вещественными коэффициентами на произведение неприводимых множителей. В дальнейшем нам придется иметь дело с многочленами от переменной, принимающей лишь вещественные значения. Поэтому эту переменную мы будем обозначать буквой x, а не z.

Пользуясь теоремой 8.3, найдем разложение многочлена с вещественными коэффициентами $f(x)$ на произведение неприводимых вещественных множителей. Пусть многочлен $f(x)$ имеет вещественные корни $b_1, b_2, \ldots, b_m$ кратности $\beta_1, \beta_2, \ldots, \beta_m$ соответственно и комплексно сопряженные пары корней a_1 и $\bar{a}_1$, a_2 и $\bar{a}_2, \ldots, a_n$ и $\bar{a}_n$ кратности $\lambda_1, \lambda_2, \ldots, \lambda_n$ каждая пара соответственно.

Тогда, согласно результатам п. 2, многочлен $f(x)$ может быть представлен в виде

$$f(x)=(x-b_1)^{\beta_1}(x-b_2)^{\beta_2}\ldots(x-b_m)^{\beta_m}\times$$
$$\times(x-a_1)^{\lambda_1}(x-\bar{a}_1)^{\lambda_1}(x-a_2)^{\lambda_2}(x-\bar{a}_2)^{\lambda_2}\ldots(x-a_n)^{\lambda_n}(x-\bar{a}_n)^{\lambda_n}. \tag{8.32}$$

Обозначим вещественную и мнимую части корня a_k $(k=1, 2, \ldots, n)$ соответственно через u_k и v_k, т. е. пусть $a_k=u_k+iv_k$. Тогда $\bar{a}_k=u_k-iv_k$. Преобразуем для любого $k=1, 2, \ldots, n$ выражение

$$(x-a_k)^{\lambda_k}(x-\bar{a}_k)^{\lambda_k}=[(x-a_k)(x-\bar{a}_k)]^{\lambda_k}=$$
$$=[(x-u_k-iv_k)(x-u_k+iv_k)]^{\lambda_k}=$$
$$=[(x-u_k)^2+v_k^2]^{\lambda_k}=(x^2+p_kx+q_k)^{\lambda_k}, \tag{8.33}$$

где $p_k=-2u_k$, $q_k=u_k^2+v_k^2$.

Используя (8.33) в (8.32), окончательно получим следующее разложение многочлена $f(x)$ на произведение вещественных неприводимых множителей:

---

* Ибо представление (8.30) при условии $\psi(\bar{a})\neq 0$ и означает по определению, что комплексное число $\bar{a}$ является корнем многочлена $f(z)$ кратности λ.

** Мы учитываем, что комплексное число равно нулю тогда и только тогда, когда равно нулю сопряженное ему комплексное число (см. п. 1 настоящего параграфа).

$$f(x) = (x-b_1)^{\beta_1}(x-b_2)^{\beta_2}\ldots(x-b_m)^{\beta_m}\times$$
$$\times(x^2+p_1x+q_1)^{\lambda_1}(x^2+p_2x+q_2)^{\lambda_2}\ldots(x^2+p_nx+q_n)^{\lambda_n}. \quad (8.34)$$

Мы приходим к выводу, что многочлен $f(x)$ с вещественными коэффициентами распадается на произведение (8.34) неприводимых вещественных множителей, причем множители, соответствующие вещественным корням, имеют вид двучленов в степенях, равных кратности корней, а множители, соответствующие комплексным парам корней, имеют вид квадратных трехчленов в степенях, равных кратности этих пар корней.

4. Разложение правильной рациональной дроби на сумму простейших дробей. Рациональной дробью называется отношение двух алгебраических многочленов.

Всюду в дальнейшем мы будем рассматривать рациональную дробь, являющуюся отношением двух алгебраических многочленов с вещественными коэффициентами (такую дробь принято называть рациональной дробью с вещественными коэффициентами).

Рациональная дробь $\frac{P(x)}{Q(x)}$ называется правильной, если степень многочлена $P(x)$, стоящего в числителе, меньше степени многочлена $Q(x)$, стоящего в знаменателе.

В противном случае рациональная дробь называется неправильной.

Докажем две вспомогательные теоремы.

Лемма 1. *Пусть $\frac{P(x)}{Q(x)}$ — правильная рациональная дробь с вещественными коэффициентами, знаменатель $Q(x)$ которой имеет вещественное число a корнем кратности α, т. е.*

$$Q(x) = (x-a)^{\alpha}\varphi(x), \quad \text{где} \quad \varphi(a) \neq 0. \quad (8.35)$$

Тогда для этой дроби справедливо следующее представление:

$$\frac{P(x)}{Q(x)} = \frac{A}{(x-a)^{\alpha}} + \frac{\psi(x)}{(x-a)^{\alpha-k}\varphi(x)}. \quad (8.36)$$

В этом представлении A — вещественная постоянная, равная $A = \frac{P(a)}{\varphi(a)}$, k — целое число, удовлетворяющее условию $k \geqslant 1$, $\psi(x)$ — некоторый многочлен с вещественными коэффициентами такой, что последняя дробь в правой части (8.36) является правильной.

Доказательство. Обозначив через A — вещественное число $A = \frac{P(a)}{\varphi(a)}$ *, рассмотрим разность

---

* Число A всегда определено, ибо $\varphi(a) \neq 0$ в силу (8.35).

$$\frac{P(x)}{Q(x)} - \frac{A}{(x-a)^{\alpha}}.$$

Приводя эту разность к общему знаменателю, будем иметь

$$\frac{P(x)}{Q(x)} - \frac{A}{(x-a)^{\alpha}} = \frac{P(x) - A\varphi(x)}{(x-a)^{\alpha}\varphi(x)} = \frac{\Phi(x)}{(x-a)^{\alpha}\varphi(x)}, \qquad (8.37)$$

где через $\Phi(x)$ обозначен многочлен с вещественными коэффициентами вида $\Phi(x) = P(x) - A\varphi(x)$.

Так как $\Phi(a) = P(a) - A\varphi(a) = P(a) - \frac{P(a)}{\varphi(a)}\varphi(a) = 0$, вещественное число a является корнем многочлена $\Phi(x)$ некоторой кратности $k \geqslant 1$.

Это означает, что справедливо представление

$$\Phi(x) = (x-a)^k \psi(x), \quad \text{где} \quad \psi(a) \neq 0, \qquad (8.38)$$

а $\psi(x)$ — некоторый многочлен с вещественными коэффициентами.

Вставляя представление (8.38) в равенство (8.37), окончательно будем иметь

$$\frac{P(x)}{Q(x)} - \frac{A}{(x-a)^{\alpha}} = \frac{\psi(x)}{(x-a)^{\alpha-k}\varphi(x)}. \qquad (8.39)$$

Тем самым представление (8.36) доказано. Остается только убедиться в том, что дробь, стоящая в правой части (8.39), является правильной, но это сразу вытекает из того, что разность двух правильных рациональных дробей является правильной рациональной дробью (чтобы убедиться в этом, достаточно привести разность правильных рациональных дробей к общему знаменателю). Лемма 1 доказана.

Лемма 2. *Пусть* $\frac{P(x)}{Q(x)}$ — *правильная рациональная дробь с вещественными коэффициентами, знаменатель которой* $Q(x)$ *имеет комплексные числа* $a = u + iv$ *и* $\bar{a} = u - iv$ *корнями кратности* λ, *т. е.*

$$Q(x) = (x^2 + px + q)^{\lambda}\varphi(x), \quad \text{где} \quad \varphi(a) \neq 0, \quad \varphi(\bar{a}) \neq 0, \qquad (8.40)$$
$$p = -2u, \quad q = u^2 + v^2.$$

Тогда для этой дроби справедливо следующее представление:

$$\frac{P(x)}{Q(x)} = \frac{Mx+N}{(x^2+px+q)^{\lambda}} + \frac{\psi(x)}{(x^2+px+q)^{\lambda-k}\varphi(x)}. \qquad (8.41)$$

В этом представлении M *и* N *— некоторые вещественные постоянные,* k *— целое число* $\geqslant 1$, *а* $\psi(x)$ *— некоторый многочлен с вещественными коэффициентами такой, что последняя дробь в правой части (8.41) является правильной.*

Доказательство. Договоримся обозначать вещественную часть комплексной величины A символом $\operatorname{Re}[A]$, мнимую часть комплексной величины A символом $\operatorname{Im}[A]$. Положим *

$$M=\frac{1}{v}\operatorname{Im}\left[\frac{P(a)}{\varphi(a)}\right],\quad N=\operatorname{Re}\left[\frac{P(a)}{\varphi(a)}\right]-\frac{u}{v}\operatorname{Im}\left[\frac{P(a)}{\varphi(a)}\right].$$

Нетрудно проверить, что указанные M и N являются решением следующего уравнения:

$$P(a)-(Ma+N)\varphi(a)=0. \tag{8.42}$$

В самом деле, поделив это уравнение на $\varphi(a)$ и приравняв нулю действительные и мнимые части, мы получим два равенства

$$\left.\begin{aligned} Mu+N&=\operatorname{Re}\left[\frac{P(a)}{\varphi(a)}\right],\\ Mv&=\operatorname{Im}\left[\frac{P(a)}{\varphi(a)}\right],\end{aligned}\right\}$$

из которых определяются написанные выше M и N. Рассмотрим теперь разность

$$\frac{P(x)}{Q(x)}-\frac{Mx+N}{(x^2+px+q)^\lambda}.$$

Приводя указанную разность к общему знаменателю, будем иметь

$$\frac{P(x)}{Q(x)}-\frac{Mx+N}{(x^2+px+q)^\lambda}=\frac{P(x)-(Mx+N)\varphi(x)}{(x^2+px+q)^\lambda\varphi(x)}=\frac{\Phi(x)}{(x^2+px+q)^\lambda\varphi(x)}. \tag{8.43}$$

Здесь через $\Phi(x)$ обозначен многочлен с вещественными коэффициентами вида $\Phi(x)=P(x)-(Mx+N)\varphi(x)$. Равенство (8.42) позволяет утверждать, что комплексное число a, а значит, в силу теоремы 8.3 и сопряженное ему число $\bar a$ являются корнями многочлена $\Phi(x)$ некоторой кратности $k\geqslant 1$. В таком случае для многочлена $\Phi(x)$ справедливо представление

$$\Phi(x)=(x^2+px+q)^k\psi(x), \tag{8.44}$$

где $\psi(x)$ — некоторый многочлен с вещественными коэффициентами, не имеющий в качестве корней числа a и $\bar a$. Вставляя представление (8.44) в формулу (8.43), получим представление (8.41). Тот факт, что последняя дробь, стоящая в правой части (8.41), является правильной, вытекает из того, что эта дробь равна разности двух правильных дробей.

Лемма 2 доказана.

Последовательное применение лемм 1 и 2 к дроби $\dfrac{P(x)}{Q(x)}$ по

---

* В силу (8.40) $\varphi(a)\neq 0$, так что отношение $\dfrac{P(a)}{\varphi(a)}$ рассматривать можно.

всем корням знаменателя приводит нас к следующему замечательному утверждению.

Теорема 8.4. *Пусть* $\frac{P(x)}{Q(x)}$ — *правильная рациональная дробь с вещественными коэффициентами, знаменатель которой имеет вид*

$$Q(x) = (x-b_1)^{\beta_1}(x-b_2)^{\beta_2}\ldots(x-b_m)^{\beta_m}(x^2+p_1x+q_1)^{\lambda}\ldots(x^2+p_nx+q_n)^{\lambda_n}. \tag{8.45}$$

Тогда для этой дроби справедливо следующее разложение на сумму простейших дробей:

$$\begin{aligned}\frac{P(x)}{Q(x)} &= \frac{B_1^{(1)}}{(x-b_1)} + \frac{B_2^{(1)}}{(x-b_1)^2} + \ldots + \frac{B_{\beta_1}^{(1)}}{(x-b_1)^{\beta_1}} + \ldots + \\ &+ \frac{B_1^{(m)}}{(x-b_m)} + \frac{B_2^{(m)}}{(x-b_m)^2} + \ldots + \frac{B_{\beta_m}^{(m)}}{(x-b_m)^{\beta_m}} + \\ &+ \frac{M_1^{(1)}x+N_1^{(1)}}{(x^2+p_1x+q_1)} + \frac{M_2^{(1)}x+N_2^{(1)}}{(x^2+p_1x+q_1)^2} + \ldots + \frac{M_{\lambda_1}^{(1)}x+N_{\lambda_1}^{(1)}}{(x^2+p_1x+q_1)^{\lambda_1}} + \\ &+ \frac{M_1^{(n)}x+N_1^{(n)}}{(x^2+p_nx+q_n)} + \frac{M_2^{(n)}x+N_2^{(n)}}{(x^2+p_nx+q_n)^2} + \ldots + \frac{M_{\lambda_n}^{(n)}x+N_{\lambda_n}^{(n)}}{(x^2+p_nx+q_n)^{\lambda_n}}.\end{aligned} \tag{8.46}$$

В этом разложении $B_1^{(1)}, B_2^{(1)}, \ldots, B_{\beta_m}^{(m)}, M_1^{(1)}, N_1^{(1)}, \ldots, M_{\lambda_n}^{(n)}, N_{\lambda_n}^{(n)}$ — *некоторые вещественные постоянные, часть из которых может быть равна нулю.*

Замечание. Для конкретного определения только что указанных постоянных следует привести равенство (8.46) к общему знаменателю и после этого сравнивать коэффициенты при одинаковых степенях x в числителе.

Примеры и разъяснения.

1°. Разложить на сумму простейших правильную дробь

$$\frac{2x^3+4x^2+x+2}{(x-1)^2(x^2+x+1)}. \tag{8.47}$$

Убедившись в том, что квадратный трехчлен x^2+x+1 имеет комплексные корни, ищем, согласно теореме 8.4, разложение дроби (8.47) в виде

$$\frac{2x^3+4x^2+x+2}{(x-1)^2(x^2+x+1)} = \frac{B_1}{(x-1)} + \frac{B_2}{(x-1)^2} + \frac{Mx+N}{x^2+x+1}. \tag{8.48}$$

Приводя равенство (8.48) к общему знаменателю, получим

$$\frac{2x^3+4x^2+x+2}{(x-1)^2(x^2+x+1)} = \frac{B_1(x^3-1)+B_2(x^2+x+1)+(Mx+N)(x^2-2x+1)}{(x-1)^2(x^2+x+1)}.$$

Сравнивая в числителях коэффициенты при x^0, x^1, x^2 и x^3, придем к системе уравнений *

$$\left.\begin{aligned} B_1+M \qquad &=2,\\ B_2+N-2M&=4,\\ B_2+M-2N&=1,\\ B_1+B_2+H&=2. \end{aligned}\right\}$$

Решая эту систему, найдем $B_1=2$, $B_2=3$, $M=0$, $N=1$. Окончательно получим

$$\frac{2x^3+4x^2+x+2}{(x-1)^2(x^2+x+1)}=\frac{2}{x-1}+\frac{3}{(x-1)^2}+\frac{1}{x^2+x+1}. \qquad (8.49)$$

Только что проиллюстрированный метод отыскания разложения правильной рациональной дроби называется **методом неопределенных коэффициентов**. Этот метод приводит к цели всегда; доказывать разрешимость полученной в результате применения этого метода системы уравнений не нужно — разрешимость вытекает из теоремы 8.4.

2°. Проиллюстрируем метод неопределенных коэффициентов еще одним примером. Требуется найти разложение правильной дроби

$$\frac{3x^4+2x^3+3x^2-1}{(x-2)(x^2+1)^2}.$$

Так как квадратный трехчлен x^2+1 имеет комплексные корни, ищем, согласно теореме 8.4, разложение в виде

$$\frac{3x^4+2x^3+3x^2-1}{(x-2)(x^2+1)^2}=\frac{B}{x-2}+\frac{M_1x+N_1}{x^2+1}+\frac{M_2x+N_2}{(x^2+1)^2}.$$

Последнее равенство приводим к общему знаменателю и после этого сопоставляем числители. Получим

$$3x^4+2x^2+3x^2-1=$$
$$=B(x^4+2x^2+1)+(M_1x+N_1)(x^3-2x^2+x-2)+(M_2x+N_2)(x-2).$$

Сравнивая коэффициенты при x^0, x^1, x^2, x^3 и x^4, придем к системе уравнений

$$\left\{\begin{aligned} &B+M_1 &&=3,\\ &N_1-2M_1 &&=2,\\ &2B+M_1-2N_1+M_2 &&=3,\\ &N_1-2M_1+N_2-2M_2 &&=0,\\ &B-2N_1-2N_2 &&=-1. \end{aligned}\right.$$

* При этом мы используем утверждение, сформулированное в сноске *** на с. 308.

Решая эту систему, найдем $B=3$, $M_1=0$, $N_1=2$, $M_2=1$, $N_2=0$. Окончательно получим

$$\frac{3x^4+2x^3+3x^2-1}{(x-2)(x^2+1)^2}=\frac{3}{x-2}+\frac{2}{x^2+1}+\frac{x}{(x^2+1)^2}. \tag{8.50}$$

3°. Метод неопределенных коэффициентов, как видно из рассмотренных примеров, является довольно громоздким. Естественно поэтому в тех случаях, когда это возможно, найти другой, более простой метод отыскания коэффициентов в разложении правильной рациональной дроби на сумму простейших. Пусть знаменатель $Q(x)$ правильной рациональной дроби $\frac{P(x)}{Q(x)}$ имеет вещественное число a корнем кратности α. Тогда среди простейших дробей, на сумму которых раскладывается дробь $\frac{P(x)}{Q(x)}$, будет фигурировать дробь

$$\frac{A}{(x-a)^\alpha}. \tag{8.51}$$

Укажем совсем простой метод вычисления коэффициента A при этой простейшей дроби. Привлекая лемму 1 и формулу (8.36), мы убедимся в том, что коэффициент A равен

$$A=\frac{P(a)}{Q(a)}, \text{ где } \varphi(x)=\frac{Q(x)}{(x-a)^\alpha}.$$

Мы приходим к следующему правилу: *для вычисления коэффициента A при простейшей дроби (8.51), соответствующей вещественному корню a многочлена $Q(x)$ кратности α, следует вычеркнуть в знаменатéле дроби $\frac{P(x)}{Q(x)}$ скобку $(x-a)^\alpha$ и в оставшемся выражении положить $x=a$.*

Указанный прием нахождения коэффициента A обычно называют **методом вычеркивания**. Отметим, что этот прием применим лишь для вычисления коэффициентов при *старших степенях простейших дробей, соответствующих вещественным корням* $Q(x)$.

Метод вычеркивания особенно эффективен в случае, когда знаменатель $Q(x)$ *имеет лишь однократные вещественные корни, т. е. когда* $Q(x)=(x-a_1)(x-a_2)\ldots(x-a_n)$. Тогда, как мы знаем, справедливо разложение

$$\frac{P(x)}{Q(x)}=\frac{A_1}{x-a_1}+\frac{A_2}{x-a_2}+\ldots+\frac{A_k}{x-a_k}+\ldots+\frac{A_n}{x-a_n},$$

все коэффициенты которого могут быть вычислены по методу вычеркивания. Для вычисления коэффициента A_k следует вычеркнуть в знаменателе дроби $\frac{P(x)}{Q(x)}$ скобку $(x-a_k)$ и в оставшемся выражении положить $x=a_k$.

Пример. Найти разложение дроби

$$\frac{x+1}{(x-1)\,x\,(x-2)}. \tag{8.52}$$

Согласно теореме 8.4 пишем:

$$\frac{x+1}{(x-1)\,x\,(x-2)}=\frac{A_1}{x-1}+\frac{A_2}{x}+\frac{A_3}{x-2}.$$

Для отыскания A_1 вычеркиваем в выражении (8.52) скобку $(x-1)$ и в оставшемся выражении берем $x=1$. Получим $A_1=-2$. Аналогично находим $A_2=\frac{1}{2}$, $A_3=3/2$.

Окончательно получим

$$\frac{x+1}{(x-1)\,x\,(x-2)}=\frac{-2}{x-1}+\frac{1}{2x}+\frac{3}{2\,(x-2)}. \tag{8.53}$$

5. Интегрируемость рациональной дроби в элементарных функциях. Теперь мы подготовлены к тому, чтобы в общем виде решить проблему об интегрировании рациональной дроби с вещественными коэффициентами.

Прежде всего отметим, что эта проблема сводится к проблеме интегрирования *только правильной* рациональной дроби, ибо всякую неправильную рациональную дробь можно (посредством деления числителя на знаменатель «столбиком») представить в виде суммы алгебраического многочлена и правильной рациональной дроби.

Пример. $\dfrac{x^4-x^3+1}{x^2+x+2}=(x^2-2x)+\dfrac{4x+1}{x^2+x+2}$, ибо

$$\begin{array}{r|l}
x^4-x^3+1 & x^2+x+2 \\ \cline{2-2}
-\;\underline{x^4+x^3+2x^2} & x^2-2x \\
-2x^3-2x^2+1 & \\
-\;\underline{-2x^3-2x^2-4x} & \\
\text{остаток }1+4x &
\end{array}$$

Интегрировать многочлен мы умеем (напомним, что неопределенный интеграл от многочлена представляет собой некоторый многочлен степени, на единицу более высокой). Остается научиться интегрировать *правильную* рациональную дробь. В силу теоремы 8.4 проблема интегрирования правильной рациональной дроби сводится к интегрированию простейших дробей *следующих четырех типов*:

$$\text{I. } \frac{B}{x-b};\quad \text{II. } \frac{B}{(x-b)^{\beta}};\quad \text{III. } \frac{Mx+N}{x^2+px+q};\quad \text{IV. } \frac{Mx+N}{(x^2+px+q)^{\lambda}}. \tag{8.54}$$

Здесь $\beta=2, 3, \ldots$; $\lambda=2, 3, \ldots$; B, M, N, b, p, q — некоторые вещественные числа, причем трехчлен x^2+px+q не имеет вещественных корней, т. е. $q-\frac{p^2}{4}>0$.

Докажем, что каждая из четырех указанных дробей интегрируема в элементарных функциях.

Дроби вида I и II элементарно интегрируются при помощи подстановки $t=x-b$. Мы получим

$$\int \frac{B}{x-b}\,dx = B\int \frac{dt}{t} = B\ln|t| + C = B\ln|x-b| + C, \qquad (8.55)$$

$$\int \frac{B}{(x-b)^{\beta}}\,dx = B\int \frac{dt}{t^{\beta}} = -\frac{B}{(\beta-1)}\,\frac{1}{t^{\beta-1}} + C = \frac{-B}{(\beta-1)}\cdot\frac{1}{(x-b)^{\beta-1}} + C. \qquad (8.56)$$

Для вычисления интеграла от дроби вида III представим квадратный трехчлен в виде $x^2+px+q=\left(x+\frac{p}{2}\right)^2+\left(q-\frac{p^2}{4}\right)$ и, учитывая, что $q-\frac{p^2}{4}>0$, введем в рассмотрение вещественную постоянную $a=+\sqrt{q-\frac{p^2}{4}}$. Сделав подстановку $t=x+\frac{p}{2}$, будем иметь

$$\int \frac{Mx+N}{x^2+px+q}\,dx = \int \frac{Mt+\left(N-\frac{Mp}{2}\right)}{t^2+a^2}\,dt =$$

$$= \frac{M}{2}\int \frac{2t\,dt}{t^2+a^2} + \left(N-\frac{Mp}{2}\right)\int \frac{dt}{t^2+a^2} =$$

$$= \frac{M}{2}\int \frac{d(t^2+a^2)}{t^2+a^2} + \left(N-\frac{Mp}{2}\right)\frac{1}{a}\int \frac{d\left(\frac{t}{a}\right)}{\left(\frac{t}{a}\right)^2+1} =$$

$$= \frac{M}{2}\ln(t^2+a^2) + \frac{2N-Mp}{2a}\operatorname{arc\,tg}\frac{t}{a} + C =$$

$$= \frac{M}{2}\ln(x^2+px+q) + \frac{2N-Mp}{2\sqrt{q-\frac{p^2}{4}}}\operatorname{arc\,tg}\frac{x+\frac{p}{2}}{\sqrt{q-\frac{p^2}{4}}} + C. \qquad (8.57)$$

Остается вычислить интеграл от дроби вида IV. Используя введенные выше обозначения $t=x+\frac{p}{2}$, $a=\sqrt{q-\frac{p^2}{4}}$, будем иметь

$$\int \frac{Mx+N}{(x^2+px+q)^\lambda}\,dx = \int \frac{Mt+\left(N-\frac{Mp}{2}\right)}{(t^2+a^2)^\lambda}\,dt =$$

$$= \frac{M}{2}\int \frac{d\,(t^2+a^2)}{(t^2+a^2)^\lambda} + \left(N-\frac{Mp}{2}\right)\int \frac{dt}{(t^2+a^2)^\lambda}.$$

Введем обозначения

$$I = \int \frac{d\,(t^2+a^2)}{(t^2+a^2)^\lambda}, \quad K_\lambda = \int \frac{dt}{(t^2+a^2)^\lambda}.$$

Интересующий нас интеграл будет вычислен, если будут вычислены интегралы I и K_λ. Интеграл I берется элементарно:

$$I = -\frac{1}{(\lambda-1)}\,\frac{1}{(t^2+a^2)^{\lambda-1}} + C = -\frac{1}{(\lambda-1)}\,\frac{1}{(x^2+px+q)^{\lambda-1}} + C.$$

Интеграл K_λ вычислен нами в примере 6 в конце § 2 настоящей главы. Там мы получили для этого интеграла рекуррентную формулу (3.12), позволяющую последовательно вычислить K_λ для любого $\lambda=2, 3, \ldots$, опираясь на то, что

$$K_1 = \int \frac{dt}{t^2+a^2} = \frac{1}{a}\operatorname{arc\,tg}\frac{t}{a} + C.$$

Итак, нами вычислены интегралы от всех четырех простейших дробей (8.54) и доказано, что каждый из этих интегралов представляет собой *элементарную функцию* *. Тем самым мы приходим к следующей теореме, исчерпывающей проблему интегрирования рациональной дроби.

Теорема 8.5. *Всякая рациональная дробь с вещественными коэффициентами интегрируема в элементарных функциях.*

В заключение этого параграфа мы остановимся на примерах вычисления неопределенных интегралов от рациональных дробей. Вычислим неопределенные интегралы от трех дробей, рассмотренных в предыдущем параграфе, (8.49), (8.50) и (8.53). Пользуясь указанными тремя формулами, а также формулами (8.55) — (8.57), будем иметь:

1. $$\int \frac{2x^3+4x^2+x+2}{(x-1)^2(x^2+x+1)}\,dx =$$

$$= \int \frac{2}{x-1}\,dx + \int \frac{3}{(x-1)^2}\,dx + \int \frac{dx}{x^2+x+1} =$$

$$= 2\ln|x-1| - \frac{3}{x-1} + \frac{2}{\sqrt{3}}\operatorname{arc\,tg}\frac{2x+1}{\sqrt{3}} + C.$$

* Точнее, выражается через логарифм, арктангенс и рациональную функцию.

2. $$\int \frac{3x^4+2x^3+3x^2-1}{(x-2)(x^2+1)^2}\,dx = \int \frac{3}{x-2}\,dx + \int \frac{2dx}{x^2+1} + \int \frac{x\,dx}{(x^2+1)^2} =$$
$$= 3\ln|x-2| + 2\operatorname{arc\,tg} x + \frac{1}{2}\int \frac{d(x^2+1)}{(x^2+1)^2} =$$
$$= 3\ln|x-2| + 2\operatorname{arc\,tg} x - \frac{1}{2(x^2+1)} + C.$$

3. $$\int \frac{x+1}{(x-1)x(x-2)}\,dx = \int \frac{-2}{x-1}\,dx + \int \frac{dx}{2x} + \int \frac{3dx}{2(x-2)} =$$
$$= -2\ln|x-1| + \frac{1}{2}\ln|x| + \frac{3}{2}\ln|x-2| + C.$$

6. Интегрируемость в элементарных функциях некоторых тригонометрических и иррациональных выражений. В рассуждениях настоящего пункта важную роль будет играть рациональная функция от двух аргументов. С определения такой функции и выяснения некоторых ее свойств и начнем наше изложение.

Многочленом степени n от двух аргументов x и y называется выражение вида

$$P_n(x, y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + \ldots + a_{0n}y^n,$$

в котором через $a_{00}, a_{10}, \ldots, a_{0n}$ обозначены некоторые постоянные вещественные числа такие, что среди чисел $a_{n0}, a_{(n-1)1}, a_{(n-2)2}, \ldots$ $\ldots, a_{0n}$ есть хотя бы одно число, отличное от нуля.

Рациональной функцией от двух аргументов x и y называется выражение вида

$$R(x, y) = \frac{P_n(x, y)}{Q_m(x, y)},$$

в котором через $P_n(x, y)$ обозначен произвольный многочлен от двух аргументов x и y степени n, а через $Q_m(x, y)$ обозначен произвольный многочлен от двух аргументов x и y степени m.

Справедливо следующее тривиальное утверждение: *если $R(x, y)$ — рациональная функция от двух аргументов x и y, а $R_1(t)$, $R_2(t)$ и $R_3(t)$ — три произвольных рациональных функции от одной переменной t*, то выражение вида*

$$R[R_1(t),\ R_2(t)]R_3(t) \tag{8.58}$$

представляет собой рациональную функцию от одной переменной.

Для доказательства этого утверждения достаточно заметить, что в результате применения к рациональным функциям одной пе-

---

* Под термином «рациональная функция от одной переменной t» мы понимаем рациональную дробь с вещественными коэффициентами от аргумента t.

ременной t операций сложения, вычитания, умножения и деления мы снова получим рациональные функции одной переменной t.

В дальнейшем для доказательства интегрируемости в элементарных функциях некоторых выражений мы будем посредством специально подобранной подстановки сводить интеграл от рассматриваемых выражений к интегралу от рациональной дроби *. При этом мы будем говорить, что интеграл от рассматриваемого выражения рационализируется указанной специальной подстановкой.

1°. Интегрирование некоторых тригонометрических выражений. Договоримся всюду в дальнейшем символом $R(x, y)$ обозначать любую рациональную функцию от двух аргументов x и y.

В этом пункте мы докажем интегрируемость в элементарных функциях любой функции вида

$$R(\sin x, \cos x). \tag{8.59}$$

Докажем, что интеграл от этой функции рационализируется подстановкой $t=\operatorname{tg}\frac{x}{2}$. Действительно,

$$\sin x=\frac{2\operatorname{tg}\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}}=\frac{2t}{1+t^2},\quad \cos x=\frac{1-\operatorname{tg}^2\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}}=\frac{1-t^2}{1+t^2},$$

$$x=2\operatorname{arc\,tg} t,\quad dx=\frac{2dt}{1+t^2},$$

так что

$$\int R(\sin x, \cos x)\,dx=\int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2}. \tag{8.60}$$

Если положить $R_1(t)=\frac{2t}{1+t^2}$, $R_2(t)=\frac{1-t^2}{1+t^2}$, $R_3(t)=\frac{2}{1+t^2}$, то в правой части (8.60) мы получим интеграл от выражения вида (8.58), который представляет собой интеграл от рациональной функции аргумента t.

Пример. Вычислить интеграл $I_1=\int\frac{dx}{1+a\cos x}$, где $a>0$, $a\neq 1$. Применяя универсальную тригонометрическую подстановку $t=\operatorname{tg}\frac{x}{2}$, получим

$$x=2\operatorname{arc\,tg} t,\quad dx=\frac{2dt}{1+t^2},\quad \cos x=\frac{1-t^2}{1+t^2},$$

---

* Интегрируемость в элементарных функциях рациональной дроби установлена в теореме 8.5.

$$I_1 = 2\int \frac{dt}{(a+1)+t^2(1-a)} = \frac{2}{a+1}\int \frac{dt}{1+\frac{1-a}{1+a}t^2}.$$

Далее нужно отдельно рассмотреть два случая: 1) $0<a<1$, 2) $a>1$.

В случае $0<a<1$

$$I_1 = \frac{2}{\sqrt{1-a^2}}\operatorname{arc\,tg}\left(t\sqrt{\frac{1-a}{1+a}}\right)+C =$$

$$= \frac{2}{\sqrt{1-a^2}}\operatorname{arc\,tg}\left(\sqrt{\frac{1-a}{1+a}}\operatorname{tg}\frac{x}{2}\right)+C.$$

В случае $a>1$

$$I_1 = \frac{1}{\sqrt{a^2-1}}\ln\left|\frac{1+t\sqrt{\frac{a-1}{a+1}}}{1-t\sqrt{\frac{a-1}{a+1}}}\right|+C=$$

$$= \frac{1}{\sqrt{a^2-1}}\ln\left|\frac{1+\sqrt{\frac{a-1}{a+1}}\operatorname{tg}\frac{x}{2}}{1-\sqrt{\frac{a-1}{a+1}}\operatorname{tg}\frac{x}{2}}\right|+C.$$

2°. Интегрирование дробно-линейных иррациональностей. В этом пункте мы докажем интегрируемость в элементарных функциях любой функции вида

$$R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right), \tag{8.61}$$

где a, b, c и d — некоторые постоянные, n — любое целое положительное число. Функцию такого вида мы будем называть дробно-линейной иррациональностью.

Докажем, что интеграл от функции (8.61) при $ad-bc\neq 0$ рационализируется подстановкой $t=\sqrt[n]{\frac{ax+b}{cx+d}}$. В самом деле,

$$t^n = \frac{ax+b}{cx+d}, \quad x = \frac{dt^n-b}{a-ct^n}, \quad dx = \frac{(ad-bc)\,nt^{n-1}}{(a-ct^n)^2}\,dt,$$

так что

$$\int R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right)dx = \int R\left(\frac{dt^n-b}{a-ct^n},\ t\right)\frac{(ad-bc)\,nt^{n-1}}{(a-ct^n)^2}\,dt. \tag{8.62}$$

Если положить $R_1(t)=\dfrac{dt^n-b}{a-ct^n}$, $R_2(t)=t$, $R_3(t)=\dfrac{(ad-bc)\,nt^{n-1}}{(a-ct^n)^2}$, то в правой части (8.62) мы получим интеграл от выражения вида (8.58), который представляет собой интеграл от рациональной функции аргумента t. Тем самым доказано, что интеграл от дробно-линейной иррациональности (8.61) рационализируется подстановкой $t=\sqrt[n]{\dfrac{ax+b}{cx+d}}$.

Пример. Вычислить интеграл $I=\int\sqrt{\dfrac{1+x}{1-x}}\dfrac{dx}{1-x}$. Сделав подстановку

$$t=\sqrt{\frac{1+x}{1-x}},\quad t^2=\frac{1+x}{1-x},\quad x=\frac{t^2-1}{t^2+1},\quad dx=\frac{4t\,dt}{(t^2+1)^2},$$

и учитывая, что $1-x=2/(t^2+1)$, получим

$$I=2\int\frac{t^2\,dt}{t^2+1}=2\int dt-2\int\frac{dt}{t^2+1}=2t-2\operatorname{arctg}t+C=$$
$$=2\sqrt{\frac{1+x}{1-x}}-2\operatorname{arctg}\sqrt{\frac{1+x}{1-x}}+C.$$

3°. Интегрирование квадратичных иррациональностей. В этом пункте мы докажем интегрируемость в элементарных функциях любой функции вида

$$R\left(x,\ \sqrt{ax^2+bx+c}\right), \tag{8.63}$$

где a, b и c — некоторые постоянные. Функцию такого вида будем называть квадратичной иррациональностью. При этом мы, конечно, считаем, что квадратичный трехчлен ax^2+bx+c не имеет равных корней (иначе корень из этого трехчлена может быть заменен рациональным выражением).

Мы докажем, что интеграл от функции (8.63) всегда рационализируется одной из так называемых подстановок Эйлера.

Сначала рассмотрим случай, когда квадратичный трехчлен ax^2+bx+c имеет комплексные корни. В этом случае знак квадратного трехчлена совпадает со знаком a, и поскольку по смыслу квадратный трехчлен (из которого извлекается квадратный корень) положителен, то $a>0$.

Таким образом, мы имеем право сделать следующую подстановку:

$$t=\sqrt{ax^2+bx+c}+x\sqrt{a}. \tag{8.64}$$

Подстановку (8.64) обычно называют первой подстановкой Эйлера. Докажем, что эта подстановка рационализирует интеграл от функции (8.63) для рассматриваемого случая. Возводя в квадрат обе части равенства $\sqrt{ax^2+bx+c}=t-x\sqrt{a}$, получим $bx+c=t^2-2\sqrt{a}\,tx$, так что

$$x=\frac{t^2-c}{2\sqrt{a}t+b},\quad \sqrt{ax^2+bx+c}=\frac{\sqrt{a}t^2+bt+c\sqrt{a}}{2\sqrt{a}t+b},$$

$$dx=2\frac{\sqrt{a}t^2+bt+c\sqrt{a}}{(2\sqrt{a}t+b)^2}\,dt.$$

Таким образом,

$$\int R\left(x,\ \sqrt{ax^2+bx+c}\right)dx=$$

$$=\int R\left(\frac{t^2-c}{2\sqrt{a}t+b},\ \frac{\sqrt{a}t^2+bt+c\sqrt{a}}{2\sqrt{a}t+b}\right)2\frac{\sqrt{a}t^2+bt+c\sqrt{a}}{(2\sqrt{a}t+b)^2}\,dt. \qquad (8.65)$$

Под знаком интеграла в правой части (8.65) стоит выражение вида (8.58) при $R_1(t)=\dfrac{t^2-c}{2\sqrt{a}t+b}$, $R_2(t)=\dfrac{\sqrt{a}t^2+bt+c\sqrt{a}}{2\sqrt{a}t+b}$, $R_3(t)=2\dfrac{\sqrt{a}t^2+bt+c\sqrt{a}}{(2\sqrt{a}t+b)^2}$. Таким образом, в правой части (8.65) мы получаем интеграл от рациональной дроби.

Рассмотрим теперь случай, когда квадратный трехчлен ax^2+bx+c имеет несовпадающие вещественные корни x_1 и x_2. В таком случае $ax^2+bx+c=a(x-x_1)(x-x_2)$. Докажем, что в этом случае интеграл от функции (8.63) рационализируется посредством подстановки

$$t=\frac{\sqrt{ax^2+bx+c}}{x-x_1}, \qquad (8.66)$$

называемой обычно **второй подстановкой Эйлера**. В самом деле, возводя в квадрат равенство $\sqrt{ax^2+bx+c}=t(x-x_1)$ и сокращая полученное равенство на $x-x_1$, получим $a(x-x_2)=$ $=t^2(x-x_1)$, так что

$$x=\frac{-ax_2+x_1t^2}{t^2-a},\quad \sqrt{ax^2+bx+c}=\frac{a(x_1-x_2)}{t^2-a}t,$$

$$dx=\frac{2a(x_1-x_2)t}{(t^2-a)^2}\,dt.$$

Таким образом,

$$\int R(x,\ \sqrt{ax^2+bx+c})\,dx=$$

$$=\int R\left(\frac{-ax_2+x_1t^2}{t^2-a},\ \frac{a(x_1-x_2)t}{t^2-a}\right)\frac{2a(x_1-x_2)t}{(t^2-a)^2}\,dt. \qquad (8.67)$$

В правой части (8.67) под знаком интеграла стоит выражение вида (8.58) при $R_1(t)=\frac{-ax_2+x_1t^2}{t^2-a}$, $R_2(t)=\frac{a(x_1-x_2)t}{t^2-a}$, $R_3(t)=\frac{2a(x_1-x_2)t}{(t^2-a)^2}$. Таким образом, в правой части (8.67) мы получаем интеграл от рациональной дроби.

Примеры. 1) Вычислить интеграл $I=\int\frac{dx}{x+\sqrt{x^2+x+1}}$. Поскольку квадратный трехчлен x^2+x+1 имеет комплексные корни, сделаем первую подстановку Эйлера

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

Возвышая в квадрат обе части равенства $\sqrt{x^2+x+1}=t-x$, получим $x^2+x+1=t^2-2tx+x^2$, или $x+1=t^2-2tx$, так что

$$x=\frac{t^2-1}{1+2t},\quad dx=2\,\frac{t^2+t+1}{(1+2t)^2}\,dt.$$

Таким образом,

$$I=2\int\frac{t^2+t+1}{t(1+2t)^2}\,dt=\int\left[\frac{A}{t}+\frac{B}{1+2t}+\frac{D}{(1+2t)^2}\right]dt.$$

Неопределенные коэффициенты A, B и D легко вычисляются: $A=2$, $B=-3$, $D=-3$. Окончательно получим

$$I=2\ln|t|-\frac{3}{2}\ln|1+2t|+\frac{3}{2(1+2t)}+C=$$

$$=2\ln|\sqrt{x^2+x+1}+x|-\frac{3}{2}\ln|1+2x+2\sqrt{x^2+x+1}|+$$

$$+\frac{3}{2(1+2x+2\sqrt{x^2+x+1})}+C.$$

2) Вычислить интеграл $I=\int\frac{dx}{1+\sqrt{1-2x-x^2}}$. Поскольку квадратный трехчлен $1-2x-x^2$ имеет вещественные корни $x_1=-1+\sqrt{2}$ и $x_2=-1-\sqrt{2}$, сделаем вторую подстановку Эйлера (8.66)

$$t=\frac{\sqrt{1-2x-x^2}}{x+1+\sqrt{2}}.$$

Возвышая в квадрат обе части равенства

$\sqrt{1-2x-x^2}=t(x+1+\sqrt{2})$, будем иметь $(-1)(x+1-\sqrt{2})=$
$=t^2(x+1+\sqrt{2})$,

так что

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

$$1+\sqrt{1-2x-x^2}=\frac{t^2+2\sqrt{2}t+1}{t^2+1},\quad dx=-\frac{4\sqrt{2}t}{(t^2+1)^2}\,dt.$$

Таким образом,

$$I=-4\sqrt{2}\int\frac{tdt}{(t^2+1)(t^2+2\sqrt{2}+1)}.$$

Получаем интеграл от рациональной дроби, вычисление которого предоставляем читателю.

§ 4. ЭЛЛИПТИЧЕСКИЕ ИНТЕГРАЛЫ

К интегралам от квадратичных иррациональностей естественно примыкают следующие интегралы:

$$\int R(x,\ \sqrt{ax^3+bx^2+cx+d}\ )\,dx, \tag{8.68}$$

$$\int R(x,\ \sqrt{ax^4+bx^3+cx^2+dx+e}\ )\,dx, \tag{8.69}$$

подынтегральные функции которых содержат корень квадратный из многочленов третьей или четвертой степени (с вещественными коэффициентами).

Эти интегралы весьма часто встречаются в приложениях. Отметим сразу же, что интегралы (8.68) и (8.69), *вообще говоря, не являются элементарными функциями.*

Оба эти интеграла принято называть **эллиптическими** в тех случаях, когда они не выражаются через элементарные функции, и **псевдоэллиптическими** в тех случаях, когда они выражаются через элементарные функции *.

Ввиду важности для приложений интегралов (8.68) и (8.69) возникла необходимость составления таблиц и графиков функций, определяемых этими интегралами. При произвольных коэффициентах a, b, c, d и e такие таблицы и графики составить очень трудно. Поэтому возникла задача о сведении всех интегралов вида (8.68) и (8.69) к нескольким типам интегралов, содержащих по возможности меньше произвольных коэффициентов (или, как говорят, о приведении интегралов (8.68) и (8.69) к **канонической** форме).

Прежде всего заметим, что интеграл (8.68) сводится к интегралу (8.69). В самом деле, кубичный трехчлен заведомо име-

* Эти названия происходят оттого, что впервые с этими интегралами встретились при решении задачи о спрямлении эллипса (см. пример 4° п. 6 § 1 гл. 10).

ет хотя бы один вещественный корень x_0, а поэтому его можно представить в виде $ax^3+bx^2+cx+d=a(x-x_0)(x^2+px+q)$.

Сделав подстановку $x-x_0=\pm t^2$, мы, как легко видеть, преобразуем интеграл (8.68) в (8.69).

Таким образом, нам достаточно рассмотреть лишь интеграл (8.69).

В силу п. 3 § 3 многочлен четвертой степени можно разложить на произведение двух квадратных трехчленов с вещественными коэффициентами

$$ax^4+bx^3+cx^2+dx+e=a(x^2+px+q)(x^2+p'x+q').$$

Всегда найдется некоторая линейная или дробно-линейная подстановка, уничтожающая у обоих квадратных трехчленов линейные члены. Сделав такую подстановку, мы с точностью до слагаемого, представляющего собой элементарную функцию, преобразуем интеграл (8.69) к виду

$$\int \frac{R(t^2)\,dt}{\sqrt{A(1+mt^2)(1+m't^2)}}, \tag{8.70}$$

где R — некоторая рациональная функция. Далее можно показать, что при любых комбинациях абсолютных значений и знаков постоянных A, m и m' может быть найдена замена, сводящая интеграл (8.70) к так называемому каноническому интегралу

$$\int \frac{R_1(z^2)\,dz}{\sqrt{(1-z^2)(1-k^2z^2)}}, \tag{8.71}$$

в котором через k обозначена постоянная, удовлетворяющая условию $0<k<1$.

Любой канонический интеграл (8.71) с точностью до слагаемого, представляющего собой элементарную функцию, может быть приведен к следующим трем стандартным интегралам:

$$\int \frac{dz}{\sqrt{(1-z^2)(1-k^2z^2)}}, \quad \int \frac{z^2dz}{\sqrt{(1-z^2)(1-k^2z^2)}}$$

и (8.72)

$$\int \frac{dz}{(1+hz^2)\sqrt{(1-z^2)(1-k^2z^2)}} \quad (0<k<1).$$

Интегралы (8.72) принято называть **эллиптическими интегралами** соответственно 1-го, 2-го и 3-го рода. Каждый из этих интегралов, как показано Лиувиллем *, представляет собой неэлементарную функцию. Эллиптические интегралы 1-го и 2-го рода содержат только один параметр, принимающий вещественные значения из интервала $0<k<1$, а эллиптический инте-

* Жозеф Лиувилль — французский математик (1809—1882).

грал 3-го рода, кроме того, содержит параметр h, который может принимать и комплексные значения.

Лежандр * подверг интегралы (8.72) дальнейшему упрощению, сделав замену $z = \sin\varphi \left(0 \leqslant \varphi \leqslant \frac{\pi}{2}\right)$. С помощью этой замены первый из интегралов (8.72) преобразуется к виду

$$\int \frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}}. \tag{8.73}$$

При этой же замене второй из интегралов (8.73) с точностью до постоянного множителя равен разности интеграла (8.73) и следующего интеграла:

$$\int \sqrt{1-k^2\sin^2\varphi}\, d\varphi. \tag{8.74}$$

Третий из интегралов (8.72) преобразуется к виду

$$\int \frac{d\varphi}{(1+h\sin^2\varphi)\sqrt{1-k^2\sin^2\varphi}}. \tag{8.75}$$

Интегралы (8.73), (8.74) и (8.75) принято называть **эллиптическими интегралами** соответственно 1-го, 2-го и 3-го **рода в форме Лежандра**.

Особенно важную роль в приложениях играют интегралы (8.73) и (8.74). Если считать, что оба эти интеграла обращаются в нуль при $\varphi = 0$, то получатся две вполне определенные функции, которые обычно обозначают символами $F(k, \varphi)$ и $E(k, \varphi)$. Для этих функций составлены обширные таблицы и графики. Лежандром и другими математиками изучены свойства этих функций, для них установлен ряд формул.

Наряду с элементарными функциями функции E и F прочно вошли в семейство функций, часто используемых в анализе. Здесь еще раз стоит отметить *условность* понятия элементарной функции. Вместе с тем следует подчеркнуть, что задачи интегрального исчисления вовсе не ограничиваются изучением функций, интегрируемых в элементарных функциях.

---

* Андриан Мари Лежандр — французский математик (1752—1833).

Глава 9

ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ РИМАНА

Во вводной главе было показано, что к понятию определенного интеграла приводят такие важные задачи естествознания, как задача о вычислении площади криволинейной трапеции и задача об определении пути, пройденного материальной точкой, двигающейся со скоростью $f(x)$ за промежуток времени от $x=a$ до $x=b$. Целью данной главы является построение строгой теории определенного интеграла Римана.

§ 1. ОПРЕДЕЛЕНИЕ ИНТЕГРАЛА. ИНТЕГРИРУЕМОСТЬ

Рассмотрим функцию $f(x)$, определенную в каждой точке сегмента $[a, b]$. Введем понятия разбиения сегмента $[a, b]$, измельчения этого разбиения и объединения двух разбиений.

Определение 1. *Будем говорить, что задано* **разбиение** *сегмента $[a, b]$, если заданы точки $x_0, x_1, \ldots, x_n$ такие, что $a=x_0<x_1<\ldots<x_{n-1}<x_n=b$. Разбиение сегмента $[a, b]$ будем в дальнейшем обозначать символом $\{x_k\}$.*

Определение 2. *Разбиение $\{x_k'\}$ сегмента $[a, b]$ называется* **измельчением** *разбиения $\{x_k\}$ того же сегмента, если каждая точка x_p разбиения $\{x_k\}$ совпадает с одной из точек x_q' разбиения $\{x_k'\}$.*

Определение 3. *Разбиение $\{x_k\}$ сегмента $[a, b]$ называется* **объединением** *разбиений $\{x_k'\}$ и $\{x_k''\}$ того же сегмента, если все точки разбиений $\{x_k'\}$ и $\{x_k''\}$ являются точками разбиения $\{x_k\}$ и других точек разбиение $\{x_k\}$ не содержит.*

Заметим, что объединение двух разбиений является измельчением каждого из них.

Рассмотрим на сегменте $[a, b]$ функцию $f(x)$, принимающую в каждой точке сегмента конечные значения. По данному разбиению $\{x_k\}$ построим число, так называемую «интегральную сумму», $\sigma(x_k, \xi_k)=\sum_{k=1}^{n} f(\xi_k)(x_k-x_{k-1})$, где ξ_k — некоторая точка сегмента $[x_{k-1}, x_k]$. Подчеркнем, что интегральная сумма $\sigma(x_k, \xi_k)$ зависит как от разбиения $\{x_k\}$, так и от выбора точек ξ_k на сегментах $[x_{k-1}, x_k]$. Если обозначить через Δx_k разность x_k-x_{k-1}, то интегральную сумму, в дальнейшем часто обозначаемую просто через σ, можно записать и так:

$$\sigma = \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k,\quad \xi_k \in [x_{k-1},\ x_k].$$

Сегменты $[x_{k-1}, x_k]$ иногда называют **частичными сегментами**, а точки ξ_k — **промежуточными точками**.

Число $d = \max\limits_{1 \leqslant k \leqslant n} \Delta x_k$* договоримся называть **диаметром разбиения** $\{x_k\}$. Введем фундаментальные понятия предела интегральных сумм и интегрируемости функции по Риману.

Определение 4. *Число I называется* **пределом интегральных сумм** $\sigma(x_k, \xi_k)$ *при стремлении диаметра d разбиений $\{x_k\}$ к нулю, если для всякого $\varepsilon > 0$ существует такое число $\delta = \delta(\varepsilon) > 0$, что из условия $d < \varepsilon$ при любом выборе промежуточных точек ξ_k следует неравенство $|I - \sigma| < \varepsilon$.*

Легко убедиться в том, что может существовать только один предел интегральных сумм σ при $d \to 0$.

Для обозначения предела интегральных сумм употребляют символ

$$I = \lim_{d\to 0} \sigma(x_k,\ \xi_k).$$

Определение 5. *Функция $f(x)$ называется* **интегрируемой по Риману на сегменте** $[a, b]$, *если для этой функции на указанном сегменте существует предел I ее интегральных сумм σ при стремлении диаметра d разбиений $\{x_k\}$ к нулю.*

Число I называется **определенным интегралом Римана** от функции $f(x)$ в пределах от a до b и обозначается символом $\int\limits_a^b f(x)\,dx$.

Таким образом, по определению $\int\limits_a^b f(x)\,dx = \lim\limits_{d\to 0} \sigma(x_k, \xi_k)$. Отметим, что число a называют **нижним пределом** интегрирования, а число b — **верхним пределом** интегрирования. Переменную x под знаком определенного интеграла можно заменить на любую другую переменную, т. е. справедливы равенства

$$\int_a^b f(x)\,dx = \int_a^b f(y)\,dy = \int_a^b f(t)\,dt \text{ и т. д.}$$

Выясним геометрический смысл интегральной суммы. Рассмотрим криволинейную трапецию, т. е. фигуру, ограниченную графи-

---

* Т. е. длину наибольшего частичного сегмента.

ком непрерывной неотрицательной функции $y=f(x)$, заданной на сегменте $[a, b]$, двумя прямыми $x=a$ и $x=b$, перпендикулярными к оси абсцисс, и сегментом $[a, b]$ оси абсцисс (рис. 9.1). Очевидно, интегральная сумма $\sigma(x_k, \xi_k)$, отвечающая выбранному разбиению $\{x_k\}$ и данному обозначенному на рисунке выбору точек ξ_k, представляет собой площадь ступенчатой фигуры, заштрихованной на этом рисунке.

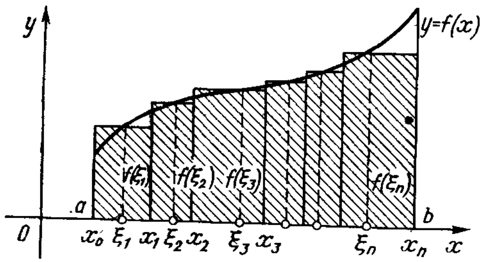


Рис. 9.1

В следующей главе будет сформулировано понятие площади произвольной плоской фигуры и установлено, что предел при $d \to 0$ площади указанной ступенчатой фигуры равен площади криволинейной трапеции.

Приведем простейший пример интегрируемой по Риману функции. Покажем, что функция $f(x)=c=\text{const}$ интегрируема на любом сегменте $[a, b]$, причем $\int\limits_a^b c\,dx=c(b-a)$. Действительно, при любом разбиении $\{x_k\}$ и любом выборе точек ξ_k на сегментах $[x_{k-1}, x_k]$ справедливо равенство $f(\xi_k)=c$. Следовательно,

$$\sigma(x_k, \xi_k)=c\Delta x_1+c\Delta x_2+\ldots+c\Delta x_n=$$
$$=c(\Delta x_1+\Delta x_2+\ldots+\Delta x_n)=c(b-a)$$

для любого разбиения $\{x_k\}$ и любого выбора точек $\xi_k \in [x_{k-1}, x_k]$. Поэтому

$$\int\limits_a^b c\,dx=\lim_{d\to 0}\sigma(x_k, \xi_k)=\lim_{d\to 0}c(b-a)=c(b-a).$$

Докажем следующее

Утверждение. *Если функция $f(x)$ не является ограниченной на сегменте $[a, b]$, то эта функция не интегрируема на этом сегменте.* Пусть $f(x)$ не ограничена на $[a, b]$. Покажем, что для любого разбиения $\{x_k\}$ интегральную сумму $\sigma(x_k, \xi_k)$ можно сделать сколь угодно большой по абсолютной величине только за счет выбора промежуточных точек ξ_k. В самом деле, если функция $f(x)$ не ограничена на сегменте $[a, b]$, а сегмент $[a, b]$ разбит на конечное число сегментов $[x_{k-1}, x_k]$, то функция $f(x)$ будет неограниченной хотя бы на одном частичном сегменте разбиения. Не нарушая общности, будем считать, что $f(x)$ не ограничена на сегменте $[x_0, x_1]$. На остальных сегментах $[x_1, x_2]$, $[x_2, x_3]$, ..., $[x_{n-1}, x_n]$ промежуточные точки $\xi_2, \xi_3, \dots, \xi_n$ выберем произвольными и фиксируем. Обозначим через $\sigma_1(x_k, \xi_k)$ величину

$$\sigma_1=\sigma_1(x_k, \xi_k)=f(\xi_2)\Delta x_2+f(\xi_3)\Delta x_3+\dots+f(\xi_n)\Delta x_n.$$

Рассмотрим теперь функцию $f(x)$ только на сегменте $[x_0, x_1]$. Поскольку фунцкия $f(x)$ не ограничена на этом сегменте, то для любого наперед заданного положительного числа M найдется точка ξ_1 из этого сегмента такая, что

$$|f(\xi_1)|\geqslant(|\sigma_1|+M)/\Delta x_1.$$

Отсюда следует, что $|f(\xi_1)|\Delta x_1\geqslant|\sigma_1|+M$, и поэтому

$$|\sigma(x_k, \xi_k)|=\left|\sum_{k=1}^{n} f(\xi_k)\,\Delta x_k\right|=|f(\xi_1)\,\Delta x_1+\sigma_1(x_k, \xi_k)|\geqslant$$
$$\geqslant|f(\xi_1)|\Delta x_1-|\sigma_1(x_k, \xi_k)|\geqslant M.$$

Теперь уже совсем просто убедиться в неинтегрируемости функции $f(x)$ на сегменте $[a, b]$. В самом деле, предположим, что $f(x)$ интегрируема на $[a, b]$, т. е. для нее существует предел I интегральных сумм при стремлении диаметра разбиения d к нулю. В силу доказанного выше для разбиения с как угодно малым диаметром d и для положительного числа $M=|I|+1$ найдется интегральная сумма σ, удовлетворяющая условию $|\sigma|>|I|+1$, из которого в силу неравенства $|\sigma-I|\geqslant|\sigma|-|I|$ вытекает, что $|\sigma-I|>1$ (для разбиения с как угодно малым диаметром d). Это противоречит определению 4 предела интегральных сумм.

Итак, мы доказали, что интегрируемыми на сегменте $[a, b]$ могут являться только ограниченные на этом сегменте функции. Естественно возникает вопрос: всякая ли ограниченная на сегменте $[a, b]$ функция является интегрируемой на этом сегменте?

Докажем, что *функция Дирихле* $D(x)$, равная нулю в иррациональных и единице в рациональных точках сегмента $[a, b]$, *представляет собой пример ограниченной и неинтегрируемой на этом сегменте функции.*

Для разбиения $[a, b]$ с как угодно малым диаметром d, взяв в качестве всех промежуточных точек иррациональные точки, мы получим интегральную сумму σ_1, равную нулю, а взяв в качестве всех промежуточных точек рациональные точки, мы получим интегральную сумму σ_2, равную $b-a$. Если бы существовал предел I интегральных сумм при стремлении d к нулю, то для положительного числа $\varepsilon=\frac{b-a}{2}$ и для достаточно малого d мы получили бы $|\sigma_1-I|<\frac{b-a}{2}$, $|\sigma_2-I|<\frac{b-a}{2}$, откуда $|\sigma_2-\sigma_1|=|(\sigma_2-I)+(I-\sigma_1)|\leqslant|\sigma_2-I|+|\sigma_1-I|<b-a$, а это противоречит тому, что $\sigma_2-\sigma_1=b-a$.

§ 2. ВЕРХНИЕ И НИЖНИЕ СУММЫ И ИХ СВОЙСТВА

1. Определение верхней и нижней сумм. Утверждение, доказанное в § 1, дает нам основание рассматривать всюду в дальнейшем только *ограниченные* на данном сегменте функции (ибо неограниченные функции заведомо не являются интегрируемыми по Риману).

Пусть $f(x)$ — ограниченная на сегменте $[a, b]$ функция и $\{x_k\}$ — произвольное разбиение этого сегмента. Так как $f(x)$ ограничена на сегменте $[a, b]$, то она ограничена и на любом частичном сегменте $[x_{k-1}, x_k]$, а поэтому у функции $f(x)$ существуют точная нижняя грань m_k и точная верхняя грань M_k на частичном сегменте $[x_{k-1}, x_k]$.

Итак, пусть $m_k = \inf\limits_{x_{k-1}\leqslant x\leqslant x_k} f(x)$, $M_k = \sup\limits_{x_{k-1}\leqslant x\leqslant x_k} f(x)$.

Введем фундаментальные понятия верхней и нижней сумм.

Определение 1. *Суммы*

$$S = M_1\Delta x_1 + M_2\Delta x_2 + \ldots + M_n\Delta x_n = \sum_{k=1}^{n} M_k\Delta x_k$$

и

$$s = m_1\Delta x_1 + m_2\Delta x_2 + \ldots + m_n\Delta x_n = \sum_{k=1}^{n} m_k\Delta x_k$$

будем называть соответственно верхней и нижней суммами функции $f(x)$ для данного разбиения $\{x_k\}$ сегмента $[a, b]$.

Выясним геометрический смысл верхней и нижней сумм. Рассмотрим снова криволинейную трапецию, т. е. фигуру, ограниченную отрезком $[a, b]$ оси Ox, графиком неотрицательной непрерывной функции $y=f(x)\geqslant 0$ и прямыми $x=a$, $x=b$, перпендикулярными к оси Ox (рис. 9.2). Пусть дано любое разбиение $\{x_k\}$ сегмента $[a, b]$. Число M_k в случае непрерывной функции $y=f(x)$ является ее максимальным значением на частичном сегменте $[x_{k-1}, x_k]$. Поэтому верхняя сумма равна площади элементарной ступенчатой фигуры, содержащей криволинейную трапецию. Эта площадь заштрихована на рис. 9.2.

Аналогично нижняя сумма равна площади элементарной ступенчатой фигуры, которая содержится в криволинейной трапеции (рис. 9.3). Отметим, что число m_k в этом случае является минимальным значением функции $y=f(x)$ на частичном сегменте $[x_{k-1}, x_k]$.

Анализируя геометрический смысл интегральной суммы, можно ожидать, что интеграл от интегрируемой по сегменту $[a, b]$ функции $y=f(x)$ должен равняться числу, которое следует принять за площадь соответствующей криволинейной трапеции. Но к этому же числу будут стремиться верхние и нижние суммы при стремлении диаметра разбиений к нулю. Поэтому представляется вероятным, что для интегрируемости функции необходимо и достаточно, чтобы разность между верхними и нижними суммами стремилась к нулю. Строго это будет доказано ниже.

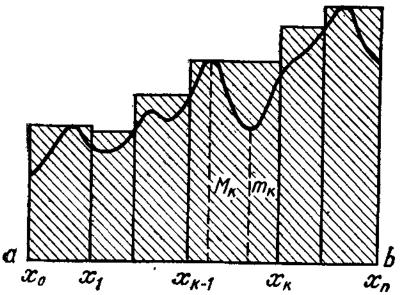


Рис. 9.2

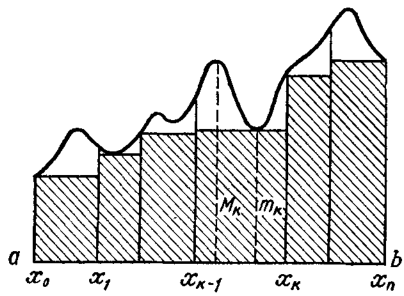


Рис. 9.3

2. Основные свойства верхних и нижних сумм. Докажем следующие леммы.

Лемма 1. *Пусть $\sigma(x_k, \xi_k)$ — интегральная сумма, отвечающая данному разбиению $\{x_k\}$. Тогда при любом выборе промежуточных точек ξ_k всегда справедливы неравенства*

$$s \leqslant \sigma \leqslant S,$$

где s и S — соответственно нижняя и верхняя суммы, отвечающие тому же разбиению.

Доказательство. По определению чисел m_k и M_k заключаем, что $m_k \leqslant f(\xi_k) \leqslant M$ для любого ξ_k из сегмента $[x_{k-1}, x_k]$. Умножая написанные неравенства на Δx_k и суммируя по всем k от 1 до n, получаем требуемое утверждение леммы.

Лемма 2. *Пусть $\{x_k\}$ — произвольное фиксированное разбиение сегмента $[a, b]$, ε — произвольное положительное число. Тогда можно выбрать промежуточные точки ξ_k так, чтобы интегральная сумма $\sigma(x_k, \xi_k)$ и верхняя сумма S удовлетворяли неравенству $0 \leqslant S - \sigma(x_k, \xi_k) < \varepsilon$. Промежуточные точки η_k можно выбрать и таким образом, чтобы для интегральной суммы $\sigma(x_k, \eta_k)$ и нижней суммы s выполнялись неравенства $0 \leqslant \sigma(x_k, \eta_k) - s < \varepsilon$.*

Доказательство. Пусть $\{x_k\}$ — фиксированное разбиение сегмента $[a, b]$ и $\varepsilon > 0$. Докажем сначала первое утверждение леммы. Поскольку $M_k = \sup\limits_{x_{k-1} \leqslant x \leqslant x_k} f(x)$, то для выбранного нами $\varepsilon > 0$ найдется точка ξ_k сегмента $[x_{k-1}, x_k]$ такая, что $0 \leqslant M_k - f(\xi_k) < \varepsilon/(b-a)$. Умножив эти неравенства на Δx_k и просуммировав по всем k от 1 до n, получим

$$0 \leqslant S - \sigma(x_k, \xi_k) < \varepsilon.$$

Аналогично в силу того, что $m_k=\inf\limits_{x_{k-1}\leqslant x\leqslant x_k} f(x)$, существует такая точка $\eta_k\in[x_{k-1}, x_k]$, что

$$0\leqslant f(\eta_k)-m_k<\varepsilon/(b-a).$$

Последние неравенства после умножения на Δx_k и суммирования приводят к оценкам

$$0\leqslant \sigma(x_k, \eta_k)-s<\varepsilon.$$

Лемма доказана.

Следствие. *Для любого фиксированного разбиения $\{x_k\}$ справедливы следующие соотношения*:

$$S=\sup_{\xi_k}\sigma(x_k, \xi_k),\quad s=\inf_{\eta_k}\sigma(x_k, \eta_k),$$

где точные верхняя и нижняя грани берутся по всевозможным промежуточным точкам.

Лемма 3. *При измельчении данного разбиения верхняя сумма может только уменьшиться, а нижняя сумма — только увеличиться.*

Доказательство. Пусть $\{x_k\}$ —данное разбиение, а разбиение $\{x_k'\}$ получается из него добавлением только одной новой точки $\bar{x}$. Легко видеть, что общий случай сводится к данному. Предположим, что $\bar{x}$ лежит внутри $[x_{k-1}, x_k]$. Тогда в выражении для S слагаемое $M_k\Delta x_k$ заменится на $M_k'(\bar{x}-x_{k-1})+M_k''(x_k-\bar{x})$, где $M_k'=\sup\limits_{x_{k-1}\leqslant x\leqslant\bar{x}} f(x)$, $M_k''=\sup\limits_{\bar{x}\leqslant x\leqslant x_k} f(x)$.

Точная верхняя грань функции на части сегмента не превосходит точной верхней грани функции на всем сегменте. Поэтому $M_k'\leqslant M_k$, $M_k''\leqslant M_k$ и

$$M_k'(\bar{x}-x_{k-1})+M_k''(x_k-\bar{x})\leqslant M_k[(\bar{x}-x_{k-1})+(x_k-\bar{x})]=M_k\Delta x_k.$$

Так как все другие слагаемые в выражении для верхней суммы сохранятся, то мы доказали, что при добавлении точки $\bar{x}$ верхняя сумма может только уменьшиться. Случай, когда к данному разбиению добавляется несколько новых точек, сводится, очевидно, к рассмотренному. Точно так же устанавливается, что при измельчении данного разбиения нижняя сумма может только увеличиться. Лемма доказана.

Лемма 4. *Для двух произвольных и, вообще говоря, различных разбиений сегмента $[a, b]$ нижняя сумма одного из этих разбиений не превосходит верхней суммы другого разбиения.*

Доказательство. Пусть $\{x_k'\}$ и $\{x_k''\}$ — два произвольных разбиения сегмента $[a, b]$, а S', s', S'', s'' — верхние и нижние суммы этих разбиений соответственно. Обозначим через $\{x_k\}$ объединение разбиений $\{x_k'\}$ и $\{x_k''\}$, а через S и s верхнюю и нижнюю суммы разбиения $\{x_k\}$. Заметим, что $\{x_k\}$ является измельчением как разбиения $\{x_k'\}$, так и разбиения $\{x_k''\}$. Согласно утверждению леммы 3 справедливы неравенства

$$\underline{S' \geqslant S}, \; S'' \geqslant S, \; s' \leqslant s, \; \underline{s'' \leqslant s}.$$

Кроме того, в силу леммы 1 получим, что $\underline{s \leqslant S}$. Пользуясь свойством транзитивности для числовых неравенств и используя три подчеркнутых выше неравенства, заключаем, что $s'' \leqslant S'$. Аналогично устанавливается, что $s' \leqslant S''$. Лемма доказана.

Следствие. *Множество верхних сумм данной функции $f(x)$, отвечающих всевозможным разбиениям сегмента $[a, b]$, ограничено снизу. Множество нижних сумм ограничено сверху.*

Действительно, любая верхняя сумма не меньше некоторой фиксированной нижней суммы, следовательно, множество верхних сумм ограничено снизу. Аналогично проводятся рассуждения для нижних сумм. В силу основной теоремы 2.1 из гл. 2 существуют точная нижняя грань множества $\{S\}$ и точная верхняя грань множества $\{s\}$.

Определение 2. ***Верхним интегралом Дарбу*** *от функции $f(x)$ называется число I^*, равное точной нижней грани множества верхних сумм $\{S\}$ данной функции $f(x)$ для всевозможных разбиений сегмента $[a, b]$.* ***Нижним интегралом Дарбу*** *от функции $f(x)$ называется число I_*, равное точной верхней грани множества нижних сумм $\{s\}$ данной функции $f(x)$ для всевозможных разбиений сегмента $[a, b]$.*

Лемма 5. *Нижний интеграл Дарбу всегда не превосходит верхнего интеграла Дарбу, т. е. $I_* \leqslant I^*$.*

Доказательство. Допустим противное, т. е. предположим, что $I_* > I^*$. Тогда $I_* - I^* = \varepsilon > 0$.

Для указанного ε, согласно определению числа I^*, найдется такое разбиение $\{x_k'\}$ сегмента $[a, b]$, что для соответствующей верхней суммы S' будет выполнено неравенство $S' < I^* + \varepsilon/2$. Точно так же можно указать такое разбиение $\{x_k''\}$ сегмента $[a, b]$, что для соответствующей нижней суммы s'' будет выполнено неравенство $s'' > I_* - \varepsilon/2$. Вычтем второе неравенство из первого. Получим $S' - s'' < I^* - I_* + \varepsilon$. Но $I^* - I_* = -\varepsilon$, поэтому $S' - s'' < 0$, т. е. $s'' > S'$. Получившееся неравенство противоречит утверждению леммы 4. Таким образом, доказываемое утверждение справедливо, т. е. $I_* \leqslant I^*$.

Пусть $M = \sup\limits_{x \in [a,b]} f(x)$, $m = \inf\limits_{x \in [a,b]} f(x)$, а $\{x_k\}$ — произвольное разбиение сегмента $[a, b]$, d — диаметр этого разбиения. Обозначим через $\{x_k'\}$ разбиение, полученное из разбиения $\{x_k\}$ путем добавления к нему l произвольных новых точек. Пусть S и s — верхняя и нижняя суммы разбиения $\{x_k\}$, а S' и s' — верхняя и нижняя суммы разбиения $\{x_k'\}$. Справедливо следующее утверждение.

Лемма 6. *Для разностей $S-S'$ и $s'-s$ выполняются следующие неравенства: $S-S'\leqslant(M-m)ld$, $s'-s\leqslant(M-m)ld$.*

Доказательство. Не ограничивая общности, мы можем провести рассуждения лишь для случая, когда к точкам разбиения $\{x_k\}$ добавляется только одна новая точка $\bar{x}$, и доказать, что в этом случае справедливы неравенства $S-S'\leqslant(M-m)d$, $s'-s\leqslant(M-m)d$.

Пусть вновь добавляемая точка $\bar{x}$ лежит внутри сегмента $[x_{k-1}, x_k]$. Тогда верхняя сумма S будет отличаться от верхней суммы S' только тем, что одно слагаемое $M_k\Delta x_k$ у суммы S заменится двумя слагаемыми $M_k'(\bar{x}-x_{k-1})+M_k''(x_k-\bar{x})$ у суммы S' (здесь через M_k, M_k' и M_k'' обозначены точные верхние грани $f(x)$ на сегментах $[x_{k-1}, x_k]$, $[x_{k-1}, \bar{x}]$ и $[\bar{x}, x_k]$ соответственно).

Все остальные слагаемые у верхних сумм S и S' будут общими. Отсюда следует, что

$$S-S'=M_k\Delta x_k-[M_k'(\bar{x}-x_{k-1})+M_k''(x_k-\bar{x})].$$

Из последнего соотношения, учитывая, что в силу свойств точных граней* $M_k\leqslant M$, $m\leqslant M_k'$, $m\leqslant M_k''$, получим, что

$$S-S'\leqslant M\Delta x_k-m[(\bar{x}-x_{k-1})+(x_k-\bar{x})]=(M-m)\Delta x_k\leqslant(M-m)d.$$

Доказательство оценки для нижних сумм аналогично.

Лемма доказана.

Определение 3. *Число A называется пределом верхних сумм S при стремлении к нулю диаметра разбиений d, если для любого положительного числа ε можно указать положительное число δ такое, что при условии $d<\delta$ выполняется неравенство $|S-A|<\varepsilon$.*

Для обозначения указанного предела естественно употреблять символ $A=\lim\limits_{d\to 0} S$.

Аналогично определяется предел B нижних сумм s при стремлении d к нулю.

Основная лемма Дарбу. *Верхний интеграл Дарбу I^* является пределом верхних сумм S при стремлении диаметра d разбиений к нулю, т. е. $I^*=\lim\limits_{d\to 0} S$. Аналогично $I_*=\lim\limits_{d\to 0} s$.*

Доказательство. Проведем доказательство первого утверждения леммы. Заметим, что если функция $f(x)=c=\text{const}$, то $S=c(b-a)=I^*$ для любого разбиения. Поэтому $\lim\limits_{d\to 0} S=I^*$. Если функция $f(x)$ непостоянна, то $M=\sup\limits_{x\in[a,b]} f(x)>m=\inf\limits_{x\in[a,b]} f(x)$. Фиксируем произвольное положительное число ε. По определению числа I^* существует такое разбиение $\{x_k^*\}$, что верхняя сумма S^* этого

* Если m_k' и m_k'' — точные нижние грани $f(x)$ на сегментах $[x_{k-1}, x]$ и $[x, x_k]$ соответственно, то, поскольку $m_k'\leqslant M_k'$, $m_k''\leqslant M_k''$ и $m<m_k'$, $m<m_k'$, мы получили, что $m\leqslant M_k'$, $m\leqslant M_k''$.

разбиения будет удовлетворять условию $S^*-I^*<\varepsilon/2$. Обозначим через l число точек разбиения $\{x_k^*\}$, не совпадающих с концами сегмента $[a, b]$.

Пусть $\{x_k\}$ — произвольное разбиение сегмента $[a, b]$, диаметр которого удовлетворяет неравенству $d<\delta=\varepsilon/[2l(M-m)]$, и пусть S — верхняя сумма этого разбиения. Произведем измельчение разбиения $\{x_k\}$, добавив к нему отмеченные выше l точек разбиения $\{x_k^*\}$. Полученное при этом разбиение обозначим символом $\{x_k'\}$. По лемме 6 верхняя сумма S' этого последнего разбиения удовлетворяет условию

$$0\leqslant S-S'\leqslant (M-m)\,ld<\varepsilon/2.$$

Но разбиение $\{x_k'\}$ можно рассматривать как измельчение разбиения $\{x_k^*\}$, к которому добавляются точки разбиения $\{x_k\}$, не совпадающие с концами сегмента $[a, b]$. Поэтому в силу определения I^* и леммы 3

$$I^*\leqslant S'\leqslant S^*, \text{ т. е. } 0\leqslant S'-I^*<S^*-I^*.$$

Выше было показано, что $S^*-I^*<\varepsilon/2$, поэтому $0\leqslant S'-I^*<\varepsilon/2$. Объединяя эти неравенства с установленными выше неравенствами $0\leqslant S-S'<\varepsilon/2$, получаем, что $0\leqslant S-I^*<\varepsilon$, если только d меньше указанного выше δ. Следовательно, $I^*=\lim\limits_{d\to 0} S$. Для нижних сумм доказательство аналогично. Основная лемма Дарбу доказана.

§ 3. ТЕОРЕМЫ О НЕОБХОДИМЫХ И ДОСТАТОЧНЫХ УСЛОВИЯХ ИНТЕГРИРУЕМОСТИ ФУНКЦИЙ. КЛАССЫ ИНТЕГРИРУЕМЫХ ФУНКЦИЙ

Доказанные выше утверждения о свойствах верхних и нижних сумм позволяют нам установить необходимые и достаточные условия интегрируемости по Риману произвольной ограниченной функции.

1. Необходимые и достаточные условия интегрируемости.

Вспомогательная теорема. *Для того чтобы ограниченная на сегменте $[a, b]$ функция $f(x)$ была интегрируема на этом сегменте, необходимо и достаточно, чтобы выполнялось равенство $I_*=I^*$.*

Доказательство. *Необходимость.* Пусть функция $f(x)$ интегрируема по Риману на сегменте $[a, b]$. Тогда существует предел I ее интегральных сумм σ при стремлении диаметра d разбиений к нулю.

По определению предела интегральных сумм для любого $\varepsilon>0$ существует такое $\delta>0$, что для любого выбора промежуточных точек ξ_k разбиения $\{x_k\}$ с диаметром $d<\delta$ выполняется неравенство

$$|I-\sigma(x_k, \xi_k)|<\varepsilon/4.$$

Согласно лемме 2 для данного разбиения $\{x_k\}$ можно так выбрать промежуточные точки ξ_k' и ξ_k'' в каждом частичном сегменте $[x_{k-1}, x_k]$, что будут справедливы неравенства

$$S-\sigma(x_k, \xi_k')<\varepsilon/4, \quad \sigma(x_k, \xi_k'')-s<\varepsilon/4.$$

Подчеркнем, что, кроме того, для данного разбиения $\{x_k\}$ одновременно выполнены неравенства

$$|I-\sigma(x_k, \xi_k')|<\varepsilon/4, \quad |I-\sigma(x_k, \xi_k'')|<\varepsilon/4.$$

Заметим теперь, что

$$S-s=[S-\sigma(x_k, \xi_k')]+[\sigma(x_k, \xi_k')-I]+[I-\sigma(x_k, \xi_k'')]+ \\ +[\sigma(x_k, \xi_k'')-s].$$

Отсюда, учитывая, что модуль суммы четырех величин не превосходит суммы их модулей, получаем, что $S-s<\varepsilon$. Итак, *для любого* $\varepsilon>0$ *существует такое* $\delta>0$, *что для любого разбиения с диаметром* $d<\delta$ *справедливо неравенство* $S-s<\varepsilon$. Поскольку для любого разбиения выполнены неравенства

$$s\leqslant I_*\leqslant I^*\leqslant S,$$

то из неравенства $S-s<\varepsilon$ вытекает, что $0\leqslant I^*-I_*<\varepsilon$, а отсюда в силу произвольности $\varepsilon>0$ вытекает, что $I^*=I_*$.

Достаточность. Пусть $I^*=I_*=A$. Согласно основной лемме Дарбу $I^*=\lim\limits_{d\to 0} S$, $I_*=\lim\limits_{d\to 0} s$, т. е. верхний интеграл является пределом верхних сумм, а нижний интеграл — пределом нижних сумм при стремлении диаметра d разбиений к нулю. Поэтому для любого $\varepsilon>0$ можно указать такое число $\delta>0$, что для любого разбиения с диаметром $d<\delta$ одновременно выполняются неравенства $I_*-s=A-s<\varepsilon$, $S-I^*=S-A<\varepsilon$. При любом указанном разбиении любая интегральная сумма $\sigma(x_k, \xi_k)$ удовлетворяет неравенствам $s\leqslant\sigma(x_k, \xi_k)\leqslant S$, а значит, и неравенствам

$$A-\varepsilon<s\leqslant\sigma(x_k, \xi_k)\leqslant S<A+\varepsilon.$$

Отсюда $|A-\sigma(x_k, \xi_k)|<\varepsilon$ (для любого разбиения с диаметром d, меньшим δ).

Таким образом, $A=\lim\limits_{d\to 0}\sigma(x_k, \xi_k)$, т. е. функция $f(x)$ интегрируема.

Докажем следующую теорему, имеющую важное значение в теории интеграла Римана.

Основная теорема. *Для того чтобы ограниченная на сегменте* $[a, b]$ *функция* $f(x)$ *была интегрируемой на этом сегменте, необходимо и достаточно, чтобы для любого* $\varepsilon>0$ *нашлось такое разбиение* $\{x_k\}$ *сегмента* $[a, b]$, *для которого* $S-s<\varepsilon$.

Доказательство. *Необходимость*. Пусть функция $f(x)$ интегрируема на сегменте $[a, b]$. При доказательстве необходимости вспомогательной теоремы установлено, что для любого $\varepsilon>0$ существует такое $\delta>0$, что для любого разбиения сегмента $[a, b]$ с диаметром d, меньшим δ, справедливо неравенство $S-s<\varepsilon$. Необходимость доказана.

Достаточность. Дано, что для любого $\varepsilon>0$ существует такое разбиние $\{x_k\}$ сегмента $[a, b]$, что для соответствующих верхней и нижней сумм выполнено соотношение $S-s<\varepsilon$. Тогда поскольку

$$s\leqslant I_*\leqslant I^*\leqslant S,$$

то $I^*-I_*<\varepsilon$. Из этого неравенства и из произвольности ε заключаем, что $I_*=I^*$, а по вспомогательной теореме получаем, что функция $f(x)$ интегрируема. Теорема доказана.

2. Классы интегрируемых функций. Выше в § 1 настоящей главы мы видели, что функция $f(x)$, постоянная на сегменте $[a, b]$, интегрируема по Риману на этом сегменте, а также что интегрируемые на данном сегменте функции обязаны быть ограниченными на этом сегменте. Естественно, возникает вопрос об описании классов функций, интегрируемых по Риману на сегменте $[a, b]$. Среди этих классов важную роль играет класс непрерывных на сегменте $[a, b]$ функций. Докажем следующую фундаментальную теорему.

Теорема 9.1. *Непрерывные на сегменте $[a, b]$ функции интегрируемы на этом сегменте по Риману.*

Доказательство. Пусть $f(x)$ непрерывна на сегменте $[a, b]$. Выберем произвольное число $\varepsilon>0$. Поскольку функция $f(x)$, будучи непрерывной на сегменте, является равномерно непрерывной на нем, то для любого данного $\varepsilon>0$ существует такое число $\delta>0$, что если ξ' и ξ'' — любые две точки сегмента $[a, b]$, для которых $|\xi'-\xi''|<\delta$, то $|f(\xi')-f(\xi'')|<\varepsilon/(b-a)$. Отсюда следует, что разность между точными верхней и нижней гранями $f(x)$ на любом сегменте, имеющем длину, меньшую δ, будет меньше числа $\varepsilon/(b-a)$. Выберем теперь разбиение $\{x_k\}$ сегмента $[a, b]$ с диаметром d, меньшим указанного числа $\delta: d<\delta$. Пусть $M_k=\sup\limits_{x\in[x_{k-1}, x_k]} f(x)$, $m_k=\inf\limits_{x\in[x_{k-1}, x_k]} f(x)$.

По определению верхней и нижней сумм

$$S-s=\sum_{k=1}^{n}(M_k-m_k)\,\Delta x_k.$$

Используя в этом соотношении установленное для выбранного нами разбиения неравенство $M_k-m_k<\varepsilon/(b-a)$, утверждающее, что разность между точными гранями на любом частичном сегменте меньше $\varepsilon/(b-a)$, мы получим, что для выбранного разбиения

$$S-s<\frac{\varepsilon}{b-a}\sum_{k=1}^{n}\Delta x_k=\varepsilon.$$

По основной теореме заключаем, что функция $f(x)$ интегрируема на $[a, b]$. Теорема доказана.

Следующая теорема дает *достаточное* условие интегрируемости некоторого класса *разрывных* функций.

Условимся говорить, что точка x *покрыта интервалом*, если она содержится в указанном интервале.

Докажем следующую теорему.

Теорема 9.2. *Пусть функция $f(x)$ определена и ограничена на сегменте $[a, b]$. Если для любого числа $\varepsilon>0$ можно указать конечное число интервалов, покрывающих все точки разрыва этой функции и имеющих общую сумму длин, меньшую ε, то функция $f(x)$ интегрируема по Риману на сегменте $[a, b]$.*

Доказательство. Пусть M и m — точные верхняя и нижняя грани функции $f(x)$ на сегменте $[a, b]$. Заметим, что если $M=m$, т. е. функция $f(x)$ постоянна, то, как мы уже доказали в § 1, она интегрируема. Поэтому будем считать, что $M>m$. Пусть $\varepsilon>0$ — произвольное число. Покроем точки разрыва функции $f(x)$ конечным числом интервалов, сумма длин которых не превосходит числа $\varepsilon_1=\varepsilon/2\cdot(M-m)$. Точки сегмента $[a, b]$, не принадлежащие указанным интервалам, очевидно, образуют множество, состоящее из конечного числа непересекающихся сегментов. Назовем эти сегменты дополнительными. На каждом из таких сегментов функция непрерывна, а следовательно, и равномерно непрерывна. Значит, существуют такие числа $\delta_i>0$, что если $|\xi'-\xi''|<\delta_i$, то $|f(\xi')-f(\xi'')|<\varepsilon/(2(b-a))$ для всех ξ' и ξ'', принадлежащих i-му дополнительному сегменту.

Пусть $\delta=\min\limits_i \delta_i$. Тогда если взять разбиения дополнительных сегментов на частичные сегменты так, чтобы диаметр каждого из частичных сегментов не превосходил δ, то разность между точными верхней гранью M_k и нижней гранью m_k функции $f(x)$ на k-м частичном сегменте будет не больше $\varepsilon/(2(b-a))$. Объединяя все разбиения дополнительных сегментов и указанные выше интервалы с присоединенными к ним концами, мы получим разбиение $\{x_k\}$ всего сегмента $[a, b]$. Для так построенного общего разбиения $[a, b]$

$$S-s=\sum_{k=1}^{n}(M_k-m_k)\,\Delta x_k=\Sigma'\,(M_k-m_k)\,\Delta x_k+\Sigma''\,(M_k-m_k)\,\Delta x_k,$$

где в сумму с одним штрихом отнесены слагаемые, отвечающие частичным сегментам, образованным из интервалов, покрывающих точки разрыва, а в сумму с двумя штрихами — все остальные. Рассмотрим первое слагаемое правой части записанного выше равенства. Поскольку $M_k - m_k \leqslant M - m$ для любого k, то $\Sigma'(M_k - m_k)\Delta x_k \leqslant (M-m)\Sigma'\Delta x_k < \varepsilon/2$.

Далее, в силу сказанного выше, из свойства равномерной непрерывности функции $f(x)$ на дополнительных сегментах получаем, что

$$\Sigma''(M_k - m_k)\Delta x < \frac{\varepsilon}{2(b-a)}\Sigma''\Delta x_k < \frac{\varepsilon}{2(b-a)}(b-a) = \frac{\varepsilon}{2}.$$

Таким образом, нами указано разбиение $\{x_k\}$, для которого $S - s < \varepsilon$. По основной теореме получаем, что функция $f(x)$ интегрируема. Теорема доказана.

Следствие. *Ограниченная на сегменте* $[a, b]$ *функция* $f(x)$, *имеющая лишь конечное число точек разрыва, интегрируема на этом сегменте. В частности, кусочно непрерывная на данном сегменте функция интегрируема на этом сегменте.*

Действительно, в условии предыдущей теоремы достаточно выбрать интервалы, покрывающие точки разрыва, одинаковой длины, меньшей чем $\varepsilon/(2p)$, где p — число точек разрыва функции $f(x)$.

Замечание. Пусть функция $f(x)$ интегрируема на сегменте $[a, b]$, а функция $g(x)$ совпадает с функцией $f(x)$ во всех точках сегмента $[a, b]$, кроме, быть может, конечного числа точек. Тогда функция $g(x)$ интегрируема на сегменте $[a, b]$ и

$$\int_a^b f(x)\,dx = \int_a^b g(x)\,dx.$$

Указание. Использовать схему доказательства следствия.

Теорема 9.3. *Монотонная на сегменте* $[a, b]$ *функция* $f(x)$ *интегрируема по Риману на этом сегменте.*

Доказательство. Случай, когда функция $f(x)$ постоянна на сегменте $[a, b]$, можно исключить. Рассмотрим, например, неубывающую на сегменте $[a, b]$ функцию $f(x)$. Пусть $\varepsilon > 0$ — произвольное число. Выберем разбиение $\{x_k\}$ сегмента $[a, b]$ с диаметром $d < \varepsilon/(f(b) - f(a))$. Заметим, что поскольку $f(x)$ не постоянна, то $f(b) > f(a)$. Оценим разность $S - s = \sum_{k=1}^{n}(M_k - m_k)\Delta x_k$, M_k и m_k — верхняя и нижняя грани $f(x)$ на $[x_{k-1}, x_k]$. Получим $S - s < \varepsilon\sum_{k=1}^{n}(M_k - m_k)/(f(b) - f(a))$. Но для неубывающей функции $\sum_{k=1}^{n}(M_k - m_k) = f(b) - f(a)$ *. Поэтому $S - s < \varepsilon$ и функция

* Ибо $M_k = f(x_k) = m_{k+1}$ при $k = 1, 2, \ldots, n-1$, $M_n = f(b)$, $m_1 = f(a)$.

$f(x)$ интегрируема. Для невозрастающей функции рассуждения аналогичны. Теорема доказана.

Докажем сейчас теорему об интегрируемости суперпозиции двух функций.

Теорема 9.4. *Пусть функция $f(x)$ интегрируема по Риману на сегменте $[a, b]$, M и m — ее точные верхняя и нижняя грани на $[a, b]$. Пусть, далее, функция $\varphi(x)$ определена на сегменте $[m, M]$ и удовлетворяет условию*: существует неотрицательное число C такое, что для любых x_1 и x_2 из сегмента $[m, M]$ выполняется неравенство*

$$|\varphi(x_1)-\varphi(x_2)|\leqslant C|x_1-x_2|;$$

тогда сложная функция $h(x)=\varphi(f(x))$ интегрируема по Риману на сегменте $[a, b]$.

Доказательство. Пусть ε — произвольное положительное число. В силу интегрируемости функции $f(x)$ на сегменте $[a, b]$ найдется такое разбиение $\{x_k\}$ этого сегмента, что $S-s<\varepsilon/C$, где S и s — верхняя и нижняя суммы функции $f(x)$ соответственно, C — постоянная из условий теоремы. Пусть M_k и m_k — точные грани функции $f(x)$ на частичном сегменте $[x_{k-1}, x_k]$ разбиения $\{x_k\}$, а M_k^* и m_k^* — соответствующие величины для функции $h(x)$. Тогда в силу условия на функцию $\varphi(x)$ для любых точек x и y, принадлежащих частичному сегменту $[x_{k-1}, x_k]$ разбиения $\{x_k\}$, справедливо неравенство

$$h(x)-h(y)\leqslant|h(x)-h(y)|=|\varphi[f(x)]-\varphi[f(y)]|\leqslant$$
$$\leqslant C|f(x)-f(y)|\leqslant C(M_k-m_k).$$

Поскольку неравенство $h(x)-h(y)\leqslant C(M_k-m_k)$ справедливо для любых точек x и y, принадлежащих сегменту $[x_{k-1}, x_k]$, то тем более $M_k^*-m_k^*\leqslant C(M_k-m_k)$, ибо найдутся две последовательности $\{x_n\}$ и $\{y_n\}$ точек сегмента $[x_{k-1}, x_k]$ такие, что $h(x_n)\to M_k^*$, $h(y_n)\to m_k^*$ при $n\to\infty$.

Пусть, далее, S^* и s^* — соответственно верхняя и нижняя суммы функции $h(x)$ для выбранного разбиения $\{x_k\}$ сегмента $[a, b]$. Тогда

$$S^*-s^*=\sum_{k=1}^{n}(M_k^*-m_k^*)\Delta x_k\leqslant C\sum_{k=1}^{n}(M_k-m_k)\Delta x_k<\varepsilon.$$

Поскольку ε — произвольное положительное число, то в силу основной теоремы функция $h(x)$ интегрируема на сегменте $[a, b]$. Теорема доказана.

Докажем теперь следующую несколько более общую теорему.

* Данное условие называется условием Липшица. Ясно, что функция, удовлетворяющая на сегменте условию Липшица, непрерывна на нем.

Теорема 9.4'. *Пусть $f(x)$ — интегрируемая по Риману на сегменте $[a, b]$ функция, M и m — ее точные верхняя и нижняя грани на $[a, b]$. Пусть, далее, функция $\varphi(x)$ непрерывна на сегменте $[m, M]$.*

Тогда сложная функция $h(x)=\varphi[f(x)]$ интегрируема по Риману на сегменте $[a, b]$.

Доказательство. Пусть $C=\max\limits_{m\leqslant t\leqslant M}|\varphi(t)|$ и ε — произвольное положительное число. Положим $\varepsilon_1=\varepsilon/(b-a+2C)$. Ввиду того, что φ равномерно непрерывна на $[m, M]$, существует такое $\delta>0$, что $|\varphi(t_1)-\varphi(t_2)|<\varepsilon_1$, если $|t_1-t_2|<\delta$ и $t_1, t_2\in[m, M]$. Выберем δ еще и таким, что $\delta<\varepsilon_1$. В силу интегрируемости функции $f(x)$ на $[a, b]$ существует такое разбиение $\{x_k\}$ сегмента $[a, b]$, для которого $S-s<\delta^2$.

Положим

$$M_k=\sup_{x\in[x_{k-1},x_k]} f(x),\quad m_k=\inf_{x\in[x_{k-1},x_k]} f(x),$$

$$M_k^*=\sup_{x\in[x_{k-1},x_k]} h(x),\quad m_k^*=\inf_{x\in[x_{k-1},x_k]} h(x).$$

Разобьем целые числа $1, \ldots, n$ на два множества A и B: число $k\in A$, если $M_k-m_k<\delta$, число $p\in B$, если $M_p-m_p\geqslant\delta$. Если индекс $k\in A$, то $M_k-m_k<\delta$, следовательно, в силу равномерной непрерывности функции $\varphi(t)$ на сегменте $[m, M]$ получим, что $M_k^*-m_k^*\leqslant\varepsilon_1$. Действительно, если рассматривается индекс $k\in A$, то мы получим, что $M_k-m_k==\sup\limits_{[x_{k-1},x_k]} f(x)-\inf\limits_{[x_{k-1},x_k]} f(x)<\delta$, т. е. при $x, y\in[x_{k-1}, x_k]$ разность $f(x)-f(y)=t_1-t_2$ по абсолютной величине не превосходит $\delta: |t_1-t_2|<\delta$, где $t_1=f(x)$, $t_2=f(y)$. Следовательно, в силу равномерной непрерывности функции φ на всем сегменте $[m, M]$ мы получим, что

$$|\varphi(f(x))-\varphi(f(y))|=|\varphi(t_1)-\varphi(t_2)|<\varepsilon_1.$$

Так как последнее неравенство справедливо для любых x и y, принадлежащих сегменту $[x_{k-1}, x_k]$, то и

$$\sup_{[x_{k-1},x_k]}\varphi(f(x))-\inf_{[x_{k-1},x_k]}\varphi(f(x))\leqslant\varepsilon_1.$$

Далее, если $k\in B$, то, очевидно, что $M_k^*-m_k^*\leqslant 2C$. Запишем теперь разность (S^* и s^* — соответственно верхняя и нижняя суммы функции $h(x)$ для рассматриваемого разбиения $\{x_k\}$)

$$S^*-s^*=\sum_{k=1}^{n}(M_k^*-m_k^*)\,\Delta x_k=\sum_{k\in A}(M_k^*-m_k^*)\,\Delta x_k+$$

$$+\sum_{k\in B}(M_k^*-m_k^*)\,\Delta x_k\leqslant\varepsilon_1(b-a)+2C\sum_{k\in B}\Delta x_k.$$

Осталось произвести оценку для величины $\sum\limits_{k\in B}\Delta x_k$. Имеем

$$\delta\sum_{k\in B}\Delta x_k\leqslant\sum_{k\in B}(M_k-m_k)\,\Delta x_k\leqslant\sum_{k=1}^{n}(M_k-m_k)\,\Delta x_k$$

(здесь мы пользуемся тем, что все слагаемые $(M_k-m_k)\Delta x_k$ являются неотрицательными). Учитывая, что при выбранном разбиении $\sum\limits_{k=1}^{n}(M_k-m_k)\,\Delta x_k=S-s<\delta^2$, получаем, что

$$\delta\sum_{k\in B}\Delta x_k\leqslant\sum_{k=1}^{n}(M_k-m_k)\,\Delta x_k<\delta^2,\ \text{т. е.}\ \sum_{k\in B}\Delta x_k<\delta.$$

Окончательно получим, что

$$S^*-s^*\leqslant\varepsilon_1(b-a)+2C\sum_{k\in B}\Delta x_k\leqslant$$

$$\leqslant\varepsilon_1(b-a)+2C\delta<\varepsilon_1(b-a+2C)=\varepsilon.$$

(Выше мы воспользовались тем, что $\delta<\varepsilon_1$.) Таким образом, функция $h(x)$ интегрируема, и теорема доказана.

Следствие. *Если функция $f(x)$ интегрируема на сегменте $[a,b]$, то для любого положительного числа α функция $|f(x)|^\alpha$ интегрируема на этом же сегменте.*

Действительно, достаточно рассмотреть непрерывную функцию $\varphi(t)=|t|^\alpha$ и применить предыдущую теорему.

Приведем несколько примеров.

Примеры. 1) *Пример интегрируемой функции, имеющей бесконечное число точек разрыва.* Пусть на сегменте $[0, 2/\pi]$ задана функция $f(x)=\operatorname{sgn}\sin\frac{1}{x}$ * (рис. 9.4). Указанная функция имеет разрывы 1-го рода во всех точках $x_k=1/\pi k$, $k=1, 2, 3, \ldots$, а также разрыв 2-го рода в точке 0. Фиксируем произвольное число $\varepsilon>0$. Покроем точку $x=0$ интервалом $(-\varepsilon/4, \varepsilon/4)$. Вне этого интервала находится лишь конечное число p точек разрыва функции. Число p зависит от заданного ε. Покроем каждую из этих точек интервалом длины меньше $\varepsilon/2p$. Тогда все точки разрыва функции $f(x)=\operatorname{sgn}\sin\frac{1}{x}$ будут покрыты конечным числом интервалов, общая сумма длин которых не превосходит $\frac{\varepsilon}{2}+p\,\frac{\varepsilon}{2p}=\varepsilon$. По теореме 9.2 функция $f(x)$ интегрируема на сегменте $[0,2/\pi]$.

---

* В точке $x=0$ эту функцию доопределяем произвольно, например, полагаем $f(0)=0$.

2) *Из интегрируемости функции $|f(x)|$ не следует, вообще говоря, интегрируемость $f(x)$*. Действительно, рассмотрим функцию $D_1(x)$, равную единице для x рациональных и минус единице для x иррациональных. Тогда $|D_1(x)| \equiv 1$ интегрируема. Точно так же, как и для функции Дирихле $D(x)$, показывается, что функция $D_1(x)$ неинтегрируема (см. пример из § 1).

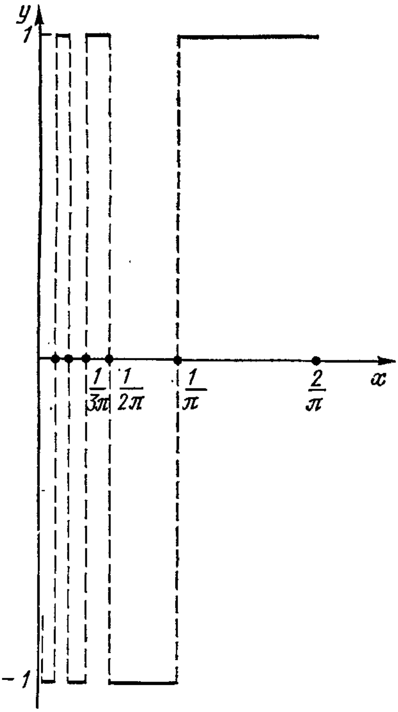


Рис. 9.4

§ 4. СВОЙСТВА ОПРЕДЕЛЕННОГО ИНТЕГРАЛА. ОЦЕНКИ ИНТЕГРАЛОВ. ТЕОРЕМЫ О СРЕДНЕМ ЗНАЧЕНИИ

1. Свойства интеграла. Выясним основные свойства интеграла Римана.

а) *Пусть функции $f(x)$ и $g(x)$ интегрируемы на сегменте $[a, b]$. Тогда функции $f(x) \pm \pm g(x)$ также интегрируемы на этом сегменте, причем*

$$\int_a^b (f(x) \pm g(x))\, dx = \int_a^b f(x)\, dx \pm \int_a^b g(x)\, dx.$$

Действительно, при любом разбиении сегмента $[a, b]$ и любом выборе промежуточных точек ξ_k справедливы следующие равенства:

$$\sum_{k=1}^{n} [f(\xi_k) \pm g(\xi_k)]\, \Delta x_k = \sum_{k=1}^{n} f(\xi_k)\, \Delta x_k \pm \sum_{k=1}^{n} g(\xi_k)\, \Delta x_k.$$

Поэтому, если существует предел правой части при стремлении диаметра разбиений к нулю, то существует предел и левой части. Из линейных свойств этого предела, которые устанавливаются точно так же, как и для предела последовательностей, вытекает доказываемое свойство.

б) *Если функция $f(x)$ интегрируема на сегменте $[a, b]$, то функция $cf(x)$, где $c=\text{const}$, также интегрируема на этом сегменте, причем*

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx.$$

В самом деле, для любого разбиения сегмента $[a, b]$ и любого выбора промежуточных точек ξ_k выполнено соотношение

$$\sum_{k=1}^{n} cf(\xi_k)\,\Delta x_k = c\sum_{k=1}^{n} f(\xi_k)\,\Delta x_k,$$

откуда, так же как и выше, получаем доказательство утверждения б).

Следствие. *Линейная комбинация $\sum_{i=1}^{n} c_i f_i(x)$ интегрируемых функций $f_i(x)$ является интегрируемой функцией.*

в) *Пусть функции $f(x)$ и $g(x)$ интегрируемы на сегменте $[a, b]$. Тогда $f(x)g(x)$ также интегрируема на этом сегменте.*

Запишем очевидное тождество:

$$4f(x)g(x) = [f(x)+g(x)]^2 - [f(x)-g(x)]^2.$$

Заметим, далее, что в силу теоремы 9.4 из интегрируемости какой-либо функции $u(x)$ следует интегрируемость ее квадрата*. Поскольку функции $f(x)+g(x)$ и $f(x)-g(x)$ по свойству а) интегрируемы, то по сказанному интегрируемы и их квадраты, а следовательно, в силу указанного, функция $f(x)g(x)$ интегрируема.

г) *Пусть функция $f(x)$ интегрируема на сегменте $[a, b]$. Тогда эта функция интегрируема на любом сегменте $[c, d]$, содержащемся в сегменте $[a, b]$.*

Выберем произвольное число $\varepsilon>0$ и такое разбиение $\{x_k\}$ сегмента $[a, b]$, что $S-s<\varepsilon$. Добавим к точкам разбиения $\{x_k\}$ точки c и d. Для верхних сумм S' и нижних s' вновь полученного разбиения $\{x_k'\}$ в силу леммы 3 § 2 тоже будет справедлива оценка $S'-s'<\varepsilon$. Рассмотрим разбиение $\{\bar{x}_k\}$ сегмента $[c, d]$, образованное точками разбиения $\{x_k'\}$ всего сегмента $[a, b]$. Для верхних и нижних сумм $\bar{S}$ и $\bar{s}$ разбиения $\{\bar{x}_k\}$ выполнено, очевидно, соотношение $\bar{S}-\bar{s}\leqslant S'-s'$, поскольку каждое неотрицательное слагаемое $(M_k-m_k)\Delta x_k$ в выражении $\bar{S}-\bar{s}$ будет слагаемым и в выражении $S'-s'$. Таким образом, $\bar{S}-\bar{s}<\varepsilon$, и функция $f(x)$ интегрируема на сегменте $[c, d]$. Свойство г) доказано.

---

* Так как $u^2(x)=\varphi[u(x)]$ при $\varphi(t)=t^2$ и функция $\varphi(t)$ удовлетворяет условию Липшица на любом сегменте $[m, M]$, ибо для любых t_1 и t_2 из сегмента $[m, M]$ справедливо неравенство $|\varphi(t_1)-\varphi(t_2)|=|t_1^2-t_2^2|=|t_1-t_2|\cdot|t_1+t_2|\leqslant C|t_1-t_2|$, где C — наибольшее из двух чисел $2|m|$ и $2|M|$.

Будем по определению всегда считать, что интеграл Римана от функции, взятый в пределах от точки* a *до точки* a, *равен нулю, т. е.* $\int\limits_a^a f(x)\,dx=0$. Это свойство следует рассматривать как соглашение. *Условимся также о том, что при* $a<b$ *по определению* $\int\limits_b^a f(x)\,dx=-\int\limits_a^b f(x)\,dx$ *для любой интегрируемой функции.* Эту формулу следует также рассматривать как соглашение.

д) *Если функция* $f(x)$ *интегрируема на сегментах* $[a, c]$ *и* $[c, b]$, *то функция* $f(x)$ *интегрируема и на сегменте* $[a, b]$, *причем*

$$\int\limits_a^b f(x)\,dx=\int\limits_a^c f(x)\,dx+\int\limits_c^b f(x)\,dx.$$

При $a=b$ это свойство справедливо в силу принятых выше соглашений.

Предположим сначала, что $a<c<b$. Выберем произвольное число $\varepsilon>0$. Пусть $\{x_k'\}$ и $\{x_k''\}$ — такие разбиения сегментов $[a, c]$ и $[c, d]$, что на каждом из этих сегментов $S-s<\varepsilon/2$. Пусть $\{x_k\}$ — разбиение сегмента $[a, b]$, состоящее из точек разбиений $\{x_k'\}$ и $\{x_k''\}$. Очевидно, что разность между верхней и нижней суммами разбиения $\{x_k\}$ не будет превосходить ε. Интегрируемость функции $f(x)$ на сегменте $[a, b]$ доказана. Пусть теперь $\{x_k\}$ — произвольное разбиение сегмента $[a, b]$, содержащее c. Тогда

$$\sum_{k=1}^{n} f(\xi_k)\,\Delta x_k=\sum{}' f(\xi_k)\,\Delta x_k+\sum{}'' f(\xi_k)\,\Delta x_k,$$

где Σ' берется по частичным сегментам, принадлежащим $[a, c]$, а Σ'' — по частичным сегментам, принадлежащим $[c, d]$. Поскольку это верно для любого разбиения, то, перейдя к пределу при стремлении диаметра разбиений к нулю, получим, что

$$\int\limits_a^b f(x)\,dx=\int\limits_a^c f(x)\,dx+\int\limits_c^d f(x)\,dx.$$

Если c не принадлежит $[a, b]$, то сегмент $[a, b]$ принадлежит либо $[c, b]$, либо $[a, c]$. Пусть, например, $c<a<b$. В силу свойства г) функция $f(x)$ интегрируема на $[a, b]$. Действительно, функция $f(x)$ интегрируема на $[c, b]$ по условию, а $[a, b]\subset[c, b]$.

Далее, поскольку $c<a<b$,

$$\int\limits_c^a f(x)\,dx+\int\limits_a^b f(x)\,dx=\int\limits_c^b f(x)\,dx.$$

* Функция в точке a определена и принимает конечное значение.

Но по принятому соглашению $\int_c^a f(x)\,dx = -\int_a^c f(x)\,dx$. Таким образом, свойство д) установлено нами и для случая, когда c лежит вне сегмента $[a, b]$. Заметим, что формулу этого свойства можно записать так:

$$\int_a^c f(x)\,dx + \int_c^b f(x)\,dx + \int_b^a f(x)\,dx = 0.$$

2. Оценки интегралов.

а) *Если функция $f(x)$ интегрируема на сегменте $[a, b]$ и $f(x) \geqslant 0$ для всех x из $[a, b]$, то интеграл от функции $f(x)$ по этому сегменту неотрицателен.*

Доказательство следует из того, что для любого разбиения $\{x_k\}$ и любого выбора ξ_k интегральная сумма

$$\sigma = \sum_{k=1}^{n} f(\xi_k)\,\Delta x_k \geqslant 0.$$

В этом случае предел интегральных сумм тоже будет неотрицателен. Действительно, допустим, что предел A этих сумм отрицателен. Пусть $\varepsilon = |A|$. Выберем разбиение $\{x_k\}$ такое, что $|\sigma - A| < |A|$. Но последнее неравенство может выполняться, только если $\sigma < 0$, что противоречит условию $\sigma \geqslant 0$. Значит, $\int_a^b f(x)\,dx \geqslant 0$.

б) Интегрирование неравенства. *Если функции $f(x)$ и $g(x)$ интегрируемы на сегменте $[a, b]$ и $f(x) \leqslant g(x)$ для всех x из $[a, b]$, то $\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx$.*

Действительно, функция $g(x) - f(x)$ интегрируема и неотрицательна на $[a, b]$, так что

$$\int_a^b (g(x) - f(x))\,dx \geqslant 0.$$

Но тогда в силу свойства а) из п. 1 $\int_a^b g(x)\,dx - \int_a^b f(x)\,dx \geqslant 0$, что и требовалось установить.

в) *Пусть функция $f(x)$ непрерывна и неотрицательна на сегменте $[a, b]$. Если существует хотя бы одна точка x_0 сегмента $[a, b]$, в которой $f(x_0) > 0$, то найдется положительное число α такое, что*

$$\int_a^b f(x)\,dx \geqslant \alpha > 0.$$

Действительно, пусть $f(x_0)=\beta>0$. Тогда в силу непрерывности функции $f(x)$ в точке x_0 найдется такая окрестность точки x_0, что для любого сегмента $[c, d]$, $c\neq d$, лежащего целиком в этой окрестности, будет выполнено неравенство $f(x)\geqslant\beta/2$. Но тогда в силу оценки из п. б)

$$\int_a^b f(x)\,dx \geqslant \int_c^d f(x)\,dx \geqslant \int_c^d \frac{\beta}{2}\,dx = \frac{\beta}{2}(d-c)=\alpha>0.$$

Следовательно, $\int_a^b f(x)\,dx \geqslant \alpha > 0.$

г) *Если функция $f(x)$ интегрируема по Риману на сегменте $[a, b]$, то функция $|f(x)|$ интегрируема на этом сегменте и*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx.$$

Рассмотрим функцию $\varphi(t)=|t|$. Согласно теореме 9.4 из интегрируемости $f(x)$ следует интегрируемость $\varphi(f(x))=|f(x)|$ (так как функция $\varphi(t)=|t|$ на любом сегменте удовлетворяет условию Липшица*). Выберем теперь число $\alpha=\pm 1$ так, чтобы $\alpha\int_a^b f(x)\,dx \geqslant 0$. Очевидно, что $\alpha f(x)\leqslant|\alpha f(x)|=|f(x)|$. Тогда в силу свойства б)

$$\left|\int_a^b f(x)\,dx\right| = \alpha\int_a^b f(x)\,dx = \int_a^b \alpha f(x)\,dx \leqslant \int_a^b |f(x)|\,dx,$$

что и требовалось.

д) **Первая формула среднего значения.** *Пусть каждая из функций $f(x)$ и $g(x)$ интегрируема на сегменте $[a, b]$ и функция $g(x)$, кроме того, неотрицательна (или неположительна) на этом сегменте.*

*Обозначим через M и m точные грани $f(x)$ на сегменте $[a, b]$**. Тогда найдется число μ, удовлетворяющее неравенствам $m\leqslant\mu\leqslant M$ и такое, что справедлива следующая формула:*

$$\int_a^b f(x)\,g(x)\,dx = \mu\int_a^b g(x)\,dx. \qquad (*)$$

При дополнительном предположении о непрерывности $f(x)$ на сег-

---

* Функция $\varphi(t)=|t|$, очевидно, удовлетворяет условию Липшица на любом сегменте, ибо $|\varphi(t_1)-\varphi(t_2)|=||t_1|-|t_2||\leqslant|t_1-t_2|$.

** Интегрируемая на $[a, b]$ функция ограничена на $[a, b]$, и потому у нее существуют на $[a, b]$ точные грани.

менте $[a, b]$ *можно утверждать, что на этом сегменте найдется точка* ξ *такая, что справедлива формула*

$$\int_a^b f(x)\,g(x)\,dx = f(\xi)\int_a^b g(x)\,dx. \qquad (**)$$

Формулу $(**)$ принято называть первой формулой среднего значения. Формулу $(*)$ иногда также называют первой формулой среднего значения.

Заметим сразу же, что формула $(**)$ сразу вытекает из формулы $(*)$ и из того, что непрерывная на сегменте $[a, b]$ функция достигает на этом сегменте как своих точных граней M и m, так и любого промежуточного значения $\mu\,(m\leqslant\mu\leqslant M)$.

Таким образом, нужно доказать только формулу $(*)$. Для доказательства формулы $(*)$ заметим, что по определению точных граней для любого значения x из $[a, b]$ справедливы неравенства

$$m\leqslant f(x)\leqslant M.$$

Предполагая ради определенности $g(x)$ неотрицательной на $[a, b]$, мы получим, умножая последние неравенства на $g(x)$, что для любого x из $[a, b]$

$$mg(x)\leqslant f(x)\,g(x)\leqslant Mg(x). \qquad (***)$$

Так как, кроме того, в силу свойств б) и в) из п. 1 каждая из функций $mg(x)$, $Mg(x)$ и $f(x)g(x)$ интегрируема на $[a, b]$, то оценка $(***)$ позволяет утверждать справедливость следующих неравенств:

$$\int_a^b mg(x)\,dx \leqslant \int_a^b f(x)\,g(x)\,dx \leqslant \int_a^b Mg(x)\,dx$$

или, что то же самое (в силу свойства б) из п. 1),

$$m\int_b^b g(x)\,dx \leqslant \int_a^b f(x)\,g(x)\,dx \leqslant M\int_a^b g(x)\,dx. \qquad (****)$$

Могут представиться два случая: 1) $\int_a^b g(x)\,dx = 0$ и 2) $\int_a^b g(x)\,dx > 0$.

В первом случае из неравенств $(****)$ вытекает, что $\int_a^b f(x)\,g(x)\,dx = 0$, и потому формула $(*)$ справедлива при любом μ.

Во втором случае, поделив неравенства $(****)$ на число $\int_a^b g(x)\,dx$, мы получим, что

$$m \leqslant \frac{\int\limits_a^b f(x)\, g(x)\, dx}{\int\limits_a^b g(x)\, dx} \leqslant M.$$

Для завершения доказательства формулы (*) остается обозначить символом μ число

$$\mu = \frac{\int\limits_a^b f(x)\, g(x)\, dx}{\int\limits_a^b g(x)\, dx}.$$

Следствие. Сформулируем отдельно доказанную нами теорему для частного случая $g(x) \equiv 1$.

Пусть функция $f(x)$ интегрируема на сегменте $[a, b]$, а символы M и m обозначают точные грани $f(x)$ на указанном сегменте. Тогда найдется число μ, удовлетворяющее неравенствам $m \leqslant \mu \leqslant$ $\leqslant M$ и такое, что справедлива формула

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

При дополнительном предположении о непрерывности $f(x)$ на $[a, b]$ можно утверждать, что на этом сегменте найдется точка ξ такая, что справедлива формула

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

Последнюю формулу обычно также называют формулой среднего значения.

е) Вторая формула среднего значения. *Пусть функция $f(x)$ интегрируема, а функция $g(x)$ монотонна на сегменте $[a, b]$. Тогда на этом сегменте найдется число ξ такое, что*

$$\int\limits_a^b f(x)\, g(x)\, dx = g(a) \int\limits_a^\xi f(x)\, dx + g(b) \int\limits_\xi^b f(x)\, dx.$$

Установим сначала следующий факт, которым мы воспользуемся ниже

Лемма Абеля*. *Пусть числа p_i удовлетворяют условиям $p_i \geqslant p_j \geqslant 0$ при $i \leqslant j$, а числа $S_i = \sum\limits_{k=1}^{i} q_k$ $(i = 1, 2, \ldots, n)$ —*

---

* Нильс Генрих Абель — норвежский математик (1802—1829).

неравенствам $m \leqslant S_i \leqslant M$, где q_k, m, M *также некоторые числа. Тогда* $mp_1 \leqslant \sum\limits_{k=1}^{n} p_k q_k \leqslant Mp_1$.

Доказательство. Легко проверить, что

$$\sum_{k=1}^{n} p_k q_k = \sum_{k=1}^{n} p_k (S_k - S_{k-1}) = \sum_{k=1}^{n} S_k (p_k - p_{k+1}),$$

где положено $S_0 = 0$, $p_{n+1} = 0$. Так как $p_k \geqslant 0$, $p_k - p_{k+1} \geqslant 0$, то, заменив в последнем равенстве каждое S_i сначала на m, а затем на M, получим

$$m \sum_{k=1}^{n} (p_k - p_{k+1}) \leqslant \sum_{k=1}^{n} p_k q_k \leqslant M \sum_{k=1}^{n} (p_k - p_{k+1}),$$

но

$$\sum_{k=1}^{n} (p_k - p_{k+1}) = p_1 - p_{n+1} = p_1.$$

Таким образом, лемма доказана.

Установим теперь вторую формулу среднего значения. Допустим, что функция $g(x)$ не возрастает на $[a, b]$ и неотрицательна. Функция $f(x)g(x)$ интегрируема как произведение двух интегрируемых функций. Пусть M_k и m_k — точные грани $f(x)$ на частичных сегментах $[x_{k-1}, x_k]$.

Тогда очевидно, что

$$\sum_{k=1}^{n} m_k g(x_{k-1}) \Delta x_k \leqslant \sum_{k=1}^{n} f(x_{k-1}) g(x_{k-1}) \Delta x_k \leqslant \sum_{k=1}^{n} M_k g(x_{k-1}) \Delta x_k.$$

В силу монотонности $g(x)$ справедлива оценка

$$\sum_{k=1}^{n} (M_k - m_k) g(x_{k-1}) \Delta x_k \leqslant g(a) \sum_{k=1}^{n} (M_k - m_k) \Delta x_k.$$

В силу интегрируемости $f(x)$ сумма в правой, а значит, и в левой части последнего неравенства стремится к нулю при стремлении диаметра d разбиений к нулю. Следовательно, при любых числах μ_k таких, что $m_k \leqslant \mu_k \leqslant M_k$, все суммы

$$\sum_{k=1}^{n} m_k g(x_{k-1}) \Delta x_k, \quad \sum_{k=1}^{n} \mu_k g(x_{k-1}) \Delta x_k, \quad \sum_{k=1}^{n} M_k g(x_{k-1}) \Delta x_k$$

стремятся при $d \to 0$ к интегралу $\int\limits_a^b f(x) g(x)\, dx$. Это следует из двусторонней оценки для интегральной суммы функции $f(x)g(x)$ *.

По свойству д) настоящего пункта числа μ_k, где $m_k \leqslant \mu_k \leqslant \leqslant M_k$, можно выбрать так, чтобы $\int\limits_{x_{k-1}}^{x_k} f(x)\,dx = \mu_k\,\Delta x_k$.

Заметим теперь, что функция $F(x) = \int\limits_a^x f(t)\,dt$ непрерывна на сегменте $[a, b]$, так как

$$\Delta F = F(x+\Delta x) - F(x) = \int\limits_x^{x+\Delta x} f(t)\,dt = \mu\Delta x,$$

$$\inf_{t\in[x,x+\Delta x]} f(t) \leqslant \mu \leqslant \sup_{t\in[x,x+\Delta x]} f(t),$$

и, следовательно, $\Delta F \to 0$ при $\Delta x \to 0$.

Рассмотрим следующие числа $S_i = \sum\limits_{k=1}^{i} \mu_k \Delta x_k = \int\limits_a^{x_i} f(t)\,dt$.

Ясно, что $m \leqslant S_i \leqslant M$, где m и M — точные грани функции $F(x)$ на сегменте $[a, b]$. Введем следующие обозначения:

$$p_k = g(x_{k-1}), \quad q_k = \mu_k \Delta x_k, \quad k = 1, 2, \ldots, n.$$

В силу монотонности и неотрицательности функции $g(x)$ выполнимо $p_i \geqslant p_j \geqslant 0$ при $i \leqslant j$. Числа p_k, S_k, q_k удовлетворяют условиям леммы Абеля. Поэтому

$$mg(a) \leqslant \sum_{k=1}^{n} g(x_{k-1})\,\mu_k \Delta x_k \leqslant Mg(a).$$

Сумма $\sum\limits_{k=1}^{n} g(x_{k-1})\,\mu_k \Delta x_k$ заключена между $mg(a)$ и $Mg(a)$. Устремим теперь диаметр d разбиений к нулю. Тогда и предел этой суммы будет заключен между $mg(a)$ и $Mg(a)$, т. е. будут справедливы неравенства

$$mg(a) \leqslant \int\limits_a^b f(x)\,g(x)\,dx \leqslant Mg(a).$$

* Т. е. из неравенств

$$\sum_{k=1}^{n} m_k g(x_{k-1})\,\Delta x_k \leqslant \sum_{k=1}^{n} f(x_{k-1})\,g(x_{k-1})\,\Delta x_k \leqslant \sum_{k=1}^{n} M_k g(x_k)\,\Delta x_k.$$

Непрерывная функция $F(x)=\int\limits_a^x f(t)\,dt$ принимает любое значение, заключенное между ее точными гранями m и M. Так как

$$m \leqslant \frac{\int\limits_a^b f(x)\,g(x)\,dx}{g(a)} \leqslant M,$$

то существует точка ξ такая, что

$$F(\xi)=\int\limits_a^\xi f(t)\,dt=\frac{\int\limits_a^b f(x)\,g(x)\,dx}{g(a)}.$$

Следовательно, *в случае, когда $g(x)$ не возрастает и неотрицательна, доказана формула*

$$\int\limits_a^b f(x)\,g(x)\,dx=g(a)\int\limits_a^\xi f(t)\,dt.$$

Рассмотрим теперь общий случай невозрастающей функции $g(x)$. В этом случае функция $h(x)=g(x)-g(b)$ не возрастает и неотрицательна. Подставив ее вместо $g(x)$ в формулу, доказанную выше, получим, что

$$\int\limits_a^b f(x)\,[g(x)-g(b)]\,dx=[g(a)-g(b)]\int\limits_a^\xi f(x)\,dx.$$

Окончательно получаем равенство

$$\int\limits_a^b f(x)\,g(x)\,dx=g(a)\int\limits_a^\xi f(x)\,dx+g(b)\int\limits_a^b f(x)\,dx-$$

$$-g(b)\int\limits_a^\xi f(x)\,dx=g(a)\int\limits_a^\xi f(x)\,dx+g(b)\int\limits_\xi^b f(x)\,dx,$$

что и требовалось доказать *.

Рассмотрим примеры на применение оценок для интегралов.

Примеры. 1) Рассмотрим функцию

$$f(x)=\begin{cases} x^x, & x\in(0,1],\\ 1, & x=0.\end{cases}$$

Эта функция $f(x)$ непрерывна на сегменте $[0,1]$. Легко убедиться

* Если $g(x)$ не убывает, то заменой $g_1(x)=-g(x)$ сводим этот случай к уже рассмотренному.

с помощью вычисления производной, что эта функция достигает локального минимума при $x_0=1/e$. При этом $f(1/e)=e^{-1/e}$, и это значение является ее наименьшим значением на сегменте $[0,1]$. Используя свойство б) настоящего пункта, получим, что $e^{-1/e}\leqslant$ $\leqslant\int\limits_0^1 x^x dx\leqslant 1$, а для числа $e^{-1/e}$ легко получить, что $e^{-1/e}=0{,}692\ldots$. Заметим, что в этом случае значение интеграла нельзя определить через значения элементарных функций.

2) Если функция $f(x)$ не является непрерывной, то формула среднего значения (**) может быть несправедливой. Рассмотрим функцию

$$f(x)=\begin{cases} 1/2, & 0\leqslant x\leqslant 1/2,\\ 3/4, & 1/2<x\leqslant 1.\end{cases}$$

Тогда $\int\limits_0^1 f(x)\,dx=5/8$. Значение 5/8 не принимается функцией $f(x)$ ни в одной точке ξ сегмента $[0,1]$. Следовательно, не существует числа $\xi\in[0,1]$, для которого $\int\limits_0^1 f(x)\,dx=f(\xi)$.

§ 5. ПЕРВООБРАЗНАЯ НЕПРЕРЫВНОЙ ФУНКЦИИ. ПРАВИЛА ИНТЕГРИРОВАНИЯ ФУНКЦИЙ

В предыдущих параграфах уже достаточно полно изучены свойства интеграла Римана. В частности, было показано, что, пользуясь определением интеграла, можно вычислить интеграл от некоторых простейших функций. Однако такое вычисление интеграла с помощью предельного перехода в интегральных суммах обладает рядом неудобств и приводит к значительным трудностям. Поэтому на повестку дня встает вопрос о простых правилах вычисления определенного интеграла Римана. Ниже нами будет дано одно из таких правил вычисления определенного интеграла, а именно будет доказана основная формула интегрального исчисления (формула Ньютона — Лейбница).

1. Первообразная. Рассмотрим интегрируемую на сегменте $[a, b]$ функцию $f(x)$. Пусть p принадлежит $[a, b]$. Тогда для любого x из $[a, b]$ функция $f(x)$ интегрируема на $[p, x]$, и поэтому на сегменте $[a, b]$ определена функция $F(x)=\int\limits_p^x f(t)\,dt$, которая называется *интегралом с переменным верхним пределом*. Аналогично определяется функция $F(x)$ на интервале (a, b) при условии, что $f(x)$ определена на интервале (a, b) и интегрируема на любом сегменте, принадлежащем этому интервалу.

Теорема 9.5. *Если функция $f(x)$ интегрируема на сегменте $[a, b]$, p — любая точка этого сегмента, то производная функции $F(x)=\int\limits_p^x f(t)\,dt$ существует в каждой точке x_0 непрерывности подынтегральной функции, причем $F'(x_0)=f(x_0)$* *.

Доказательство. В силу непрерывности функции $f(x)$ в точке x_0 для любого $\varepsilon>0$ найдется такое $\delta>0$, что $f(x_0)-\varepsilon< <f(x)<f(x_0)+\varepsilon$, если $|x-x_0|<\delta$. Для всех t из $[x_0, x]$ выполняется неравенство $|t-x_0|\leqslant|x-x_0|<\delta$. Поэтому для всех таких t

$$f(x_0)-\varepsilon\leqslant f(t)\leqslant f(x_0)+\varepsilon.$$

Согласно свойству д) п. 2 § 4 (независимо от знака разности $x-x_0$) получим из последних неравенств

$$f(x_0)-\varepsilon\leqslant\frac{1}{x-x_0}\int\limits_{x_0}^{x} f(t)\,dt\leqslant f(x_0)+\varepsilon \text{ при } |x-x_0|<\delta.$$

(Значение $\mu=\dfrac{1}{x-x_0}\int\limits_{x_0}^{x} f(t)\,dt$ не меняется при перестановке чисел x и x_0, так как при этом одновременно меняется знак у величины $x-x_0$ и у интеграла $\int\limits_{x_0}^{x} f(t)\,dt$.). Но $\dfrac{1}{x-x_0}\int\limits_{x_0}^{x} f(t)\,dt=$ $=\dfrac{F(x)-F(x_0)}{x-x_0}$, следовательно, при $|x-x_0|<\delta$

$$f(x_0)-\varepsilon\leqslant\frac{F(x)-F(x_0)}{x-x_0}\leqslant f(x_0)+\varepsilon,$$

т. е. $F'(x)$ существует и равна $f(x_0)$. Теорема доказана.

Следствие. *Любая непрерывная на сегменте $[a, b]$ функция $f(x)$ имеет на этом сегменте первообразную. Одной из первообразных является функция $F(x)=\int\limits_a^x f(t)\,dt$.*

Замечание 1. Теорема остается справедливой, если $f(x)$ непрерывна на интервале (a, b). В этом случае в качестве нижнего предела надлежит взять любую точку p этого интервала. Все рассуждения сохраняются.

---

* Если точка x_0 совпадает с одним из концов сегмента $[a, b]$, то под производной в точке x_0 функции $F(x)$ понимается левая или правая производная соответственно. Доказательство теоремы при этом не меняется.

Замечание 2. Можно рассматривать и функцию нижнего предела интеграла от $f(x)$, т. е. функцию $\Phi(x)=\int\limits_x^q f(t)\,dt$. Для такой функции

$$\Phi'(x)=-f(x).$$

Замечание 3. Если функция $f(x)$ интегрируема на любом сегменте, содержащемся в интервале (a, b), то интеграл с переменным верхним пределом есть непрерывная на (a, b) функция верхнего предела.

Действительно, пусть $F(x)=\int\limits_p^x f(t)\,dt$, $p\in(a, b)$. Тогда

$$\Delta F=F(x+\Delta x)-F(x)=\int\limits_x^{x+\Delta x} f(t)\,dt=\mu\Delta x,$$

где

$$\inf_{x\in[x,x+\Delta x]} f(x)\leqslant\mu\leqslant\sup_{x\in[x,x+\Delta x]} f(x)$$

по первой формуле среднего значения. Если функция $f(x)$ интегрируема, то она ограничена, а поэтому для всех достаточно малых Δx ограничена и величина μ, зависящая от x и Δx. Более точно, $\inf\limits_{x\in[c,d]} f(x)\leqslant\mu\leqslant\sup\limits_{x\in[c,d]} f(x)$*. Поэтому $\Delta F\to 0$ при $\Delta x\to 0$.

Замечание 4. Интегралы с переменным верхним (или нижним) пределом можно использовать для *определения* новых функций, не выражающихся через элементарные функции.

Так, например, интеграл $\int\limits_0^x e^{-t^2}\,dt$, как уже отмечалось, называется интегралом Пуассона, интеграл $\int\limits_0^x \frac{dt}{\sqrt{(1-t^2)(1-k^2t)}}$, $0<k<1$ называется эллиптическим интегралом, интеграл $\int\limits_0^x \frac{\sin t}{t}\,dt$ — интегральным синусом, $\int\limits_1^x \frac{\cos t\,dt}{t}$ — интегральным косинусом и т. д.

2. Основная формула интегрального исчисления. Мы уже знаем из предыдущих рассмотрений, что любые две первообразные функции $f(x)$, заданной на сегменте $[a, b]$, отличаются на постоянную. Поэтому если $F(x)=\int\limits_a^x f(t)\,dt$, а $\Phi(x)$ — любая дру-

* Здесь $[c, d]$ — какой-либо фиксированный сегмент, принадлежащий интервалу (a, b) и такой, что $x\in[c, d]$, $x+\Delta x\in[c, d]$.

гая первообразная непрерывной функции $f(x)$, то $\Phi(x)-F(x)=$ $=C=\text{const}$, т. е. $\Phi(x)=\int\limits_a^x f(t)\,dt+C$ (см. теорему 9.5). Положим в последней формуле сначала $x=a$, а затем $x=b$. Как мы условились (см. п. 1 предыдущего параграфа), $\int\limits_a^a f(t)\,dt=0$ для любой функции, принимающей конечное значение в точке a, поэтому $\Phi(a)=C$, $\Phi(b)=\int\limits_a^b f(x)\,dx+C$. Отсюда

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

и нами получена основная формула интегрального исчисления.

Сформулируем ее в виде теоремы.

Теорема (основная теорема интегрального исчисления). *Для того чтобы вычислить определенный интеграл по сегменту* $[a, b]$ *от непрерывной функции* $f(x)$, *следует вычислить значение произвольной ее первообразной в точке* b *и в точке* a *и вычесть из первого значения второе.*

Теперь мы имеем правило вычисления определенного интеграла от широкого класса интегрируемых функций. Задача вычисления определенного интеграла свелась к задаче нахождения первообразной непрерывной функции. Естественно, что не для всякой функции найти первообразную просто. Мы уже неоднократно указывали функции, первообразные которых не выражаются через элементарные функции (см. например, п. 1 этого параграфа). В этом случае, естественно, возникает вопрос о приближенном вычислении определенных интегралов, о чем пойдет речь ниже.

Основную формулу интегрального исчисления часто записывают в форме $\int\limits_a^b f(x)\,dx=\Phi(x)\,|_a^b$, где введено обозначение

$$\Phi(x)\,|_a^b=\Phi(b)-\Phi(a).$$

3. Важные правила, позволяющие вычислять определенные интегралы. При вычислении определенных интегралов очень часто используется правило замены переменной под знаком определенного интеграла.

Пусть функция $x=g(t)$ *имеет непрерывную производную на* *сегменте* $[m, M]$ * и $\min\limits_{t\in[m,M]} g(t)=a$, $\max\limits_{t\in[m,M]} g(t)=b$, *причем* $g(m)=a$, $g(M)=b$. *Тогда* $\int\limits_a^b f(x)\,dx=\int\limits_m^M f[g(t)]\,g'(t)\,dt$ (*при условии, что функция* $f(x)$ *непрерывна на сегменте* $[a, b]$).

Указанная формула называется формулой замены переменной под знаком определенного интеграла.

Доказательство. *Пусть* $\Phi(x)$ — некоторая первообразная функции $f(x)$. Функции $\Phi(x)$ и $x=g(t)$ дифференцируемы на сегментах $[a, b]$ и $[m, M]$ соответственно. Поэтому, согласно правилу вычисления производной сложной функции, для всех t из $[m, M]$

$$\frac{d}{dt}\Phi(g(t))=\Phi'(g(t))\,g'(t).$$

Заметим, что производная Φ' в выражении справа вычислена по аргументу $x: \Phi'(g(t))=\Phi'(x)$, $x=g(t)$.

Заметим также, что $\Phi'(x)=f(x)$. Подставив в правую часть формулы для $\frac{d}{dt}\Phi(g(t))$ это равенство, получаем

$$\frac{d}{dt}\Phi(g(t))=f(g(t))\,g'(t).$$

Таким образом, функция $\Phi(g(t))$ является на сегменте $m\leqslant t\leqslant M$ первообразной для функции $f(g(t))g'(t)$, т. е.

$$\int\limits_m^M f(g(t))\,g'(t)\,dt=\Phi(g(M))-\Phi(g(m))=\Phi(b)-\Phi(a)$$

согласно условию. Следовательно, с одной стороны, $\int\limits_a^b f(x)\,dx=\Phi(b)-\Phi(a)$, а с другой стороны, $\Phi(b)-\Phi(a)=\int\limits_m^M f(g(t))\,g'(t)\,dt$, что и требовалось.

Сформулируем и установим теперь правило интегрирования по частям.

Пусть функции $f(x)$ *и* $g(x)$ *имеют непрерывные производные на сегменте* $[a, b]$, *тогда*

$$\int\limits_a^b f(x)\,g'(x)\,dx=f(x)\,g(x)\Big|_a^b-\int\limits_a^b g(x)\,f'(x)\,dx.$$

---

* Будем говорить, что $g(t)$ имеет непрерывную производную на сегменте $[m, M]$, если производная $g'(t)$ существует и непрерывна в любой внутренней точке $[m, M]$ и существуют конечные пределы $\lim\limits_{t\to m+0} g'(t)$ и $\lim\limits_{t\to M-0} g'(t)$.

Действительно, $\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x)$. Поэтому функция $f(x)g(x)$ является первообразной функции $[f(x)g'(x) + f'(x)\cdot g(x)]$. Следовательно,

$$\int_a^b [f(x)g'(x) + f'(x)g(x)]\,dx = f(x)g(x)|_a^b,$$

и наше утверждение доказано. Последнюю формулу удобно записывать в виде

$$\int_a^b f dg = [fg]|_a^b - \int_a^b g df.$$

4. Остаточный член формулы Тейлора в интегральной форме. Целью этого пункта является получение формулы Тейлора для функции $f(x)$ в окрестности произвольной точки a с остаточным членом в так называемой интегральной форме.

Пусть функция $f(x)$ имеет в некоторой ε-окрестности точки a непрерывную производную $(n+1)$-го порядка. Пусть x принадлежит указанной окрестности. Рассмотрим равенство

$$f(x) - f(a) = \int_a^x f'(t)\,dt.$$

Полагая $u(t) = f'(t)$, $v(t) = -(x-t)$, применим к интегралу $\int_a^x f'(t)\,dt = \int_a^x u(t)\,dv(t)$ формулу интегрирования по частям. Получим

$$f(x) - f(a) = \int_a^x f'(t)\,dt = -[f'(t)(x-t)]|_a^x + \int_a^x f''(t)(x-t)\,dt =$$

$$= f'(a)(x-a) + \int_a^x f''(t)(x-t)\,dt.$$

Таким образом, последовательно интегрируя по частям, получим

$$f(x) - f(a) = \int_a^x f'(t)\,dt = f'(a)(x-a) + \int_a^x f''(t)(x-t)\,dt =$$

$$= f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{2}\int_a^x f'''(t)(x-t)^2\,dt = \dots$$

$$\ldots = f'(a)(x-a) + 1/2\, f''(a)(x-a)^2 + \ldots$$

$$\ldots + \frac{1}{n!} f^{(n)}(a)(x-a)^n + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt =$$

$$= \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(a)(x-a)^k + R_{n+1}(x),$$

где

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt.$$

Мы видим, что $R_{n+1}(x)$ является остаточным членом разложения Тейлора для функции $f(x)$ в окрестности точки a. Эта форма остаточного члена и называется **интегральной формой**.

Если применить первую формулу среднего значения (см. свойство д) п. 2 § 4), то

$$R_{n+1}(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt = \frac{f^{(n+1)}(\xi)}{n!}\int_a^x (x-t)^n\,dt =$$

$$= -\frac{f^{(n+1)}(\xi)}{n!}\,\frac{(x-t)^{n+1}}{n+1}\Bigg|_a^x = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1},$$

где ξ — некоторая точка сегмента $[a, x]$. Следовательно, при тех же предположениях мы получим остаточный член в форме Лагранжа. На самом деле, легко заключить (используя теорему Дарбу о прохождении производной через все промежуточные значения), что $R_{n+1}(x) = \frac{(x-a)^{n+1} f^{(n+1)}(\xi)}{(n+1)!}$ лишь при условии существования и интегрируемости $f^{(n+1)}(x)$.

Рассмотрим теперь примеры, иллюстрирующие изложенный в этом параграфе материал.

Примеры. 1) Вычислить интегралы, применяя основную формулу интегрального исчисления:

а) $\int_a^b x^n dx = \frac{x^{n+1}}{n+1}\Big|_a^b = \frac{b^{n+1} - a^{n+1}}{n+1}$, $n \neq -1$.

б) $\int_a^b \sin x\,dx = -\cos x\,|_a^b = \cos a - \cos b$,

$$\int_a^b \cos x\,dx = \sin b - \sin a.$$

в) $\int\limits_0^a \frac{dx}{1+x^2} = \operatorname{arctg} x \,|_0^a = \operatorname{arctg} a.$

г) $\int\limits_0^a \frac{x^2}{a^3+x^3}\,dx = \frac{1}{3}\ln(a^3+x^3)\,|_0^a = \frac{1}{3}\ln 2,\ a>0.$

2) Вычислить интегралы, применяя правило замены переменной:

а) $\int\limits_0^{\pi/4} \frac{\sin^4 x}{\cos^6 x}\,dx = \int\limits_0^{\pi/4} \frac{\operatorname{tg}^4 x}{\cos^2 x}\,dx = \int\limits_0^1 t^4 dt = \frac{1}{5}t^5\Big|_0^1 = 1/5,$

где $t=\operatorname{tg} x$.

б) $\int\limits_0^1 x\sqrt{1+x^2}\,dx = \int\limits_1^{\sqrt{2}} t^2 dt = \frac{1}{3}t^3\Big|_1^{\sqrt{2}} = (2\sqrt{2}-1)/3,$

где $t=\sqrt{1+x^2}$.

3) Вычислить интеграл, применяя правило интегрирования по частям:

$$I_m = \int\limits_0^{\pi/2} \sin^m x dx = -\int\limits_0^{\pi/2} \sin^{m-1} x d(\cos x) =$$

$$= -\sin^{m-1} x\cos x\,|_0^{\pi/2} + (m-1)\int\limits_0^{\pi/2} \cos^2 x \sin^{m-2} x dx =$$

$$= (m-1)\int\limits_0^{\pi/2} (1-\sin^2 x)\sin^{m-2} x dx = (m-1)\,[I_{m-2}-I_m].$$

Отсюда: $I_m = \frac{m-1}{m} I_{m-2}$ при $m \geqslant 2$.

Легко видеть, что $I_0=\pi/2$, $I_1=1$. По индукции получаем, что для $m \geqslant 2$

$$I_{2m} = \frac{2m-1}{2m}\,\frac{2m-3}{2m-2}\cdots\frac{3}{4}\,\frac{1}{2}\,\frac{\pi}{2} = \frac{(2m-1)!!}{(2m)!!}\,\frac{\pi}{2},$$

$$I_{2m+1} = \frac{2m}{2m+1}\,\frac{2m-2}{2m-1}\cdots\frac{4}{5}\,\frac{2}{3}\,1 = \frac{(2m)!!}{(2m+1)!!}.$$

4) Доказать, что для функции $f(x)=(1+x)^\alpha$ остаточный член R_{n+1} в интегральной форме стремится к нулю, если $|x|<$ <1. Заметим, что

$$R_{n+1}=\frac{\alpha(\alpha-1)\ldots(\alpha-n)}{n!}\int_0^x(1+t)^{\alpha-n-1}(x-t)^n\,dt.$$

Из очевидных неравенств $t/x\geqslant0$, $1-x>0$ следует, что

$$\frac{t}{x}-t=\frac{t}{x}(1-x)\geqslant0 \text{ или } \frac{x-t}{x}=1-\frac{t}{x}\leqslant1-t.$$

Далее, поскольку x и $x-t$ — числа одного знака, то

$$\left|\frac{x-t}{x}\right|=1-\frac{t}{x}\leqslant1-t=|1-t| \text{ или } \left|\frac{x-t}{1-t}\right|\leqslant|x|.$$

Следовательно,

$$\left|\int_0^x\frac{(x-t)^n}{(1+t)^n}(1+t)^{\alpha-1}\,dt\right|\leqslant|x|^n\int_0^x(1+t)^{\alpha-1}\,dt=C(x,\alpha)|x|^n,$$

где $C(x,\alpha)$ не зависит от n. Иными словами,

$$|R_{n+1}|\leqslant C(x,\alpha)\cdot|\alpha(\alpha-1)\ldots(\alpha-n)|\,|x|^n/n!=p_n.$$

Рассмотрим какое-либо число q, удовлетворяющее условию $|x|<q<1$. Так как

$$\frac{p_{n+1}}{p_n}=\frac{|\alpha-n-1|\,|x|}{n+1}\to|x|\quad(n\to\infty),$$

то найдется такой номер N, что $\frac{p_{n+1}}{p_n}<q$ при $n\geqslant N$. Отсюда вытекает, что $p_n\leqslant p_Nq^{n-N}$ при $n\geqslant N$. Устремляя в этом неравенстве n к ∞, убеждаемся в том, что p_n, а следовательно, и R_{n+1} стремится к нулю.

§ 6. НЕРАВЕНСТВА ДЛЯ СУММ И ИНТЕГРАЛОВ

В этом параграфе мы получим некоторые важные неравенства для сумм и интегралов.

1. Неравенство Юнга*. Рассмотрим два неотрицательных числа a и b и два числа p и q, превосходящие единицу и такие, что $1/p+1/q=1$. Докажем следующее *неравенство Юнга*:

$$ab\leqslant\frac{1}{p}a^p+\frac{1}{q}b^q.$$

Доказательство. Рассмотрим функцию $f(x)=x^{1/p}-x/p$ при $x\geqslant0$. Поскольку $f'(x)=\frac{1}{p}(x^{-1/q}-1)$, то $f'(x)>0$

* Вилиам Юнг — английский математик (1882—1946).

при $0<x<1$ и $f'(x)<0$ при $x>1$. В точке $x=1$ функция $f(x)$ принимает наибольшее значение, причем $f(1)=1-\frac{1}{p}=\frac{1}{q}$. Следовательно, $x^{1/p}-x/p\leqslant 1/q$ при всех $x\geqslant 0$. В последнем неравенстве полагаем $x=a^p/b^p$, $b\neq 0$. Неравенство Юнга при $b\neq 0$, таким образом, установлено. При $b=0$ оно очевидно.

2. Неравенство Гёльдера* для сумм. *Пусть* $a_1, a_2, \ldots, a_n$ *и* $b_1, b_2, \ldots, b_n$ — *какие угодно неотрицательные числа. Тогда*

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} b_i^q\right)^{1/q},$$

где $1/p+1/q=1$, $p\geqslant 1$, $q\geqslant 1$. Это неравенство называется **неравенством Гёльдера для сумм**.

Доказательство. Заметим, что указанное неравенство однородно в том смысле, что если оно выполнено для чисел a_i, b_i, то оно справедливо и для чисел ta_i, kb_i. Поэтому достаточно установить, что $\sum_{i=1}^{n} a_i b_i \leqslant 1$ при условии $\sum_{i=1}^{n} a_i^p = 1$, $\sum_{i=1}^{n} b_i^q = 1$, так как мы всегда можем разделить числа a_i и b_i соответственно на величины $\left(\sum_{i=1}^{n} a_i^p\right)^{1/p}$ и $\left(\sum_{i=1}^{n} b_i^q\right)^{1/q}$**. Записав неравенство Юнга для таких чисел a_i и b_i и просуммировав эти неравенства по i, получим

$$\sum_{i=1}^{n} a_i b_i \leqslant \frac{1}{p}\sum_{i=1}^{n} a_i^p + \frac{1}{q}\sum_{i=1}^{n} b_i^q.$$

Поэтому $\sum_{i=1}^{n} a_i b_i \leqslant \frac{1}{p}+\frac{1}{q}=1$, что и требовалось.

Замечание. В случае $p=2$, $q=2$ неравенство Гёльдера превращается в неравенство

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \left(\sum_{i=1}^{n} b_i^2\right)^{1/2},$$

называемое **неравенством Коши — Буняковского***** для сумм.

---

* О. Гёльдер — немецкий математик (1859—1937).

** Мы предполагаем, что хотя бы одно из чисел a_i и хотя бы одно из чисел b_i отлично от нуля. В противном случае неравенство очевидно.

*** Виктор Яковлевич Буняковский — русский математик (1804—1889).

3. Неравенство Минковского* для сумм. *Пусть* $a_1, a_2, \ldots, a_n$ и $b_1, b_2, \ldots, b_n$ — *произвольные неотрицательные числа, число* $p>1$. *Тогда справедливо следующее* *неравенство Минковского для сумм:*

$$\left(\sum_{i=1}^{n}(a_i+b_i)^p\right)^{1/p} \leqslant \left(\sum_{i=1}^{n}a_i^p\right)^{1/p} + \left(\sum_{i=1}^{n}b_i^p\right)^{1/p}.$$

Доказательство. Запишем равенство

$$\sum_{i=1}^{n}(a_i+b_i)^p = \sum_{i=1}^{n}a_i(a_i+b_i)^{p-1} + \sum_{i=1}^{n}b_i(a_i+b_i)^{p-1}.$$

К каждой из сумм, стоящих в правой части, применимо неравенство Гёльдера. Если $1/p+1/q=1$, $p>1$, $q>1$, то $(p-1)q=$ $=p$, $\frac{1}{q}=\frac{p-1}{p}$. Поэтому

$$\sum_{i=1}^{n}(a_i+b_i)^p \leqslant \left(\sum_{i=1}^{n}a_i^p\right)^{1/p}\left(\sum_{i=1}^{n}(a_i+b_i)^{(p-1)q}\right)^{1/q} +$$

$$+\left(\sum_{i=1}^{n}b_i^p\right)^{1/p}\left(\sum_{i=1}^{n}(a_i+b_i)^{(p-1)q}\right)^{1/q} =$$

$$=\left\{\left(\sum_{i=1}^{n}a_i^p\right)^{1/p} + \left(\sum_{i=1}^{n}b_i^p\right)^{1/p}\right\}\left[\sum_{i=1}^{n}(a_i+b_i)^p\right]^{\frac{p-1}{p}}.$$

Поделив обе части последнего неравенства на $\left[\sum_{i=1}^{n}(a_i+\right.$ $\left.+b_i)^p\right]^{\frac{p-1}{p}}$, получим доказываемое неравенство.

4. Неравенство Гёльдера для интегралов. Пусть $f(x)$ и $g(x)$ — две произвольные интегрируемые на сегменте $[a, b]$ функции, пусть p и q — два числа, превосходящие единицу, и $1/p+1/q=1$. Тогда справедливо неравенство Гёльдера для интегралов

$$\left|\int_a^b f(x)g(x)\,dx\right| \leqslant \left(\int_a^b |f(x)|^p\,dx\right)^{1/p}\left(\int_a^b |g(x)|^q\,dx\right)^{1/q}$$

(все написанные интегралы существуют в силу следствия из теоремы 9.4′).

* Герман Минковский — немецкий математик и физик (1864—1909).

Заметим, что, как и в п. 2, достаточно рассмотреть случай, когда $\int\limits_a^b |f(x)|^p dx \leqslant 1$ и $\int\limits_a^b |g(x)|^q dx \leqslant 1$, и доказать неравенство $\left|\int\limits_a^b f(x)g(x)\,dx\right| \leqslant 1$. Запишем неравенство Юнга в любой точке x для функций $|f(x)|$ и $|g(x)|$. Получим

$$|f(x)|\,|g(x)| \leqslant \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q.$$

Интегрируя это неравенство, получим

$$\int\limits_a^b |f(x)|\,|g(x)|\,dx \leqslant 1,$$

но, согласно свойству д) п. 2 § 4

$$\left|\int\limits_a^b f(x)g(x)\,dx\right| \leqslant \int\limits_a^b |f(x)|\,|g(x)|\,dx.$$

Поэтому доказательство завершено.

Как и при выводе неравенства Гёльдера для сумм, мы предполагали, что $\int\limits_a^b |f(x)|\,dx \neq 0$ и $\int\limits_a^b |g(x)|\,dx \neq 0$. В противном случае неравенство очевидно.

Замечание. В случае $p=2$, $q=2$ неравенство Гёльдера *превращается в неравенство*

$$\left|\int\limits_a^b f(x)g(x)\,dx\right| \leqslant \left(\int\limits_a^b |f(x)|^2 dx\right)^{1/2} \left(\int\limits_a^b |g(x)|^2 dx\right)^{1/2},$$

которое называется неравенством Коши—Буняковского для интегралов.

5. Неравенство Минковского для интегралов. *Пусть $f(x)$ и $g(x)$ — любые две неотрицательные и интегрируемые на сегменте $[a, b]$ функции и число $p>1$. Тогда справедливо неравенство Минковского для интегралов*

$$\left\{\int\limits_a^b (f(x)+g(x))^p dx\right\}^{1/p} \leqslant \int\limits_a^b \left(f^p(x)\,dx\right)^{1/p} + \left(\int\limits_a^b g^p(x)\,dx\right)^{1/p}.$$

Заметим, что согласно следствию из теоремы 9.4′ все подынтегральные функции интегрируемы.

Доказательство. Точно так же, как и при доказательстве неравенства Минковского для сумм, запишем равенство

$$\int_a^b (f(x)+g(x))^p\,dx=$$

$$=\int_a^b f(x)\,(f(x)+g(x))^{p-1}\,dx+\int_a^b g(x)\,(f(x)+g(x))^{p-1}\,dx.$$

Далее применяем к интегралам, стоящим справа, неравенство Гёльдера и, как и в п. 3, получаем доказательство.

По индукции можно доказать и более общее неравенство для функций $f_1(x), f_2(x), \ldots, f_n(x)$, неотрицательных и интегрируемых на сегменте $[a, b]$:

$$\left\{\int_a^b (f_1(x)+f_2(x)+\ldots+f_n(x))^p\,dx\right\}^{1/p}\leqslant$$

$$\leqslant\left(\int_a^b f_1^p(x)\,dx\right)^{1/p}+\left(\int_a^b f_2^p(x)\,dx\right)^{1/p}+\ldots+\left(\int_a^b f_n^p(x)\,dx\right)^{1/p}.$$

§ 7. ДОПОЛНИТЕЛЬНЫЕ СВЕДЕНИЯ ОБ ОПРЕДЕЛЕННОМ ИНТЕГРАЛЕ РИМАНА

1. Предел интегральных сумм по базису фильтра. Рассмотрим множество всех разбиений $\{x_k\}$ сегмента $[a, b]$. Обозначим это множество* всевозможных разбиений символом P^*. Каждому фиксированному разбиению $\{x_k\}\in P^*$ отвечает некоторая интегральная сумма σ. Таким образом, получается отображение множества разбиений в множество интегральных сумм. Выберем в P^* базис фильтра (базу) $B^*=\{B_\delta{}^*\}$ с элементами

$$B_\delta{}^*=\{\{x_k\}\in P^*: d<\delta\}.$$

Эта запись означает: элемент $B_\delta{}^*$ есть разбиение $\{x_k\}$ с диаметром d, меньшим δ. Вспоминая общее определение предела по базе (базису фильтра), мы можем записать, что

$$\int_a^b f(x)\,dx=\lim_{d\to 0}\sigma,$$

где предел означает предел числовой функции σ (интегральной суммы функции $f(x)$) по базису фильтра $(d\to 0)$. В силу свойств предела по базе такой предел является единственным. Ясно, что все теоремы предыдущих параграфов, где использовался предел сумм при стремлении диаметра разбиений к нулю,

* Предполагается, что каждому разбиению отвечает и некоторый выбор промежуточных точек.

можно сформулировать, используя понятие предела по базису фильтра.

2. Критерий интегрируемости Лебега. Будем говорить, что множество A точек сегмента $[a, b]$ имеет меру нуль, если для любого $\varepsilon>0$ можно указать не более чем счетную систему интервалов, покрывающую все точки множества A и имеющую сумму длин, не большую чем ε. Заметим, что число интервалов может быть и бесконечным, однако они имеют длины a_n такие, что если

$$S_n = a_1 + a_2 + \ldots + a_n, \text{ то } \lim_{n\to\infty} S_n = I < \varepsilon.$$

В теории, изучающей меру множеств, доказывается следующий критерий интегрируемости функции $f(x)$ на сегменте $[a, b]$ по Риману.

Теорема 9.6 (критерий Лебега). *Для того чтобы ограниченная на сегменте $[a, b]$ функция $f(x)$ была интегрируемой по Риману на этом сегменте, необходимо и достаточно, чтобы множество точек разрыва этой функции имело меру нуль.*

Доказательство этой теоремы можно найти в книге В. А. Ильина, Э. Г. Позняка «Основы математического анализа», II, с. 265—266.

ДОПОЛНЕНИЕ 1

НЕСОБСТВЕННЫЕ ИНТЕГРАЛЫ

Изученное в гл. 9 понятие определенного интеграла Римана существенно использовало два обстоятельства: 1) тот факт, что промежуток $[a, b]$, по которому требуется произвести интегрирование, является конечным; 2) тот факт, что подынтегральная функция $f(x)$ является на рассматриваемом промежутке ограниченной.

В настоящем дополнении будет дано обобщение понятия определенного интеграла Римана на два случая: 1) на случай, когда промежуток, по которому требуется произвести интегрирование, является бесконечным*; 2) на случай, когда подынтегральная функция $f(x)$ является неограниченной в окрестности некоторых точек области интегрирования.

Возникающее при таком обобщении понятие интеграла принято называть несобственным интегралом соответственно первого или второго рода.

* Т. е. представляет собой полупрямую или всю бесконечную прямую.

§ 1. *Несобственные интегралы первого рода*

1. Понятие несобственного интеграла первого рода. Перенесем понятие определенного интеграла на случай, когда область, по которой производится интегрирование, является бесконечной. На прямой $-\infty<x<+\infty$ существует три типа бесконечных связных замкнутых множеств: 1) полупрямая $a\leqslant x<+\infty$; 2) полупрямая $-\infty<x\leqslant b$; 3) вся бесконечная прямая $-\infty<x<+\infty$.

Ради определенности рассмотрим подробно первое из указанных множеств, т. е. полупрямую $a\leqslant x<+\infty$.

Предположим, что функция $f(x)$ определена на полупрямой $a\leqslant x<+\infty$ и что для любого числа A, удовлетворяющего неравенству $A\geqslant a$, существует определенный интеграл Римана $\int\limits_a^A f(x)\,dx$.

Этот определенный интеграл мы обозначим символом

$$F(A)=\int\limits_a^A f(x)\,dx. \qquad (9.1.1)$$

Естественно, возникает вопрос о существовании предела функции $F(A)$ при $A\to+\infty$

$$\lim_{A\to+\infty} F(A)=\lim_{A\to+\infty}\int\limits_a^A f(x)\,dx. \qquad (9.1.2)$$

Определение. *Предел* (9.1.2) *в случае, если он существует, называется* ***несобственным интегралом первого рода*** *от функции* $f(x)$ *по полупрямой* $[a, +\infty)$ *и обозначается символом*

$$\int\limits_a^{+\infty} f(x)\,dx. \qquad (9.1.3)$$

При этом говорят, что несобственный интеграл (9.1.3) *сходится, и пишут равенство*

$$\int\limits_a^{+\infty} f(x)\,dx=\lim_{A\to+\infty}\int\limits_a^A f(x)\,dx.$$

Впрочем, *символ* (9.1.3) *употребляют и в случае, если указанного выше предела* (9.1.2) *не существует, но в этом случае говорят, что несобственный интеграл* (9.1.3) *расходится.*

Совершенно аналогично определяются несобственные интегралы по полупрямой $-\infty<x\leqslant b$ и по всей бесконечной прямой $-\infty<x<+\infty$.

Первый из этих интегралов определяется как предел $\lim\limits_{A\to-\infty}\int\limits_A^b f(x)\,dx$ и обозначается символом $\int\limits_{-\infty}^b f(x)\,dx$.

Что же касается интеграла $\int\limits_{-\infty}^{\infty} f(x)\,dx$, то он определяется как предел

$$\lim_{\substack{A'\to-\infty\\ A''\to+\infty}}\int\limits_{A'}^{A''} f(x)\,dx$$

при независимом друг от друга стремлении A' к $-\infty$ и A'' к $+\infty$.

Из этих определений следует, что если для некоторого вещественного числа a сходится каждый из несобственных интегралов $\int\limits_{-\infty}^{a} f(x)\,dx$ и $\int\limits_{a}^{+\infty} f(x)\,dx$, то сходится и несобственный интеграл $\int\limits_{-\infty}^{+\infty} f(x)\,dx$, причем справедливо равенство

$$\int\limits_{-\infty}^{+\infty} f(x)\,dx=\int\limits_{-\infty}^{a} f(x)\,dx+\int\limits_{a}^{+\infty} f(x)\,dx.$$

Заметим еще, что если сходится несобственный интеграл $\int\limits_{a}^{+\infty} f(x)\,dx$ и b — любое число, превосходящее a, то сходится и несобственный интеграл $\int\limits_{b}^{+\infty} f(x)\,dx$, причем $\int\limits_{a}^{+\infty} f(x)\,dx=\int\limits_{a}^{b} f(x)\,dx+$ $+\int\limits_{b}^{+\infty} f(x)\,dx$. Это утверждение непосредственно вытекает из определения сходимости несобственного интеграла.

Примеры. 1) Изучим вопрос о сходимости несобственного интеграла

$$\int\limits_0^{+\infty}\frac{dx}{1+x^2}.$$

Поскольку функция $f(x)=\frac{1}{1+x^2}$ при любом $A>0$ интегрируема на сегменте $[0, A]$, причем для нее

$$F(A)=\int\limits_0^A\frac{dx}{1+x^2}=\operatorname{arctg} x\,|_0^A=\operatorname{arctg} A,$$

то

$$\lim_{A\to+\infty} F(A) = \lim_{A\to+\infty} \operatorname{arctg} A = \pi/2.$$

Поэтому несобственный интеграл $\int\limits_0^{+\infty} \frac{dx}{1+x^2}$ сходится и для него справедливо равенство $\int\limits_0^{+\infty} \frac{dx}{1+x^2} = \frac{\pi}{2}$.

2) Изучим вопрос о сходимости несобственного интеграла $\int\limits_a^{+\infty} \frac{dx}{x^\lambda}$, где a и λ — произвольные вещественные числа, первое из которых положительно $(a>0)$.

Так как функция $f(x) = 1/x^\lambda$ при любом $A>0$ интегрируема на сегменте $[a, A]$, причем

$$F(A) = \int\limits_a^A \frac{dx}{x^\lambda} = \left\{ \frac{x^{1-\lambda}}{1-\lambda} \right|_a^A = \frac{A^{1-\lambda} - a^{1-\lambda}}{1-\lambda} \quad \text{при } \lambda \neq 1,$$

то при $\lambda>1$ предел $F(A)$ при $A\to+\infty$ существует и равен $\frac{a^{1-\lambda}}{\lambda-1}$, а при $\lambda\leqslant 1$ указанный предел не существует.

Таким образом, при $\lambda>1$ несобственный интеграл $\int\limits_a^{+\infty} \frac{dx}{x^\lambda}$ сходится и равен $\frac{a^{1-\lambda}}{\lambda-1}$, а при $\lambda\leqslant 1$ несобственный интеграл $\int\limits_a^{+\infty} \frac{dx}{x^\lambda}$ расходится.

2. Критерий Коши сходимости несобственного интеграла первого рода. Достаточные признаки сходимости. Вопрос о сходимости несобственного интеграла 1-го рода эквивалентен вопросу о существовании предельного значения функции $F(A) = \int\limits_0^A f(x)\,dx$ при $A\to+\infty$. Как известно *, для существования предельного значения функции $F(A)$ при $A\to+\infty$ необходимо и достаточно, чтобы она удовлетворяла следующему условию Коши: для любого $\varepsilon>0$ можно указать такое $B>a$, что для любых A_1 и A_2, превосходящих B, выполняется неравенство

* Гл. 3, § 4, п. 3.

$$|F(A_2)-F(A_1)| = \left|\int_{A_1}^{A_2} f(x)\,dx\right| < \varepsilon.$$

Таким образом, справедливо следующее утверждение.

Утверждение 1 (критерий Коши сходимости несобственного интеграла). *Для сходимости несобственного интеграла* (9.1.3) *необходимо и достаточно, чтобы для любого* $\varepsilon>0$ *можно было указать такое* $B>a$, *что для любых* A_1 *и* A_2, *превосходящих* B,

$$\left|\int_{A_1}^{A_2} f(x)\,dx\right| < \varepsilon.$$

Замечание. Отметим, что из сходимости несобственного интеграла не вытекает даже ограниченность подынтегральной функции. Например, интеграл $\int\limits_0^{+\infty} f(x)\,dx$, где функция $f(x)$ равна нулю для всех нецелых x и равна n при $x=n$, где n — целое число, очевидно, сходится, хотя подынтегральная функция не ограничена.

Поскольку критерий Коши мало удобен для практических применений, целесообразно указать различные достаточные признаки сходимости несобственных интегралов.

В дальнейшем *мы все время будем считать, что функция* $f(x)$ *задана на полупрямой* $a\leqslant x<+\infty$ *и для любого* $A\geqslant a$ *существует обычный интеграл* $\int\limits_a^A f(x)\,dx$.

Докажем следующее утверждение.

Утверждение 2 (общий признак сравнения). *Пусть на полупрямой* $a\leqslant x<+\infty$

$$|f(x)|\leqslant g(x). \tag{9.1.4}$$

Тогда из сходимости интеграла $\int\limits_a^{+\infty} g(x)\,dx$ *вытекает сходимость интеграла* $\int\limits_a^{+\infty} f(x)\,dx$.

Доказательство. Пусть $\int\limits_a^{+\infty} g(x)\,dx$ сходится. Тогда, согласно критерию Коши, для любого $\varepsilon>0$ найдется такое $B>a$, что для любых $A_1>B$ и $A_2>B$ выполняется неравенство

$$\left|\int_{A_1}^{A_2} g(x)\,dx\right| < \varepsilon. \tag{9.1.5}$$

Согласно известным неравенствам для интегралов и неравенству (9.1.4) получим $\left|\int\limits_{A_1}^{A_2} f(x)\,dx\right| \leqslant \int\limits_{A_1}^{A_2} |f(x)|\,dx \leqslant \int\limits_{A_1}^{A_2} g(x)\,dx$. Отсюда и из неравенства (9.1.5) вытекает, что для любых $A_1 > B$ и $A_2 > B$ справедливо неравенство $\left|\int\limits_{A_1}^{A_2} f(x)\,dx\right| < \varepsilon$. Следовательно, интеграл $\int\limits_a^{+\infty} f(x)\,dx$ сходится.

Утверждение 3 (частный признак сравнения). *Пусть на полупрямой $0 < a \leqslant x < +\infty$ функция $f(x)$ удовлетворяет соотношению $|f(x)| \leqslant c/x^\lambda$, где c и λ — постоянные, $\lambda > 1$. Тогда интеграл $\int\limits_a^{+\infty} f(x)\,dx$ сходится. Если же существует такая постоянная $c > 0$, что на полупрямой $0 < a \leqslant x < +\infty$ справедливо соотношение $f(x) \geqslant c/x^\lambda$, в котором $\lambda \leqslant 1$, то интеграл $\int\limits_a^{+\infty} f(x)\,dx$ расходится.*

Утверждение этой теоремы вытекает из утверждения 2 и примера, рассмотренного в предыдущем пункте (достаточно положить $g(x) = c/x^\lambda$).

Следствие (частный признак сравнения в предельной форме). *Если при $\lambda > 1$ существует конечный предел $\lim\limits_{x\to+\infty} |f(x)|\,x^\lambda = c$, то интеграл $\int\limits_a^{+\infty} f(x)\,dx$ сходится. Если же при $\lambda \leqslant 1$ существует положительный предел $\lim\limits_{x\to+\infty} f(x)\,x^\lambda = c > 0$, то интеграл $\int\limits_a^{+\infty} f(x)\,dx$ расходится.*

Убедимся в справедливости первой части следствия. Для этого заметим, что из существования предела при $x \to +\infty$ вытекает ограниченность функции $|f(x)|\,x^\lambda$, т. е. с некоторой постоянной $c_0 > 0$ выполняется неравенство $|f(x)| \leqslant c_0/x^\lambda$.

После этого применяется первая часть утверждения 3. Справедливость второй части следствия вытекает из следующих рассуждений. Так как $c > 0$, то можно указать столь малое $\varepsilon > 0$, что $c - \varepsilon > 0$. Этому ε отвечает такое $B > 0$, что при $x \geqslant B$ выполняется неравенство $c - \varepsilon < f(x)\,x^\lambda$ (это неравенство следует из определения предела). Поэтому $f(x) > (c - \varepsilon)/x^\lambda$, и в этом случае действует вторая часть утверждения 3.

3. Абсолютная и условная сходимость несобственных интегралов. Введем понятия абсолютной и условной сходимости интегралов. Пусть $f(x)$ интегрируема по любому сегменту $[a, A]$ *.

Определение 1. *Несобственный интеграл* $\int\limits_a^{+\infty} f(x)\,dx$ *называется* ***абсолютно сходящимся,*** *если сходится* $\int\limits_a^{+\infty} |f(x)|\,dx$.

Определение 2. *Несобственный интеграл* $\int\limits_a^{+\infty} f(x)\,dx$ *называется* ***условно сходящимся,*** *если он сходится, а интеграл* $\int\limits_a^{+\infty} |f(x)|\,dx$ *расходится.*

Замечание. Положив в утверждении 2 $g(x)=|f(x)|$, мы получим, что из абсолютной сходимости несобственного интеграла вытекает его сходимость.

Отметим, что утверждения 2 и 3 позволяют установить лишь абсолютную сходимость исследуемых несобственных интегралов.

Приведем еще один признак сходимости несобственного интеграла первого рода, пригодный для установления и условной сходимости этого интеграла.

Утверждение 4 (признак Дирихле—Абеля). *Пусть выполнены следующие три условия:*

1) *функция $f(x)$ непрерывна на полупрямой $a\leqslant x<+\infty$ и имеет на этой полупрямой ограниченную первообразную $F(x)$* **;

2) *функция $g(x)$ определена и монотонно не возрастает на полупрямой $a\leqslant x<+\infty$ и имеет равный нулю предел при $x\to+\infty$;*

3) *производная $g'(x)$ функции $g(x)$ существует и непрерывна в каждой точке полупрямой $a\leqslant x<+\infty$. При выполнении этих трех условий несобственный интеграл*

$$\int\limits_a^{+\infty} f(x)\,g(x)\,dx \qquad (9.1.6)$$

сходится.

Доказательство. Воспользуемся критерием Коши сходимости несобственных интегралов. Предварительно проведем

---

* Тогда и функция $|f(x)|$ интегрируема по любому сегменту $[a, A]$.

** Это означает, что первообразная $F(x)$, которую можно определить как $\int\limits_a^x f(t)\,dt$, удовлетворяет для всех $x\geqslant a$ неравенству $|F(x)|\leqslant K$, где K — постоянная.

интегрирование по частям интеграла $\int\limits_{A_1}^{A_2} f(x)\,dx$ на произвольном сегменте $[A_1, A_2]$, $A_2>A_1$ полупрямой $a\leqslant x<+\infty$. Получим

$$\int\limits_{A_1}^{A_2} f(x)\,g(x)\,dx = F(x)\,g(x)\Big|_{A_1}^{A_2} - \int\limits_{A_1}^{A_2} F(x)\,g'(x)\,dx. \qquad (9.1.7)$$

По условию теоремы $F(x)$ ограничена: $|F(x)|\leqslant K$. Так как $g(x)$ не возрастает и стремится к нулю при $x\to+\infty$, то $g(x)\geqslant 0$, а $g'(x)\leqslant 0$. Таким образом, оценивая (9.1.7), получим

$$\left|\int\limits_{A_1}^{A_2} f(x)\,g(x)\,dx\right| \leqslant K[g(A_1)+g(A_2)] + K\int\limits_{A_1}^{A_2} (-g'(x))\,dx.$$

Так как интеграл в правой части этого неравенства равен $g(A_1)-g(A_2)$, то, очевидно,

$$\left|\int\limits_{A_1}^{A_2} f(x)\,g(x)\,dx\right| \leqslant 2Kg(A_1). \qquad (9.1.8)$$

Используя это неравенство, нетрудно завершить доказательство теоремы. Пусть ε — произвольное положительное число. Так как $g(x)\to 0$ при $x\to+\infty$, то по данному ε можно выбрать B так, что при $A_1\geqslant B$ выполняется неравенство $g(A_1)<\varepsilon/(2K)$. Отсюда и из неравенства (9.1.8) следует, что для любых A_1 и A_2, бо́льших B, выполняется неравенство $\left|\int\limits_{A_1}^{A_2} f(x)\,g(x)\,dx\right|<\varepsilon$, которое, согласно критерию Коши, гарантирует сходимость интеграла (9.1.6). Утверждение доказано.

Замечание. Требование 3) в утверждении 4 является лишним и обусловлено лишь методом доказательства (желанием применить формулу интегрирования по частям). Для того чтобы доказать утверждение 4 без требования 3), достаточно применить для оценки интеграла $\int\limits_{A_1}^{A_2} f(x)\,g(x)\,dx$ вторую формулу среднего значения (см. свойство п. 2 § 4 гл. 9). Убедитесь в этом сами.

Пример 1. Рассмотрим интеграл

$$\int\limits_{1}^{+\infty} \frac{\sin x}{x^{\alpha}}\,dx, \quad \alpha>0. \qquad (9.1.9)$$

Полагая $f(x)=\sin x$, $g(x)=1/x^{\alpha}$, легко убедиться что для этого интеграла выполнены все условия утверждения 4. Поэтому интеграл (9.1.9) сходится *.

Пример 2. Рассмотрим интеграл Френеля $\int\limits_0^{+\infty} \sin x^2 dx$. Согласно п. 1 этого дополнения его сходимость вытекает из сходимости интеграла $\int\limits_1^{+\infty} \sin x^2 dx = \int\limits_1^{+\infty} x \sin x^2 \frac{1}{x}\, dx$. Полагая $f(x) = x\sin x^2$, $g(x)=1/x$, легко убедимся, что выполнены все условия утверждения 4. Поэтому интеграл Френеля сходится.

4. Замена переменных под знаком несобственного интеграла и формула интегрирования по частям. В этом пункте мы сформулируем условия, при которых действуют формулы замены переменных и интегрирования по частям для несобственных интегралов первого рода. Рассмотрим сначала вопрос о замене переменной под знаком несобственного интеграла.

Мы будем предполагать выполненными следующие условия:

1) *функция $f(x)$ непрерывна на полуоси $a \leqslant x < +\infty$;*

2) *полуось $a \leqslant x < +\infty$ является множеством значений некоторой строго монотонной функции $x=g(t)$, заданной на полуоси $\alpha \leqslant t < +\infty$ (или $-\infty < t \leqslant \alpha$) и имеющей на этой полуоси непрерывную производную;*

3) $g(\alpha)=a$.

При этих условиях из сходимости одного из следующих несобственных интегралов

$$\int\limits_a^{+\infty} f(x)\,dx \quad \text{и} \quad \int\limits_{\alpha}^{+\infty} f(g(t))\,g'(t)\,dt \left(\text{или } -\int\limits_{-\infty}^{\alpha} f(g(t))\,g'(t)\,dt\right) \qquad (9.1.10)$$

вытекает сходимость другого и равенство этих интегралов.

Сформулированное утверждение устанавливается с помощью следующих рассуждений.

Рассмотрим произвольный сегмент $[a, A]$. Этому сегменту отвечает, согласно строгой монотонности функции $g(t)$, сегмент $[\alpha, \rho]$ (или $[\rho, \alpha]$) оси t такой, что при изменении на сегменте $[\alpha, \rho]$ значения функции $x=g(t)$ заполняют сегмент $[a, A]$, причем $g(\rho)=A$. Таким образом, для указанных сегментов выполне-

---

* Так как $|\sin x| \geqslant \sin^2 x = \frac{1}{2} - \frac{1}{2}\cos 2x$ и интеграл $\int\limits_1^{+\infty} \frac{dx}{2x^{\alpha}}$ при $0 \leqslant \alpha \leqslant 1$ расходится (см. пример 2 из п. 1 § 1), а интеграл $\int\limits_1^{+\infty} \frac{\cos 2x}{2x^{\alpha}}\, dx$ сходится (в силу утверждения 4 при $f(x)=\cos 2x$, $g(x)=1/(2x^{\alpha})$), то при $0<\alpha\leqslant 1$ интеграл (9.1.9) сходится условно.

ны все условия п. 3 § 5 гл. 9, при которых действует формула замены переменной под знаком определенного интеграла. Поэтому имеет место равенство

$$\int_a^A f(x)\,dx = \int_\alpha^\rho f(g(t))\,g'(t)\,dt \left(\text{или} \quad = -\int_\rho^\alpha f(g(t))\,g'(t)\,dt\right). \quad (9.1.11)$$

В силу строгой монотонности функции $x=g(t)$, $A\to+\infty$ при $\rho\to+\infty$ и, обратно, $\rho\to+\infty$ при $A\to+\infty$ (или $A\to+\infty$ при $\rho\to-\infty$ и $\rho\to-\infty$ при $A\to+\infty$). Поэтому из формулы (9.1.11) вытекает справедливость сформулированного выше утверждения.

Перейдем теперь к вопросу об интегрировании по частям несобственных интегралов первого рода.

Докажем следующее

Утверждение. *Пусть функции $u(x)$ и $v(x)$ имеют непрерывные производные на полупрямой $a\leqslant x<+\infty$ и, кроме того, существует предельное значение* $\lim\limits_{x\to+\infty}[u(x)\,v(x)]=L$. *При этих условиях из сходимости одного из интегралов*

$$\int_a^{+\infty} u(x)\,v'(x)\,dx \quad \text{и} \quad \int_a^{+\infty} v(x)\,u'(x)\,dx \qquad (9.1.12)$$

вытекает сходимость другого из этих интегралов и справедливость формулы

$$\int_a^{+\infty} u(x)\,v'(x)\,dx = L - u(a)\,v(a) - \int_a^{+\infty} v(x)\,u'(x)\,dx. \qquad (9.1.13)$$

Для доказательства сформулированного утверждения рассмотрим произвольный сегмент $[a, A]$. На этом сегменте действует обычная формула интегрирования по частям. Поэтому

$$\int_a^A u(x)\,v'(x)\,dx = [u(x)\,v(x)]_a^A - \int_a^A v(x)\,u'(x)\,dx.$$

Так как при $A\to+\infty$ выражение $[u(x)v(x)]_a{}^A$ стремится к $L-u(a)v(a)$, то из последнего равенства следует одновременная сходимость или расходимость интегралов (9.1.12) и справедливость формулы (9.1.13) в случае сходимости одного из интегралов (9.1.12).

§ 2. Несобственные интегралы второго рода

В этом параграфе будет дано обобщение понятия определенного интеграла на случай неограниченных функций.

Пусть на полусегменте $[a, b)$ задана функция $f(x)$. Точку b мы будем называть *особой*, если функция не ограничена на полусегменте $[a, b)$, но ограничена на любом сегменте $[a, b-\alpha]$ $\alpha>0$, заключенном в полусегменте $[a, b)$. Будем также предполагать, что на любом таком сегменте функция $f(x)$ интегрируема. При наших предположениях на полусегменте $(0, b-a)$ задана функция аргумента α

$$F(\alpha)=\int_a^{b-\alpha} f(x)\,dx.$$

Исследуем вопрос о правом пределе функции $F(\alpha)$ в точке $\alpha=0$:

$$\lim_{\alpha\to+0}\int_a^{b-\alpha} f(x)\,dx. \tag{9.1.14}$$

Определение. *Правый предел (9.1.14) в случае, если он существует, называется несобственным интегралом второго рода от функции $f(x)$ по сегменту $[a, b]$ и обозначается символом*

$$\int_a^b f(x)\,dx. \tag{9.1.15}$$

При этом говорят, что несобственный интеграл (9.1.15) сходится и пишут равенство $\int\limits_a^b f(x)dx=\lim\limits_{\alpha\to+0}\int\limits_a^{b-\alpha} f(x)\,dx.$

Символ (9.1.15) употребляют и в случае, если указанного выше предела (9.1.14) не существует, но в этом случае говорят, что несобственный интеграл (9.1.15) *расходится.*

Замечание. Понятие несобственного интеграла второго рода легко переносится на случай, когда функция $f(x)$ имеет конечное число особых точек.

Пример. Рассмотрим на полусегменте $[a, b)$ функцию $f(x)=1/(b-x)^\lambda$, $\lambda>0$. Ясно, что точка b является особой точкой для этой функции. Кроме того, очевидно, что функция интегрируема на любом сегменте $[a, b-\alpha]$, $\alpha>0$, причем

$$\int_a^{b-\alpha}\frac{dx}{(b-x)^\lambda}=\begin{cases} -\dfrac{(b-x)^{1-\lambda}}{1-\lambda}\Big|_a^{b-\alpha}=\dfrac{(b-a)^{1-\lambda}-\alpha^{1-\lambda}}{1-\lambda} & \text{при } \lambda\neq 1,\\ -\ln(b-x)\big|_a^{b-\alpha}=\ln(b-a)-\ln\alpha & \text{при } \lambda=1.\end{cases}$$

Очевидно, предел $\lim\limits_{\alpha\to+0}\int\limits_a^{b-\alpha}\dfrac{dx}{(b-x)^\lambda}$ существует и равен $\dfrac{(b-a)^{1-\lambda}}{1-\lambda}$ при $\lambda<1$ и не существует при $\lambda\geqslant1$. Следовательно, рассматриваемый несобственный интеграл сходится при $\pi<1$ и расходится при $\lambda\geqslant1$.

Сформулируем *критерий Коши* сходимости несобственного интеграла второго рода. При этом мы будем предполагать, что функция $f(x)$ задана на полусегменте $[a, b)$ и b — особая точка этой функции.

Утверждение 5 (критерий Коши). *Для сходимости несобственного интеграла второго рода* (9.1.15) *необходимо и достаточно, чтобы для любого* $\varepsilon>0$ *можно было указать такое* $\delta>0$, *что для любых* α' *и* α'', *удовлетворяющих условию* $0<\alpha''<\alpha'<\delta$, *справедливо неравенство*

$$\left|\int_{b-\alpha'}^{b-\alpha''} f(x)\,dx\right| < \varepsilon.$$

Справедливость этой теоремы вытекает из того, что понятие сходимости интеграла по определению эквивалентно понятию существования предельного значения функции $F(\alpha)$, введенной в начале этого пункта. Мы не будем подробно развивать теорию несобственных интегралов второго рода. Это объясняется тем, что основные выводы и теоремы предыдущего параграфа без труда могут быть перенесены на случай интегралов второго рода. Поэтому мы ограничимся некоторыми замечаниями.

1°. При некоторых ограничениях на подынтегральные функции интегралы второго рода сводятся к интегралам первого рода. Именно: пусть функция $f(x)$ непрерывна на полусегменте $[a, b)$ и b — особая точка этой функции. При этих условиях в интеграле $\int\limits_a^{b-\alpha} f(x)\,dx$, $\alpha>0$, мы можем произвести следующую замену переменных:

$$x=b-\frac{1}{t}, \quad dx=\frac{dt}{t^2}, \quad \frac{1}{b-a}\leqslant t\leqslant\frac{1}{\alpha}.$$

В результате этой замены переменных мы получим равенство

$$\int_a^{b-\alpha} f(x)\,dx=\int_{1/(b-a)}^{1/\alpha} f\left(b-\frac{1}{t}\right)\frac{1}{t^2}\,dt. \qquad (9.1.16)$$

Пусть интеграл $\int\limits_a^b f(x)\,dx$ сходится. Это означает, что существует предел $\lim\limits_{\alpha\to+0}\int\limits_a^{b-\alpha} f(x)\,dx$. Обращаясь к равенству (9.1.16), мы видим, что существует также и предел при $1/\alpha\to+\infty$ выражения в правой части (9.1.16). Тем самым доказана сходимость несобственного интеграла первого рода

$$\int\limits_{1/(b-a)}^{+\infty} f\left(b-\frac{1}{t}\right)\frac{1}{t^2}\,dt$$

и равенство этого интеграла интегралу $\int\limits_a^b f(x)\,dx$. Очевидно, сходимость только что указанного несобственного интеграла первого рода влечет сходимость интеграла $\int\limits_a^b f(x)\,dx$ и равенство этих интегралов. Итак, *из сходимости одного из интегралов*

$$\int\limits_a^b f(x)\,dx \quad \text{и} \quad \int\limits_{1/(b-a)}^{+\infty} f\left(b-\frac{1}{t}\right)\frac{1}{t^2}\,dt$$

следует сходимость другого и равенство этих интегралов.

2°. Для несобственных интегралов второго рода легко доказываются утверждения, аналогичные утверждениям п. 2 предыдущего параграфа, которые можно объединить общим наименованием «признаки сравнения». Отметим, что во всех формулировках функция $f(x)$ рассматривается на полусегменте $[a, b)$, где b — особая точка функции.

Частный признак сравнения будет иметь следующий вид.

Если $|f(x)| \leqslant c(b-x)^{-\lambda}$, где $\lambda<1$, то несобственный интеграл (9.1.15) сходится. Если же $f(x) \geqslant c(b-x)^{-\lambda}$, где $c>0$ и $\lambda \geqslant 1$, то несобственный интеграл (9.1.15) расходится. Доказательство вытекает из общего признака сравнения и примера, рассмотренного выше.

В полной аналогии с п. 4 предыдущего параграфа для несобственных интегралов второго рода формулируются правила интегрирования путем замены переменной и интегрирования по частям.

§ 3. Главное значение несобственного интеграла

Определение. *Пусть функция $f(x)$ определена на прямой $-\infty<x<+\infty$ и интегрируема на каждом сегменте, принадлежащем этой прямой. Будем говорить, что функция $f(x)$ интегрируема по Коши, если существует предел* $\lim\limits_{A\to+\infty}\int\limits_{-A}^{A} f(x)\,dx$.

Этот предел мы будем называть главным значением несобственного интеграла от функции $f(x)$ (в смысле Коши) и обозначать символом *

* V. p. — начальные буквы французских слов «Valeur principal», обозначающих «главное значение». Подчеркнем, что, в отличие от понятия несобственного

$$\text{V. p.} \int_{-\infty}^{+\infty} f(x)\,dx = \lim_{A\to+\infty} \int_{-A}^{A} f(x)\,dx.$$

Пример 1. Найдем главное значение интеграла от функции x. Поскольку в силу нечетности x,

$$\int_{-A}^{A} x\,dx = 0, \text{ то V. p.} \int_{-\infty}^{+\infty} x\,dx = 0.$$

Точно так же заключаем, что $\text{V. p.} \int_{-\infty}^{+\infty} \sin x\,dx = 0.$

Справедливо следующее

Утверждение. *Пусть функция $f(x)$ интегрируема на каждом сегменте прямой $-\infty<x<+\infty$. Если эта функция $f(x)$ нечетна, то она интегрируема по Коши и главное значение интеграла от нее равняется нулю.*

Если функция $f(x)$ четна, то она интегрируема по Коши тогда и только тогда, когда сходится несобственный интеграл

$$\int_{0}^{+\infty} f(x)\,dx. \tag{9.1.17}$$

Первая часть этого утверждения является очевидной. Для доказательства второй части достаточно воспользоваться равенством $\int_{A}^{A} f(x)\,dx = 2\int_{0}^{A} f(x)\,dx$, справедливым для любой четной функции, и определением сходимости несобственного интеграла (9.1.17).

Понятие интегрируемости по Коши можно ввести и для несобственных интегралов второго рода в случае, когда особая точка является внутренней точкой сегмента, по которому производится интегрирование.

Определение. *Пусть функция $f(x)$ определена на сегменте $[a, b]$, кроме, быть может, точки c, $a<c<b$, и интегрируема на любом сегменте, принадлежащем либо $[a, c)$, либо $(c, b]$. Будем говорить, что функция $f(x)$ интегрируема по Коши, если существует предел*

интеграла $\int_{-\infty}^{\infty} f(x)\,dx$, определяемого как предел $\lim_{A'\to-\infty,\ A''\to+\infty} \int_{A'}^{A''} f(x)dx$ при независимом стремлении A' к $-\infty$, A'' к $+\infty$, интеграл по Коши определяется как предел при $A\to+\infty$ интеграла в симметричных пределах $\int_{-A}^{A} f(x)dx$.

$$\lim_{\alpha\to+0}\left(\int_a^{c-\alpha} f(x)\,dx+\int_{c+\alpha}^{b} f(x)\,dx\right)=\text{V. p.}\int_a^b f(x)\,dx,\quad \alpha>0,$$

называемый главным значением интеграла в смысле Коши.

Пример 2. Функция $1/(x-c)$ не интегрируема на сегменте $[a, b]$, $a<c<b$, в несобственном смысле, однако она интегрируема по Коши. При этом

$$\text{V. p.}\int_a^b \frac{dx}{x-c}=\lim_{\alpha\to+0}\left(\int_a^{c-\alpha}\frac{dx}{x-c}+\int_{c+\alpha}^{b}\frac{dx}{x-c}\right)=\ln\frac{b-c}{c-a}.$$

ДОПОЛНЕНИЕ 2

Интеграл Стилтьеса *

Понятие интеграла Стилтьеса является непосредственным обобщением понятия интеграла Римана, построению которого была посвящена гл. 9. В настоящем дополнении мы приведем основные сведения об интеграле Стилтьеса.

1. Определение интеграла Стилтьеса и условия его существования. Понятие интеграла Стилтьеса реализует идею интегрирования функции $f(x)$ относительно другой функции $u(x)$.

Пусть функции $f(x)$ и $u(x)$ определены и ограничены на сегменте $[a, b]$ и $\{x_k\}$ — разбиение этого сегмента **:

$$a=x_0<x_1<\ldots<x_{i-1}<x_i<\ldots<x_n=b.$$

Сумму вида

$$\sigma=f(\xi_1)[u(x_1)-u(x_0)]+\ldots+f(\xi_i)[u(x_i)-u(x_{i-1})]+\ldots$$
$$\ldots+f(\xi_n)[u(x_n)-u(x_{n-1})], \tag{9.2.1}$$

где $x_{i-1}\leqslant\xi_i\leqslant x_i$, $i=1, 2, \ldots, n$, называют **интегральной суммой Стилтьеса**.

Число I называют **пределом** интегральных сумм (9.2.1) при $\max\limits_{1\leqslant i\leqslant n}\Delta x_i\to 0$, если для любого $\varepsilon>0$ найдется $\delta>0$ такое, что при $\max\limits_{1\leqslant i\leqslant n}\Delta x_i<\delta$ справедливо неравенство $|\sigma-I|<\varepsilon$.

Определение. *Функция $f(x)$ называется **интегрируемой по функции** $u(x)$ на сегменте $[a, b]$, если существует конечный предел интегральных сумм* (9.2.1) *при* $\max\limits_{1\leqslant i\leqslant n}\Delta x_i\to 0$. *Указанный предел называется **интегралом***

* Т. Стилтьес — голландский математик (1856—1894).

** Мы предполагаем, что $a<b$. Случай $a>b$ сводится к рассматриваемому.

Стилтьеса (или интегралом Римана—Стилтьеса) от функции $f(x)$ по функции $u(x)$ на сегменте $[a, b]$ и обозначается символом

$$I = \int_a^b f(x)\, du(x). \tag{9.2.2}$$

Функцию $u(x)$ иногда называют интегрирующей функцией.

Т. Стилтьес пришел к идее такого интеграла, рассматривая положительное «распределение масс» на прямой, заданное возрастающей функцией $u(x)$, точки разрыва которой соответствуют массам, «сконцентрированным в одной точке».

Интеграл Римана представляет собой частный случай интеграла Стилтьеса, когда в качестве интегрирующей функции взята функция $x+c$, где $c=\text{const}$.

Укажем *ряд условий существования интеграла Стилтьеса* (т. е. условий, когда функция $f(x)$ интегрируема по функции $u(x)$).

Предположим, что интегрирующая функция $u(x)$ является возрастающей. Отсюда следует, что, поскольку $a=x_0<x_1<\ldots<x_{i-1}<x_i<\ldots<x_n=b$, все $\Delta u(x_i)=u(x_i)-u(x_{i-1})>0$. Это позволяет, заменяя Δx_i на $\Delta u(x_i)$, повторить почти все построения, проводимые для интеграла Римана.

Аналогично суммам Дарбу для обычного интеграла Римана вводятся верхняя и нижняя суммы Дарбу—Стилтьеса

$$S=\sum_{i=1}^{n} M_i\,[u(x_i)-u(x_{i-1})],\quad s=\sum_{i=1}^{n} m_i\,[u(x_i)-u(x_{i-1})], \tag{9.2.3}$$

где M_i и m_i — точные верхняя и нижняя грани функции $f(x)$ на сегменте $[x_{i-1}, x_i]$.

Суммы (9.2.3) *называются соответственно* ***верхней*** *и* ***нижней*** *суммами Дарбу—Стилтьеса.*

Как и в случае сумм Дарбу (т. е. в простейшем случае $u(x)=x+c$, $c=\text{const}$) при одном и том же разбиении выполнены неравенства $s\leqslant\sigma\leqslant S$, причем s и S служат точными гранями для стилтьесовых сумм σ, отвечающих всевозможным выборам промежуточных точек ξ_i на частичных сегментах.

Суммы Дарбу—Стилтьеса обладают (как и в простейшем случае) свойствами:

а) *если к имеющимся точкам разбиения добавить новые точки, то нижняя сумма Дарбу—Стилтьеса может от этого лишь возрасти, а верхняя сумма — лишь уменьшиться;*

б) *каждая нижняя сумма Дарбу—Стилтьеса не превосходит любой верхней суммы, отвечающей тому же или другому разбиению сегмента* $[a, b]$.

Аналогично тому, как это сделано при построении интеграла Римана, вводятся верхний и нижний интегралы Дарбу—Стилтьеса:

$$I^* = \inf\{S\}, \qquad I_* = \sup\{s\},$$

где нижняя и верхняя грани берутся по всевозможным разбиениям сегмента $[a, b]$.

Легко проверить, что справедливы соотношения

$$s \leqslant I_* \leqslant I^* \leqslant S.$$

Точно так же, как и в случае обычного **интеграла Римана**, в случае интеграла Стилтьеса доказывается, что верхний интеграл Дарбу—Стилтьеса является пределом верхних сумм S при стремлении диаметра разбиений к нулю. Аналогично нижний интеграл Дарбу—Стилтьеса есть предел нижних сумм s (см. п. 2, § 2, основную лемму Дарбу).

Сформулируем теперь теорему, которая является обобщением основной теоремы п. 1 § 3 и справедлива в случае интеграла Римана—Стилтьеса.

Основная теорема. *Для того чтобы ограниченная на сегменте $[a, b]$ $(a<b)$ функция $f(x)$ была интегрируемой на этом сегменте по возрастающей функции $u(x)$, необходимо и достаточно, чтобы для любого $\varepsilon>0$ нашлось такое разбиение $\{x_k\}$ сегмента $[a, b]$, для которого $S-s<\varepsilon$.*

Доказательство этой теоремы (как, впрочем, и других упомянутых выше фактов и свойств) является дословным повторением рассуждений, проведенных для интеграла Римана.

Укажем теперь некоторые классы интегрируемых по Риману — Стилтьесу функций.

1. *Если функция $f(x)$ непрерывна, а $u(x)$ возрастает на сегменте $[a, b]$, то интеграл Стилтьеса $\int\limits_a^b f(x)\,du(x)$ существует.*

Доказательство этого факта полностью аналогично доказательству теоремы 9.1 (см. п. 2 § 3).

Замечание. Указанный выше факт справедлив и в том случае, когда функция $u(x)$ является функцией ограниченной вариации*. Так называются функции $u(x)$, определенные на сегменте $[a, b]$, $a<b$ и обладающие тем свойством, что для любого разбиения

$$x_0 = a < x_1 < x_2 < \ldots < x_i < x_{i+1} < \ldots < x_n = b$$

числовое множество $V[\{x_k\}] = \sum\limits_{i=1}^{n-1} |u(x_{i+1}) - u(x_i)|$ ограничено сверху.

* Или ограниченного изменения.

Точная верхняя грань множества $\{V[\{x_k\}]\}$ называется полным изменением или полной вариацией функции $u(x)$ на сегменте $[a, b]$ и обозначается символом $\overset{b}{\underset{a}{V}} u(x) = \sup\{V[\{x_k\}]\}$. Для функций ограниченной вариации справедлив следующий основной критерий:

Для того чтобы функция $u(x)$ имела на сегменте $[a, b]$ ограниченную вариацию, необходимо и достаточно, чтобы она представлялась на этом сегменте в виде разности двух возрастающих и ограниченных функций:

$$u(x) = g(x) - h(x).$$

Таким образом, в случае, когда $u(x)$ — функция ограниченной вариации, сумму Стилтьеса, отвечающую функции $u(x)$, можно записать в виде

$$\sigma = \sum_{i=1}^{n} f(\xi_i)\,\Delta u(x_i) = \sum_{i=1}^{n} f(\xi_i)\,\Delta g(x_i) - \sum_{i=1}^{n} f(\xi_i)\,\Delta h(x_i) = \sigma_1 - \sigma_2,$$

где $\Delta u(x_i) = u(x_i) - u(x_{i-1})$, $\Delta g(x_i) = g(x_i) - g(x_{i-1})$, $\Delta h(x_i) = h(x_i) - h(x_{i-1})$.

Суммы σ_1 и σ_2 стремятся к конечным пределам при стремлении диаметра разбиений к нулю, так как $g(x)$ и $h(x)$ — возрастающие функции. Поэтому существует конечный предел и сумм σ при стремлении диаметра разбиений к нулю.

Следовательно, теорию интеграла Стилтьеса можно строить и в случае, когда интегрирующая функция $u(x)$ имеет ограниченную вариацию, вполне аналогично случаю возрастающей функции $u(x)$.

Выделим еще один класс функций, для которых существует интеграл Стилтьеса.

2. *Интеграл Стилтьеса* (9.2.2) *существует при условии, что функция $f(x)$ интегрируема на сегменте $[a, b]$ по Риману, а функция $u(x)$ удовлетворяет на этом сегменте условию Липшица, т. е. условию*

$$|u(x') - u(x'')| \leqslant c\,|x' - x''|,$$

где $c = \text{const}$, для любых x' и x'' из $[a, b]$.

Так как функция, удовлетворяющая условию Липшица, является функцией с ограниченной вариацией, то для доказательства этого критерия, очевидно, достаточно рассмотреть лишь случай возрастающей функции $u(x)$, удовлетворяющей условию Липшица, и заметить, что

$$S - s = \sum_{i=1}^{n} (M_i - m_i)\,\Delta u(x_i) \leqslant C \sum_{i=1}^{n} (M_i - m_i)\,\Delta x_i, \tag{9.2.4}$$

где $M_i=\sup f(x)$, $m_i=\inf f(x)$, $x\in[x_{i-1}, x_i]$, $\Delta u(x_i)=u(x_i)-u(x_{i-1})$, c — постоянная из условия Липшица. Выражение $\sum\limits_{i=1}^{n}(M_i-m_i)\Delta x_i$ в неравенстве (9.2.4) может быть сделано, в силу интегрируемости по Риману функции $f(x)$, сколь угодно малой величиной за счет выбора разбиения сегмента $[a, b]$. Следовательно, величина $S-s$ также может быть сделана меньше наперед заданного числа $\varepsilon>0$, если выбрать диаметр разбиения достаточно малым.

Согласно утверждению основной теоремы функция $f(x)$ интегрируема по Стилтьесу.

В общем случае функции $u(x)$, удовлетворяющей условию Липшица, также можно рассмотреть представление

$$u(x)=cx-[cx-u(x)]=u_1(x)-u_2(x).$$

В этом представлении обе функции $u_1(x)$ и $u_2(x)$ удовлетворяют условию Липшица и возрастают*. В таком случае доказательство завершается так же, как и выше.

Укажем, наконец, еще один класс интегрируемых по Стилтьесу функций.

3. *Если функция $f(x)$ интегрируема на сегменте $[a, b]$ по Риману, а функция $u(x)$ допускает представление в виде интеграла с переменным верхним пределом*

$$u(x)=A+\int_a^x \varphi(\xi)\,d\xi,$$

где $\varphi(\xi)$ — интегрируемая на сегменте $[a, b]$ по Риману функция, то интеграл (9.2.2) существует.

Действительно, так как $\varphi(\xi)$ интегрируема по Риману, то она ограничена: $|\varphi(\xi)|\leqslant K=\text{const}$. Следовательно,

$$|u(x')-u(x'')|=\left|\int_{x'}^{x''}\varphi(\xi)\,d\xi\right|\leqslant K|x'-x''|.$$

Поэтому справедливость этого критерия вытекает из справедливости предыдущего.

Заметим, что в ряде случаев интеграл Стилтьеса $I=\int\limits_a^b f(x)\,du(x)$ сводится к интегралу Римана по формуле

$$\int_a^b f(x)\,du(x)=\int_a^b f(x)\,\varphi(x)\,dx. \qquad (9.2.5)$$

---

* Для функции $u_2(x)$ при $x'>x''$, очевидно, можно записать соотношение $u_2(x')-u_2(x'')\geqslant c(x'-x'')+[u(x')-u(x'')]\geqslant 0$.

В частности, равенство (9.2.5) имеет место в случае, если $u(x)$ имеет ограниченную и интегрируемую в смысле Римана на сегменте $[a, b]$ производную $u'(x)$. В этом случае $\varphi(x) = u'(x)$.

2. Свойства интеграла Стилтьеса. Сформулируем ряд свойств интеграла Стилтьеса, непосредственно вытекающих из его определения.

а) *Линейное свойство как относительно интегрируемой, так и относительно интегрирующей функции (при условии существования каждого из интегралов Стилтьеса в правой части)*:

$$\int_a^b (\alpha f_1 + \beta f_2)\, du = \alpha \int_a^b f_1 du + \beta \int_a^b f_2 du,$$

$$\int_a^b f d[\alpha u_1 + \beta u_2] = \alpha \int_a^b f du_1 + \beta \int_a^b f du_2,$$

здесь α, β — произвольные числа.

б) *Если выполнено условие $a<c<b$, то*

$$\int_a^b f(x)\, du(x) = \int_a^c f(x)\, du(x) + \int_c^b f(x)\, du(x),$$

в предположении, что существуют все три интеграла.

Подчеркнем, что из существования обоих интегралов $\int_a^c f(x)\, du(x)$ и $\int_c^b f(x)\, du(x)$, вообще говоря, не вытекает существование интеграла $\int_a^b f(x)\, du(x)$. Вот соответствующий пример:

$$f(x) = \begin{cases} 0, & \text{если } -1 \leqslant x \leqslant 0, \\ A, & \text{если } 0 < x \leqslant 1, \ A \neq 0; \end{cases}$$

$$u(x) = \begin{cases} 0, & \text{если } -1 \leqslant x < 0, \\ B, & \text{если } 0 \leqslant x \leqslant 1, \ B \neq 0. \end{cases}$$

Интегралы $\int_{-1}^0 f(x)\, du(x)$, $\int_0^1 f(x)\, du(x)$ оба существуют и равны нулю, так как соответствующие им суммы Стилтьеса все равны нулю. Действительно, в первом интеграле $f(x) = 0$, $-1 \leqslant x \leqslant 0$, во втором $\Delta u(x_i) = u(x_i) - u(x_{i-1}) = 0$ для любого разбиения сегмента $[-1, 1]$. Однако интеграл $\int_{-1}^1 f(x)\, du(x)$

не существует. В самом деле, пусть $\{x_k\}$ — разбиение сегмента $[-1, 1]$, не содержащее в качестве точки разбиения точку 0. Тогда в сумме Стилтьеса $\sigma=\sum\limits_{i=1}^{n} f(\xi_i)\,\Delta u(x_i)$ остается лишь одно слагаемое, а именно слагаемое

$$f(\xi_k)\,[u(x_k)-u(x_{k-1})]=Bf(\xi_k), \qquad B\neq 0,$$

для которого точка нуль содержится в сегменте $[x_{k-1}, x_k]$. В зависимости от того, будет ли ξ_k удовлетворять условию $\xi_k<0$ или $\xi_k>0$, мы получим, что $\sigma=0$ или $\sigma=A\cdot B\neq 0$, так что σ не имеет предела при стремлении диаметра разбиений к нулю.

Указанный факт связан с тем, что как у функции $f(x)$, так и у функции $u(x)$ точка 0 является точкой разрыва.

в) *Для интеграла Стилтьеса* (9.2.2) *справедлива формула среднего значения.*

Пусть функция $f(x)$ ограничена на сегменте $[a, b]$, так что $m\leqslant f(x)\leqslant M$, а функция $u(x)$ возрастает на этом сегменте. Тогда найдется такое число μ, удовлетворяющее неравенствам $m\leqslant\mu\leqslant M$, что для интеграла Стилтьеса справедлива формула среднего значения

$$\int\limits_a^b f(x)\,du(x)=\mu\,[u(b)-u(a)].$$

В частности, если дополнительно предположить непрерывность $f(x)$ на сегменте $[a, b]$, то найдется точка $\xi\in[a, b]$ такая, что $\mu=f(\xi)$.

Доказательство этой формулы вполне аналогично доказательству формулы среднего значения для интеграла Римана (см. п. 2, § 4).

Глава 10

ГЕОМЕТРИЧЕСКИЕ ПРИЛОЖЕНИЯ ОПРЕДЕЛЕННОГО ИНТЕГРАЛА

§ 1. ДЛИНА ДУГИ КРИВОЙ

1. Понятие простой кривой. Начнем наше рассмотрение с выяснения понятия кривой. Пусть на сегменте $[\alpha, \beta]$ заданы две непрерывные функции $\varphi(t)$ и $\psi(t)$. Аргумент этих функций будет в дальнейшем называться параметром. Рассмотрим плоскость (x, y), т. е. совокупность всевозможных упорядоченных пар (x, y) чисел x и y. Каждая такая пара называется точкой плоскости, а числа x и y называются координатами этой точки. Точка (x, y) может обозначаться также одной буквой M, при этом запись $M(x, y)$ означает, что точка M имеет координаты x и y.

Если рассматривать параметр t как время, то уравнения

$$x=\varphi(t), \quad y=\psi(t), \quad \alpha \leqslant t \leqslant \beta \tag{10.1}$$

определяют закон движения точки M с координатами x и y на плоскости. Множество $\{M\}$ точек M, отвечающих всевозможным значениям параметра t из $[\alpha, \beta]$ можно рассматривать как след точки M, движущейся по закону (10.1).

В общем случае даже для непрерывных функций $\varphi(t)$ и $\psi(t)$ этот закон движения может быть очень сложным. Например, можно так подобрать непрерывные на сегменте $\alpha \leqslant t \leqslant \beta$ функции $\varphi(t)$ и $\psi(t)$, что при изменении параметра t на этом сегменте точки $\{M(x, y)\}$ заполняют целый квадрат *.

Введем понятие простой плоской кривой.

Определение. *Множество $\{M\}$ всех точек M, координаты x и y которых определяются уравнениями $x=\varphi(t)$, $y=\psi(t)$ при t из $[\alpha, \beta]$, называется простой плоской кривой L, если различным значениям параметра t из $[\alpha, \beta]$ отвечают разные точки множества $\{M\}$.*

Каждую точку множества $\{M\}$, определяющего простую плоскую кривую, называют точкой этой кривой, причем точки, отвечающие граничным значениям α и β параметра t, называются граничными точками простой кривой.

При этом говорят, что «уравнения (10.1) определяют простую плоскую кривую L» или «простая плоская кривая L параметризована при помощи уравнений (10.1)».

---

* См. Бибербах. Дифференциальное и интегральное исчисление, ч. 1, с. 156.

Примером простой кривой является график полуокружности радиуса r, лежащей в верхней полуплоскости с центром в начале координат:

$$x=r\cos t, \quad y=r\sin t, \quad \text{при} \quad 0\leqslant t\leqslant \pi.$$

Более общим примером простой кривой является график непрерывной на сегменте $[\alpha, \beta]$ функции $y=f(x)$; параметризацию этой кривой вводят по правилу

$$x=t, \quad y=f(t) \quad \text{при} \quad t\in[\alpha, \beta].$$

Заметим, что простые кривые не исчерпывают всего множества кривых, которые могут быть определены уравнениями (10.1).

В следующем пункте мы рассмотрим более общие кривые, определяемые этими уравнениями.

В заключение этого пункта сделаем два замечания.

Замечание 1. Одна и та же простая кривая L может быть параметризована различными способами. Целесообразно рассматривать только те параметризации, которые получаются из данной путем представления параметра t в виде непрерывных строго монотонных функций другого параметра s. При таких преобразованиях параметра сохраняется порядок следования точек на кривой L.

Замечание 2. *Пусть L_1 и L_2 — две простые кривые, причем граничные точки кривой L_1 совпадают с граничными точками кривой L_2, а любые не граничные точки кривых L_1 и L_2 различны. Кривая L, полученная объединением кривых L_1 и L_2, называется **простой замкнутой кривой**.*

2. Понятие параметризуемой кривой. В предыдущем пункте мы рассматривали простые кривые. Следует заметить, что в математическом анализе часто приходится рассматривать кривые, не являющиеся простыми, например кривые, имеющие точки самопересечения или целые участки самоналегания. В связи с этим возникает необходимость ввести в рассмотрение понятие так называемой параметризуемой кривой.

Будем считать, что множество $\{t\}$ представляет собой либо сегмент, либо полусегмент, либо интервал, либо числовую прямую, либо открытую или замкнутую полупрямую.

Введем понятие разбиения множества $\{t\}$. *Будем говорить, что конечная или бесконечная система сегментов $\{[t_{i-1}, t_i]\}$ **разбивает множество** $\{t\}$, если, во-первых, объединение всех этих сегментов представляет собой все множество $\{t\}$ и, во-вторых, общими точками любых двух сегментов системы могут быть лишь их концы.*

Рассмотрим примеры разбиений некоторых из указанных выше множеств $\{t\}$.

1) Система сегментов $[0, 1/2]$, $[1/2, 1]$, очевидно, разбивает сегмент $[0, 1]$.

2) Система сегментов $[n-1, n]$, где $n=1, 2, \ldots$, разбивает полупрямую $[0, \infty)$.

3) Система сегментов $[n-1, n]$, где n — любое целое число, очевидно, разбивает всю числовую прямую.

Пусть множество $\{t\}$ представляет собой одно из указанных выше множеств, а функции $\varphi(t)$ и $\psi(t)$ непрерывны на этом множестве. Введем понятие параметризуемой кривой.

Определение. *Будем говорить, что уравнения*

$$x=\varphi(t), \quad y=\psi(t) \tag{10.2}$$

задают параметризуемую кривую L, если существует такая система сегментов $\{[t_{i-1}, t_i]\}$, разбивающих множество $\{t\}$, что для значений t из каждого сегмента этой системы уравнения (10.2) *определяют простую кривую.*

Между точками кривой L можно ввести некоторое отношение порядка. Пусть точка M_1 соответствует значению параметра t_1, а точка M_2 — значению t_2.

Мы скажем, что точка M_1 *предшествует точке* M_2 (и запишем $M_1 \prec M_2$) если $t_1 < t_2$.

Заметим, что *точки, отвечающие различным значениям параметра, всегда считаются различными.*

Таким образом, параметризуемую кривую можно рассматривать как объединение простых кривых, причем эти простые кривые последовательно пробегаются точкой M, координаты которой определяются уравнениями (10.2), когда параметр t, монотонно возрастая, пробегает множество $\{t\}$.

Замечание 1. Простую кривую можно рассматривать как частный случай параметризуемой кривой. В этом случае система сегментов, разбивающих сегмент $[\alpha, \beta]$, состоит из одного сегмента, а именно из самого сегмента $[\alpha, \beta]$.

Замечание 2. Про параметризуемую кривую, определяемую уравнениями (10.2), говорят также, что эта кривая параметризована при помощи уравнений (10.2). Заметим, что одна и та же кривая L может быть параметризована *различными* способами. Мы будем рассматривать всевозможные параметризации кривой L, получающиеся из любой данной параметризации путем представления параметра t в виде непрерывных, строго возрастающих функций другого параметра s. Только при таких преобразованиях параметра сохраняется порядок следования точек на кривой L.

Замечание 3. Понятие *пространственной* кривой вводится в полной аналогии с понятием *плоской* кривой. Так же, как и для плоской простой кривой, *пространственная простая кривая — это множество $\{M\}$ точек пространства, координаты x, y, z которых определяются уравнениями*

$$x=\varphi(t), \quad y=\psi(t), \quad z=\chi(t), \quad \alpha \leqslant t \leqslant \beta, \tag{10.3}$$

при условии непрерывности функций $\varphi(t)$, $\psi(t)$, $\chi(t)$ *на сегменте* $[\alpha, \beta]$ *и при условии несовпадения точек множества* $\{M\}$, *отвечающих различным значениям параметра* t.

Используя понятие простой пространственной кривой и понятие разбиения множества $\{t\}$ изменения параметра, так же как и в плоском случае, дается определение параметризуемой пространственной кривой (задаваемой уравнениями (10.3)).

Приведем примеры параметризуемых кривых.

1) Пусть плоская кривая L задается уравнениями

$$x=\cos t, \quad y=\sin t, \quad 0\leqslant t\leqslant 3\pi.$$

Очевидно, сегмент $[0, 3\pi]$ можно разбить на сегменты $[0, \pi]$, $[\pi, 2\pi]$, $[2\pi, 3\pi]$ такие, что для значений t из каждого указанного сегмента выписанные уравнения определяют простую кривую, а именно *полуокружность*. В данном случае кривая L представляет собой окружность, у которой полуокружность, лежащая в верхней полуплоскости, проходится дважды.

2) Уравнения

$$x=\cos t, \quad y=\sin t, \quad z=t, \quad -\infty<t<\infty,$$

задают простую пространственную кривую, называемую *винтовой линией*.

3. Длина дуги кривой. Понятие спрямляемой кривой. В этом пункте мы введем понятие длины дуги параметризуемой кривой и рассмотрим некоторые свойства кривых, имеющих длину дуги (такие кривые принято называть спрямляемыми).

Условимся называть прямой линию, определяемую параметрическими уравнениями $x=at+b$, $y=ct+d$. Всегда можно выбрать так постоянные a, b, c и d, чтобы прямая проходила через две данные точки $M_1(x_1, y_1)$ и $M_2(x_2, y_2)$. Участок прямой между точками M_1 и M_2 называется отрезком, соединяющим эти точки, а совокупность конечного числа примыкающих друг к другу отрезков называется ломаной.

Пусть плоская кривая L параметризуется уравнениями

$$x=\varphi(t), \quad y=\psi(t), \quad t\in[\alpha, \beta].$$

Пусть, далее, T — произвольное разбиение сегмента $[\alpha, \beta]$ точками $\alpha=t_0<t_1<t_2< \dots <t_n=\beta$*. Обозначим через $M_0, M_1, \dots, M_n$ соответствующие точки кривой L, т. е. точки с координатами

$$M_0(\varphi(t_0), \psi(t_0)), M_1(\varphi_1(t_1), \psi(t_1)), M_2(\varphi(t_2), \psi(t_2)), \dots$$
$$\dots, M_n(\varphi(t_n), \psi(t_n)).$$

Возникающую при этом ломаную $l(t_i)=M_0M_1M_2\dots M_n$ будем на-

* В гл. 9 разбиение сегмента мы обозначали символом $\{t_k\}$.

зывать **ломаной, вписанной в кривую** L и отвечающей разбиению T сегмента $[\alpha, \beta]$. Длина $|l_i|$ звена $l_i = M_{i-1}M_i$ этой ломаной есть расстояние между точками $M_{i-1}(\varphi(t_{i-1}), \psi(t_{i-1}))$ и $M_i(\varphi(t_i), \psi(t_i))$. Поэтому $|l_i| = \sqrt{[\varphi(t_i)-\varphi(t_{i-1})]^2+[\psi(t_i)-\psi(t_{i-1})]^2}$ и длина $|l|$ всей ломаной $l = M_0M_1M_2 \ldots M_n$ будет равна

$$|l| = \sum_{i=1}^{n} |l_i| = \sum_{i=1}^{n} \sqrt{[\varphi(t_i)-\varphi(t_{i-1})]^2+[\psi(t_i)-\psi(t_{i-1})]^2}.$$

Введем понятие спрямляемой кривой.

Определение. *Кривая L называется* ***спрямляемой***, *если множество $\{|l|\}$ длин вписанных в кривую L ломаных $l = l(t_i)$, отвечающих всевозможным разбиениям T сегмента $[\alpha, \beta]$, ограничено. При этом точная верхняя грань множества $\{|l|\}$ называется* ***длиной дуги кривой*** *L и обозначается символом $|L|$.*

Из сформулированного определения нетрудно заключить, что длина $|L|$ дуги L кривой всегда положительна.

Замечание 1. *Существуют неспрямляемые кривые.* Пример неспрямляемой кривой можно найти в книге В. А. Ильина и Э. Г. Позняка «Основы математического анализа», 1 (М., 1982, с. 382—386).

Докажем следующее вспомогательное утверждение.

Лемма. *Пусть $|l_0|$ — длина ломаной, вписанной в кривую L и отвечающей разбиению T_0 сегмента $[\alpha, \beta]$, а $|l_1|$ — длина ломаной, вписанной в кривую L и отвечающей разбиению T_1, полученному из разбиения T_0 посредством добавления одной или нескольких новых точек. Тогда $|l_0| \leqslant |l_1|$.*

Доказательство. Достаточно рассмотреть случай, когда к разбиению T_0 добавляется одна точка γ. В этом случае ломаная, отвечающая разбиению T_0, отличается от ломаной, отвечающей разбиению T_1, только тем, что одно звено M_kM_{k+1} ломаной, отвечающей разбиению T_0, заменяется двумя звеньями M_kN и NM_{k+1} * ломаной, отвечающей разбиению T_1 (все остальные звенья у ломаных, отвечающих разбиениям T_0 и T_1, являются общими). Так как длина одной стороны треугольника $M_kM_{k+1}N$ не превосходит суммы длин двух других его сторон, то $|M_kM_{k+1}| \leqslant |M_kN| + |NM_{k+1}|$, а это и означает, что $|l_0| \leqslant |l_1|$. Лемма доказана.

Приведем некоторые **свойства** спрямляемых кривых:

1°. *Если кривая L спрямляема, то длина $|L|$ ее дуги не зависит от параметризации этой кривой.*

Действительно, пусть имеются две параметризации кривой L, а t и s — параметры этих параметризаций, определенные соответственно на сегментах $[\alpha, \beta]$ и $[a, b]$. Так как t представляет собой строго монотонную и непрерывную функцию от s, а

* Мы обозначили через N точку, отвечающую значению параметра $t = \gamma$.

s — строго монотонную и непрерывную функцию от t, то каждому разбиению T сегмента $[\alpha, \beta]$ соответствует определенное разбиение P сегмента $[a, b]$ и наоборот. Очевидно, что вписанные в L ломаные, отвечающие соответствующим разбиениям сегментов $[\alpha, \beta]$ и $[a, b]$, тождественны, и поэтому их длины равны. Но тогда и точные верхние грани двух тождественных числовых множеств равны, т. е. равны длины кривой L при двух различных параметризациях.

2°. *Если спрямляемая кривая L разбита при помощи конечного числа точек $M_0, M_1, \ldots, M_n$ на конечное число кривых L_i, причем точки $M_0, M_1, \ldots, M_n$ соответствуют значениям $t_0, t_1, \ldots, t_n$ параметра t и $\alpha = t_0 < t_1 < \ldots < t_n = \beta$, то каждая из кривых L_i спрямляема и сумма длин $|L_i|$ всех кривых L_i равна длине $|L|$ кривой L.*

Очевидно, это свойство достаточно доказать для случая, когда кривая L разбита точкой C на две кривые L_1 и L_2. Обозначим через γ значение параметра t, которому отвечает точка C. Тогда точки кривой L_1 соответствуют значениям параметра t из сегмента $[\alpha, \gamma]$, а точки кривой L_2 соответствуют значениям параметра t из сегмента $[\gamma, \beta]$. Пусть T_1 и T_2 — произвольные разбиения указанных сегментов, а T — разбиение сегмента $[\alpha, \beta]$, полученное объединением разбиений T_1 и T_2. Если $|l_1|$, $|l_2|$, $|l|$ — длины ломаных, вписанных в кривые L_1, L_2 и L и отвечающих разбиениям T_1, T_2 и T указанных выше сегментов, то, очевидно,

$$|l_1| + |l_2| = |l|. \tag{10.4}$$

Поскольку числа $|l_1|$, $|l_2|$ и $|l|$ положительны, то из равенства (10.4) и спрямляемости кривой L следует, что множества длин вписанных в кривые L_1 и L_2 ломаных, отвечающих всевозможным разбиениям сегментов $[\alpha, \gamma]$ и $[\gamma, \beta]$, ограничены, т. е. кривые L_1 и L_2 спрямляемы. Отметим, что из равенства (10.4) и из определения длины дуги кривой следует, что длины $|L_1|$, $|L_2|$ и $|L|$ дуг кривых L_1, L_2 и L удовлетворяют неравенству

$$|L_1| + |L_2| \leqslant |L|. \tag{10.5}$$

Действительно, из равенства (10.4) вытекает, что для любых разбиений T_1 и T_2 сегментов $[\alpha, \gamma]$ и $[\gamma, \beta]$ справедливо неравенство $|l_1| + |l_2| \leqslant |l|$. Из этого неравенства и определения точной верхней грани получим неравенство (10.5).

Покажем, что в неравенстве (10.5) на самом деле знак неравенства можно заменить на знак равенства. Предположим противное, т. е. предположим, что $|L_1| + |L_2| < |L|$. Тогда число

$$|L| - (|L_1| + |L_2|) = \varepsilon \tag{10.6}$$

положительно. Из определения длины $|L|$ дуги кривой L вытекает, что для положительного числа ε можно указать такое разбиение T_0 сегмента $[\alpha, \beta]$, что длина $|l_0|$ ломаной l_0, вписанной в кривую L и отвечающей этому разбиению, удовлетворяет неравенству $|L|-|l_0|<\varepsilon$. Добавим к разбиению T_0 точку γ и обозначим полученное при этом разбиение через T. Тогда, в силу доказанной выше леммы, длина $|l|$ ломаной, отвечающей разбиению T, тем более удовлетворяет неравенству $|L|-|l|<\varepsilon$. Так как разбиение T сегмента $[\alpha, \beta]$ образовано объединением некоторых разбиений T_1 и T_2 сегментов $[\alpha, \gamma]$ и $[\gamma, \beta]$, то длины $|l_1|$ и $|l_2|$ ломаных, отвечающих этим разбиениям, удовлетворяют соотношению (10.4). Поэтому справедливо неравенство $|L|-(|l_1|+|l_2|)<\varepsilon$. Так как $|l_1|+|l_2|\leqslant|L_1|+|L_2|$, то тем более справедливо неравенство $|L|-(|L_1|+|L_2|)<\varepsilon$. Но это неравенство противоречит равенству (10.6). Полученное противоречие доказывает, что предположение о том, что $|L_1|+|L_2|<|L|$ является неверным, и, следовательно, $|L_1|+|L_2|=|L|$. Свойство 2° полностью установлено.

Замечание 2. Понятие длины дуги пространственной кривой, заданной параметрическими уравнениями (10.3), вводится точно так же, как и понятие длины дуги плоской кривой. Точно так же, как и в плоском случае, рассматриваются длины $|l|$ ломаных, вписанных в кривую L, причем

$$|l|=\sum_{i=1}^{n}\sqrt{[(\varphi(t_i)-\varphi(t_{i-1})]^2+[\psi(t_i)-\psi(t_{i-1})]^2+[\chi(t_i)-\chi(t_{i-1})]^2}.$$

Пространственная кривая L, определяемая уравнениями (10.3), *называется спрямляемой, если множество $\{|l|\}$ длин ломаных l, вписанных в эту кривую, ограничено. Точная верхняя грань $|L|$ этого множества называется длиной дуги L.*

Пространственные спрямляемые кривые обладают свойствами 1° и 2°, приведенными выше для плоских кривых. Доказательство этих свойств аналогично доказательствам для плоских кривых.

4. Критерий спрямляемости кривой. Вычисление длины дуги кривой. Приведем достаточное условие спрямляемости кривой и формулу для вычисления длины ее дуги.

Договоримся об употреблении следующей терминологии.

1°. *Будем говорить, что функция $f(t)$ имеет на сегменте $[\alpha, \beta]$ непрерывную первую производную, если производная $f'(t)$ существует и непрерывна в любой внутренней точке этого сегмента и если, кроме того, существуют конечные пределы $\lim\limits_{t\to\alpha+0} f'(t)$ и $\lim\limits_{t\to\beta-0} f'(t)$.*

При таком определении функция $f'(t)$ окажется непрерывной на сегменте $[\alpha, \beta]$, если значения этой функции на концах указанного сегмента положить равными пределам $\lim\limits_{t\to\alpha+0} f'(t)$ и $\lim\limits_{t\to\beta-0} f'(t)$ соответственно *.

* Если в условиях определения 1° дополнительно потребовать существования односторонних производных $f'(\alpha+0)$ и $f'(\beta-0)$, то в силу п. 3 § 4 гл. 6 $\lim\limits_{x\to\alpha+0} f'(x)=f'(\alpha+0)$, $\lim\limits_{x\to b-0} f'(x)=f'(\beta-0)$.

2°. *Будем говорить, что* ***функция*** $f(t)$ ***имеет на сегменте*** $[\alpha, \beta]$ ***ограниченную первую производную,*** *если* $f'(t)$ *существует и удовлетворяет соотношению* $|f'(t)| \leqslant M$, *где* M — *некоторая постоянная, для всех внутренних точек сегмента* $\alpha < t < \beta$.

3°. *Будем говорить, что* ***производная функции*** $f(t)$ ***интегрируема*** *на сегменте* $[\alpha, \beta]$, *если* $f'(t)$ *существует для всех внутренних точек этого сегмента и после доопределения произвольными конечными значениями на концах этого сегмента представляет собой интегрируемую на этом сегменте функцию.*

Теорема 10.1. *Пусть функции* $x=\varphi(t)$ *и* $y=\psi(t)$ *непрерывны и имеют непрерывные первые производные на сегменте* $[\alpha, \beta]$. *Тогда кривая* L, *определяемая параметрическими уравнениями* $x=\varphi(t)$, $y=\psi(t)$ *при* t *из* $[\alpha, \beta]$, *спрямляема и длина* $|L|$ *ее дуги может быть вычислена по формуле*

$$|L| = \int_{\alpha}^{\beta} \sqrt{\varphi'^2(t) + \psi'^2(t)}\, dt. \tag{10.7}$$

Доказательство. Сначала докажем, что кривая L спрямляема. Рассмотрим формулу для длины $|l|$ ломаной l, вписанной в кривую L и отвечающей произвольному разбиению T сегмента $[\alpha, \beta]$:

$$|l| = \sum_{i=1}^{n} \sqrt{[\varphi(t_i) - \varphi(t_{i-1})]^2 + [\psi(t_i) - \psi(t_{i-1})]^2}\,.$$

Для каждой из функций $\varphi(t)$ и $\psi(t)$ выполнены на каждом частичном сегменте $[t_{i-1}, t_i]$ (при $i=1, 2, \ldots, n$) все условия теоремы 6.4 Лагранжа *. В силу этой теоремы между t_{i-1} и t_i найдутся точки ξ_i и η_i, такие, что будут справедливы равенства

$$\varphi(t_i) - \varphi(t_{i-1}) = \varphi'(\xi_i)\Delta t_i, \qquad \psi(t_i) - \psi(t_{i-1}) = \psi'(\eta_i)\Delta t_i,$$

где $\Delta t_i = t_i - t_{i-1}$.

Следовательно,

$$|l| = \sum_{i=1}^{n} \sqrt{\varphi'^2(\xi_i) + \psi'^2(\eta_i)}\, \Delta t_i. \tag{10.8}$$

По условию теоремы функции $\varphi(t)$ и $\psi(t)$ имеют на сегменте $[\alpha, \beta]$ непрерывные, а потому и ограниченные первые производные, т. е. для всех t, лежащих внутри сегмента $[\alpha, \beta]$, справедливы неравенства $|\varphi'(t)| \leqslant M$, $|\psi'(t)| \leqslant M$. Поэтому из формулы (10.8) следует, что

---

* Т. е. каждая из функций $\varphi(t)$ и $\psi(t)$ непрерывна на любом частичном сегменте $[t_{i-1}, t_i]$ и дифференцируема во внутренних точках этого сегмента.

$$0<|l|\leqslant\sum_{i=1}^{n}\sqrt{M^2+M^2}\,\Delta t_i=M\sqrt{2}\sum_{i=1}^{n}\Delta t_i=M\sqrt{2}\,(\beta-\alpha).$$

Таким образом, множество $\{|l|\}$ длин вписанных в кривую L ломаных, отвечающих всевозможным разбиениям T сегмента $[\alpha, \beta]$, ограничено, и по определению кривая L спрямляема.

Докажем теперь, что длина $|L|$ кривой L может быть вычислена по формуле (10.7).

Введем в рассмотрение следующую конкретную интегральную сумму интегрируемой функции $\sqrt{\varphi'^2(t)+\psi'^2(t)}$ *:

$$\sigma(t_i, \xi_i)=\sum_{i=1}^{n}\sqrt{\varphi'^2(\xi_i)+\psi'^2(\xi_i)}\,\Delta t_i,$$

отвечающую разбиению T сегмента $[\alpha, \beta]$ и выбору промежуточных точек ξ_i, определенному в формуле (10.8). Пусть d — диаметр разбиения T, т. е. $d=\max\limits_{1\leqslant i\leqslant n}\Delta t_i$. Докажем, что для любого положительного числа ε можно указать такое $\delta>0$, что при $d<\delta$ выполняется неравенство

$$||l|-I|<\frac{\varepsilon}{2}, \tag{10.9}$$

где I — предел при $d\to 0$ интегральных сумм $\sigma(t_i, \xi_i)$, т. е. $I=\int\limits_{\alpha}^{\beta}\sqrt{\varphi'^2(t)+\psi'^2(t)}\,dt$. Другими словами, мы покажем, что можно выбрать столь малым диаметр разбиения T, что длина $|l|$ ломаной l, вписанной в кривую L и отвечающей этому разбиению T, отличается от интеграла I на величину, меньшую, чем наперед заданное число $\varepsilon/2$. Заметим, что **

---

* Интегрируемость этой функции вытекает из ее непрерывности на сегменте $[\alpha, \beta]$.

** Первое из этих неравенств вытекает из следующих оценок, справедливых для любых чисел a, b, b_1:

$$\left|\sqrt{a^2+b_1^2}-\sqrt{a^2+b^2}\right|=\frac{|b_1^2-b^2|}{\sqrt{a^2+b_1^2}+\sqrt{a^2+b^2}}\leqslant$$

$$\leqslant\frac{|b_1-b|\cdot|b_1+b|}{|b_1|+|b|}\leqslant|b_1-b|.$$

Второе из этих неравенств очевидно, так как разность любых значений функции не больше разности ее точных граней.

$$\left|\sqrt{\varphi'^2(\xi_i)+\psi'^2(\eta_i)}-\sqrt{\varphi'^2(\xi_i)+\psi'^2(\xi_i)}\right| \leqslant$$
$$\leqslant |\psi'(\eta_i)-\psi'(\xi_i)| \leqslant M_i-m_i,$$

где M_i и m_i — точные грани функции $\psi(t)$ на частичном сегменте $[t_{i-1}, t_i]$.

Поэтому

$$||l|-\sigma| = \left|\sum_{i=1}^{n}\left(\sqrt{\varphi'^2(\xi_i)+\psi'^2(\eta_i)}-\sqrt{\varphi'^2(\xi_i)+\psi'^2(\xi_i)}\right)\Delta t_i\right| \leqslant$$
$$\leqslant \sum_{i=1}^{n}\left|\sqrt{\varphi'^2(\xi_i)+\psi'^2(\eta_i)}-\sqrt{\varphi'^2(\xi_i)+\psi'^2(\xi_i)}\right|\Delta t_i \leqslant$$
$$\leqslant \sum_{i=1}^{n}(M_i-m_i)\,\Delta t_i = S-s, \tag{10.10}$$

где S и s — соответственно верхняя и нижняя суммы функции $\psi'(t)$ для разбиения T сегмента $[\alpha, \beta]$.

Функции $\sqrt{\varphi'^2(t)+\psi'^2(t)}$ и $\psi'(t)$ непрерывны, а значит, и интегрируемы на сегменте $[\alpha, \beta]$, поскольку по условию на $[\alpha, \beta]$, непрерывны $\varphi'(t)$ и $\psi'(t)$.

Из определения интегрируемости и из основной теоремы § 3 гл. 9 вытекает, что для любого $\varepsilon>0$ можно указать такое $\delta>0$, что при диаметре разбиения $d<\delta$ выполняются неравенства

$$|\sigma(t_i, \xi_i)-I| < \frac{\varepsilon}{4}, \quad S-s < \frac{\varepsilon}{4}. \tag{10.11}$$

Поэтому при $d<\delta$ в силу (10.10) и (10.11) справедливы неравенства

$$||l|-I| = ||l|-\sigma+\sigma-I| \leqslant ||l|-\sigma| + |\sigma-I| < \frac{\varepsilon}{4}+\frac{\varepsilon}{4} = \frac{\varepsilon}{2},$$

и справедливость неравенства (10.9) доказана.

Докажем теперь, что среди всевозможных ломаных l, длины $|l|$ которых удовлетворяют неравенству (10.9), имеются ломаные, длины которых отличаются от длины $|L|$ дуги кривой L меньше чем на $\varepsilon/2$.

Действительно, $|L|$ — точная верхняя грань множества $\{|l|\}$ длин ломаных l, вписанных в кривую L и отвечающих всевозможным разбиениям сегмента $[\alpha, \beta]$. Поэтому найдется такое разбиение T^*, что длина $|l^*|$ соответствующей этому разбиению ломаной l^* удовлетворяет неравенству

$$0 \leqslant |L|-|l^*| < \frac{\varepsilon}{2}. \tag{10.12}$$

Подвергаем теперь разбиение T^* измельчению, добавляя к нему новые точки разбиения так, чтобы в результате добавления этих точек получилось разбиение T с диаметром d, меньшим δ. При этом, как мы показали, длина $|l|$ ломаной l, отвечающей этому разбиению T, удовлетворяет неравенству (10.9). Так как все вершины ломаной, отвечающей разбиению T^*, являются также вершинами ломаной, отвечающей разбиению T, то, согласно доказанной в п. 3 лемме, $0<|l^*|\leqslant|l|\leqslant|L|$. Поэтому неравенства (10.12) дают право утверждать, что

$$0\leqslant|L|-|l|<\frac{\varepsilon}{2}. \tag{10.13}$$

Итак, мы доказали, что среди множества ломаных $\{l\}$, длины которых удовлетворяют неравенству (10.9), имеются ломаные, длины которых удовлетворяют и неравенству (10.13). Из неравенств (10.9) и (10.13) получаем, что

$$||L|-I|<\varepsilon.$$

Поскольку ε — произвольное положительное число, то $|L|=I$ и теорема полностью доказана.

Замечание 1. *Если функции $\varphi(t)$ и $\psi(t)$ непрерывны и имеют на сегменте $[\alpha, \beta]$, ограниченные первые производные, то кривая L, определяемая уравнениями $x=\varphi(t)$, $y=\psi(t)$, спрямляема.*

Действительно, в ходе доказательства теоремы 10.1 мы установили, что если функции $\varphi(t)$ и $\psi(t)$ непрерывны на $[\alpha, \beta]$, то при условии ограниченности на сегменте $[\alpha, \beta]$ первых производных функций $\varphi(t)$ и $\psi(t)$ длины $|l|$ ломаных, вписанных в кривую L и отвечающих всевозможным разбиениям T сегмента $[\alpha, \beta]$, ограничены.

Замечание 2. *Формула* (10.7) *для вычисления длины дуги справедлива, если функции $\varphi(t)$ и $\psi(t)$ непрерывны, а производные $\varphi'(t)$ и $\psi'(t)$ только интегрируемы на сегменте $[\alpha, \beta]$.*

В самом деле, из интегрируемости этих производных следует их ограниченность, а поэтому, в силу замечания 1, и спрямляемость кривой L. Для вывода неравенств (10.10), (10.11), а следовательно, и неравенства (10.9) достаточно лишь непрерывности $\varphi(t)$ и $\psi(t)$ и интегрируемости $\varphi'(t)$ и $\psi'(t)$ на сегменте $[\alpha, \beta]$, так как отсюда, согласно теореме 9.4 (см. гл. 9), вытекает интегрируемость на сегменте $[\alpha, \beta]$ функции $\sqrt{\varphi'^2(t)+\psi'^2(t)}$. Все остальные рассуждения такие же, как и при доказательстве теоремы (10.1).

Замечание 3. *Если кривая L является графиком функции $y=f(x)$, непрерывной и имеющей на сегменте $[a, b]$ непрерывную* *производную $f'(x)$, то кривая L спрямляема и ее длина $|L|$ может быть найдена по формуле*

$$|L|=\int_a^b \sqrt{1+f'^2(x)}\,dx. \tag{10.14}$$

Действительно, график рассматриваемой функции представляет собой кривую, определяемую параметрическими уравнениями $x=t$, $y=f(t)$, $a\leqslant t\leqslant b$. При этом все условия теоремы 10.1 выполнены. Полагая в формуле (10.7) $\varphi(t)=t$, $\psi(t)=f(t)$ и заменяя переменную интегрирования t на x, получим формулу (10.14).

Замечание 4. *Если кривая L определяется так называемым полярным уравнением $r=r(\theta)$, $\theta_1\leqslant\theta\leqslant\theta_2$ и функция $r(\theta)$ непрерывна и имеет на сегменте $[\theta_1, \theta_2]$ непрерывную производную, то кривая L спрямляема и ее длина определяется равенством*

$$|L|=\int_{\theta_1}^{\theta_2} \sqrt{r^2(\theta)+r'^2(\theta)}\,d\theta. \tag{10.15}$$

Для доказательства надо воспользоваться формулами перехода от полярных координат к декартовым $x=r(\theta)\cos\theta$, $y=r(\theta)\sin\theta$. Таким образом, кривая L определяется параметрическими уравнениями $\varphi=r(\theta)\cos\theta$, $\psi=r(\theta)\sin\theta$, $\theta\in[\theta_1, \theta_2]$, причем выполнены все условия теоремы 10.1. Простые вычисления приводят к формуле (10.15).

Замечание 5. *Если рассматривается пространственная параметризуемая кривая L, заданная уравнениями $x=\varphi(t)$, $y=\psi(t)$, $z=\chi(t)$, и функции $\varphi(t)$, $\psi(t)$, $\chi(t)$ непрерывны и имеют непрерывные первые производные на $[\alpha, \beta]$, то кривая спрямляема и длина $|L|$ ее дуги может быть найдена по формуле*

$$|L|=\int_\alpha^\beta \sqrt{\varphi'^2(t)+\psi'^2(t)+\chi'^2(t)}\,dt. \tag{10.16}$$

Доказательство аналогично доказательству теоремы 10.1.

Замечание 6. *Если функции $\varphi(t)$, $\psi(t)$, $\chi(t)$ непрерывны и имеют ограниченные на сегменте $[\alpha, \beta]$ первые производные, то кривая L, определяемая уравнениями* (10.3), *спрямляема. Если при этом производные указанных функций интегрируемы на сегменте $[\alpha, \beta]$, то длина $|L|$ дуги кривой L может быть также вычислена по формуле* (10.16) (см. замечания 1 и 2).

5. Дифференциал дуги. Пусть функции $x=\varphi(t)$, $y=\psi(t)$ непрерывны и имеют на сегменте $[\alpha, \beta]$ непрерывные первые производные. В этом случае переменная длина дуги $L(t)$, отвечающая значениям параметра из сегмента $[\alpha, t]$, как это следует из теоремы 10.1, представляется в виде

$$L(t)=\int_{\alpha}^{t}\sqrt{\varphi'^2(\tau)+\psi'^2(\tau)}\,d\tau. \qquad (10.17)$$

Подынтегральная функция в правой части формулы (10.17) непрерывна, поэтому функция $L(t)$ дифференцируема и справедливо равенство

$$L'(t)=\sqrt{\varphi'^2(t)+\psi'^2(t)}.$$

Возводя полученное равенство в квадрат и умножая на $(dt)^2$, будем иметь

$$[L'(t)\,dt]^2=[\varphi'(t)\,dt]^2+[\psi'(t)\,dt]^2. \qquad (10.18)$$

Поскольку $L'(t)\,dt=dL$, $\varphi'(t)\,dt=dx$, $\psi'(t)\,dt=dy$, из формулы (10.18) получаем

$$(dL)^2=(dx)^2+(dy)^2. \qquad (10.19)$$

Если рассматривается пространственная кривая, определяемая уравнениями (10.3), то при условии непрерывности функций $\varphi(t)$, $\psi(t)$ и $\chi(t)$ и их производных первого порядка на сегменте $[\alpha, t]$ для дифференциала dL пути пространственной кривой справедлива формула

$$(dL)^2=(dx)^2+(dy)^2+(dz)^2. \qquad (10.20)$$

6. Примеры. 1) Найдем длину $|L|$ части дуги астроиды $x=a\cos^3 t$, $y=a\sin^3 t$, лежащей в первой четверти. Этой части, как нетрудно видеть, соответствует изменение параметра t от 0 до $\pi/2$. В рассматриваемом случае $\varphi'(t)=-3a\cos^2 t\sin t$, $\psi'(t)=3a\sin^2 t\cos t$. Поэтому по формуле (10.7) получим

$$|L|=\int_0^{\pi/2}\sqrt{9a^2\cos^4 t\sin^2 t+9a^2\sin^4 t\cos^2 t}\,dt=\frac{3a}{2}\int_0^{\pi/2}\sin 2t\,dt=\frac{3a}{2}.$$

2) Вычислим длину $|L|$ дуги параболы $y=ax^2$, $0\leqslant x\leqslant 1$. Поскольку $y'=2ax$, по формуле (10.14) получим

$$|L|=\int_0^1\sqrt{1+4a^2x^2}\,dx.$$

Неопределенный интеграл $I=\int\sqrt{1+4a^2x^2}\,dx$ вычислим следующим образом. Сначала проинтегрируем его по частям, затем к числителю дроби, получившейся под знаком интеграла, прибавим и вычтем единицу, произведем деление и проинтегрируем одну из получившихся дробей. Получим

$$I=\int\sqrt{1+4a^2x^2}\,dx=x\sqrt{1+4a^2x^2}-\int\frac{4a^2x^2dx}{\sqrt{1+4a^2x^2}}=$$
$$=x\sqrt{1+4a^2x^2}-\int\sqrt{1+4a^2x^2}\,dx+\int\frac{dx}{\sqrt{1+4a^2x^2}}=$$
$$=x\sqrt{1+4a^2x^2}-I+\frac{1}{2a}\ln\left|2ax+\sqrt{1+4a^2x^2}\right|.$$

Нами получено уравнение относительно величины I. Решив его, находим, что

$$I=\frac{1}{2}x\sqrt{1+4a^2x^2}+\frac{1}{4a}\ln|2ax+\sqrt{1+4a^2x^2}\,|+C.$$

Таким образом,

$$|L|=\int_0^1\sqrt{1+4a^2x^2}\,dx=\frac{1}{2}\sqrt{1+4a^2}+\frac{1}{4a}\ln|2a+\sqrt{1+4a^2}|.$$

3) Найдем длину дуги логарифмической спирали $r=ae^{b\varphi}$ от точки $(\varphi_0,\ r_0)$ до точки $(\varphi,\ r)$. По формуле (10.15) имеем, что

$$|L|=\int_{\varphi_0}^{\varphi}\sqrt{a^2e^{2b\varphi}+a^2b^2e^{2b\varphi}}\,d\varphi=a\sqrt{1+b^2}\int_{\varphi_0}^{\varphi}e^{b\varphi}d\varphi=$$
$$=\frac{a}{b}\sqrt{1+b^2}\,(e^{b\varphi}-e^{b\varphi_0})=\frac{\sqrt{1+b^2}}{b}\,(r-r_0).$$

4) Найдем переменную длину дуги эллипса $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, $a>b$, отсчитываемую от точки $M_0(0,b)$. Рассмотрим параметрическое уравнение эллипса $x=a\cdot\sin t$, $y=b\cos t$, $t\in[0,\ 2\pi]$. По формуле (10.17) получим

$$L(t)=\int_0^t\sqrt{a^2\cos^2\tau+b^2\sin^2\tau}\,d\tau=a\int_0^t\sqrt{1-\varepsilon^2\sin^2\tau}\,d\tau=aE(\varepsilon,\ t). \tag{10.21}$$

Число $\varepsilon=\frac{\sqrt{a^2-b^2}}{a}$ называется эксцентриситетом эллипса.

Первообразная функции $\sqrt{1-\varepsilon^2\sin^2 t}$, обращающаяся в нуль при $t=0$, называется эллиптическим интегралом 2-го рода (см. § 4 гл. 8). Этот интеграл обозначается символом $E(\varepsilon,\ t)$ и не выражается через элементарные функции.

§ 2. ПЛОЩАДЬ ПЛОСКОЙ ФИГУРЫ

В этом параграфе мы изучим вопрос об определении и о существовании площади так называемой плоской фигуры, под которой мы фактически будем понимать произвольное ограниченное множество точек плоскости.

Мы начнем наше рассмотрение с уточнения понятия такой фигуры и ее границы.

1. Понятия границы множества и плоской фигуры. Рассмотрим множество всех точек плоскости и фиксируем одну из этих точек A.

Договоримся называть ε-окрестностью точки A множество тех точек плоскости, которые расположены внутри круга радиуса ε с центром в точке A.

Пусть теперь $\{M\}$ — какое угодно множество точек плоскости.

Точку M множества $\{M\}$ назовем внутренней точкой этого множества, если найдется $\varepsilon>0$ такое, что ε-окрестность точки M также принадлежит множеству $\{M\}$.

*Точку M, не принадлежащую множеству $\{M\}$, назовем внешней точкой множества $\{M\}$, если найдется $\varepsilon>0$ такое, что ε-окрестность точки M также не принадлежит множеству $\{M\}$**.

*Точку M назовем граничной точкой множества $\{M\}$, если эта точка не является ни внутренней, ни внешней точкой этого множества**.*

Совокупность всех граничных точек множества $\{M\}$ назовем границей этого множества.

Замечание. Для простейших типов множеств $\{M\}$, представляющих собой часть плоскости, ограниченную простым замкнутым контуром или несколькими такими контурами, введенное нами понятие границы множества укладывается в наглядное интуитивное представление о границе. Для множеств произвольной природы граница в определенном нами смысле может иметь весьма причудливый вид и не укладываться в интуитивное представление о граничном многообразии. Так, для множества $\{M\}$ тех точек круга, абсцисса и ордината которых являются рациональными числами, границей в определенном нами смысле является весь указанный круг.

---

* Внешняя точка множества $\{M\}$ является, очевидно, внутренней точкой дополнения этого множества.

** Заметим, что *точка M является граничной точкой множества $\{M\}$ тогда и только тогда, когда для любого $\varepsilon>0$ в ε-окрестности точки M содержатся как точки, принадлежащие множеству $\{M\}$, так и точки, ему не принадлежащие.*

В самом деле, если бы в некоторой ε-окрестности точки M не нашлось либо точек, принадлежащих множеству $\{M\}$, либо точек, ему не принадлежащих, то точка M оказалась бы либо внешней, либо внутренней точкой множества $\{M\}$ и не являлась бы граничной точкой этого множества.

Множество $\{M\}$ точек плоскости будем называть ограниченным, если существует круг, содержащий все точки этого множества.

В дальнейшем мы будем рассматривать произвольное ограниченное множество F точек плоскости и будем называть это множество плоской фигурой.

Границу плоской фигуры F будем обозначать символом ∂F.

2. Площадь плоской фигуры. Для введения понятия площади плоской фигуры будем отправляться от специального частного вида плоских фигур, так называемых многоугольных фигур.

Многоугольной фигурой на плоскости мы назовем множество, составленное из конечного числа лежащих на этой плоскости ограниченных многоугольников.

Из курса средней школы известно понятие площади многоугольной фигуры.

В дальнейшем мы будем обозначать символом $\mu(P)$ площадь многоугольной фигуры P.

Напомним, что площадь многоугольной фигуры является неотрицательным числом, обладающим следующими тремя свойствами:

1° (Аддитивность). *Если P_1 и P_2 — две многоугольные фигуры без общих внутренних точек и символ $P_1 \cup P_2$ означает объединение этих фигур, то*

$$\mu(P_1 \cup P_2) = \mu(P_1) + \mu(P_2). \tag{10.22}$$

2° (Инвариантность). *Если многоугольные фигуры P_1 и P_2 равны между собой*, то*

$$\mu(P_1) = \mu(P_2). \tag{10.23}$$

3° (Монотонность). *Если многоугольная фигура P_1 содержится в многоугольной фигуре P_2, то $\mu(P_1) \leqslant \mu(P_2)$.*

Заметим, что свойство монотонности является логическим следствием свойства аддитивности и свойства неотрицательности площади. В самом деле, если P_1 содержится в P_2 то $P_2 = P_1 \cup (P_2 \setminus P_1)$**, а потому в силу того, что P_1 и $P_2 \setminus P_1$ не содержат общих внутренних точек, и в силу свойства аддитивности $\mu(P_2) = \mu(P_1) + \mu(P_2 \setminus P_1)$. Остается заметить, что $\mu(P_2 \setminus P_1) \geqslant 0$.

Замечание. Полезно подчеркнуть, что площадь многоугольной фигуры естественно считать равной одному и тому же числу независимо от того, с границей или без границы рассматривается эта многоугольная фигура. При рассмотрении разности двух многоугольных фигур $P_2 \setminus P_1$ можно договориться считать

---

* Напомним, что две фигуры F_1 и F_2 называются равными, если существует взаимно однозначное соответствие с сохранением расстояния между точками, при котором фигура F_1 отображается на F_2.

** Причем разность $P_2 \setminus P_1$ двух многоугольных фигур представляет собой также многоугольную фигуру.

фигуру P_2 взятой с границей, а фигуру P_1 взятой без границы. При такой договоренности разность $P_2 \setminus P_1$ будет представлять собой некоторую многоугольную фигуру, взятую с границей.

Перейдем теперь к определению площади некоторой произвольной плоской фигуры F (т. е. некоторого произвольного ограниченного множества точек плоскости).

Рассмотрим всевозможные многоугольные фигуры P, целиком содержащиеся в F, и многоугольные фигуры Q, целиком содержащие F. Фигуры P будем называть в п и с а н н ы м и, а фигуры Q — о п и с а н н ы м и. Числовое множество $\{\mu(P)\}$ площадей всех вписанных многоугольных фигур P *ограничено сверху* (например, площадью любой описанной многоугольной фигуры Q). Числовое множество $\{\mu(Q)\}$ площадей всех описанных вокруг фигуры Q многоугольных фигур Q *ограничено снизу* (например, нулем). Поэтому существуют точная верхняя грань

$$\mu_* = \mu_*(F) = \sup_{P \subset F} \mu(P) \tag{10.24}$$

площадей всех многоугольных фигур, вписанных в фигуру F, и и точная нижняя грань

$$\mu^* = \mu^*(F) = \inf_{Q \supset F} \mu(Q) \tag{10.25}$$

площадей всех многоугольных фигур, описанных вокруг F.

Заметим, что если в фигуру F нельзя вписать ни одного многоугольника, то по определению полагается $\mu_* = 0$.

Величину μ_* называют н и ж н е й п л о щ а д ь ю фигуры F, а μ^* — в е р х н е й п л о щ а д ь ю этой фигуры. Из того, что площадь любой вписанной фигуры не больше, чем площадь любой описанной фигуры, следует, что

$$\mu_*(F) \leqslant \mu^*(F).$$

О п р е д е л е н и е 1. *Плоская фигура F называется* к в а д р и р у е м о й *(или* и м е ю щ е й п л о щ а д ь*), если верхняя площадь μ^* этой фигуры совпадает с ее нижней площадью μ_*. При этом число $\mu = \mu(F) = \mu^* = \mu_*$ называется* п л о щ а д ь ю *фигуры F.*

Ясно, что всякая многоугольная фигура F является квадрируемой в смысле данного нами определения и для нее площадь $\mu(F) = \mu^*(F) = \mu_*(F)$, являющаяся точной нижней гранью площадей описанных многоугольных фигур и точной верхней гранью площадей вписанных фигур, совпадает с исходной величиной площади, заимствованной из элементарного курса.

Таким образом, мы распространили понятие площади многоугольников на некоторый более широкий класс фигур.

Сохранение свойств аддитивности, инвариантности и монотонности будет доказано ниже.

Начнем с доказательства следующего критерия квадрируемости плоской фигуры.

Теорема 10.2. *Для квадрируемости плоской фигуры F необходимо и достаточно, чтобы для любого $\varepsilon>0$ нашлись такая описанная вокруг F многоугольника фигура Q и такая вписанная в F многоугольная фигура P, для которых*

$$\mu(Q)-\mu(P)<\varepsilon. \tag{10.26}$$

Доказательства. *Необходимость.* Пусть фигура F квадрируема, т. е. $\mu^*=\mu_*$. По определению точных граней (10.24) и (10.25) для любого фиксированного нами $\varepsilon>0$ найдутся вписанная многоугольная фигура P и описанная многоугольная фигура Q такие, что

$$\mu_*-\frac{\varepsilon}{2}<\mu(P)\leqslant\mu_*,\quad \mu^*<\mu(Q)\leqslant\mu^*+\frac{\varepsilon}{2}.$$

Из этих неравенств и из равенств $\mu^*=\mu_*$ заключаем, что $\mu(Q)-\mu(P)<\varepsilon$. Необходимость доказана.

Достаточность. Пусть для любого $\varepsilon>0$ существуют многоугольные фигуры Q и P, указанные в формулировке теоремы. Тогда из неравенства (10.26) и из соотношений

$$\mu(P)\leqslant\mu_*\leqslant\mu^*\leqslant\mu(Q)$$

получаем, что

$$0\leqslant\mu^*-\mu_*\leqslant\mu(Q)-\mu(P)<\varepsilon.$$

Поскольку ε — произвольное положительное число, то из условия $0\leqslant\mu^*-\mu_*<\varepsilon$ вытекает, что $\mu^*=\mu_*$, т. е. доказано, что фигура F квадрируема. Теорема доказана.

Теорема 10.2 допускает простое, но важное обобщение: в ее формулировке вместо описанной и вписанной многоугольных фигур Q и P можно взять произвольные описанную и вписанную *квадрируемые* плоские фигуры Q и P. Именно справедлива теорема.

Теорема 10.2′. *Для квадрируемости плоской фигуры F необходимо и достаточно, чтобы для любого $\varepsilon>0$ нашлись такие содержащая F квадрируемая плоская фигура Q и такая содержащаяся в F квадрируемая плоская фигура P, для которых*

$$\mu(Q)-\mu(P)<\varepsilon.$$

Необходимость доказательства не требует, ибо многоугольные фигуры Q и P являются квадрируемыми.

Докажем *достаточность.*

Фиксируем произвольное $\varepsilon>0$ и построим по нему квадрируемые плоские фигуры Q и P, первая из которых содержит F, а вторая содержится в F, такие, что

$$\mu(Q)-\mu(P)<\frac{\varepsilon}{2}. \tag{10.26′}$$

Так как Q и P — *квадрируемые* плоские фигуры, то найдется многоугольная фигура $\hat{Q}$, содержащая Q, и многоугольная фигура $\hat{P}$, содержащаяся в P, такие, что

$$\mu(\widehat{Q})-\mu(Q)<\frac{\varepsilon}{4}, \quad \mu(P)-\mu(\widehat{P})<\frac{\varepsilon}{4}.$$

Из двух последних неравенств и из (10.26′) вытекает, что $\mu(\hat{Q})-\mu(\hat{P})<\varepsilon$. Но тогда, поскольку многоугольная фигура $\hat{Q}$ содержит F, а многоугольная фигура $\hat{P}$ содержится в F, фигура F квадрируема в силу теоремы 10.2.

Установим теперь еще одну эквивалентную формулировку теоремы 10.2.

Пусть F — произвольная плоская фигура, Q — многоугольная фигура, взятая вместе с границей и содержащая фигуру F, а P — многоугольная фигура, содержащаяся в фигуре F и взятая без границы. Тогда разность $Q\setminus P$ представляет собой многоугольную фигуру, взятую вместе с границей и *содержащую все точки ∂F фигуры F**.

В силу свойства аддитивности площади многоугольной фигуры справедливо равенство $\mu(Q\setminus P)=\mu(Q)-\mu(P)$, из которого следует, что неравенство (10.26) в формулировке теоремы 10.2 может быть переписано в виде

$$\mu(Q\setminus P)<\varepsilon. \tag{10.26″}$$

Договоримся о следующей терминологии.

Определение 2. *Множество точек плоскости назовем множеством площади нуль, если оно содержится в многоугольной фигуре сколь угодно малой площади.*

Неравенство (10.26″) и тот факт, что многоугольная фигура $Q\setminus P$ содержит все точки границы ∂F плоской фигуры F, дают нам право следующим образом переформулировать теорему 10.2.

Теорема 10.2″. *Плоская фигура F квадрируема тогда и только тогда, когда ее граница ∂F имеет площадь нуль.*

Необходимость условия теоремы очевидна.

Остановимся на доказательстве *достаточности.*

Впишем плоскую фигуру F в квадрат E со сторонами, параллельными координатным осям, и прямыми, параллельными этим осям, разобьем квадрат E на элементарные квадраты со стороной h. Это разбиение квадрата E договоримся называть сеткой с шагом h.

Докажем сначала, что *если граница ∂F фигуры F содержится в многоугольной фигуре площади, меньшей ε, то при доста-*

* Это следует из того, что любая внутренняя точка многоугольной фигуры P является внутренней точкой F, а любая внешняя точка многоугольной фигуры Q является внешней точкой F. Достаточно учесть, что разность $Q\setminus P$ содержит все точки плоскости, кроме внешних точек Q и внутренних точек P.

точно малом шаге h сетки граница ∂F фигуры F содержится в объединении элементарных квадратов сетки, общая площадь которых меньше $32\,\varepsilon$.

В самом деле, достаточно заметить, что любая многоугольная фигура, площади меньшей ε, представляет собой сумму конечного числа треугольников, не имеющих общих внутренних точек; каждый треугольник равен объединению двух прямоугольных треугольников (без общих внутренних точек); каждый прямоугольный треугольник содержится во вдвое большем по площади прямоугольнике; каждый прямоугольник содержится в объединении не более чем вдвое большей по площади сумме конечного числа квадратов; каждый квадрат содержится во вдвое большем по площади квадрате со сторонами, параллельными осям координат.

Итак, любая многоугольная фигура площади, меньшей ε, содержится в объединении конечного числа квадратов со сторонами, параллельными координатным осям общей площади, меньшей 8ε.

Из указанного конечного числа квадратов выберем квадрат с наименьшей стороной (если таких квадратов несколько, то выберем один из них) и возьмем шаг h сетки равным половине длины стороны этого квадрата.

При таком выборе h каждый указанный квадрат (со сторонами, параллельными координатным осям) будет содержаться в объединении элементарных квадратов сетки, общая площадь которых не больше учетверенной площади квадрата.

Поэтому вся многоугольная фигура, площади меньшей ε, содержится в объединении элементарных квадратов сетки, общая площадь которых меньше $32\,\varepsilon$.

Значит, если граница ∂F плоской фигуры F имеет площадь нуль, то для любого $\varepsilon>0$ при указанном выше выборе шага сетки h вся эта граница ∂F будет содержаться в объединении элементарных квадратов сетки, общая площадь которых меньше $32\,\varepsilon$.

Для завершения доказательства достаточности заметим, что объединение всех элементарных квадратов, состоящих только из внутренних точек фигуры F, представляет собой многоугольную фигуру P, содержащуюся в F, а объединение этой фигуры P со всеми элементарными квадратами сетки, содержащими точки границы ∂F фигуры F, представляет собой многоугольную фигуру Q, содержащую фигуру F, причем $\mu(Q)-\mu(P)<32\,\varepsilon$.

Пользуясь этой теоремой, установим квадрируемость широкого класса плоских фигур.

Докажем следующую лемму.

Лемма. *Всякая спрямляемая кривая имеет площадь нуль.*

Доказательство. Пусть L — спрямляемая кривая, а $|L|$ ее длина. Разобьем эту кривую с помощью $n+1$ точек на части, длина каждой из которых равна $|L|/n$. (Возможность такого разбиения не вызывает сомнений.) Примем каждую из этих $n+1$ точек за центр квадрата со стороной $2|L|/n$. Сумма этих квадратов представляет собой многоугольную фигуру, описанную вокруг кривой L, а площадь этой многоугольной фигуры не превосходит суммы площадей составляющих ее квадратов, т. е. числа $\frac{4|L|^2}{n^2}(n+1)$. Так как $|L|$ фиксировано, а n можно выбирать произвольно большим, то число $\frac{4|L|^2}{n^2}(n+1)$ может быть сделано меньшим любого наперед заданного числа $\varepsilon>0$. Следовательно, кривую L действительно можно заключить внутрь многоугольной фигуры сколь угодно малой площади.

Лемма доказана.

Из этой леммы и теоремы 10.2″ вытекает следующая теорема.

Теорема 10.3. *Всякая плоская фигура, граница которой состоит из одной или нескольких спрямляемых кривых, квадрируема.*

Покажем теперь, что введенное нами понятие площади плоской фигуры обладает свойствами аддитивности (см. (10.22)), инвариантности (см. (10.23)) и монотонности. Убедимся сначала в аддитивности площади.

Пусть F_1 и F_2 — квадрируемые фигуры без общих внутренних точек и F — их объединение. Тогда F квадрируема и

$$\mu(F)=\mu(F_1)+\mu(F_2). \tag{10.27}$$

Квадритируемость фигуры F следует из теоремы 10.2″ и из того, что ее граница ∂F составлена из множества площади нуль, поскольку ∂F является частью объединения границ ∂F_1 и ∂F_2 фигур F_1 и F_2. (Очевидно, что всякая часть множества площади нуль сама является множеством площади нуль.)

Докажем справедливость равенства (10.27). Рассмотрим многоугольные фигуры P_1 и P_2, вписанные в F_1 и F_2 соответственно, и многоугольные фигуры Q_1 и Q_2, описанные соответственно вокруг F_1 и F_2. Фигуры P_1 и P_2 составляют фигуру P и не имеют общих внутренних точек. Поэтому, согласно (10.22),

$$\mu(P)=\mu(P_1\cup P_2)=\mu(P_1)+\mu(P_2).$$

Многоугольные фигуры Q_1 и Q_2, возможно, пересекающиеся, в сумме составляют многоугольную фигуру Q, площадь которой не превосходит $\mu(Q_1)+\mu(Q_2)$. Поэтому

$$\mu(P)=\mu(P_1)+\mu(P_2)\leqslant\mu(F)\leqslant\mu(Q)\leqslant\mu(Q_1)+\mu(Q_2).$$

С другой стороны, в силу определения квадрируемости, для фигур F_1 и F_2 справедливы неравенства $\mu(P_1)\leqslant\mu(F_1)\leqslant\mu(Q_1)$ и $\mu(P_2)\leqslant\mu(F_2)\leqslant\mu(Q_2)$, из которых следует, что

$$\mu(P_1)+\mu(P_2)\leqslant\mu(F_1)+\mu(F_2)\leqslant\mu(Q_1)+\mu(Q_2).$$

Таким образом, обе величины $\mu(F)$ и $\mu(F_1)+\mu(F_2)$ заключены между двумя числами $[\mu(Q_1)+\mu(Q_2)]$ и $[\mu(P_1)+\mu(P_2)]$, разность между которыми

$$[\mu(Q_1)+\mu(Q_2)]-[\mu(P_1)+\mu(P_2)]=$$
$$=[\mu(Q_1)-\mu(P_1)]+[\mu(Q_2)-\mu(P_2)]$$

может быть сделана как угодно малой.

Следовательно, указанные две величины равны, т. е. справедливо равенство (10.27).

Свойство инвариантности площади произвольной плоской фигуры непосредственно вытекает из инвариантности площади для многоугольных фигур (см. (10.23)) и из самого способа определения площади квадрируемой фигуры через площади многоугольных фигур.

Наконец, свойство монотонности площади непосредственно вытекает из определения квадрируемости плоской фигуры.

Замечание. *Пересечение двух квадрируемых фигур есть квадрируемая фигура.*

Действительно, пусть $F=F_1\cap F_2$, и F_1 и F_2 квадрируемы. Каждая точка, граничная для F, является граничной либо для F_1, либо для F_2. Поэтому наше утверждение следует из теоремы 10.2″ и того факта, что объединение двух множеств площади нуль само имеет площадь нуль.

Введенное в этом пункте понятие площади называют понятием площади по Жордану* или мерой Жордана.

Выше мы убедились, что площадь по Жордану обладает свойством аддитивности, т. е. если $F=F_1\cup F_2$, а F_1 и F_2 — квадрируемые фигуры без общих точек, то F квадрируема и $\mu(F)=\mu(F_1)+\mu(F_2)$. Указанное свойство, очевидно, справедливо и для объединения любого конечного числа F_1, $F_2, \ldots, F_n$ квадрируемых фигур без общих внутренних точек. Если

$$F=\bigcup_{i=1}^{n}F_i,$$

* Камилл Жордан — французский математик (1838—1922).

то F квадрируема и $\mu(F)=\sum_{i=1}^{n}\mu(F_i)$ (свойство конечной аддитивности).

Однако *площадь по Жордану (мера Жордана) не обладает свойством счетной аддитивности, т. е. объединение счетной совокупности квадрируемых фигур $F_1, F_2, \ldots$ без общих внутренних точек не обязано быть квадрируемой фигурой.*

Проиллюстрируем этот факт примером. Рассмотрим на плоскости квадрат D: $0<x<1$, $0<y<1$. Отметим в квадрате D точки, у которых обе координаты рациональны. Нетрудно показать, что таких точек счетное множество. Расположим их в виде последовательности

$$z_1=(x_1, y_1),\quad z_2=(x_2, y_2), \ldots z_n=(x_n, y_n), \ldots.$$

Фиксируем число $\varepsilon>0$ и построим круг O_1 с центром в точке z_1 радиуса $r_1<\varepsilon/2$, целиком содержащийся в квадрате D.

Первую из точек $z_2, z_3, \ldots$, не попавшую в круг O_1, обозначим через Z_{n_2} и построим круг O_2 с центром в точке Z_{n_2} радиуса $r_2<\varepsilon/2^2$, не пересекающийся с кругом O_1 и целиком лежащий в квадрате D.

Продолжая эти рассуждения далее, мы построим последовательность содержащихся в квадрате D непересекающихся кругов $O_1, O_2, \ldots, O_n, \ldots$ радиусов $r_1, r_2, \ldots, r_n, \ldots$.

Каждый из этих кругов квадрируем и имеет площадь (меру Жордана), равную πr_n^2 $(n=1, 2, \ldots)$.

Убедимся в том, что объединение F счетного числа указанных кругов $F=O_1\cup O_2\cup\ldots$ представляет собой фигуру, не квадрируемую по Жордану. Пусть Q — любая многоугольная фигура, содержащая фигуру F. Заметим, что в любой ε-окрестности каждой точки квадрата D есть точки последовательности $\{z_n\}$, т. е. есть точки фигуры F. Но это означает, что любая точка квадрата D является внутренней либо граничной точкой фигуры F, т. е. многоугольная фигура Q содержит весь квадрат D и, значит,

$$\mu(Q)\geqslant\mu(D)=1.$$

Пусть, далее, P — любая многоугольная фигура, содержащаяся в F. Тогда площадь $\mu(P)$ не превосходит сумму площадей всех кругов $Q_1, Q_2, \ldots$, т. е.

$$\mu(P)\leqslant\pi(r_1^2+r_2^2+\ \ldots)<\pi\varepsilon^2\left(\frac{1}{2^2}+\frac{1}{2^4}+\ \ldots\right)=\frac{\pi\cdot\varepsilon^2}{3}.$$

Итак, $\mu(Q)>1$ и $\mu(P)\leqslant\frac{\pi\varepsilon^2}{3}$ для любой многоугольной фигуры Q, содержащей F, и любой многоугольной фигуры P, содержащейся в F. Но это и означает, что при малом ε разность $\mu(Q)-\mu(P)$ больше $1-\frac{\pi\varepsilon^2}{3}$ и не может быть сделана как угодно малой, т. е. фигура F не квадрируема по Жордану.

Отметим, что можно ввести другое обобщение понятия площади, так называемую меру Лебега*, которая уже *будет обладать и свойством счетной аддитивности.* Такое обобщение понятия площади выходит за рамки приложений интеграла Римана и его естественно рассматривать при изучении так называемого интеграла Лебега.

3. Площадь криволинейной трапеции и криволинейного сектора. *Криволинейной трапецией называется фигура, ограниченная графиком заданной на сегменте $[a, b]$ непрерывной и неотрицательной функции $f(x)$, перпендикулярными к оси Ox прямыми $x=a$ и $x=b$ и отрезком оси Ox между точками a и b* (рис. 10.1).

Справедливо следующее

Утверждение. *Криволинейная трапеция представляет собой квадрируемую фигуру F, площадь которой $\mu(F)$ вычисляется по формуле*

$$\mu(F)=\int_a^b f(x)\,dx. \tag{10.28}$$

Доказательство. Непрерывная на сегменте $[a, b]$ функция $f(x)$ интегрируема, поэтому для любого положительного числа ε можно указать такое разбиение сегмента $[a, b]$, для которого разность между верхней суммой S и нижней суммой s будет меньше ε. Но S и s равны соответственно $\mu(Q)$ и $\mu(P)$, где $\mu(Q)$ и $\mu(P)$ — площади многоугольных фигур, первая из которых содержит криволинейную трапецию, а вторая содержится в криволинейной трапеции (на рис. 10.1 изображены также и ука-

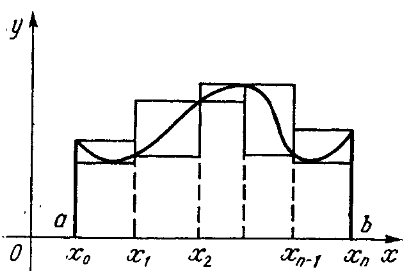


Рис. 10.1

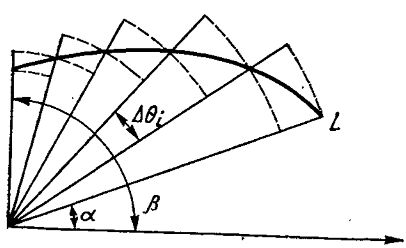


Рис. 10.2

* Анри Лебег — французский математик (1875—1941).

занные многоугольные фигуры). Таким образом, $\mu(Q)-\mu(P)<\varepsilon$, и в силу теоремы 10.2 криволинейная трапеция квадрируема. Поскольку для любой интегрируемой функции предел при стремлении диаметра разбиения к нулю как верхних S, так и нижних сумм s равен $\int\limits_a^b f(x)\,dx$ и $s\leqslant\mu(F)\leqslant S$, то площадь $\mu(F)$ криволинейной трапеции находится по формуле (10.28).

Замечание. Если функция $f(x)$ непрерывна и неположительна на сегменте $[a, b]$, то значение интеграла $\int\limits_a^b f(x)\,dx$ равно взятой с отрицательным знаком площади криволинейной трапеции, ограниченной графиком функции $|f(x)|$, ординатами в точках a и b и отрезком оси Ox между точками a и b. Поэтому если $f(x)$ *меняет знак*, то $\int\limits_a^b f(x)\,dx$ равен сумме взятых с определенным знаком площадей криволинейных трапеций, расположенных выше и ниже оси Ox, причем площади первых берутся со знаком $+$, а вторых — со знаком $-$.

Перейдем теперь к рассмотрению площади так называемого криволинейного сектора. Пусть кривая L задана в полярной системе координат уравнением $r=r(\theta)$, $\alpha\leqslant\theta\leqslant\beta$ (рис. 10.2), причем функция $r(\theta)$ непрерывна и неотрицательна на сегменте $[\alpha, \beta]$.

Назовем криволинейным сектором плоскую фигуру, ограниченную кривой L и двумя лучами, составляющими с полярной осью углы α и β.

Докажем следующее

Утверждение. *Криволинейный сектор представляет собой квадрируемую фигуру F, площадь $\mu(F)$ которой может быть вычислена по формуле*

$$\mu(F)=\frac{1}{2}\int\limits_\alpha^\beta r^2(\theta)\,d\theta. \tag{10.29}$$

Доказательство. Рассмотрим разбиение сегмента $[\alpha, \beta]$ точками $\alpha=\theta_0<\theta_1<\ldots<\theta_n=\beta$ и для каждого частичного сегмента $[\theta_{i-1}, \theta_i]$ построим круговые секторы, радиусы которых равны минимальному r_i и максимальному R_i значениям функции $r(\theta)$ на сегменте $[\theta_{i-1}, \theta_i]$. В результате получатся две квадрируемые фигуры, первая фигура A содержится в криволинейном секторе, а вторая B содержит этот сектор (см. рис. 10.2).

Площади $\mu(A)$ и $\mu(B)$ указанных квадрируемых фигур A и B соответственно равны $\frac{1}{2}\sum\limits_{i=1}^n r_i^2(\theta_i-\theta_{i-1})$ и $\frac{1}{2}\sum\limits_{i=1}^n R_i^2(\theta_i-\theta_{i-1})$.

Обратим внимание на то, что первая из этих сумм является нижней суммой s, а вторая — верхней суммой S функции $\frac{1}{2}r^2(\theta)$ на сегменте $[\alpha, \beta]$ для указанного разбиения этого сегмента. Так как непрерывная на $[\alpha, \beta]$ функция $\frac{1}{2}r^2(\theta)$ интегрируема на этом сегменте, то для любого $\varepsilon > 0$ найдется разбиение, для которого разность $S - s = \mu(B) - \mu(A)$ меньше ε.

Так как A и B — две квадрируемые фигуры, первая из которых содержится в криволинейном секторе F, а вторая содержит F, то в силу теоремы 10.2′ криволинейный сектор квадрируем.

Справедливость для его площади формулы (10.29) вытекает из того, что эта площадь $\mu(F)$ заключена между $s = \mu(A)$ и $S = \mu(B)$, а обе суммы s и S стремятся к интегралу, стоящему в правой части (10.29), при стремлении диаметра разбиения к нулю.

4. Примеры вычисления площадей. 1.° Найти площадь $\mu(F)$ фигуры F, ограниченной графиками функции $y = x^\alpha$ и $x = y^\alpha$, $\alpha \geqslant 1$ (рис. 10.3). Поскольку фигура F симметрична относительно биссектрисы первого координатного угла, то ее площадь может быть получена посредством вычитания из единицы (площади квадрата) удвоенной площади криволинейной трапеции, задаваемой графиком функции $y = x^\alpha$, $\alpha \geqslant 1$, на сегменте $[0, 1]$.

Таким образом, по формуле (10.28) мы получим, что

$$\mu(F) = 1 - 2\int_0^1 x^\alpha dx = 1 - 2\left[\frac{x^{\alpha+1}}{\alpha+1}\right]_0^1 = \frac{\alpha-1}{\alpha+1}.$$

2°. Через три точки с координатами $(-h, y_0)$, $(0, y_1)$, (h, y_2) проходит только одна парабола $y = Ax^2 + Bx + D$ (или прямая, если эти точки лежат на одной прямой).

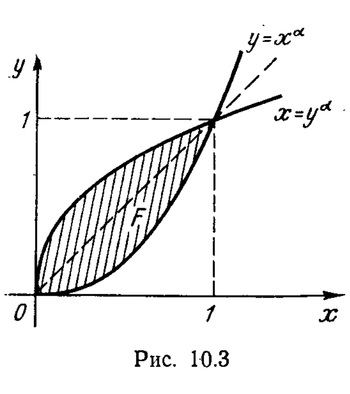


Рис. 10.3

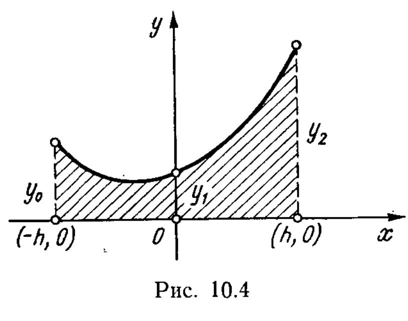


Рис. 10.4

В самом деле, условия расположения точек $(-h, y_0)$, $(0, y_1)$, (h, y_2) на параболе приводят к системе уравнений относительно A, B, D

$$\begin{cases} Ah^2 - Bh + D = y_0, \\ D = y_1, \\ Ah^2 + Bh + D = y_2. \end{cases}$$

Эта система имеет единственное решение

$$A = \frac{y_0 - 2y_1 + y_2}{2h^2}, \quad B = \frac{y_2 - y_0}{2h}, \quad D = y_1.$$

Найдем площадь $\mu(F)$ криволинейной трапеции F, определяемой указанной параболой, ординатами в точках $(-h, 0)$ и $(h, 0)$ отрезком оси Ox между этими точками (рис. 10.4).

По формуле (10.28)

$$\mu(F) = \int_{-h}^{h} (Ax^2 + Bx + D)\,dx =$$

$$= \left[\frac{Ax^3}{3} + \frac{Bx^2}{2} + Dx\right]_{-h}^{h} = \frac{2Ah^3}{3} + 2Dh.$$

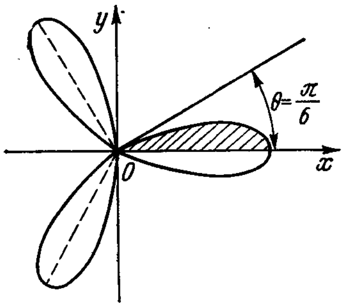


Рис. 10.5

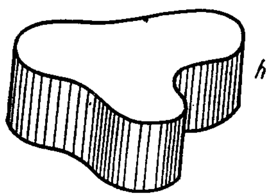


Рис. 10.6

Подставляя найденные значения A и D через ординаты y_0, y_1 и y_2 и величину h, получим $\mu(F) = \frac{h}{3}(y_0 + 4y_1 + y_2)$.

3°. Найти площадь $\mu(F)$ трилистника $r = a\cos 3\theta$ (рис. 10.5). Из рисунка можно заключить, что достаточно вычислить площадь части трилистника, отвечающей изменению θ от 0 до $\pi/6$, и полученный результат умножить на шесть. Поэтому по формуле (10.29) получаем, что

$$\mu(F) = 6\frac{a^2}{2}\int_0^{\pi/6} \cos^2 3\theta d\theta =$$

$$= 3a^2 \int_0^{\pi/6} \frac{1 + \cos 6\theta}{2} d\theta = 3a^2 \left| \frac{\pi}{12} + \frac{\sin 6\theta}{12} \Big|_0^{\pi/6} \right| = \frac{\pi a^2}{4}.$$

§ 3. ОБЪЕМ ТЕЛА В ПРОСТРАНСТВЕ

Основные определения и утверждения настоящего параграфа аналогичны соответствующим определениям и утверждениям § 2. Это позволяет нам ограничиться основными формулировками.

1. Объем тела. Рассмотрим множество всех точек пространства и фиксируем одну из этих точек A.

ε-окрестностью точки A будем называть множество всех тех точек пространства, которые расположены внутри шара радиуса ε с центром в точке A.

Точку A будем называть внутренней [внешней] точкой произвольного множества точек пространства $\{M\}$, если найдется $\varepsilon>0$ такое, что ε-окрестность точки A целиком принадлежит [целиком не принадлежит] множеству $\{M\}$.

Точки множества $\{M\}$, не являющиеся ни внутренними, ни внешними, назовем граничными точками множества $\{M\}$, а совокупность всех граничных точек назовем границей множества $\{M\}$.

Множество $\{M\}$ точек пространства назовем ограниченным множеством или телом, если найдется шар, содержащий все точки этого множества.

Среди всех тел выделим так называемые многогранные тела, представляющие собой объединение конечного числа ограниченных многогранников. Объем многогранного тела заимствуем из курса средней школы. Подчеркнем, что этот объем (как и площадь многоугольной фигуры) обладает свойствами аддитивности, инвариантности и монотонности.

Рассмотрим произвольное тело F, а также всевозможные многогранные тела P, содержащиеся в F, и всевозможные многогранные тела Q, содержащие F.

Назовем верхним объемом тела F точную нижнюю грань числового множества $\{\mu(Q)\}$ объемов всех многогранных тел Q, содержащих F т. е. число

$$\mu^* = \mu^*(F) = \inf_{Q \supset F} \mu(Q).$$

Аналогично назовем нижним объемом тела F точную верхнюю грань числового множества $\{\mu(P)\}$ объемов всех многогранных тел P, содержащихся в F, т. е. число

$$\mu_* = \mu_*(F) = \sup_{P \subset F} \mu(P).$$

Из этих определений очевидно, что $\mu_* \leqslant \mu^*$.

Определение 1. *Тело F называется кубируемым (или имеющим объем), если $\mu^* = \mu_*$.*

При этом число $\mu = \mu(F) = \mu^* = \mu_*$ называется объемом тела F.

В полной аналогии с теоремой 10.2 доказывается следующее утверждение.

Теорема 10.4. *Для кубируемости тела F необходимо и достаточно, чтобы для любого $\varepsilon>0$ нашлись такое содержащиеся в F многогранное тело P и такое содержащее F многогранное тело Q, для которых $\mu(Q)-\mu(P)<\varepsilon$.*

Замечание. В формулировке теоремы 10.4 вместо многогранных тел P и Q могут быть взяты произвольные кубируемые тела P и Q, удовлетворяющие всем другим условиям этой теоремы.

Определение 2. *Множество точек пространства назовем множеством объема нуль, если это множество содержится в многогранном теле сколь угодно малого объема.*

Теорема 10.4 может быть переформулирована.

Теорема 10.4′. *Тело F кубируемо тогда и только тогда, когда его граница имеет объем нуль.*

Введенное нами понятие объема тела обладает свойствами аддитивности, инвариантности и монотонности.

2. Некоторые классы кубируемых тел. *Цилиндрическим телом будем называть тело, ограниченное цилиндрической поверхностью с образующими, параллельными некоторой оси, и двумя плоскостями, перпендикулярными этой оси.*

Эти плоскости в пересечении с цилиндрической поверхностью образуют плоские фигуры, называемые основаниями цилиндрического тела, а расстояние h между основаниями цилиндрического тела называется его высотой (рис. 10.6).

Справедливо следующее

Утверждение. *Если основанием цилиндрического тела F является плоская квадрируемая фигура G, то тело F кубируемо, причем объем $\mu(F)$ этого тела равен $\mu(G)h$, где $\mu(G)$ — площадь основания G, а h — высота этого цилиндрического тела.*

Доказательство. Поскольку плоская фигура G квадрируема, то для любого $\varepsilon>0$ можно указать такие описанную и вписанную в эту фигуру многоугольные фигуры Q и P, что $\mu(Q)-\mu(P)<\varepsilon/h$.

Объемы цилиндрических многогранных тел F_Q и F_P, основанием которых служат многоугольные фигуры Q и P, а высота которых равна h, равны соответственно $\mu(Q)h$ и $\mu(P)h$. Поэтому

$$\mu(Q)h-\mu(P)h=\left[\mu(Q)-\mu(P)\right]h<\frac{\varepsilon}{h}h=\varepsilon.$$

Так как многогранное тело F_Q содержит F, а многогранное тело F_P содержится в F, то в силу теоремы 10.4 тело F кубируемо. Поскольку $\mu(P)h\leqslant\mu(G)h\leqslant\mu(Q)h$, то объем цилиндрического тела F равен $\mu(G)h$.

Из свойства аддитивности объема и из доказанного утверждения вытекает кубируемость ступенчатых тел (*ступенчатым телом называется объединение конечного числа цилиндрических тел, расположенных так, что верхнее основание каждого* *предыдущего из этих тел находится в одной плоскости с нижним основанием последующего*; см. рис. 10.7).

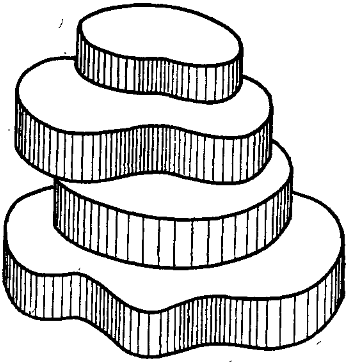


Рис. 10.7

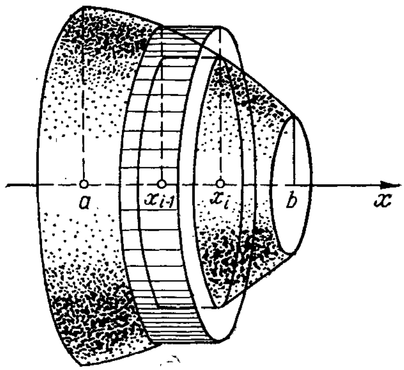


Рис. 10.8

Из предыдущих рассуждений непосредственно вытекает утверждение. *Если для любого положительного числа* ε *можно указать такое содержащее* F *ступенчатое тело* F_1 *и такое содержащееся в* F *ступенчатое тело* F_2, *что* $\mu(F_1)-\mu(F_2)<\varepsilon$, *то тело* F *кубируемо.*

Пользуясь этим, докажем кубируемость тела вращения.

Утверждение. *Пусть функция* $y=f(x)$ *непрерывна на сегменте* $[a, b]$. *Тогда тело* F, *образованное вращением вокруг оси* Ox *криволинейной трапеции, ограниченной графиком функции* $|f(x)|$ *ординатами в точках* a *и* b *и отрезком оси* Ox *от* a *до* b, *кубируемо и его объем* $\mu(F)$ *может быть найден по формуле*

$$\mu(F)=\pi\int_a^b f^2(x)\,dx. \tag{10.30}$$

Доказательство. Разобьем сегмент $[a, b]$ на частичные сегменты точками $a=x_0<x_1<\ldots<x_n=b$. Пусть m_i и M_i — точные грани $f(x)$ на частичном сегменте $[x_{i-1}, x_i]$. На каждом таком сегменте построим два прямоугольника с высотами m_i и M_i (на рис. 10.8 эти прямоугольники изображены только на одном сегменте $[x_{i-1}, x_i]$). В результате получатся две ступенчатые фигуры, одна из которых *содержится* в криволинейной трапеции, а другая *содержит* ее. При вращении криволинейной трапеции и этих ступенчатых фигур мы получим тело F и два ступенчатых тела, одно из которых Q содержит F, а другое P содержится в Q. Объемы этих тел Q и P равны соответственно

$$\mu(Q) = \pi \sum_{i=1}^{n} M_i^2 \Delta x_i, \quad \mu(P) = \pi \sum_{i=1}^{n} m_i^2 \Delta x_i.$$

Легко видеть, что эти выражения представляют собой верхнюю и нижнюю суммы для функции $\pi f^2(x)$. Поскольку эта функция интегрируема, то разность указанных сумм для некоторого разбиения сегмента $[a, b]$ будет меньше наперед взятого положительного числа ε. Следовательно, тело кубируемо. Поскольку предел указанных сумм при стремлении диаметра разбиения сегмента $[a, b]$ к нулю равен $\pi \int_a^b f^2(x)\,dx$, то объем $\mu(F)$ тела F вычисляется по формуле (10.30).

3. Примеры. 1) *Найти объем* $\mu(F)$ *шара* F *радиуса* r. Рассмотрим этот шар как результат вращения полуокружности $y = \sqrt{r^2 - x^2}$, $-r \leqslant x \leqslant r$, вокруг оси Ox (рис. 10.9). По формуле (10.30) получим

$$\mu(F) = \pi \int_{-r}^{r} (r^2 - x^2)\,dx = \pi r^2 x \Big|_{-r}^{r} - \frac{\pi x^3}{3} \Big|_{-r}^{r} = \frac{4}{3} \pi r^3.$$

2) Найдем *объем* $\mu(F)$ *прямого кругового конуса с высотой, равной* h, *и радиусом основания* r. Рассматривая указанный конус как тело, полученное вращением треугольника с вершинами в точках $(0, 0)$, $(h, 0)$ и (h, r) вокруг оси Ox (рис. 10.10), получим, согласно формуле (10.30),

$$\mu(F) = \frac{\pi r^2}{h^2} \int_0^h x^2 dx = \frac{\pi r^2 x^3}{3h^2} \Big|_0^h = \frac{\pi r^2 h}{3}.$$

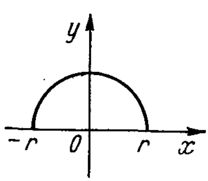


Рис. 10.9

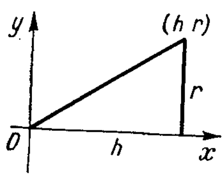


Рис. 10.10

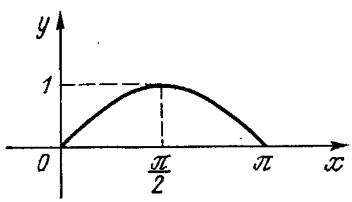


Рис. 10.11

3) Найдем *объем тела* F, *полученного вращением вокруг оси* Ox *синусоиды* $y = \sin x$ *на сегменте* $[0, \pi]$. Имеем (рис. 10.11):

$$\mu(F) = \pi \int_0^{\pi} \sin^2 x\,dx = \pi \int_0^{\pi} \frac{1 - \cos 2x}{2}\,dx = \frac{\pi^2}{2}.$$

Глава 11

ПРИБЛИЖЕННЫЕ МЕТОДЫ ВЫЧИСЛЕНИЯ КОРНЕЙ УРАВНЕНИЙ И ОПРЕДЕЛЕННЫХ ИНТЕГРАЛОВ

В этой главе рассматриваются приближенные методы нахождения корней алгебраических и трансцендентных уравнений и вычисления определенных интегралов.

§ 1. ПРИБЛИЖЕННЫЕ МЕТОДЫ ВЫЧИСЛЕНИЯ КОРНЕЙ УРАВНЕНИЙ

В этом параграфе мы займемся приближенным вычислением одного из корней уравнения $f(x)=0$, где $y=f(x)$ — некоторая, во всяком случае, непрерывная функция. Мы будем считать, что интересующий нас корень этого уравнения *изолирован на некотором сегменте* $[a, b]$, т. е. будем считать, что этот корень является внутренней точкой сегмента $[a, b]$, не содержащего других корней рассматриваемого уравнения.

На практике обычно путем грубой прикидки определяют размеры указанного сегмента $[a, b]$ *.

1. Метод «вилки». Мы начнем наше знакомство с метода, который часто используется для приближенного вычисления корней на современных быстродействующих математических машинах. Основой этого метода служит новое доказательство теоремы 4.12 о прохождении непрерывной функции через нуль при смене знака. Изложим это доказательство.

Требуется доказать следующее

Утверждение. *Если функция $f(x)$ непрерывна на сегменте $[a, b]$ и если значения этой функции $f(a)$ и $f(b)$ на концах сегмента $[a, b]$ суть числа разных знаков, то внутри сегмента $[a, b]$ найдется такая точка c, в которой значение функции $f(c)$ равно нулю, т. е. c является корнем уравнения* $f(x)=0$.

Договоримся называть «вилкой» любой сегмент, на концах которого функция $f(x)$ имеет значения разных знаков. По условию сегмент $[a, b]$ является «вилкой». Пусть ради определенности $f(a)<0$, $f(b)>0$. Разделим сегмент $[a, b]$ пополам. При этом может представиться два случая: 1) значение функции в середине сегмента $[a, b]$ равно нулю (в этом случае теорема доказана), 2) указанное значение не равно нулю. В этом случае одна

* При этом может быть использована вытекающая из физического содержания задачи дополнительная информация о расположении корня.

из половин сегмента $[a, b]$ является «вилкой». Эту половину мы обозначим $[a_1, b_1]$. Очевидно, что $f(a_1)<0$, $f(b_1)>0$. С сегментом $[a_1, b_1]$ поступим точно так же, как с сегментом $[a, b]$, т. е. разделим сегмент $[a_1, b_1]$ пополам.

Продолжая аналогичные рассуждения далее, мы будем иметь две возможности: 1) либо описанный выше процесс оборвется вследствие того, что значение функции в середине некоторого из сегментов окажется равным нулю (в этом случае теорема доказана), 2) либо описанный процесс можно продолжать неограниченно, и мы получим стягивающуюся систему сегментов — «вилок» $[a_1, b_1]$, $[a_2, b_2]$, ..., $[a_n, b_n]$, ..., причем для любого номера n $f(a_n)<0$, $f(b_n)>0$. Согласно следствию из теоремы 3.15 указанная стягивающаяся система сегментов имеет одну общую точку c, к которой сходятся каждая из последовательностей $\{a_n\}$ и $\{b_n\}$. Докажем, что $f(c)=0$. Поскольку функция $f(x)$ непрерывна в точке c, то каждая из последовательностей $f(a_n)$ и $f(b_n)$ сходится к $f(c)$. Но тогда из условий $f(a_n)<0$ и $f(b_n)>0$ в силу теоремы 3.13 получим, что одновременно справедливы неравенства $f(c)\leqslant 0$ и $f(c)\geqslant 0$, т. е. $f(c)=0$. Утверждение доказано.

Предположим теперь, что в условиях доказанного выше утверждения сегмент $[a, b]$ содержит только один корень c уравнения $f(x)=0$ *. Тогда за приближенное значение этого корня можно взять точку $\frac{a_n+b_n}{2}$, т. е. середину сегмента $[a_n, b_n]$. Поскольку длина сегмента $[a_n, b_n]$ равна $\frac{b-a}{2^n}$, то число $\frac{a_n+b_n}{2}$ отличается от точного значения корня не более чем на $\frac{b-a}{2^n}$. Таким образом, описанный выше процесс последовательного деления сегментов — «вилок» пополам позволяет вычислить искомый корень c с любой наперед заданной степенью точности. Так как описанный процесс приводит к многократному повторению однотипных вычислительных операций, он особенно удобен для проведения вычислений на быстродействующих математических машинах.

2. Метод итераций **. Излагаемый в этом пункте метод лежит в основе многих других приближенных методов. Этот метод применяется для решения уравнения

$$x=F(x). \tag{11.1}$$

---

* Т. е. предположим, что корень c является изолированным на сегменте $[a, b]$.

** Этот метод называют методом последовательных приближений.

Введем понятие итерационной последовательности.

Последовательность $x_0, x_1, \ldots, x_n, \ldots$ будем называть ***итерационной,*** *если для любого $n \geqslant 1$ элемент x_n выражается через элемент x_{n-1} по рекуррентной формуле $x_n = F(x_{n-1})$, а в качестве x_0 взято любое число из области задания функции $F(x)$.*

Мы докажем, что при определенных условиях итерационная последовательность сходится к корню уравнения (11.1) и, значит, ее элементы могут быть взяты за приближенные значения этого корня.

Справедливо следующее.

Утверждение 1. *Пусть функция $F(x)$ непрерывна на сегменте $[a, b]$, и пусть все элементы итерационной последовательности $x_0, x_1, \ldots, x_n, \ldots$ лежат на этом сегменте. Тогда, если эта последовательность сходится к некоторому числу c, то указанное число c является корнем уравнения* (11.1).

Доказательство. Так как последовательность $\{x_n\}$ сходится к c и все ее элементы принадлежат сегменту $[a, b]$, то и предел c принадлежит сегменту $[a, b]$ (см. следствие 2 из теоремы 3.13). По условию функция $F(x)$ непрерывна в точке c, и поэтому последовательность $\{F(x_{n-1})\}$ сходится к $F(c)$. Таким образом, равенство $x_n = F(x_{n-1})$ в пределе при $n \to \infty$ переходит в равенство $c = F(c)$, т. е. c является корнем уравнения (11.1). Доказанное утверждение будет существенно использовано нами в п. 3 для обоснования метода хорд и касательных.

Докажем еще одно утверждение, часто используемое для приближенного вычисления корня уравнения (11.1) с помощью итерационной последовательности.

Утверждение 2. *Пусть c — корень уравнения* (11.1), *и пусть в некотором симметричном относительно точки c сегменте $[c-\varepsilon, c+\varepsilon]$ производная функции $F(x)$ удовлетворяет условию $|F'(x)| \leqslant \alpha < 1$. Тогда итерационная последовательность $x_0, x_1, \ldots, x_n, \ldots$, у которой в качестве x_0 взято любое число из сегмента $[c-\varepsilon, c+\varepsilon]$, сходится к указанному корню c.*

Доказательство. Прежде всего докажем, что все элементы итерационной последовательности $\{x_n\}$ принадлежат указанному сегменту $[c-\varepsilon, c+\varepsilon]$. В самом деле, x_0 принадлежит этому сегменту по условию. Поэтому достаточно, предположив, что x_{n-1} принадлежит этому сегменту, доказать, что ему принадлежит и x_n. Для этого применим формулу Лагранжа к разности $F(x_{n-1}) - F(c)$ и учтем, что $F(c) = c$, $x_n = F(x_{n-1})$. Получим

$$x_n - c = F(x_{n-1}) - F(c) = F'(\xi)(x_{n-1} - c), \tag{11.2}$$

где ξ — некоторая точка, лежащая между x_{n-1} и c и, значит, принадлежащая сегменту $[c-\varepsilon, c+\varepsilon]$. Так как $|F'(\xi)| \leqslant \alpha < 1$, то из равенства (11.2) получим

$$|x_n - c| \leqslant \alpha |x_{n-1} - c|. \tag{11.3}$$

Из (11.3), поскольку $0<\alpha<1$, в свою очередь, получим

$$|x_n-c|<|x_{n-1}-c|. \tag{11.4}$$

Неравенство (11.4) устанавливает, что *каждый последующий элемент* x_n *расположен к* c *ближе, чем предыдущий элемент* x_{n-1}, *и, значит, так как* x_{n-1} *принадлежит сегменту* $[c-\varepsilon, c+\varepsilon]$ *и так как этот сегмент симметричен относительно точки* c, *то и* x_n *принадлежит этому сегменту*. Остается доказать, что последовательность $\{x_n\}$ сходится к c. Поскольку неравенство (11.3) справедливо *для всех номеров* n, то с помощью этого неравенства получим

$$|x_n-c|\leqslant\alpha^n|x_0-c|. \tag{11.5}$$

Из последнего неравенства очевидно, что $x_n\to c$, ибо $\alpha^n\to 0$. Утверждение 2 доказано.

Сделаем практические замечания относительно только что доказанного утверждения. Предположим, что путем предварительной прикидки мы установили, что интересующий нас корень уравнения (11.1) изолирован на некотором сегменте $[a, b]$, на котором производная функции $F(x)$ удовлетворяет условию $|F'(x)|\leqslant\alpha<1$. Так как сегмент $[a, b]$, вообще говоря, не является симметричным относительно искомого корня, то, естественно, возникает вопрос о том, как выбрать нулевое приближение x_0, с тем, чтобы можно было применить доказанное выше утверждение 2.

Заметим, что где бы внутри сегмента $[a, b]$ ни находился искомый корень c, хотя бы один из двух симметричных относительно c сегментов $[a, 2c-a]$, $[2c-b, b]$ (рис. 11.1) *целиком принадлежит сегменту* $[a, b]$. Поэтому хотя бы одна из точек a или b принадлежит симметричному относительно корня c сегменту, всюду на котором $|F'(x)|\leqslant\alpha<1$.

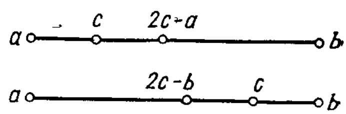


Рис. 11.1

Значит, по крайней мере одну из точек a или b можно, согласно доказанному выше утверждению 2 выбрать за x_0. Конкретно за x_0 следует выбрать ту из двух точек a или b, для которой приближение $x_1=F(x_0)$ не выходит за пределы сегмента $[a, b]$.

На практике чаще всего встречается случай, когда производная $F'(x)$ имеет на сегменте $[a, b]$ определенный знак. Если этот знак положителен, то из формулы (11.2) следует, что последовательность $\{x_n\}$ монотонна. Этот случай приводит к так называемой *ступенчатой диаграмме*, изображенной на рис. 11.2. Если же производная $F'(x)$ отрицательна на сегменте $[a, b]$, то из той же формулы (11.2) видно, что любые два последовательных элемента x_{n-1} и x_n лежат по разные стороны от корня c.

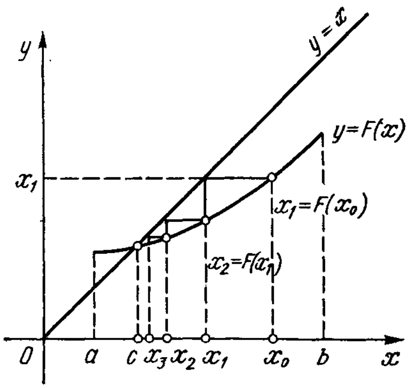


Рис. 11.2

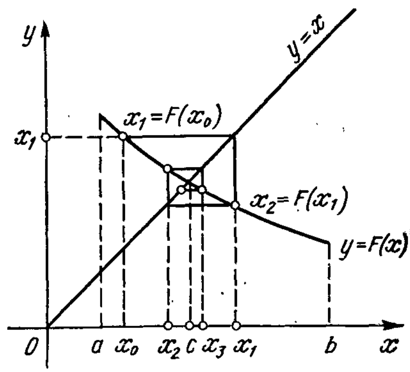


Рис. 11.3

Этот случай приводит к так называемой *спиралеобразной диаграмме*, изображенной на рис. 11.3.

Замечание. Возникает вопрос об оценке погрешности метода итераций, т. е. об оценке отклонения n-го приближения x_n от точного значения корня c. Из формулы (11.5) непосредственно вытекает следующая оценка:

$$|x_n - c| \leqslant \alpha^n (b-a),$$

где α — точная верхняя грань функции $|F'(x)|$ на сегменте $[a, b]$, на котором изолирован рассматриваемый корень.

Если производная $F'(x)$ отрицательна на сегменте $[a, b]$, то, как указано выше, x_{n-1} и x_n лежат по разные стороны от корня c, и поэтому справедлива следующая оценка:

$$|x_n - c| \leqslant |x_n - x_{n-1}|.$$

Если же в рассматриваемом случае взять за приближенное значение корня полусумму двух последовательных приближений

$$\overset{*}{x}_n = \frac{x_n + x_{n-1}}{2},$$

то получим следующую оценку погрешности:

$$|\overset{*}{x}_n - c| \leqslant \frac{|x_n - x_{n-1}|}{2}.$$

3. Методы хорд и касательных. К числу широко распространенных приближенных методов решения уравнения $f(x)=0$ относятся *метод хорд и метод касательных*, каждый из которых является одним из конкретных вариантов метода итераций.

Прежде всего рассмотрим метод хорд. Пусть искомый корень уравнения

$$f(x)=0 \tag{11.6}$$

изолирован на некотором сегменте $[a, b]$. Предположим, что *функция $y=f(x)$ имеет на сегменте $[a, b]$ монотонную и непрерывную производную, сохраняющую определенный знак.*

При этом возможны четыре случая: 1°. $f'(x)$ не убывает и положительна на $[a, b]$; 2°. $f'(x)$ не возрастает и отрицательна на $[a, b]$; 3°. $f'(x)$ не возрастает и положительна на $[a, b]$; 4°. $f'(x)$ не убывает и отрицательна на $[a, b]$.

Ради определенности подробно рассмотрим случай 1°.

Рассмотрим вместо уравнения (11.6) уравнение вида

$$x=F(x),\ F(x)=x-\frac{(b-x)f(x)^{*}}{f(b)-(x)}. \tag{11.7}$$

Легко видеть, что изолированные на сегменте $[a, b]$ корни уравнений (11.6) и (11.7) совпадают, и поэтому на сегменте $[a, b]$ эти уравнения эквивалентны. Для решения уравнения (11.7) применим к этому уравнению метод итераций, выбрав за нулевое приближение x_0 точку a. Как обычно, определим последовательность $\{x_n\}$ по рекуррентной формуле $x_n=F(x_{n-1})$, $n=1, 2, \ldots$. *Докажем, что последовательность $\{x_n\}$ сходится к искомому корню c.* Для этого, в силу утверждения 1 из п. 2, достаточно доказать, что все x_n лежат на сегменте $[a, b]$ и что последовательность $\{x_n\}$ сходится.

Применяя метод индукции, *докажем, что все x_n лежат на сегменте $[a, b]$, точнее, на сегменте $[a, c]$, где c — искомый корень.* Так как x_0 лежит на сегменте $[a, c]$, то для проведения индукции достаточно, предположив, что x_n лежит на указанном сегменте, доказать, что x_{n+1} также лежит на этом сегменте. Поскольку

$$x_{n+1}=F(x_n)=x_n-\frac{(b-x_n)f(x_n)}{f(b)-f(x_n)}, \tag{11.8}$$

то, учитывая, что $f(c)=0$, будем иметь **

$$x_{n+1}-x_n=-\frac{(b-x_n)f(x_n)}{f(b)-f(x_n)}=\frac{(b-x_n)[f(c)-f(x_n)]}{[f(b)-f(c)]+[f(c)-f(x_n)]}.$$

Применяя к выражениям в квадратных скобках формулу Лагранжа, получим

---

* При этом мы считаем, что $F(b)=b-\dfrac{f(b)}{f'(b)}$. Тогда функция $F(x)$ будет непрерывна на всем сегменте $[a, b]$.

** В дальнейшем мы предполагаем, что $x_n<c$, ибо если $x_n=c$, то $f(x_n)=f(c)=0$ и, значит, $x_{n+1}=x_n=c$, т. е. принадлежность x_{n+1} сегменту $[a, c]$ установлена.

$$x_{n+1}-x_n=\frac{(b-x_n)\,f'(\xi_n)(c-x_n)}{(b-c)\,f'(\xi_n^*)+(c-x_n)\,f'(\xi_n)}, \tag{11.9}$$

где $x_n<\xi_n<c$, $c<\xi_n^*<b$, т. е. $\xi_n<\xi_n^*$.

В силу неубывания и положительности производной $f'(x)$ можем записать $0<f'(\xi_n)\leqslant f'(\xi_n^*)$. Отсюда, так как $b-c>0$ и $c-x_n>0$, получим

$$(b-c)f'(\xi_n^*)+(c-x_n)f'(\xi_n)\geqslant[(b-c)+(c-x_n)]f'(\xi_n)= \\ =(b-x_n)f'(\xi_n).$$

Таким образом, из равенства (11.9) найдем $x_{n+1}-x_n\leqslant c-x_n$, или $x_{n+1}\leqslant c$, т. е. индукция проведена.

Докажем теперь, что последовательность $\{x_n\}$ *является неубывающей.* Для этого достаточно доказать, что дробь, стоящая в правой части равенства (11.8), является неположительной. Так как производная $f'(x)$ положительна на сегменте $[a, b]$, то функция $f(x)$ возрастает на этом сегменте, и поэтому из неравенств $x_n\leqslant c<b$ следует, что $f(x_n)\leqslant f(c)=0$, $f(b)-f(x_n)>0$. Отсюда и вытекает неположительность указанной дроби.

Итак, последовательность $\{x_n\}$ не убывает и ограничена сверху числом c. По теореме 3.15 эта последовательность сходится. В силу утверждения 1 п. 2 пределом ее является искомый корень.

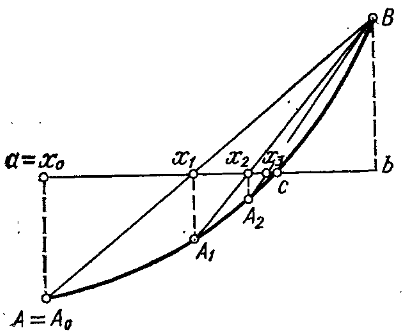


Рис. 11.4

Дадим геометрическую иллюстрацию рассмотренного выше случая 1°. Из формулы (11.8) вытекает, что x_{n+1} является абсциссой точки пересечения хорды, соединяющей точки $A_n(x_n, f(x_n))$ и $B(b, f(b))$ графика функции $y=f(x)$ с осью Ox (на рис. 11.4 изображены точки A_1 и A_2).

Как уже указано выше, кроме рассмотренного выше случая 1° возможны еще следующие три случая: 2° производная $f'(x)$ не возрастает и отрицательна на сегменте $[a, b]$, 3° производная $f'(x)$ не возрастает и положительна на сегменте $[a, b]$, 4° производная не убывает и отрицательна на сегменте $[a, b]$. Эти случаи изображены соответственно на рис. 11.5, 11.6, 11.7.

В случае 2° уравнение (11.6), так же как и выше, заменяется уравнением (11.7) и в качестве нулевого приближения берется точка $x_0=a$ (при этом последовательность $\{x_n\}$ также оказывается неубывающей). В случаях 3° и 4° уравнение (11.6) заменяется не уравнением (11.7), а следующим уравнением

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

и в качестве нулевого приближения берется точка $x_0 = b$ (при этом последовательность $\{x_n\}$ оказывается невозрастающей).

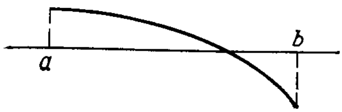


Рис. 11.5

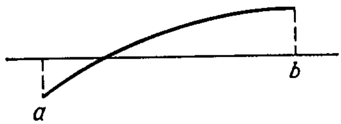


Рис. 11.6

Приведенная выше геометрическая иллюстрация является источником наименования метода хорд.

Перейдем теперь к изложению метода касательных или метода Ньютона.

Пусть, как и выше, искомый корень c уравнения (11.6) изолирован на сегменте $[a, b]$, на котором $f(x)$ имеет непрерывную и монотонную первую производную, сохраняющую определенный знак. При этом возможны те же самые четыре случая, которые отмечены при изложении метода хорд.

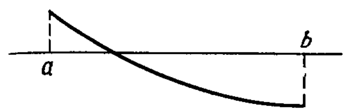


Рис. 11.7

Ради определенности рассмотрим подробно случай 1°, т. е. предположим, что производная $f'(x)$ не убывает и положительна на сегменте $[a, b]$.

Заменим уравнение (11.6) эквивалентным ему на сегменте $[a, b]$ уравнением

$$x = F(x), \text{ где } F(x) = x - \frac{f(x)}{f'(x)}, \tag{11.10}$$

и будем решать последнее уравнение методом итераций, приняв за нулевое приближение x_0 точку b и определив последовательность $\{x_n\}$ рекуррентной формулой

$$x_{n+1} = F(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{11.11}$$

Чтобы доказать, что последовательность $\{x_n\}$ сходится к искомому корню c, достаточно в силу утверждения 1 п. 2, доказать, что все x_n лежат на сегменте $[a, b]$ и что последовательность $\{x_n\}$ сходится.

Применяя метод индукции, докажем, что все x_n лежат на сегменте $[a, b]$, точнее, на сегменте $[c, b]$, где c — искомый корень. Так как $x_0 = b$ лежит на сегменте $[c, b]$, то для проведения индукции достаточно, предположив, что x_n лежит на сегменте $[c, b]$, доказать, что и x_{n+1} также лежит на этом сегменте. Если $x_n = c$, то $f(x_n) = f(c) = 0$, и из формулы (11.11) следует, что $x_{n+1} = x_n = c$, т. е. индукция проведена. Пусть теперь $x_n > c$. Тогда из формулы (11.11), учитывая, что $f(c) = 0$, получим

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

Применяя к выражению, стоящему в числителе последней дроби, формулу Лагранжа, найдем

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

где $c < \xi_n < x_n$. В силу неубывания и положительности производной дробь $\dfrac{f'(\xi_n)}{f'(x_n)}$ положительна и не превосходит единицы, т. е. $x_n - x_{n+1} \leqslant x_n - c$ или $x_{n+1} \geqslant c$.

Таким образом, индукция проведена. Из положительности производной $f'(x)$ следует возрастание функции $f(x)$, а поэтому из неравенства $c \leqslant x_n$ следует, что $0 = f(c) \leqslant f(x_n)$. Таким образом, $f(x_n)/f'(x_n) \geqslant 0$. Отсюда в силу формулы (11.11) $x_{n+1} \leqslant x_n$, т. е. последовательность $\{x_n\}$ не возрастает. Так как эта последовательность, кроме того, ограничена снизу числом c, то по теореме 3.15 она сходится. В силу утверждения 1 из п. 2 пределом ее является искомый корень c.

Дадим геометрическую иллюстрацию рассмотренного нами случая 1°. Из формулы (11.11) вытекает, что x_{n+1} является абсциссой точки пересечения с осью Ox касательной к графику функции $y = f(x)$ в точке $B_n(x_n, f(x_n))$ (на рис. 11.8 изображены точки B_0, B_1 и B_2). Приведенная геометрическая иллюстрация является источником наименования метода касательных. Предлагаем читателю самостоятельно разобрать метод касательных для случаев 2°, 3°, 4°, указанных при изложении метода хорд.

Замечание 1. Возникает вопрос об оценке погрешности метода хорд и касательных, т. е. об оценке отклонения n-го приближения от точного значения корня c. Применяя к выражению $f(x_n) = f(x_n) - f(c)$ формулу Лагранжа, будем иметь $f(x_n) = (x_n - c) f'(\xi_n)$. Отсюда получим следующую оценку:

$$|x_n - c| \leqslant \frac{|f(x_n)|}{m}, \qquad (11.12)$$

где m — минимальное значение $|f'(x)|$ на сегменте $[a, b]$. Формула (11.12) позволяет оценить отклонение x_n от точного значения корня c через значение модуля заданной функции $y = f(x)$ в точке x_n.

Замечание 2. На практике часто используют комбинированный метод, заключающийся в поочередном применении метода хорд и метода касательных. Ради определенности остановимся на подробно рассмотренном выше случае 1°, т. е. предположим, что $f'(x)$ не убывает и положительна на сегменте $[a, b]$ (рис. 11.9). Определим x_1 по методу касательных, взяв за нулевое приближение точку b. После этого определим x_2, применяя

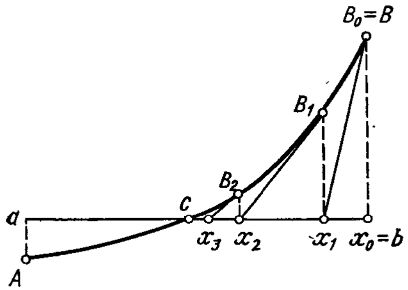


Рис. 11.8

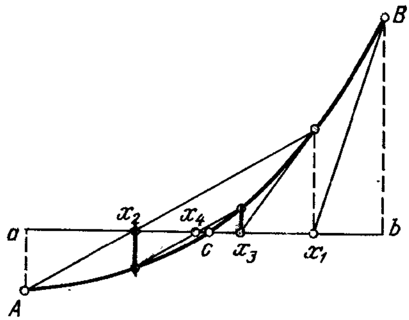


Рис. 11.9

метод хорд, но не к сегменту $[a, b]$, а к сегменту $[a, x_1]$. Далее, определим x_3 по методу касательных, исходя из уже найденного x_1; а x_4 по методу хорд, применяя его к сегменту $[x_2, x_3]$. Указанный процесс иллюстрируется на рис. 11.9.

Преимущества комбинированного метода состоят в следующем: во-первых, он дает более быструю сходимость, чем метод хорд, и, во-вторых, поскольку последовательные приближения x_n и x_{n+1} комбинированного метода с разных сторон приближается к корню, то разность $|x_{n+1}-x_n|$ дает оценку погрешности этого метода. Если за приближенное значение корня взять $x_n^* = (x_n+x_{n+1})/2$, то для погрешности получим оценку $|x_n^*-c| < |x_{n+1}-x_n|/2$.

§ 2. ПРИБЛИЖЕННЫЕ МЕТОДЫ ВЫЧИСЛЕНИЯ ОПРЕДЕЛЕННЫХ ИНТЕГРАЛОВ

1. Вводные замечания. При решении ряда актуальных физических и технических задач встречаются определенные интегралы от функций, *первообразные которых не выражаются через элементарные функции.* Кроме того, в приложениях приходится иметь дело с определенными интегралами, *сами подынтегральные функции которых не являются элементарными.* Это приводит к необходимости разработки приближенных методов вычисления определенных интегралов *.

В этом параграфе мы познакомимся с тремя наиболее употребительными приближенными методами вычисления определенных интегралов: методом прямоугольников, методом трапеций и методом парабол.

Основная идея этих методов заключается в замене подынтегральной функции $f(x)$ функцией более простой природы — многочленом, совпадающим с $f(x)$ в некоторых точках. Для уяснения этой идеи рассмотрим при малых c интеграл $\int_{-c}^{c} f(x)dx$, представляющий собой площадь узкой криволинейной трапеции, лежащей под графиком функции $y=f(x)$ на сегменте $[-c, c]$ (рис. 11.10).

Заменим функцию $f(x)$ многочленом нулевого порядка, а именно константой $f(0)$. При этом интеграл $\int_{-c}^{c} f(x)\,dx$, приближенно заменится *площадью прямоугольника*, заштрихованного на рис. 11.11. Ниже мы покажем, что при определенных требованиях на $f(x)$ ошибка, совершаемая при такой замене имеет порядок c^3.

Заменим, далее, функцию $f(x)$ многочленом первого порядка, а именно линейной функцией $y=kx+b$, совпадающей с $f(x)$ в точках $-c$ и c. При этом интеграл $\int_{-c}^{c} f(x)\,dx$ приближенно за-

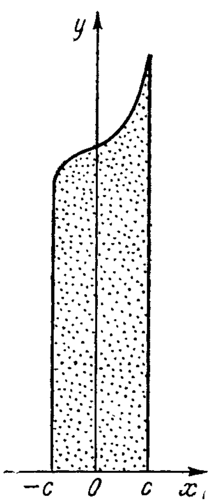


Рис. 11.10

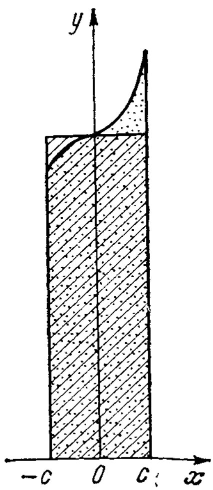


Рис. 11.11

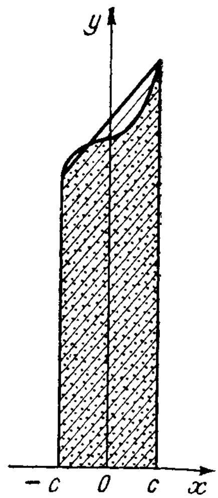


Рис. 11.12

* Заметим, что приближенными методами часто пользуются и для вычисления интегралов, выражающихся через элементарные функции.

менится *площадью прямоугольной трапеции*, заштрихованной на рис. 11.12. Ниже мы покажем, что ошибка, совершаемая при такой замене, также имеет порядок c^3.

Заменим, наконец, функцию $f(x)$ многочленом второго порядка, т. е. параболой $y=Ax^2+Bx+C$, совпадающей с $f(x)$ в точках $-c$, 0 и c. При этом интеграл $\int_{-c}^{c} f(x)\,dx$ приближенно заменится *площадью, лежащей под параболой фигуры*, заштрихованной на рис. 11.13. Ниже мы покажем, что при определенных требованиях на функцию $f(x)$ ошибка, совершаемая при такой замене, имеет порядок c^5.

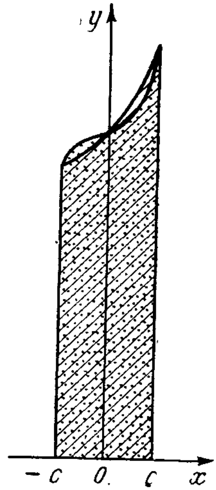


Рис. 11.13

Если потребуется вычислить интеграл $\int_a^b f(x)\,dx$ по любому сегменту $[a, b]$, естественно разбить этот сегмент на достаточно большое число малых сегментов и к каждому из этих сегментов применить изложенные выше рассуждения. При этом мы и придем к методам прямоугольников, трапеций и парабол в их общем виде. Для того чтобы оценить ошибку, возникающую при применении методов прямоугольников, трапеций и парабол, мы подойдем к изложению этих методов с другой точки зрения.

Прежде всего введем понятие **усреднения** n **чисел**.

Пусть $\lambda_1, \lambda_2, \ldots, \lambda_n$ — какие угодно положительные числа. *Любое число* c *вида*

$$c=\frac{\lambda_1\cdot f(x_1)+\lambda_2 f(x_2)+\ldots+\lambda_n f(x_n)}{\lambda_1+\lambda_2+\ldots+\lambda_n} \qquad (11.13)$$

назовем ***усреднением*** n *чисел* $f(x_1), f(x_2), \ldots, f(x_n)$.

Очевидно, что если все числа $f(x_1), f(x_2), \ldots, f(x_n)$ заключены между числами m и $M (m \leqslant M)$, то и любое усреднение c этих чисел удовлетворяет неравенствам $m \leqslant c \leqslant M$.

Предположим далее, что функция $f(x)$ непрерывна на сегменте $[a, b]$ и все значения $x_1, x_2, \ldots, x_n$ лежат на этом сегменте. Тогда, какое бы усреднение n чисел $f(x_1), f(x_2), \ldots, f(x_n)$ мы ни взяли, на сегменте $[a, b]$ найдется точка ξ такая, что это усреднение равно значению $f(\xi)$ в точке ξ. В самом деле, так как функция $f(x)$ непрерывна на сегменте $[a, b]$, то все значения этой функции на указанном сегменте заключены между ее наибольшим значением M и наименьшим значением m. Значит, и любое усреднение c чисел $f(x_1), f(x_2), \ldots, f(x_n)$ заключено между m и M. Но, каково бы ни было это промежуточное значение c, согласно теореме 4.12 на сегменте $[a, b]$ найдется точка ξ такая, что $c=f(\xi)$.

Таким образом, для непрерывной на сегменте $[a, b]$ функции формулу (11.13) можно переписать в виде

$$\frac{\lambda_1 f(x_1)+\lambda_2 f(x_2)+\ldots+\lambda_n f(x_n)}{\lambda_1+\lambda_2+\ldots+\lambda_n}=f(\xi) \qquad (11.14)$$

или в виде

$$\frac{\lambda_1 f(x_1)+\lambda_2 f(x_2)+\ldots+\lambda_n f(x_n)}{\lambda_1+\lambda_2+\ldots+\lambda_n}(b-a)=f(\xi)(b-a). \qquad (11.15)$$

С другой стороны, для непрерывной на сегменте $[a, b]$ функции, согласно п. 2 § 4 гл. 9, найдется точка ξ' из сегмента $[a, b]$ такая, что справедлива формула среднего значения

$$\int_a^b f(x)\,dx=f(\xi')(b-a). \qquad (11.16)$$

Сопоставление формул (11.15) и (11.16) позволяет сделать предположение о том, что при некотором разумном выборе чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ и точек $x_1, x_2, \ldots, x_n$ вычисление интеграла $\int_a^b f(x)\,dx$ можно с большой точностью заменить вычислением суммы, стоящей в левой части формулы (11.15). Именно на этой идее разумного выбора чисел $\lambda_1, \lambda_2, \ldots, \lambda_n$ и точек $x_1, x_2, \ldots, x_n$ и основаны приближенные методы вычисления интеграла. Переходим к изложению этих методов.

2. **Метод прямоугольников.** Будем считать, что функция $f(x)$, интеграл от которой нам требуется приближенно вычислить, имеет на рассматриваемом сегменте *непрерывную вторую производную.*

Начнем с рассмотрения интеграла в симметричных пределах $\int_{-c}^{c} f(x)\,dx$. Для вычисления этого интеграла будем исходить из формул (11.15) и (11.16), в которых положим $n=1$, $a=-c$, $b=c$, $x_1=0$, $\lambda_1=1$. Тогда, очевидно, $\xi=0$ и правая часть (11.15) равна $f(0)\cdot 2c$. Таким образом,

$$\int_{-c}^{c} f(x)\,dx=f(0)\cdot 2c+R, \qquad (11.17)$$

где символом R обозначен остаточный член (т. е. отклонение числа $f(0)\cdot 2c$ от точного значения интеграла). Для того чтобы оценить величину остаточного члена R, обозначим через $F(x)$ первообразную функции $f(x)$. Поскольку в силу формулы Ньютона—Лейбница $\int_{-c}^{c} f(x)\,dx=F(c)-F(-c)$, то

$$R=F(c)-F(-c)-f(0)\cdot 2c. \qquad (11.18)$$

Разложим по формуле Маклорена функцию $\psi(x)=F(x)-F(-x)$. Беря остаточный член в форме Лагранжа и обозначая через ξ' возникающее при этом промежуточное значение аргумента из интервала $(0, c)$, будем иметь

$$\psi(c)=F(c)-F(-c)=\psi(0)+\frac{\psi'(0)}{1!}c+\frac{\psi''(0)}{2!}c^2+\frac{\psi'''(\xi)}{3!}c^3. \quad (11.19)$$

Подсчитаем входящие в эту формулу значения $\psi(0)$, $\psi'(0)$, $\psi''(0)$, $\psi'''(\xi')$. Имеем

$$\psi(x)=F(x)-F(-x); \quad \psi(0)=F(0)-F(0)=0;$$
$$\psi'(x)=F'(x)+F'(-x)=f(x)+f(-x); \quad \psi'(0)=f(0)+f(0)=2f(0);$$
$$\psi''(x)=f'(x)-f'(-x); \quad \psi''(0)=f'(0)-f'(0)=0;$$
$$\psi'''(x)=f''(x)+f''(-x); \quad \psi'''(\xi')=\frac{2[f''(\xi')+f''(-\xi')]}{2}=2f''(\xi).$$

(В последнем равенстве мы воспользовались формулой (11.14) при $n=2$, $\lambda_1=\lambda_2=1$, $x_1=\xi'$, $x_2=-\xi'$ и обозначили через ξ некоторую точку из интервала $(-c, c)$ на котором по предположению непрерывна функция $f''(x)$.)

Вставляя вычисленные значения в формулу (11.19), будем иметь

$$\psi(c)=F(c)-F(-c)=2f(0)c+2\frac{f''(\xi)}{3!}c^3. \quad (11.20)$$

Сопоставляя последнюю формулу с формулой (11.18), окончательно получим

$$R=2\frac{f''(\xi)}{3!}c^3=\frac{f''(\xi)}{24}(2c)^3, \quad -c<\xi<c. \quad (11.21)$$

Из полученной оценки остаточного члена видно, что формула (11.17) тем точнее, чем меньше величина $2c$. Поэтому для вычисления интеграла $\int\limits_a^b f(x)\,dx$ удобно разбить сегмент $[a, b]$ на достаточно большое число n частей и к каждой из этих частей применить формулу приближенного интегрирования (11.17). Считая, что функция $f(x)$ имеет на сегменте $[a, b]$ непрерывную вторую производную, разобьем этот сегмент на $2n$ *равных* частей при помощи точек $a=x_0<x_2<\ldots<x_{2n}=b$. Обозначим через x_{2k+1} среднюю точку сегмента $[x_{2k}, x_{2k+2}]$. Тогда

$$\int\limits_a^b f(x)\,dx=\frac{b-a}{n}[f(x_1)+f(x_3)+\ldots+f(x_{2n-1})]+R, \quad (11.22)$$

где

$$R=R_1+R_2+\ldots+R_n=\frac{(b-a)^3}{24n^3}[f''(\xi_1)+f''(\xi_2)+\ldots+f''(\xi_n)]=$$

$$= \frac{(b-a)^3}{24n^2} f''(\xi), \ a < \xi < b. \qquad (11.23)$$

(Здесь мы воспользовались для $f''(x)$ формулой усреднения (11.14) при $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 1$ и обозначили через ξ некоторое промежуточное значение аргумента из интервала (a, b).)

Формула (11.22) называется *формулой прямоугольников*. Ее геометрический смысл ясен из рис. 11.14: площадь криволинейной трапеции, лежащей под графиком $f(x)$ на сегменте $[a, b]$, приближенно заменяется суммой площадей, указанных на этом чертеже прямоугольников.

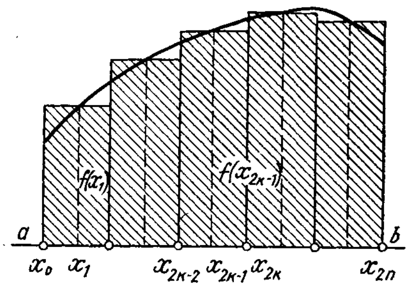


Рис. 11.14

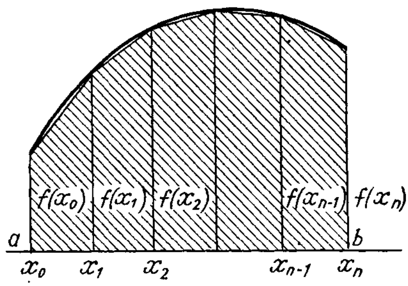


Рис. 11.15

3. Метод трапеций. Пусть, как и выше, функция $f(x)$ имеет на рассматриваемом сегменте непрерывную вторую производную.

Снова начнем с вычисления интеграла $\int_{-c}^{c} f(x)\,dx$, но на этот раз будем исходить из формул (11.15) и (11.16), считая, что $n=2$, $a=-c$, $b=c$, $x_1=-c$, $x_2=c$, $\lambda_1=\lambda_2=1$. Тогда

$$\int_{-c}^{c} f(x)\,dx = \frac{1}{2}[f(-c)+f(c)]\cdot 2c + R, \qquad (11.24)$$

где R — остаточный член, подлежащий оценке.

Обозначая, как и в п. 2, через $F(x)$ первообразную функции $f(x)$ и учитывая, что $\int_{-c}^{c} f(x)\,dx = F(c) - F(-c)$, будем иметь

$$R = F(c) - F(-c) - \frac{1}{2}[f(c)+f(-c)]\cdot 2c. \qquad (11.25)$$

Пусть, как и в методе прямоугольников, $\psi(x) = F(x) - F(-x)$. Разлагая функции $\psi(x)$ и $\psi'(x)$ по формуле Маклорена с остаточным членом в интегральной форме (см. п. 4 § 5 гл. 9) и полагая $x=c$, будем иметь

$$\psi(c) = F(c) - F(-c) =$$

$$= \psi(0) + \frac{\psi'(0)}{1!} c + \frac{\psi''(0)}{2!} c^2 + \frac{1}{2} \int_0^c \psi'''(x)(c-x)^2\, dx,$$

$$\psi'(c) = f(c) + f(-c) = \psi'(0) + \frac{\psi''(0)}{1!} c + \frac{1}{1!} \int_0^c \psi'''(x)(c-x)\, dx.$$

Подставляя в эти формулы значения $\psi(0)$, $\psi'(0)$, $\psi''(0)$, вычисленные в п. 2, получим

$$F(c) - F(-c) = 2f(0)\, c + \frac{1}{2} \int_0^c \psi'''(x)(c-x)^2 dx,$$

$$f(c) + f(-c) = 2f(0) + \int_0^c \psi'''(x)(c-x)\, dx.$$

Подставляя последние два выражения в (11.25), получим

$$R = \int_0^c \psi'''(x) \left[\frac{1}{2} (c-x)^2 - c(c-x) \right] dx = -\frac{1}{2} \int_0^c \psi'''(x)(c^2 - x^2)\, dx.$$

Имея в виду, что функция $c^2 - x^2$ неотрицательна на сегменте $[0, c]$, применим к последнему интегралу первую формулу среднего значения (см. п. 2 § 4 гл. 9). Учитывая, что $\psi'''(x) = f''(x) + f''(-x)$, и обозначая через ξ' некоторое значение аргумента из сегмента $[0, c]$, получим

$$R = -\frac{f''(\xi') + f''(-\xi')}{2} \int_0^c (c^2 - x^2)\, dx = -\left[\frac{f''(\xi') + f''(-\xi')}{2} \right] \frac{2c^3}{3}.$$

Применяя к выражению в квадратных скобках формулу усреднения (11.14) при $n=2$, $\lambda_1 = \lambda_2 = 1$ и обозначая через ξ некоторое значение аргумента из сегмента $[-c, c]$, окончательно получим

$$R = -f''(\xi) \cdot \frac{2c^3}{3} = -\frac{f''(\xi)}{12} (2c)^3.$$

Для вычисления интеграла $\int_a^b f(x)\, dx$, как и в методе прямоугольников, разобьем сегмент $[a, b]$ на n равных частей при помощи точек $a = x_0 < x_1 < \ldots < x_n = b$ и применим формулу (11.24) к каждому из частичных сегментов. Получим

$$\int_a^b f(x)\,dx=\sum_{k=0}^{n-1}\int_{x_k}^{x_{k+1}} f(x)\,dx=$$

$$=\sum_{k=0}^{n-1}\left\{\frac{x_{k+1}-x_k}{2}\left[f(x_k)+f(x_{k+1})\right]+R_k\right\}=$$

$$=\frac{b-a}{2n}\left\{\left[f(x_0)+f(x_1)\right]+\left[f(x_1)+f(x_2)\right]+\ldots\right.$$

$$\left.\ldots+\left[f(x_{n-1})+f(x_n)\right]\right\}+R=$$

$$=\frac{b-a}{2n}\left\{f(a)+f(b)+2\sum_{k=1}^{n-1} f(x_k)\right\}+R, \qquad (11.26)$$

где

$$R=R_0+R_1+\ldots+R_{n-1}=-\frac{f''(\xi_0)+f''(\xi_1)+\ldots+f''(\xi_{n-1})}{12n^3}(b-a)^3=$$

$$=-\frac{f''(\xi)}{12n^2}(b-a)^3,\quad a\leqslant\xi\leqslant b. \qquad (11.27)$$

(Мы воспользовались формулой усреднения (11.11).)

Формула (11.26) называется *формулой трапеций*. Геометрический смысл этой формулы ясен из рис. 11.15: площадь криволинейной трапеции, лежащей под графиком функции $f(x)$ на сегменте $[a, b]$ приближенно заменяется суммой площадей, указанных на этом чертеже прямолинейных трапеций. Сравнение остаточного члена (11.27) с остаточным членом (11.23) показывает, что метод трапеций не дает увеличения точности по сравнению с методом прямоугольников.

4. Метод парабол. На этот раз предположим, что функция имеет на рассматриваемом сегменте *непрерывную четвертую производную*, и снова начнем с вычисления интеграла $\int_{-c}^{c} f(x)\,dx$.

Как и выше, будем исходить из формул (11.15) и (11.16), но при этом положим в этих формулах $n=3$, $a=-c$, $b=c$, $\lambda_1=\lambda_3=1$, $\lambda_2=\lambda$ (числом λ распорядимся в дальнейшем!), $x_1=-c$, $x_2=0$, $x_3=c$. Тогда

$$\int_{-c}^{c} f(x)\,dx=\frac{f(-c)+\lambda f(0)+f(c)}{2+\lambda}\,2c+R,$$

где R — подлежащий определению остаточный член. Для оценки остаточного члена обозначим, как и выше, через $F(x)$ первообразную функции $f(x)$ и учтем, что $\int\limits_{-c}^{c} f(x)\,dx = F(c) - F(-c)$. Получим, что

$$R = F(c) - F(-c) - \frac{f(-c) + \lambda f(0) + f(c)}{2+\lambda}\, 2c. \qquad (11.28)$$

Пусть, как и выше, $\psi(x) = F(x) - F(-x)$. Разложим функции $\psi(x)$ и $\psi'(x)$ по формуле Маклорена с остаточным членом в интегральной форме. Подставляя в эти разложения значения $\psi(0)$, $\psi'(0)$, $\psi''(0)$, вычисленные в п. 2, и учитывая, что $\psi^{(4)}(0) = 0$, будем иметь

$$\psi(c) = F(c) - F(-c) =$$

$$= 2f(0)\,c + \frac{2f''(0)}{3!}\,c^3 + \frac{1}{4!}\int\limits_0^c \psi^5(x)\,(c-x)^4\,dx, \qquad (11.29)$$

$$\psi'(c) = f(c) + f(-c) = 2f(0) + \frac{2f''(0)}{2!}\,c^2 + \frac{1}{3!}\int\limits_0^c \psi^{(5)}(x)\,(c-x)^3\,dx.$$

Из последней формулы вытекает, что

$$\frac{f(-c) + \lambda f(0) + f(c)}{2+\lambda}\, 2c =$$

$$= 2f(0)\,c + \frac{f''(0)}{2+\lambda}\, 2c^3 + \frac{2c}{2+\lambda}\,\frac{1}{3!}\int\limits_0^c \psi^{(5)}(x)\,(c-x)^3\,dx. \qquad (11.30)$$

Из формулы (11.28) видно, что остаточный член R равен разности выражений (11.29) и (11.30). Чтобы сделать этот остаточный член более высоким по порядку малости, выберем значение λ так, чтобы вторые члены в правых частях формул (11.29) и (11.30) совпадали, т. е. положим $\frac{2}{3!} = \frac{2}{2+\lambda}$, т. е. $\lambda = 4$. При таком значении разность формул (11.29) и (11.30) дает

$$R = \int\limits_0^c \psi^{(5)}(x)\left[\frac{1}{24}(c-x)^4 - \frac{c}{18}(c-x)^3\right] dx =$$

$$= -\frac{1}{24}\int\limits_0^c \psi^{(5)}(x)\left[(c-x)^3\left(\frac{c}{3} + x\right)\right] dx.$$

Имея в виду, что функция $\left[(c-x)^3\left(\frac{c}{3} + x\right)\right]$ неотрицательна на сегменте $[0, c]$, применим к последнему интегралу первую формулу среднего значения. Учитывая, что $\psi^{(5)}(x)=f^{(4)}(x)+f^{(4)}(-x)$, и обозначая через ξ' некоторое значение аргумента из сегмента $[0, c]$, получим

$$R=-\frac{f^{(4)}(\xi')+f^{(4)}(-\xi')}{24}\int_0^c (c-x)^3\left(\frac{c}{3}+x\right)dx=$$

$$=-\left[\frac{f^{(4)}(\xi')+f^{(4)}(-\xi')}{2}\right]\frac{(2c)^5}{2880}.$$

Применяя к выражению в квадратных скобках формулу усреднения (11.14) при $n=2$, $\lambda_1=\lambda_2=1$ и обозначая через ξ некоторое значение аргумента из сегмента $[-c, c]$, окончательно получим

$$R=-\frac{f^{(4)}(\xi)}{2880}(2c)^5.$$

Для вычисления интеграла $\int_a^b f(x)\,dx$ разделим сегмент $[a, b]$ на n равных частей точками $a=x_0<x_2<x_4<\ldots<x_{2n}=b$ и положим $x_{2k+1}=\frac{x_{2k}+x_{2k+2}}{2}$. Получим

$$\int_0^b f(x)\,dx=\sum_{k=0}^{n-1}\int_{x_{2k}}^{x_{2k+2}} f(x)\,dx=$$

$$=\sum_{k=0}^{n-1}\left\{(x_{2k+2}-x_{2k})\frac{f(x_{2k})+4f(x_{2k+1})+f(x_{2k+2})}{6}+R_k\right\}=$$

$$=\frac{b-a}{6n}\left[f(a)+f(b)+2\sum_{k=1}^{n-1}f(x_{2k})+4\sum_{k=0}^{n-1}f(x_{2k+1})\right]+R, \qquad (11.31)$$

где

$$R=-\frac{f^{(4)}(\xi_0)+f^{(4)}(\xi_1)+\ldots+f^{(4)}(\xi_{n-1})}{2\,880n^5}(b-a)^5=$$

$$=-\frac{f^{(4)}(\xi)}{2880n^4}(b-a)^5,$$

$$a<\xi<b. \qquad (11.32)$$

(Здесь мы применим формулу усреднения (11.14).)

Формула (11.31) называется **формулой Симпсона** или **формулой парабол**. Геометрический смысл этой формулы ясен из рис. 11.16: площадь криволинейной трапеции, лежащей под графиком функции $f(x)$ на сегменте $[a, b]$ приближенно заменяется суммой площадей, заштрихованных на этом чертеже фигур, лежащих под параболами. Для того, чтобы убедиться в этом достаточно заметить, что выражение, стоящее в фигурных скобках в формуле (11.31), численно равно площади фигуры, лежащей на сегменте $[x_{2k}, x_{2k+2}]$ под параболой $y=Ax^2+Bx+C$, совпадающей с $f(x)$ в точках x_{2k}, x_{2k+1}, x_{k+2} (см. пример 2 п. 4 § 2 гл. 10).

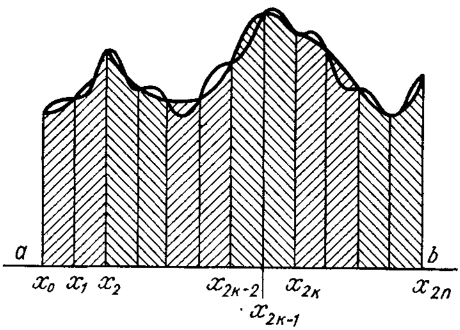


Рис. 11.16

Сравнивая остаточный член (11.32) с остаточными членами (11.23) и (11.27), мы убеждаемся в том, что формула Симпсона дает **большую точность**, чем формулы прямоугольников и трапеций.

В качестве иллюстрации применения формулы Симпсона обратимся к вычислению интеграла $I(x_0)=\int\limits_0^{x_0} e^{-x^2}dx$*, ограничиваясь для простоты значениями x_0 из сегмента $0 \leqslant x_0 \leqslant 1$. Полагая $f(x)=e^{-x^2}$ и вычисляя производную $f^{(4)}(x)=4(4x^4-12x^2+3)e^{-x^2}$ без труда убедимся в том, что для всех x из сегмента $0 \leqslant x \leqslant 1$ во всяком случае $|f^{(4)}(x)|<20$. Исходя из оценки (11.32), можем утверждать, что $R<\frac{1}{144n^4}$. Значит, разбив сегмент $[0, x_0]$ всего на 5 равных частей и заменив рассматриваемый интеграл суммой, стоящей в правой части формулы Симпсона, мы вычислим этот интеграл с точностью до $\frac{1}{144\cdot 5^4}<\frac{1}{90\,000}$**.

---

* Рассматриваемый интеграл, как уже неоднократно отмечалось, не выражается через элементарные функции, но имеет большое значение в статистической физике, теории теплопроводности и диффузии.

** На ручном электронном калькуляторе, вычисляющем значения элементарных функций, авторы вычислили за несколько минут указанный интеграл для $x_0=1$, $n=5$ и получили результат $I(1)=0{,}7468251$, который, в силу сказанного выше, содержит пять верных десятичных знаков после запятой.

Глава 12

ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Многие вопросы естествознания приводят к рассмотрению такой зависимости между несколькими переменными величинами, при которой значения одной из переменных величин полностью определяются значениями остальных переменных.

Так, например, при рассмотрении каких-либо физических характеристик тела (например, его плотности ρ или температуры T) нам приходится учитывать изменение этих характеристик при переходе от одной точки тела к другой. Поскольку каждая точка тела определяется тремя декартовыми координатами x, y и z, то рассматриваемые характеристики (плотность ρ или температура T) определяются значениями трех переменных x, y и z.

При рассмотрении физических процессов, меняющихся во времени, значения физических характеристик определяются значениями четырех переменных, трех координат точки x, y, z и времени t. Например, при изучении звуковых колебаний газа плотность ρ этого газа и его давление p определяются значениями четырех переменных x, y, z и t.

Для изучения такого рода зависимостей в этой главе вводится понятие функции нескольких переменных и развивается аппарат для исследования таких функций.

Первая часть настоящей главы посвящена построению дифференциального исчисления функций нескольких переменных.

На случай функции нескольких переменных будут распространены понятия и утверждения, установленные нами в гл. 3—7 для функции одной переменной.

В одном из дополнений к настоящей главе изучаются элементы дифференциального исчисления для абстрактных функций, представляющих собой результат отображения одного нормированного пространства в другое. Частным случаем такого отображения является довольно часто встречающееся отображение евклидова пространства размерности m в другое евклидово пространство размерности n.

§ 1. ПОНЯТИЕ ФУНКЦИИ m ПЕРЕМЕННЫХ

1. Понятие m-мерного координатного и m-мерного евклидова пространств. При изложении теории функций m переменных удобно использовать геометрическую терминологию, обобщающую и формализующую наши представления о плоскости и о реальном (трехмерном) геометрическом пространстве.

Назовем m-**мерным координатным пространством** множество всевозможных упорядоченных совокупностей $(x_1, x_2, \ldots$ $\ldots, x_m)$ вещественных чисел $x_1, x_2, \ldots, x_m$.

Будем обозначать m-мерное координатное пространство символом A^m.

Каждую упорядоченную совокупность $(x_1, x_2, \ldots, x_m)$ мы будем называть **точкой** m-мерного координатного пространства и обозначать одной буквой M.

При этом числа $x_1, x_2, \ldots, x_m$ мы будем называть **координатами** точки M.

Запись $M(x_1, x_2, \ldots, x_m)$ означает, что точка M имеет координаты $x_1, x_2, \ldots x_m$.

Замечание 1. Если рассматривать координатное пространство A^m как множество всех векторов $\mathbf{x}$ с координатами $(x_1, x_2, \ldots, x_m)$ и назвать **суммой** векторов $\mathbf{x}=(x_1, x_2, \ldots, x_m)$ и $\mathbf{y}=(y_1, y_2, \ldots, y_m)$ вектор с координатами $(x_1+y_1, x_2+y_2, \ldots, x_m+$ $+y_m)$, а **произведением** вектора $\mathbf{x}=(x_1, x_2, \ldots, x_m)$ на вещественное число λ — вектор с координатами $(\lambda x_1, \lambda x_2, \ldots, \lambda x_m)$, то координатное пространство A^m превращается в линейное пространство *.

Из курса аналитической геометрии читатель хорошо знаком с понятиями координатной плоскости и трехмерного координатного пространства. Обобщением этих понятий и является m-мерное координатное пространство A^m. Понятия координатной плоскости и трехмерного координатного пространства являются источниками удобной геометрической терминологии, употребляемой при изучении m-мерного координатного пространства A^m.

Заметим теперь, что для наших целей оказывается недостаточно понятия m-мерного координатного пространства A^m. Мы не обойдемся без измерения расстояний между точками этого пространства. Для введения понятия расстояния между точками координатного пространства A^m естественно отправляться от понятия расстояния между двумя точками координатной плоскости и двумя точками трехмерного координатного пространства (соответствую-

---

* Напомним, что линейным пространством называется совокупность векторов $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, ... любой природы при условии, что для элементов этой совокупности определены операция сложения векторов и операция умножения вектора на вещественное число, причем эти операции удовлетворяют восьми аксиомам: 1) $\mathbf{x}+$ $+\mathbf{y}=\mathbf{y}+\mathbf{x}$; 2) $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$; 3) существует нулевой вектор $\mathbf{0}$ такой, что $\mathbf{x}+\mathbf{0}=\mathbf{x}$ для любого вектора $\mathbf{x}$; 4) для любого вектора $\mathbf{x}$ существует противоположный ему вектор $\mathbf{x}'$ такой, что $\mathbf{x}+\mathbf{x}'=\mathbf{0}$; 5) $\lambda(\mathbf{x}+\mathbf{y})=\lambda\mathbf{x}+\lambda\mathbf{y}$; 6) $(\lambda+\mu)\mathbf{x}=$ $=\lambda\mathbf{x}+\mu\mathbf{x}$; 7) $\lambda(\mu\mathbf{x})=(\lambda\mu)\mathbf{x}$; 8) $1\cdot\mathbf{x}=\mathbf{x}$.

Легко проверить, что координатное пространство A^m с определениями операции сложения векторов и операции умножения вектора на число, данными в замечании 1, удовлетворяет восьми указанным аксиомам и, значит, является линейным пространством.

щие формулы хорошо известны читателю из курса аналитической геометрии).

Определение. *Координатное пространство A^m называется* **m-мерным евклидовым пространством,** *если между двумя любыми точками $M'(x_1', x_2', \ldots, x_m')$ и $M''(x''_1, x_2'', \ldots, x_m'')$ пространства A^m определено расстояние, обозначаемое символом $\rho(M', M'')$ и выражающееся соотношением*

$$\rho(M', M'') = \sqrt{(x_1'-x_1'')^2+(x_2'-x_2'')^2+\ldots+(x_m'-x_m'')^2}. \quad (12.1)$$

Будем обозначать m-мерное евклидово пространство символом E^m.

Введенное нами понятие m-мерного евклидова пространства E^m является естественным обобщением понятий евклидовой плоскости и трехмерного евклидова пространства, изученных в курсе аналитической геометрии.

Замечание 2. В курсе линейной алгебры дается общее определение **евклидова пространства** как такого линейного пространства, для которого указано правило, ставящее в соответствии любым двум элементам $\mathbf{x}$ и $\mathbf{y}$ этого пространства вещественное число, называемое **скалярным произведением** этих элементов и обозначаемое символом $(\mathbf{x}, \mathbf{y})$, при условии, что это правило удовлетворяет четырем аксиомам: 1) $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$; 2) $(\mathbf{x}+\mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}) + (\mathbf{y}, \mathbf{z})$; 3) $(\lambda\mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$; 4) $(\mathbf{x}, \mathbf{x}) \geqslant 0$, причем $(\mathbf{x}, \mathbf{x}) = 0$ только для нулевого элемента $\mathbf{x} = \mathbf{o}$.

Легко проверить, что если в пространстве A^m, элементы которого рассматриваются как векторы $\mathbf{x}$ с координатами $(x_1, x_2, \ldots, x_m)$, определить скалярное произведение двух элементов $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ и $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ соотношением

$$(\mathbf{x}, \mathbf{y}) = x_1y_1 + x_2y_2 + \ldots + x_my_m, \quad (12.2)$$

то будут выполнены четыре указанные аксиомы и пространство A^m превратится в евклидово пространство (с точки зрения общего определения евклидова пространства).

Напомним, что линейное пространство называется **нормированным**, если указано правило, ставящее в соответствие каждому элементу $\mathbf{x}$ вещественное число, называемое **нормой** этого элемента и обозначаемое символом $\|\mathbf{x}\|$, причем указанное правило удовлетворяет трем аксиомам: 1) $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$; 2) $\|\lambda\mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$; 3) $\|\mathbf{x}\| \geqslant 0$, причем $\|\mathbf{x}\| = 0$ только для нулевого элемента $\mathbf{x}$ *.

---

* В аксиоме 3) можно опустить требование $\|\mathbf{x}\| \geqslant 0$. В самом деле, это требование является логическим следствием аксиом 1) и 2) и предположения о том, что $\|\mathbf{x}\| = 0$ только для нулевого элемента $\mathbf{x}$ (достаточно в аксиоме 1) положить $\mathbf{y} = -\mathbf{x}$).

В курсе линейной алгебры доказывается, что всякое евклидово пространство является нормированным: достаточно определить норму любого элемента $\mathbf{x}$ соотношением $\|\mathbf{x}\| = \sqrt{(x,\ x)}$ *.

Нормированное пространство всегда является так называемым **метрическим пространством**, т. е. таким пространством, в котором указано правило, ставящее в соответствие любым двум элементам $\mathbf{x}'$ и $\mathbf{x}''$ вещественное число, называемое **расстоянием** между этими элементами и обозначаемое символом $\rho(\mathbf{x}', \mathbf{x}'')$, при условии, что это правило удовлетворяет трем аксиомам: 1) $\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{y}, \mathbf{x})$; 2) $\rho(\mathbf{x}, \mathbf{y}) \leqslant \rho(\mathbf{x}, \mathbf{z}) + \rho(\mathbf{z}, \mathbf{x})$; 3) $\rho(\mathbf{x}, \mathbf{y}) \geqslant 0$, причем $\rho(\mathbf{x}, \mathbf{y}) = 0$, только когда $\mathbf{x} = \mathbf{y}$ **.

Достаточно определить расстояние $\rho(\mathbf{x}', \mathbf{x}'')$ соотношением

$$\rho(\mathbf{x}', \mathbf{x}'') = \|\mathbf{x}' - \mathbf{x}''\| = \sqrt{(\mathbf{x}' - \mathbf{x}'',\ \mathbf{x}' - \mathbf{x}'')}.$$

Если учесть, что координатное пространство A^m является евклидовым пространством со скалярным произведением, определяемым соотношением (12.2), то мы придем к следующему выражению для расстояния $\rho(\mathbf{x}', \mathbf{x}'')$ между двумя элементами $\mathbf{x}' = (x_1', x_2', \ldots, x_m')$ и $\mathbf{x}'' = (x_1'', x_2'', \ldots, x_m'')$ пространства A^m.

$$\rho(\mathbf{x}', \mathbf{x}'') = \sqrt{(\mathbf{x}' - \mathbf{x}'',\ \mathbf{x}' - \mathbf{x}'')} =$$
$$= \sqrt{(x_1' - x_1'')^2 + (x_2' - x_2'')^2 + \ldots + (x_m' - x_m'')^2}.$$

Полученное выражение в точности совпадает с величиной, стоящей в правой части (12.1).

2. Множества точек m-мерного евклидова пространства. Если у функции $y = f(x)$ одной независимой переменной x, областью определения которой является некоторое множество $\{x\}$ точек одномерного евклидова пространства E^1, заменить это множество $\{x\}$ некоторым множеством $\{M\}$ точек m-мерного евклидова пространства E^m, то мы естественно придем к понятию функции m независимых переменных.

Отсюда ясно, что введению функции m переменных должно предшествовать описание важнейших типов множеств точек m-мерного евклидова пространства E^m.

Перейдем к описанию таких множеств.

1°. Множество $\{M\}$ всевозможных точек M пространства E^m координаты $x_1, x_2, \ldots, x_m$ которых удовлетворяют неравенству

$$(x_1 - \overset{\circ}{x}_1)^2 + (x_2 - \overset{\circ}{x}_2)^2 + \ldots + (x_m - \overset{\circ}{x}_m)^2 < R^2,$$

называется **открытым m-мерным шаром** радиуса R с центром в точке $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \ldots, \overset{\circ}{x}_m)$.

---

* См., например: В. А. Ильин, Э. Г. Позняк. Линейная алгебра.

** В аксиоме 3) можно опустить требование $\rho(\mathbf{x}, \mathbf{y}) \geqslant 0$. В самом деле, это требование является логическим следствием аксиом 1) и 2) и предположения о

Иными словами, открытый m-мерный шар радиуса R с центром в точке M_0 — это множество всех точек M, для каждой из которых расстояние от фиксированной точки M_0 удовлетворяет неравенству $\rho(M, M_0) < R$.

2°. Множество $\{M\}$ всевозможных точек M пространства E^m, координаты $x_1, x_2, \ldots, x_m$ которых удовлетворяют неравенству $(x_1-\overset{\circ}{x}_1)^2 + \ldots + (x_m-\overset{\circ}{x}_m)^2 \leqslant R^2$, называется замкнутым m-мерным шаром радиуса R с центром в M_0.

3°. Множество $\{M\}$ всевозможных точек M пространства E^m, координаты $x_1, x_2, \ldots, x_m$ которых удовлетворяют равенству $(x_1-\overset{\circ}{x}_1)^2 + (x_2-\overset{\circ}{x}_2)^2 + \ldots + (x_m-\overset{\circ}{x}_m)^2 = R^2$, называется m-мерной сферой радиуса R с центром в точке M_0.

Замечание 1. Отметим, что если к открытому m-мерному шару радиуса R с центром в точке M_0 присоединить m-мерную сферу радиуса R с центром в точке M_0, то получится замкнутый m-мерный шар радиуса R с центром в точке M_0.

4°. Открытый m-мерный шар радиуса $\varepsilon > 0$ с центром в точке M_0 будем называть ε-окрестностью точки M_0.

5°. Множество $\{M\}$ всех точек M, координаты $x_1, x_2, \ldots, x_m$ которых удовлетворяют неравенствам

$$|x_1-\overset{\circ}{x}_1| < d_1, \quad |x_2-\overset{\circ}{x}_2| < d_2, \ldots, |x_m-\overset{\circ}{x}_m| < d_m,$$

где $d_1, d_2, \ldots, d_m$ — некоторые положительные числа, называется открытым m-мерным координатным параллелепипедом с центром в точке $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \ldots, \overset{\circ}{x}_m)$ или прямоугольной окрестностью точки M_0.

Справедливо следующее элементарное *утверждение*: *любая ε-окрестность точки M_0 содержит некоторую прямоугольную окрестность этой точки; любая прямоугольная окрестность точки M_0, содержит некоторую ε-окрестность точки M_0* *.

6°. Точка M множества $\{M\}$ точек пространства E^m называется внутренней точкой этого множества, если существует некоторая ε-окрестность точки M, все точки которой принадлежат множеству $\{M\}$.

7°. Точка M пространства E^m называется внешней точкой множества $\{M\}$, если существует некоторая ε-окрестность точки M, все точки которой не принадлежат множеству $\{M\}$.

том, что $\rho(\mathbf{x}, \mathbf{y}) = 0$, только когда $\mathbf{x}=\mathbf{y}$ (достаточно в аксиоме 2) положить $\mathbf{y}=\mathbf{x}$).

* В самом деле, если для фиксированного $\varepsilon > 0$ положить $d_1 = d_2 = \ldots = d_m = \varepsilon/\sqrt{m}$, то прямоугольная окрестность точки M_0 с указанными $d_1, d_2, \ldots, d_m$ будет содержаться в ε-окрестности точки M_0. Если для фиксированных $d_1 > 0$, $d_2 > 0, \ldots, d_m > 0$ положить $\varepsilon = \min\{d_1, d_2, \ldots, d_m\}$, то ε-окрестность точки M_0 будет содержаться в прямоугольной окрестности точки M_0 с указанными $d_1, d_2, \ldots, d_m$.

8°. Точка M пространства E^m называется граничной точкой множества $\{M\}$, если эта точка не является ни внутренней, ни внешней точкой указанного множества *.

Замечание 2. Граничная точка M множества $\{M\}$ может как принадлежать, так и не принадлежать этому множеству. Так любая точка m-мерной сферы радиуса R с центром в точке M_0 является граничной как для открытого m-мерного шара радиуса R с центром в точке M_0, так и для замкнутого m-мерного шара радиуса R с центром в точке M_0. Но любая точка указанной сферы не принадлежит открытому шару радиуса R с центром в M_0 и принадлежит замкнутому шару радиуса R с центром в M_0.

9° Произвольное множество M точек пространства E^m называется открытым, если любая точка этого множества является его внутренней точкой **.

10°. Произвольное открытое множество, содержащее данную точку M_0, принято называть окрестностью точки M_0.

11°. Произвольное множество $\{M\}$ точек пространства E^m называется замкнутым, если это множество содержит все свои граничные точки.

Чтобы сформулировать другое эквивалентное определение замкнутого множества, введем понятие предельной точки произвольного множества M в пространстве E^m.

12°. Точку A пространства E^m назовем предельной точкой множества $\{M\}$, если в любой ε-окрестности точки A содержится хотя бы одна точка этого множества, отличная от A.

Убедимся в том, что *множество M замкнуто тогда и только тогда, когда оно содержит все свои предельные точки.*

В самом деле, если множество M замкнуто, то оно содержит все точки пространства E^m, кроме своих внешних точек. Поскольку среди внешних точек множества M нет его предельных точек, то множество M содержит все свои предельные точки. Если же множество $\{M\}$ не содержит хотя бы одной своей граничной точки M_0, то эта точка M_0, очевидно, является предельной точкой множества $\{M\}$.

Доказанное утверждение позволяет дать другое *эквивалентное* определение замкнутого множества: *множество M называется замкнутым, если оно содержит все свои предельные точки* ***.

---

* Заметим, что точка M является граничной точкой множества $\{M\}$ тогда и только тогда, когда в любой ε-окрестности точки M найдутся как точки, принадлежащие множеству $\{M\}$, так и точки, ему не принадлежащие. В самом деле, если в некоторой ε-окрестности точки M не найдется точек, принадлежащих $\{M\}$ [не принадлежащих $\{M\}$], то точка M является внешней [внутренней] точкой M и не является граничной точкой этого множества.

** Т. е. если любая точка M множества $\{M\}$ принадлежит этому множеству вместе с некоторой своей ε-окрестностью.

*** Еще одним *эквивалентным определением* замкнутого множества является следующее определение: *множество M пространства E^m называется замкнутым, если его дополнение является открытым множеством.*

13°. Множество $\{M\}$ точек пространства E^m называется ограниченным, если найдется m-мерный шар, содержащий все точки этого множества.

14°. Введем понятие непрерывной кривой в m-мерном евклидовом пространстве E^m.

Непрерывной кривой L в пространстве E^m мы будем называть множество $\{M\}$ точек этого пространства, координаты $x_1, x_2, \ldots, x_m$ которых представляют собой непрерывные функции параметра t:

$$x_1=\varphi_1(t),\ x_2=\varphi_2(t), \ldots, x_m=\varphi_m(t),\ \alpha\leqslant t\leqslant\beta. \tag{12.3}$$

Мы будем говорить, что точки $M'(x_1', x_2', \ldots, x_m')$ и $M''(x_1'', x_2'', \ldots$ $\ldots, x_m'')$ пространства E^m можно соединить непрерывной кривой L, если существует такая непрерывная кривая L, определяемая параметрическими уравнениями (12.3), что

$$x_1'=\varphi_1(\alpha),\ x_2'=\varphi_2(\alpha), \ldots, x_m'=\varphi_m(\alpha),$$

$$x_1''=\varphi_1(\beta),\ x_2''=\varphi_2(\beta), \ldots, x_m''=\varphi_m(\beta).$$

Понятие непрерывной кривой в пространстве E^m позволяет нам ввести понятие связного множества.

15°. Множество $\{M\}$ точек пространства E^m называется связным, если любые две точки этого множества можно соединить непрерывной кривой, все точки которой принадлежат этому множеству.

Убедимся в том, что всякое связное множество $\{M\}$ в пространстве E^m обладает следующим свойством: *если множество $\{M\}$ связно, то не существует двух непустых непересекающихся открытых множеств G' и G'' таких, что пересечение каждого из этих множеств с $\{M\}$ не пусто и множество $\{M\}$ содержится в объединении G' и G''.*

Проведем доказательство этого свойства от противного. Предположим, что два указанных множества G' и G'' существуют. Так как пересечение каждого из этих множеств с $\{M\}$ не пусто, то найдутся две точки M' и M'' множества $\{M\}$, первая из которых принадлежит G', а вторая G''.

Чтобы получить противоречие, завершающее наше доказательство, нам достаточно установить, что точки M' и M'' нельзя соединить непрерывной кривой, все точки которой принадлежат множеству $\{M\}$.

Пусть L — любая непрерывная кривая в пространстве E^m, определяемая уравнениями (12.3) и соединяющая указанные точки M' и M''. Упорядочим все точки кривой L по возрастанию параметра t, который изменяется в пределах сегмента $\alpha\leqslant t\leqslant\beta$. Обозначим через γ точную верхнюю грань тех значений параметра t, для которых отвечающие им точки кривой L принадлежат множеству G', а через N точку кривой L, отвечающую значению параметра $t=\gamma$.

Достаточно доказать, что указанная точка N не принадлежит множеству $\{M\}$, а для этого достаточно убедиться, что эта точка N не принадлежит ни множеству G', ни множеству G''.

Если бы точка N принадлежала G', то, поскольку множество G' является открытым, нашлась бы некоторая ε-окрестность точки N, также принадлежащая G', т. е. нашлись бы точки кривой L отвечающие значениям параметра t, превосходящим γ, принадлежащие G', а это противоречит тому, что γ является точной верхней гранью значений параметра t, для которых соответствующие точки кривой L принадлежат G'.

Аналогично если бы точка N принадлежала G'', то она принадлежала бы этому множеству вместе с некоторой своей ε-окрестностью, т. е. нашлись бы точки кривой L, отвечающие значениям параметра t, меньшим γ, принадлежащие G'' и потому не принадлежащие G', а это противоречит тому, что γ является точной верхней гранью значений параметра t, для которых соответствующие точки кривой L принадлежат множеству G'.

Итак, точка N не принадлежит ни G', ни G'', и сформулированное свойство доказано.

Мы не будем останавливаться на доказательстве обратного утверждения о том, что если множество $\{M\}$ в пространстве E^m обладает указанным выше свойством, то оно является связным, а лишь отметим, что установленное нами свойство может быть положено в основу другого «топологического» определения понятия связного множества.

16°. Всякое открытое и связное множество в пространстве E^m принято называть о б л а с т ь ю.

17°. Если множество $\{M\}$ представляет собой область, то множество $\overline{\{M\}}$, полученное присоединением к множеству $\{M\}$ всех его граничных точек, называется з а м к н у т о й о б л а с т ь ю.

Открытый m-мерный шар и открытый m-мерный координатный параллелепипед являются ограниченными, связными и открытыми множествами, т. е. дают примеры ограниченных областей в пространстве E^m.

m-мерная сфера в пространстве E^m дает пример замкнутого и ограниченного множества.

Замкнутый m-мерный шар представляет собой ограниченную замкнутую область в E^m.

Дополнение к открытому m-мерному шару представляет собой неограниченное замкнутое множество.

Совокупность двух непересекающихся областей в пространстве E^m дает пример несвязного множества.

3. Понятие функции m переменных. Теперь мы подготовлены для того, чтобы ввести понятие функции m переменных.

Если каждой точке M из множества $\{M\}$ точек m-мерного евклидова пространства E^m ставится в соответствие по известному закону некоторое число u, то говорят, что на множестве $\{M\}$ задана функция *$u=u(M)$ или $u=f(M)$. При этом множество $\{M\}$ называется* областью задания *функции $u=f(M)$.*

Число u, соответствующее данной точке M из множества $\{M\}$, будем называть частным значением функции в точке M. Совокупность всех частных значений функции $u=f(M)$ называется множеством значений этой функции. Так как точка M определяется координатами $x_1, x_2, \ldots, x_m$, то для функции $u=f(M)$ m переменных используется также обозначение $u=f(x_1, x_2, \ldots, x_m)$.

Рассмотрим примеры функций m переменных.

Начнем с примеров функций двух переменных.

1°. $u=\sqrt{4-x^2-y^2}$. Областью задания этой функции является круг радиуса 2 с центром в начале координат, а множество значений представляет собой сегмент $0\leqslant u\leqslant 2$.

2°. $u=\left[\sqrt{x^2+y^2-4}\right]^{-1}$. Областью задания этой функции является множество точек, лежащих вне круга радиуса 2 с центром в начале координат, а множество значений представляет собой открытую полупрямую $u>0$.

3°. $u=\sqrt{\cos(x^2+y^2)}$. Областью задания этой функции является множество $\{M\}$ точек, координаты которых удовлетворяют неравенству $\cos(x^2+y^2)\geqslant 0$. Это неравенство эквивалентно неравенствам $0\leqslant x^2+y^2\leqslant \frac{\pi}{2}$, $2k\pi-\frac{\pi}{2}\leqslant x^2+y^2\leqslant 2k\pi+\frac{\pi}{2}$ при $k=1, 2, \ldots$. Таким образом, $\{M\}$ состоит из круга радиуса $\sqrt{\pi/2}$ с центром в точке O $(0, 0)$ и кольцеобразных областей (рис. 12.1).

Приведем теперь примеры функций m переменных.

4°. Пусть $u=\sqrt{1-x_1^2-x_2^2-\ldots-x_m^2}$. Областью задания этой функции служит, очевидно, m-мерный шар радиуса 1 с центром в точке O $(0, 0, \ldots, 0)$. Множеством значений рассматриваемой функции является сегмент $[0, 1]$.

5. Пусть $u=\dfrac{1}{\sqrt{1-\dfrac{x_1^2}{a_1^2}-\dfrac{x_2^2}{a_2^2}-\ldots-\dfrac{x_m^2}{a_m^2}}}$. Областью задания этой функции является множество $\{M\}$ всех точек M пространства E^m, координаты $x_1, x_2, \ldots, x_m$ которых удовлетворяют неравенству

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}+\ldots+\frac{x_m^2}{a_m^2}<1$$

(при этом предполагается, что $a_1, a_2, \ldots, a_m$ — некоторые положительные числа).

Очевидно, указанное множество $\{M\}$ будет представлять собой область в пространстве E^m, все точки которой лежат внутри так называемого m-мерного эллипсоида, представляющего собой множество точек, координаты которых удовлетворяют уравнению

$$\frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}+\ldots+\frac{x_m^2}{a_m^2}=1.$$

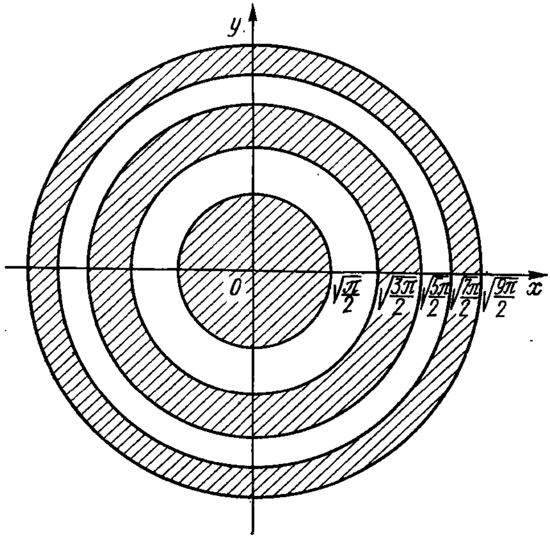


Рис. 12.1

Множеством всех значений указанной функции является полупрямая $u \geqslant 1$.

Мы видим, что область задания функции m переменных представляет собой некоторое множество точек m-мерного евклидова пространства E^m, а множество всех значений этой функции представляет собой некоторое множество одномерного евклидова пространства E^1.

Таким образом, введенную нами функцию m переменных можно рассматривать как отображение некоторого множества в пространстве E^m в некоторое множество в пространстве E^1.

§ 2. ПРЕДЕЛ ФУНКЦИИ m ПЕРЕМЕННЫХ

1. Последовательности точек пространства E^m. Введем понятие последовательности точек m-мерного евклидова пространства E^m. *Пусть каждому числу n натурального ряда чисел* 1, 2, ... *ставится в соответствие точка M_n евклидова пространства E^m. Возникающий при этом ряд точек $M_1, M_2, \ldots, M_n, \ldots$, рассматриваемый в указанном порядке, называется* ***последовательностью точек*** *евклидова пространства E^m.* Мы будет кратко обозначать эту последовательность символом $\{M_n\}$.

Введем, далее, понятие сходящейся последовательности точек пространства E^m и ее предела.

Последовательность $\{M_n\}$ точек евклидова пространства E^m называется ***сходящейся****, если существует точка A пространст-*

ва E^m такая, что для любого положительного числа ε можно указать отвечающий ему номер N такой, что при $n \geqslant N$ выполняется неравенство $\rho(M_n, A) < \varepsilon$. При этом точка A называется пределом последовательности $\{M_n\}$.

Для обозначения предела A последовательности $\{M_n\}$ используется следующая символика:

$$\lim_{n\to\infty} M_n = A, \text{ или } M_n \to A \text{ при } n \to \infty.$$

Установим следующую лемму.

Лемма 1. *Последовательность $\{M_n\}$ точек m-мерного евклидова пространства E^m сходится к точке A этого пространства тогда и только тогда, когда числовые последовательности $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, ..., $\{x_m^{(n)}\}$ координат точек M_n сходятся соответственно к числам $a_1, a_2, ..., a_m$, представляющим собой координаты точки A.*

Доказательство. Сначала докажем первую часть леммы. Так как последовательность $\{M_n\}$ сходится к точке A, то для любого $\varepsilon > 0$ можно указать номер N такой, что при $n \geqslant N$ выполняется неравенство $\rho(M_n, A) < \varepsilon$. Пусть $\{x_1^{(n)}, x_2^{(n)}, \ldots, x_m^{(n)}\}$ — координаты точки M_n, а $(a_1, a_2, ..., a_m)$ — координаты точки A. Тогда неравенство $\rho(M_n, A) < \varepsilon$ можно записать следующим образом:

$$\sqrt{(x_1^{(n)}-a_1)^2+(x_2^{(n)}-a_2)^2+\ \ldots\ +(x_m^{(n)}-a_m)^2} < \varepsilon. \qquad (12.4)$$

Отсюда следует, что при $n \geqslant N$ выполняются неравенства

$$|x_1^{(n)}-a_1| < \varepsilon, \quad |x_2^{(n)}-a_2| < \varepsilon, \ \ldots, \ |x_m^{(n)}-a_m| < \varepsilon.$$

Иными словами, последовательности $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, ..., $\{x_m^{(n)}\}$ координат точек M_n сходятся соответственно к числам $a_1, a_2, ..., a_m$.

Первая часть леммы доказана.

Перейдем к доказательству второй части леммы. Предположим, что указанные последовательности координат точек M_n сходятся соответственно к числам $a_1, a_2, ..., a_m$. Тогда для любого $\varepsilon > 0$ можно указать номера $N_1, N_2, ..., N_m$ такие, что при $n \geqslant N_1, n \geqslant N_2, ..., n \geqslant N_m$ соответственно выполняются неравенства

$$|x_1^{(n)}-a_1| < \frac{\varepsilon}{\sqrt{m}}, \quad |x_2^{(n)}-a_2| < \frac{\varepsilon}{\sqrt{m}}, \ \ldots, \ |x_m^{(n)}-a_m| < \frac{\varepsilon}{\sqrt{m}}.$$

Отсюда следует, что при $n \geqslant N = \max\{N_1, N_2, ..., N_m\}$ выполняется неравенство (12.4). Иными словами, при $n \geqslant N$ выполняется неравенство $\rho(M_n, A) < \varepsilon$, где A — точка E^m с координатами $a_1, a_2, ...$ ..., a_m. Таким образом, последовательность $\{M_n\}$ сходится к точке A. Лемма доказана.

Введем теперь понятие фундаментальной последовательности точек пространства E^m. *Последовательность $\{M_n\}$ точек m-мерного евклидова пространства называется* ***фундаментальной*** *или* ***последовательностью Коши****, если для любого положи-*

тельного числа ε *можно указать отвечающий ему номер* N *такой, что при* $n \geqslant N$ *и при любом целом* $p \geqslant 0$ *выполняется неравенство* $\rho(M_{n+p}, M_n) < \varepsilon$.

В полной аналогии с леммой 1 может быть доказано следующее утверждение.

Лемма 2. *Последовательность* $\{M_n\}$ *точек* m*-мерного евклидова пространства* E^m *является фундаментальной тогда и только тогда, когда является фундаментальной каждая из числовых последовательностей* $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, $\ldots$, $\{x_m^{(n)}\}$ *соответствующих координат точек* M_n.

Доказательство. Для доказательства первой части леммы предположим, что последовательность точек $\{M_n\}$ является фундаментальной, т. е. для любого $\varepsilon > 0$ найдется номер N такой, что при $n \geqslant N$ и для любого целого $p \geqslant 0$ справедливо неравенство $\rho(M_{n+p}, M_n) < \varepsilon$ или, что то же самое, неравенство

$$\sqrt{(x_1^{(n+p)} - x_1^{(n)})^2 + (x_2^{(n+p)} - x_2^{(n)}) + \ldots + (x_m^{(n+p)} - x_m^{(n)})^2} < \varepsilon. \quad (12.4^*)$$

Из этого неравенства вытекает, что при $n \geqslant N$ и для любого целого $p \geqslant 0$ справедливы неравенства

$$|x_1^{(n+p)} - x_1^{(n)}| < \varepsilon, \quad |x_2^{(n+p)} - x_2^{(n)}| < \varepsilon, \quad \ldots, \quad |x_m^{(n+p)} - x_m^{(n)}| < \varepsilon,$$

которые и устанавливают фундаментальность каждой из числовых последовательностей соответствующих координат точек M_n.

Для доказательства второй части леммы предположим фундаментальность каждой из числовых последовательностей соответствующих координат точек M_n.

Тогда для любого $\varepsilon > 0$ можно указать номера $N_1, N_2, \ldots, N_m$ такие, что соответственно при $n \geqslant N_1$, $n \geqslant N_2$, ..., $n \geqslant N_m$ и для любых целых $p \geqslant 0$ будут справедливы неравенства

$$|x_1^{(n+p)} - x_1^{(n)}| < \frac{\varepsilon}{\sqrt{m}}, \quad |x_2^{(n+p)} - x_2^{(n)}| < \frac{\varepsilon}{\sqrt{m}}, \quad \ldots, \quad |x_m^{(n+p)} - x_m^{(n)}| < \frac{\varepsilon}{\sqrt{m}}.$$

Отсюда следует, что при $n \geqslant N = \max\{N_1, N_2, \ldots, N_m\}$ и для любого целого $p \geqslant 0$ будет справедливо неравенство (12.4*), которое и означает фундаментальность последовательности точек $\{M_n\}$. Лемма 2 доказана.

С помощью лемм 1 и 2 легко доказывается **критерий Коши сходимости последовательности точек пространства** E^m: *для того чтобы последовательность* $\{M_n\}$ *точек пространства* E^m *была сходящейся, необходимо и достаточно, чтобы она была фундаментальной.*

В самом деле, если последовательность точек $\{M_n\}$ является фундаментальной, то в силу леммы 2 является фундаментальной и каждая из числовых последовательностей соответствующих координат точек $\{M_n\}$. В силу критерия Коши сходимости числовой последовательности указанные числовые последовательности соответствующих координат сходятся к некоторым числам $a_1, a_2, \ldots, a_m$ соответственно. Но тогда в силу леммы 1 последовательность точек $\{M_n\}$ сходится к точке $A(a_1, a_2, \ldots, a_m)$.

Если, наоборот, последовательность точек $\{M_n\}$ сходится к некоторой точке A пространства E^m, то в силу леммы 1 каждая из числовых последовательностей соответствующих координат точек $\{M_n\}$ сходится к соответствующей координате точки A. Но тогда (в силу критерия Коши сходимости числовой последовательности) каждая из числовых последовательностей соответствующих координат точек $\{M_n\}$ является фундаментальной и, значит, в силу леммы 2, является фундаментальной и последовательность точек $\{M_n\}$.

2. Свойство ограниченной последовательности точек E^m. Введем понятие ограниченной последовательности точек пространства E^m.

Последовательность $\{M_n\}$ точек m-мерного евклидова пространства называется ***ограниченной****, если существует такое число $a>0$, что для всех n выполняется неравенство $\rho(O, M_n) \leqslant a$, где O — точка, все координаты которой равны нулю.*

Иными словами, ограниченность последовательности точек $\{M_n\}$ означает, что все точки этой последовательности принадлежат замкнутому шару достаточно большого радиуса с центром в начале координат O.

Установим следующее важное свойство ограниченной последовательности точек пространства E^m.

Если $n_1, n_2, \ldots, n_k, \ldots$ — произвольная строго возрастающая последовательность целых положительных чисел, то мы будем называть последовательность точек

$$M_{n_1},\ M_{n_2},\ \ldots,\ M_{n_k},\ \ldots$$

подпоследовательностью последовательности точек $M_1, M_2, \ldots, M_n, \ldots$.

Теорема 12.1 (теорема Больцано — Вейерштрасса). *Из любой ограниченной последовательности $\{M_n\}$ точек m-мерного евклидова пространства можно выделить сходящуюся подпоследовательность.*

Доказательство. Убедимся, во-первых, что последовательности $\{x_1^{(n)}\}, \{x_2^{(n)}\}, \ldots, \{x_m^{(n)}\}$ координат точек $\{M_n\}$ являются ограниченными. Действительно, так как последовательность $\{M_n\}$ ограничена, то для всех n выполняется неравенство $\rho(O, M_n) \leqslant a$. Поскольку $\rho(O, M_n) = \sqrt{[x_1^{(n)}]^2 + [x_2^{(n)}]^2 + \ldots + [x_m^{(n)}]^2}$, отсюда следует, что для всех n выполняются неравенства $|x_1^{(n)}| \leqslant a$, $|x_2^{(n)}| \leqslant a$, $\ldots$, $|x_m^{(n)}| \leqslant a$. Иными словами, последовательности $\{x_1^{(n)}\}, \{x_2^{(n)}\}, \ldots, \{x_m^{(n)}\}$ координат точек M_n ограничены. В силу теоремы Больцано — Вейерштрасса для числовых последовательностей (см. п. 1 § 3 гл. 3) из последовательности $\{x_1^{(n)}\}$ можно выделить подпоследовательность $\{x_1^{(n_{k_1})}\}$, сходящуюся к некоторому числу a_1. Рассмотрим соответствующую подпоследовательность $\{x_2^{(n_{k_1})}\}$ последовательности вторых координат точек M_n. В силу той же теоремы из подпоследовательности $\{x_2^{(n_{k_1})}\}$ можно выделить подпоследовательность $\{x_2^{(n_{k_2})}\}$, сходящуюся к некоторому числу a_2. Заметим, что подпоследовательность $\{x_1^{(n_{k_2})}\}$ последовательности $\{x_1^{(n_{k_1})}\}$ сходится к числу a_1. Итак, подпоследовательности $\{x_1^{(n_{k_2})}\}$ и $\{x_2^{(n_{k_2})}\}$ сходятся к числам a_1 и a_2 соответственно.

Очевидно, что если мы из подпоследовательности $\{x_3^{(n_{k_2})}\}$ последовательности третьих координат точек M_n выделим сходящуюся к числу a_3 подпоследовательность $\{x_3^{(n_{k_3})}\}$, то подпоследовательности $\{x_1^{(n_{k_3})}\}$, $\{x_2^{(n_{k_3})}\}$, $\{x_3^{(n_{k_3})}\}$ сходятся соответственно к числам a_1, a_2, a_3. Продолжая эти рассуждения, мы, наконец, получим сходящуюся к некоторому числу a_m подпоследовательность $\{x_m^{(n_{k_m})}\}$ последовательности $m=x$ координат точек M_n, причем подпоследовательности $\{x_1^{(n_{k_m})}\}$, $\{x_2^{(n_{k_m})}\}$, ..., $\{x_m^{(n_{k_m})}\}$ сходятся к числам a_1, a_2, ..., a_m соответственно.

Но тогда в силу леммы 1 подпоследовательность $\{M_{n_{k_m}}\}$ последовательности точек $\{M_n\}$ сходится к точке A с координатами a_1, a_2, ..., a_m. Теорема доказана.

З а м е ч а н и е. *Предел A последовательности $\{M_n\}$ точек, принадлежащих замкнутому множеству $\{M\}$, также принадлежит этому множеству.* Чтобы убедиться в этом, достаточно заметить, что в любой ε-окрестности точки A имеются точки $\{M_n\}$, т. е. точки множества $\{M\}$, и поэтому точка A является либо внутренней, либо граничной точкой $\{M\}$, а следовательно, принадлежит $\{M\}$.

3. Предел функции m переменных. Рассмотрим функцию $u=f(M)$, определенную на множество $\{M\}$ точек m-мерного евклидова пространства E^m, и точку A пространства E^m, быть может и не принадлежащую множеству $\{M\}$, но обладающую тем свойством, что в любой ε-окрестности этой точки A содержится хотя бы одна точка множества $\{M\}$, отличная от A *.

О п р е д е л е н и е 1 (п р е д е л ф у н к ц и и в т о ч к е A п о Г е й н е). *Число b называется п р е д е л о м (или п р е д е л ь н ы м з н а ч е н и е м) функции $u=f(M)$ в точке A (или при $M \to A$), если для любой сходящейся к A последовательности $\{M_n\}$ точек множества $\{M\}$ задания этой функции, все элементы M_n которой отличны от A, соответствующая числовая последовательность значений функции $\{f(M_n)\}$ сходится к числу b.*

---

* Это означает, что точка A является п р е д е л ь н о й т о ч к о й множества $\{M\}$ (см. п. 2 § 1 настоящей главы, определение 12°).

Определение 1* (предел функции в точке A по Коши). *Число b называется пределом (или предельным значением) функции $u=f(M)$ в точке A (или при $M \to A$), если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для любой точки M из множества $\{M\}$ задания этой функции, удовлетворяющей условию $0<\rho(M, A)<\delta$, справедливо неравенство $|f(M)-b|<\varepsilon$.*

Для обозначения предела функции $u=f(M)$ в точке A используется следующая символика:

$$\lim_{M \to A} f(M) = b \quad \text{или} \quad \lim_{\substack{x_1 \to a_1 \\ x_2 \to a_2 \\ \dots\dots \\ x_m \to a_m}} f(x_1, x_2, \dots, x_m) = b.$$

(Здесь $(a_1, a_2, \dots, a_m)$ — координаты точки A.)

Доказательство *эквивалентности определения* 1 *и* 1* проводится точно так же, как и для функции одной переменной. Следует лишь в рассуждениях теоремы 3.19 (см. п. 2 § 4 гл. 3) заменить последовательность $\{x_n\}$ последовательностью точек $\{M_n\}$, точку a — точкой $A(a_1, a_2, \dots, a_m)$, разности $|x-a|$ и $|x_n-a|$ — расстояниями $\rho(M, A)$ и $\rho(M_n, A)$ соответственно, а числовую последовательность $\{f(x_n)\}$ заменить числовой последовательностью $\{f(M_n)\}$.

Введем теперь понятие предела функции $u=f(M)$ при $M \to \infty$. Для этого предположим, что множество $\{M\}$, на котором задана функция $u=f(M)$, для любого $\delta>0$ имеет хотя бы один элемент M, лежащий вне шара радиуса δ с центром в точке $O(0, 0, \dots, 0)$.

Ограничимся определением соответствующего предела функции по Коши.

Определение 2. *Число b называется* **пределом** *функции $u=f(M)$ при $M \to \infty$, если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для всех точек M из множества $\{M\}$ задания функции, удовлетворяющих условию $\rho(O, M)>\delta$, справедливо неравенство $|f(M)-b|<\varepsilon$.*

Для обозначения предела функции $u=f(M)$ при $M \to \infty$ используется символ

$$\lim_{M \to \infty} f(M) = b.$$

Так же, как и для функции одной переменной, легко убедиться в том, что арифметические операции над функциями m переменных, имеющими предел в данной точке A [или при $M \to \infty$], приводят к функциям, также имеющим предел в точке A [соответственно при $M \to \infty$].

Сформулируем соответствующее утверждение для случая предела в точке A.

Пусть две функции $f(M)$ и $g(M)$ заданы на одном и том же множестве $\{M\}$ и имеют в точке A пределы, соответственно рав-

ные b и c. Тогда функции $f(M)+g(M)$, $f(M)-g(M)$, $f(M)\cdot g(M)$ и *$f(M)/g(M)$ имеют в точке A пределы, соответственно равные $b+c$, $b-c$, $b\cdot c$* и *b/c (в случае частного нужно дополнительно требовать, чтобы c было отлично от нуля).*

Доказательство этого утверждения совершенно аналогично доказательству теоремы 3.21 (см. п. 4 § 4 гл. 3), только вместо определения по Гейне предела функции одной переменной следует использовать определение по Гейне предела функции m переменных.

Установим теперь критерий Коши существования предела функции m переменных.

Определение 3. *Будем говорить, что функция m переменных $f(M)$ удовлетворяет* **условию Коши** *в точке $M=A$ [соответственно при $M\to\infty$], если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для любых двух точек M' и M'' из множества $\{M\}$ задания функции, удовлетворяющих условиям $0<\rho(M', A)<\delta$, $0<\rho(M'', A)<\delta$ [соответственно условиям $\rho(M', O)>\delta$, $\rho(M'', O)>\delta$], справедливо неравенство $|f(M')-f(M'')|<\varepsilon$.*

Теорема 12.2 (критерий Коши существования предела функции m переменных). *Для того чтобы функция $u=f(M)$ имела конечный предел в точке $M=A$ [при $M\to\infty$], необходимо и достаточно, чтобы эта функция удовлетворяла в точке $M=A$ [при $M\to\infty$] условию Коши.*

Доказательство этой теоремы полностью идентично доказательству теоремы 3.20 (см. п. 3 § 4 гл. 3) и может быть получено из него формальной заменой букв x и a буквами M и A и выражений типа $|x-a|$ символом $\rho(M, A)$.

Весьма полезно заметить, что определение предела функции m переменных $u=f(M)$ в точке A и при $M\to\infty$ укладывается в общее определение **предела по базе**, введенное в § 5 гл. 3.

Рассмотрим сначала предел функции $u=f(M)$ в точке A.

Договоримся называть **проколотой δ-окрестностью точки** A открытый m-мерный шар радиуса δ с центром в точке A, из которого удалена сама точка A.

Подчеркнем, что $f(M)$ задана на таком множестве $\{M\}$, которое для любого $\delta>0$ имеет хотя бы одну точку M, лежащую в проколотой δ-окрестности точки A.

Обозначим символом C_δ проколотую δ-окрестность точки A и положим $B_\delta=\{M\}\cap C_\delta$. Очевидно, совокупность $B=\{B_\delta\}$ множеств B_δ при всех $\delta>0$ образует **базу множества** $\{M\}$, ибо каждое множество B_δ (при любом $\delta>0$) не является пустым и пересечение любых двух множеств совокупности $\{B_\delta\}$ представляет собой множество той же совокупности *.

Эту базу естественно обозначить символом $M\to A$, ибо, как легко проверить, определение предела по такой базе совпадает

* Определения понятий базы и предела функции по базе см. в § 5 гл. 3.

с определением 1* предела функции $u=f(M)$ в точке A по Коши.

Для случая предела функции $u=f(M)$ при $M\to\infty$ эта функция должна быть задана на таком множестве $\{M\}$, которое для любого $\delta>0$ имеет хотя бы одну точку M, удовлетворяющую условию $\rho(M, O)>\delta$, где O — точка с координатами $(0, 0, \ldots, 0)$.

Обозначим символом C_δ множество всех точек M пространства E^m, удовлетворяющих условию $\rho(M, O)>\delta$, и положим $B_\delta=\{M\}\cap C_\delta$. Легко проверить, что совокупность $\{B_\delta\}$ множеств B_δ при всех $\delta>0$ образует базу множества $\{M\}$, ибо каждое множество B_δ не является пустым и пересечение любых двух множеств совокупности $\{B_\delta\}$ представляет собой множество той же совокупности.

Эту базу естественно обозначить символом $M\to\infty$, ибо, как легко проверить, определение предела по такой базе совпадает с определением 2 предела функции $u=f(M)$ при $M\to\infty$.

В заключение отметим, что теорема 12.2 (т. е. критерий Коши существования предела функции $u=f(M)$ в точке A и при $M\to\infty$) вытекает как частный случай из теоремы 3.22, устанавливающей критерий Коши существования общего предела функции по базе.

4. Бесконечно малые функции m переменных. *Функция $u=f(M)$ называется* ***бесконечно малой*** *в точке A* (при $M\to A$), *если* $\lim\limits_{M\to A} f(M)=0$.

Легко убедиться, что функция $f(M)=(x_1-a_1)^{n_1}+(x_2-a_2)^{n_2}+\ldots+(x_m-a_m)^{n_m}$, где $n_1, n_2, \ldots, n_m$ — положительные числа, является бесконечно малой в точке $A(a_1, a_2, \ldots, a_m)$ *.

Если функция $u=f(M)$ имеет равное b предельное значение в точке A, то функция $\alpha(M)=f(M)-b$ является бесконечно малой в точке A.

Имеем: $\lim\limits_{M\to A}\alpha(M)=\lim\limits_{M\to A}[f(M)-b]=\lim\limits_{M\to A}f(M)-\lim\limits_{M\to A}b=0$. Используя этот результат, мы получим **специальное представление** для функции, имеющей равный b предел в точке A: $f(M)=b+\alpha(M)$, где $\lim\limits_{M\to A}\alpha(M)=0$.

Сравнение бесконечно малых функций нескольких переменных производится точно так же, как это указано в п. 5 § 4 гл. 3 для бесконечно малых функций одной переменной. Отметим, что, как и в случае одной переменной, под символом $o(\beta)$ мы будем понимать любую бесконечно малую в данной точке A функцию более высокого порядка малости, чем бесконечно малая в данной точке A функция $\beta(M)$.

---

* Достаточно учесть, что каждая из функций одной переменной $f(x_k)=(x_k-a_k)^{n_k}$ является бесконечно малой в точке $x_k=a_k$.

5. Повторные пределы. Для функции $u=f(x_1, x_2, ..., x_m)$ нескольких переменных можно определить понятие предела по одной из переменных x_k при фиксированных значениях остальных переменных. В связи с этим возникает понятие повторного предела. Уясним это понятие на примере функции $u=f(x, y)$ двух переменных x и y. Пусть функция $u=f(x, y)$ задана в некоторой прямоугольной окрестности $|x-x_0|<d_1$, $|y-y_0|<d_2$ точки $M_0(x_0, y_0)$, за исключением, быть может, самой точки M_0. Пусть для каждого фиксированного y, удовлетворяющего условию $0<|y-y_0|<d_2$, существует предел функции $u=f(x, y)$ одной переменной x в точке $x=x_0$

$$\lim_{\substack{x\to x_0 \\ y\text{—фикс}}} f(x, y)=\varphi(y),$$

и пусть, кроме того, существует предел b функции $\varphi(y)$ в точке $y=y_0$

$$\lim_{y\to y_0}\varphi(y)=b.$$

В этом случае говорят, что существует повторный предел b для функции $u=f(x, y)$ в точке M_0, который обозначается следующим образом:

$$\lim_{y\to y_0}\lim_{x\to x_0} f(x, y)=b.$$

Аналогично определяется повторный предел

$$\lim_{x\to x_0}\lim_{y\to y_0} f(x, y).$$

Установим достаточные условия равенства двух введенных повторных пределов.

Теорема 12.3. *Пусть функция $u=f(x, y)$ определена в некоторой прямоугольной окрестности $|x-x_0|<d_1$, $|y-y_0|<d_2$ точки $M_0(x_0, y_0)$ и имеет в этой точке предел, равный b. Пусть, кроме того, для любого фиксированного x, $0<|x-x_0|<d_1$, существует предел $\psi(x)=\lim\limits_{y\to y_0} f(x, y)$ и для любого фиксированного y, $0<|y-y_0|<d_2$, существует предел $\varphi(y)=\lim\limits_{x\to x_0} f(x, y)$. Тогда повторные пределы $\lim\limits_{x\to x_0}\lim\limits_{y\to y_0} f(x, y)$ и $\lim\limits_{y\to y_0}\lim\limits_{x\to x_0} f(x, y)$ существуют и оба равны b.*

Доказательство. Так как функция $u=f(x, y)$ имеет в точке $M_0(x_0, y_0)$ предел b, то для любого $\varepsilon>0$ можно указать такое $\delta>0$, что при $|x-x_0|<\delta$ и $|y-y_0|<\delta$ выполняется неравенство $|f(x, y)-b|<\varepsilon$. Таким образом, в прямоугольной окрестности $|x-x_0|<\delta$ и $|y-y_0|<\delta$ точки M_0 значение функции $f(x, y)$ отличается от b не больше чем на ε. Но тогда пределы $\psi(x)$ и $\varphi(y)$, указанные в формулировке теоремы при x и y, удовлетворяющих неравенствам $|x-x_0|<\delta$ и $|y-y_0|<\delta$, также отличаются от b не больше чем на ε. Следовательно, и пределы этих функций в точках x_0 и y_0 соответственно существуют и равны b. Теорема доказана.

Можно определить понятие повторного предела для так называемых двойных последовательностей $\{a_{mn}\}$, элементы a_{mn} которых определяются двумя индексами m и n. Именно символ $\lim\limits_{m\to\infty}\lim\limits_{n\to\infty} a_{mn}$ означает, что сначала определяется последовательность $\{b_n\}$, $b_n=\lim\limits_{m\to\infty} a_{mn}$, а затем находится предел этой последовательности $\{b_n\}$.

Рассмотрим, например, двойную последовательность $\{a_{mn}\}$, где $a_{mn}=\cos^m 2\pi n!x$, x — фиксированное число. Докажем, что

$$\lim_{n\to\infty}\lim_{m\to\infty}\cos^m 2\pi n!x=\begin{cases}1, & \text{если } x \text{ — рациональное число,}\\ 0, & \text{если } x \text{ — иррациональное число.}\end{cases}$$

В самом деле, если $x=p/q$, где p и q — целые числа, второе из которых положительно, то при $n\geqslant q$ имеем $\cos 2\pi n!x=1$, и поэтому $\lim\limits_{m\to\infty n\geqslant b}\cos^m 2\pi n!x=1$. Иными словами, если x — рациональное число, то $\lim\limits_{n\to\infty}\lim\limits_{m\to\infty}\cos^m 2\pi n!x=1$. Если же x — иррациональное число, то при любом n справедливо неравенство $|\cos 2\pi n!x|<1$, и поэтому

$$\lim_{m\to\infty}\cos^m 2\pi n!x=0,\quad \text{т. е.}\quad \lim_{m\to\infty}\lim_{n\to\infty}\cos^m 2\pi n!x=0.$$

Замечание. Используя полученный результат, мы можем аналитическим способом задать функцию Дирихле (см. п. 1 § 4 гл. 3) как повторный предел $\lim\limits_{n\to\infty}\lim\limits_{m\to\infty}\cos^m 2\pi n!x$.

§ 3. НЕПРЕРЫВНОСТЬ ФУНКЦИИ m ПЕРЕМЕННЫХ

1. Понятие непрерывности функции m переменных. Рассмотрим функцию m переменных $u=f(M)$, заданную на некотором множестве $\{M\}$ пространства E^m. Пусть A — некоторая точка E^m, принадлежащая множеству $\{M\}$ и такая, что в любой δ-окрестности точки A содержатся точки множества $\{M\}$, отличные от A. *

Формальное определение непрерывности функции в точке. *Функция $u=f(M)$ называется непрерывной в точке A, если предел этой функции в точке A существует и равен частному значению $f(A)$.*

* Это означает, что A является предельной точкой множества $\{M\}$.

Заметим, что поскольку $A = \lim\limits_{M \to A} M$, то условие непрерывности функции $f(M)$ в точке A можно символически записать в виде

$$\lim_{M \to A} f(M) = f(\lim_{M \to A} M).$$

Таким образом, для непрерывной в данной точке A функции символ $\lim$ предела и символ f характеристики функции можно менять местами.

Точки пространства E^m, в которых функция $u = f(M)$ не обладает свойством непрерывности, называются **точками разрыва** этой функции.

Используя определение предела функции в точке A по Гейне и по Коши, мы придем к определениям непрерывности функции в данной точке по Гейне и по Коши.

Определение 1* (непрерывность функции в данной точке по Гейне). *Функция $u = f(M)$ называется **непрерывной** в точке A, если для любой сходящейся к A последовательности $\{M_n\}$ точек множества $\{M\}$ задания этой функции соответствующая числовая последовательность $\{f(M_n)\}$ значений этой функции сходится к числу $f(A)$.*

Определение 1* (непрерывность функции в данной точке по Коши). *Функция $u = f(M)$ называется **непрерывной в точке** A, если для любого положительного числа ε найдется отвечающее ему положительное число δ такое, что для любой точки M из множества $\{M\}$ задания этой функции, удовлетворяющей условию $\rho(M, A) < \delta$, справедливо неравенство $|f(M) - f(A)| < \varepsilon$.*

Замечание. В отличие от определения предела по Гейне, мы опускаем в определении 1 требование $M_n \neq A$. Это можно сделать в силу того, что функция $f(M)$ определена в точке A и добавление к последовательности $\{f(M_n)\}$, сходящейся к числу $f(A)$, любого количества новых элементов, равных $f(A)$, не нарушит сходимости этой последовательности к $f(A)$.

Аналогично, в отличие от определения предела по Коши, мы опускаем в определении 1* требование $\rho(M, A) > 0$ или, что то же самое, $M \neq A$. Это можно сделать в силу того, что функция $f(M)$ определена в точке A и при $M = A$ неравенство $|f(M) - f(A)| < \varepsilon$ справедливо для любого $\varepsilon > 0$.

Можно сформулировать и еще одно эквивалентное определение непрерывности функции $f(M)$ в данной точке A.

Определение 1.** *Функция $f(M)$ называется **непрерывной в точке** A, если для любой окрестности точки $f(A)$ пространства E^1 найдется такая окрестность точки A пространства E^m, что образ всех точек множества задания функции, лежащих в этой окрестности точки A, при отображении, осуществляемом функцией f, целиком лежит в указанной окрестности точки $f(A)$.*

Пусть теперь $\{M\}$ — множество точек пространства E^m, в любой δ-окрестности каждой точки M которого содержатся другие точки этого множества *. Такое множество $\{M\}$ называется **плотным в себе****.

Определение 2. *Функция $u=f(M)$, определенная на множестве $\{M\}$, называется **непрерывной на этом множестве**, если она непрерывна в каждой точке M этого множества.*

Назовем **приращением** или **полным приращением** функции $u=f(M)$ в точке A функцию Δu, определяемую формулой

$$\Delta u=f(M)-f(A), \tag{12.5}$$

где M — любая точка из области задания функции. Пусть точки A и M имеют соответственно координаты $a_1, a_2, \ldots, a_m$ и $x_1, x_2, \ldots, x_m$. Обозначим $x_1-a_1=\Delta x_1$, $x_2-a_2=\Delta x_2, \ldots, x_m-a_m=\Delta x_m$. Используя эти обозначения, получим для приращения функции Δu, соответствующего приращениям аргументов $\Delta x_1, \Delta x_2, \ldots, \Delta x_m$, следующее выражение:

$$\Delta u=f(a_1+\Delta x_1,\ a_2+\Delta x_2, \ldots, a_m+\Delta x_m)-f(a_1,\ a_2, \ldots, a_m). \tag{12.6}$$

Очевидно, *для непрерывности функции $u=f(M)$ в точке A необходимо и достаточно, чтобы ее приращение Δu представляло собой бесконечно малую в точке A функцию, т. е. необходимо и достаточно, чтобы*

$$\lim_{M\to A}\Delta u=\lim_{M\to A}(f(M)-f(A))=0, \text{ или } \lim_{\substack{\Delta x_1\to 0,\\ \Delta x_2\to 0,\\ \cdots\\ \Delta x_m\to 0}}\Delta u=0. \tag{12.7}$$

Условие (12.7) естественно назвать **разностной формой условия непрерывности функции** $u=f(M)$ **в точке** A.

2. Непрерывность функции m переменных по одной переменной. Для функции $u=f(x_1, x_2, \ldots, x_m)$ нескольких переменных можно определить понятие непрерывности **по одной из переменных** при фиксированных значениях остальных переменных. Для введения этого понятия рассмотрим так называемые **частные приращения** функции $u=f(x_1, \ldots, x_m)$ в точке $M(x_1, x_2, \ldots, x_m)$, принадлежащей области определения функции. Зафиксируем все аргументы, кроме первого, а первому аргументу придадим произвольное приращение Δx_1 такое, чтобы точка с координатами $x_1+\Delta x_1$, $x_2, \ldots, x_m$ находилась в области задания функции. Соответствующее

* Это означает, что любая точка M множества $\{M\}$ является предельной точкой этого множества.

** Примером плотного в себе множества могут служить любое непустое открытое множество и любая замкнутая область, а также множество всех точек, содержащихся в открытом множестве и таких, что все координаты этих точек являются рациональными числами.

приращение функции называется **частным приращением*** функции в точке $M(x_1, x_2, ..., x_m)$, соответствующим приращению Δx_1 аргумента x_1, и обозначается $\Delta_{x_1} u$. Таким образом,

$$\Delta_{x_1} u = f(x_1 + \Delta x_1, x_2, \ldots, x_m) - f(x_1, x_2, \ldots, x_m). \quad (12.8)$$

Аналогично определяются частные приращения функции, соответствующие приращениям других аргументов:

$$\begin{aligned} &\Delta_{x_2} u = f(x_1, x_2 + \Delta x_2, x_3, \ldots, x_m) - f(x_1, x_2, \ldots, x_m), \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (12.9) \\ &\Delta_{x_m} u = f(x_1, x_2, \ldots, x_{m-1}, x_m + \Delta x_m) - f(x_1, x_2, \ldots, x_m). \end{aligned}$$

Введем теперь понятие непрерывности функции $u = f(x_1, x_2, ...$ $..., x_m)$ **по одной из переменных.**

*Функция $u = f(x_1, x_2, ..., x_m)$ называется **непрерывной в точке** $M(x_1, x_2, ..., x_m)$ **по переменной** x_k, если частное приращение $\Delta_{x_k} u$ этой функции в точке M представляет собой бесконечно малую функцию от Δx_k, т. е. если*

$$\lim_{\Delta x_k \to 0} \Delta_{x_k} u = 0.$$

При фиксированных значениях всех переменных, кроме переменной x_k, функция $u = f(x_1, x_2, ..., x_m)$ представляет собой функцию одной этой переменной. Отметим, что непрерывность функции по переменной x_k означает непрерывность указанной функции одной переменной.

Очевидно, из условия непрерывности функции $u = f(x_1, x_2, ..., x_m)$ в данной точке M вытекает непрерывность этой функции в точке M по каждой из переменных $x_1, x_2, .., x_m$. Однако из непрерывности функции в точке M по каждой из переменных $x_1, x_2, ..., x_m$ не вытекает, вообще говоря, непрерывность функции в этой точке. Чтобы убедиться в этом, рассмотрим следующие примеры.

1°. Мы будем говорить, что функция $u = f(M) = f(x, y)$ **непрерывна в точке** M **на прямой, проходящей через эту точку**, если для любой последовательности точек $\{M_n\}$ этой прямой, сходящейся к точке M, соответствующая последовательность $\{f(M_n)\}$ значений функции имеет пределом частное значение $f(M)$ функции в точке M. Так как на прямой функция $u = f(x, y)$ представляет собой функцию одной переменной, то понятие непрерывности функции на прямой совпадает, очевидно, с понятием непрерывности указанной функции одной переменной. В частности, непрерывность функции в точке M по отдельным переменным x и y представляет собой непрерывность ее на прямых,

* Термин «частное приращение» употребляется для того, чтобы отличить это приращение от полного приращения (12.6), соответствующего произвольным приращениям $\Delta x_1, \Delta x_2, ..., \Delta x_m$ всех аргументов $x_1, x_2, ..., x_m$.

проходящих через точку M и параллельных координатным осям.

Докажем, что функция

$$u=\begin{cases}\dfrac{xy}{x^2+y^2} & \text{при} \quad x^2+y^2\neq 0,\\ 0 & \text{при} \quad x^2+y^2=0\end{cases} \tag{12.10}$$

непрерывна в точке $O(0, 0)$ по каждой из переменных x и y, т. е. непрерывна на каждой из координатных осей, но не является непрерывной на всех остальных прямых, проходящих через эту точку, и поэтому не является непрерывной в точке O. Каждая прямая, отличная от координатных осей и проходящая через точку $O(0, 0)$, может быть представлена уравнением $y=kx$, где $k\neq 0$.

В каждой точке прямой $y=kx$ при $k\neq 0$, за исключением точки $O(0, 0)$ функция (12.10) принимает одно и то же постоянное значение $\dfrac{k}{1+k^2}$.

Отсюда следует, что если последовательность $\{M_n\}$ отличных от O точек такой прямой сходится к точке O, то соответствующая последовательность значений функции имеет предел $\dfrac{k}{1+k^2}$. Так как при $k\neq 0$ этот предел отличен от нуля и не совпадает с частным значением функции в точке O, то функция разрывна в этой точке на рассматриваемой прямой. Непрерывность функции на координатных осях вытекает из того, что ее значения на этих осях равны нулю.

Может сложиться впечатление, что если функция двух переменных непрерывна на любой прямой, проходящей через данную точку, то эта функция непрерывна в указанной точке. Следующий пример показывает, что это, вообще говоря, не так.

2°. Рассмотрим функцию

$$u=f(M)=\begin{cases}\dfrac{x^2y}{x^4+y^2} & \text{при} \quad x^4+y^2\neq 0,\\ 0 & \text{при} \quad x^4+y^2=0.\end{cases}$$

Докажем, что, хотя указанная функция непрерывна на любой прямой, проходящей через точку $O(0, 0)$, она не является непрерывной в этой точке. В самом деле, значения этой функции на прямой $y=kx$ равны $\dfrac{kx}{x^2+k^2}$, и поэтому $u\to 0$ при $x\to 0$. Непрерывность этой функци на оси Oy вытекает из того, что ее значения на этой оси равны нулю. С другой стороны, значения функции на параболе $y=px^2$ постоянны и равны $\dfrac{p}{1+p^2}$, и поэтому предельное значение функции при стремлении точки M к точке O по указанной параболе также равно $\dfrac{p}{1+p^2}$. Так как при $p\neq 0$ этот предел отличен от нуля и не совпадает с частным значением функции в точке O, то функция разрывна в этой точке.

3. Основные свойства непрерывных функций нескольких переменных. В этом пункте мы перечислим основные свойства непрерывных функций нескольких переменных. Поскольку доказательства этих свойств в основном аналогичны доказательствам соответствующих свойств функций одной переменной, то, как правило, мы будем давать лишь краткие пояснения, предоставляя детали доказательства читателю.

1°. Арифметические операции над непрерывными функциями. Если функции $f(M)$ и $g(M)$ заданы на одном и том же множестве $\{M\}$ и непрерывны в некоторой точке A этого множества, то функции $f(M)+g(M)$, $f(M)-g(M)$, $f(M)\cdot g(M)$ и $\frac{f(M)}{g(M)}$ также непрерывны в точке A (в случае частного нужно дополнительно потребовать, чтобы $g(A)\neq 0$).

Это утверждение сразу же вытекает из соответствующего утверждения об арифметических операциях над функциями, имеющими предел (см. п. 3 § 2 настоящей главы).

2°. Непрерывность сложной функции. Введем понятие сложной функции нескольких переменных. Пусть функции

$$\left\{\begin{array}{l} x_1=\varphi_1(t_1,\ t_2,\ \ldots,\ t_k), \\ x_2=\varphi_2(t_1,\ t_2,\ \ldots,\ t_k), \\ \cdots\cdots\cdots\cdots \\ x_m=\varphi_m(t_1,\ t_2,\ \ldots,\ t_k) \end{array}\right. \tag{12.11}$$

заданы на множестве $\{N\}$ евклидова пространства E^k (t_1, $t_2, \ldots, t_k$ — координаты точек в этом пространстве). Тогда каждой точке $N(t_1, t_2, \ldots, t_k)$ множества $\{N\}$ ставится в соответствие с помощью формул (12.11) точка $M(x_1, x_2, \ldots, x_m)$ евклидова пространства E^m. Обозначим через $\{M\}$ множество всех таких точек. Пусть $u=f(x_1, \ldots, x_m)$ — функция m переменных, заданная на указанном множестве $\{M\}$. В этом случае мы будем говорить, что на множестве $\{N\}$ евклидова пространства E^k определена сложная функция $u=f(x_1, x_2, \ldots, x_m)$, где $x_1, x_2, \ldots, x_m$ являются функциями переменных $t_1, t_2, \ldots, t_k$, причем эти функции определяются соотношениями (12.11). Справедливо следующее

Утверждение. *Пусть функции $x_1=\varphi_1(t_1, \ldots, t_k)$, $x_2=\varphi_2(t_1, \ldots, t_k), \ldots, x_m=\varphi_m(t_1, \ldots, t_k)$ непрерывны в точке $A(a_1, a_2, \ldots, a_k)$, а функция $u=f(x_1, x_2, \ldots, x_m)$ непрерывна в точке $B(b_1, b_2, \ldots, b_m)$, где $b_i=\varphi_i(a_1, a_2, \ldots, a_k)$, $i=1, 2, \ldots, m$. Тогда сложная функция $u=f(x_1, x_2, \ldots, x_m)$, где $x_1, x_2, \ldots, x_m$ представляют собой определенные выше функции аргументов $t_1, t_2, \ldots, t_k$, непрерывна в точке $A(a_1, a_2, \ldots, a_k)$.*

Для доказательства этого утверждения рассмотрим произвольную сходящуюся к A последовательность точек $N_n=(t_1^{(n)}, t_2^{(n)}, \ldots, t_k^{(n)})$ множества $\{N\}$. Обозначим через $\{M_n\}$ соответствующую последовательность точек пространства E^m, координаты $x_1^{(n)}, x_2^{(n)}, \ldots, x_m^{(n)}$ которых равны

$$x_i^{(n)}=\varphi_i=(N_n)=\varphi_i(t_1^{(n)}, t_2^{(n)}, \ldots, t_k^{(n)}), \quad i=1, 2, \ldots, m.$$

Из непрерывности функций $\varphi_i(N)$ в точке A и из определения непрерывности по Гейне вытекает сходимость последовательности точек $\{M_n\}$ к точке B $(b_1, b_2, \ldots, b_m)$.

Далее из непрерывности функции $u=f(M)$ в точке B и из определения непрерывности по Гейне вытекает сходимость последовательности $\{f(M_n)\}$ к числу $f(B)$. Но это и означает, что последовательность $f[\varphi_1(N_n), \varphi_2(N_n), \ldots, \varphi_m(N_n)]$ значений этой сложной функции сходится к частному значению этой сложной функции $f[\varphi_1(A), \varphi_2(A), \ldots, \varphi_m(A)]$, т. е. непрерывность сложной функции в точке A.

3°. Теорема об устойчивости знака непрерывной функции.

Теорема 12.4. *Если функция $u=f(M)$ непрерывна в точке A евклидова пространства E^m и если $f(A)\neq 0$, то существует такая δ-окрестность точки A, в пределах которой $f(M)$ не обращается в нуль и имеет знак, совпадающий со знаком $f(A)$.*

Справедливость этой теоремы почти непосредственно вытекает из определения непрерывности функции $f(M)$ в точке A по Коши. В самом деле, из этого определения вытекает, что для любого $\varepsilon>0$ найдется отвечающее ему $\delta>0$ такое, что $f(A)-\varepsilon<f(M)<f(A)+\varepsilon$ всюду в δ-окрестности точки A. Если в этих рассуждениях взять в качестве ε положительное число $|f(A)|/2$, то мы получим, что все три числа $f(A)-\varepsilon$, $f(A)$ и $f(A)+\varepsilon$ будут одного знака, и потому $f(M)$ имеет тот же знак, что и $f(A)$ всюду в δ-окрестности точки A.

4°. Теорема о прохождении непрерывной функции через любое промежуточное значение.

Теорема 12.5. *Пусть функция $u=f(M)$ непрерывна во всех точках связного множества $\{M\}$ евклидова пространства E^m, причем $f(A)$ и $f(B)$ — значения этой функции в точках A и B этого множества. Пусть, далее, C — любое число, заключенное между $f(A)$ и $f(B)$. Тогда на любой непрерывной кривой L, соединяющей точки A и B и целиком располагающейся в $\{M\}$, найдется точка N такая, что $f(N)=C$.*

Доказательство. *Пусть* $x_1=\varphi_1(t)$, $x_2=\varphi_2(t), \ldots, x_m=\varphi_m(t)$, $\alpha\leqslant t\leqslant\beta$, — уравнения непрерывной кривой L, соединяющей точки A и B множества $\{M\}$ и целиком располагающейся в $\{M\}$ (См. п. 2 § 1). На сегменте $[\alpha, \beta]$ определена сложная функция $u=f(x_1, x_2, \ldots, x_m)$, где $x_i=\varphi_i(t)$, $i=1, 2, \ldots, m$, $\alpha\leqslant t\leqslant\beta$.

Очевидно, значения этой функции на сегменте $[\alpha, \beta]$ совпадают со значениями функции $u=f(M)$ на кривой L. Указанная сложная функция одной переменной t, в силу утверждения раздела 2° этого пункта, непрерывна на сегменте $[\alpha, \beta]$ и, согласно теореме 4.13, в некоторой точке ξ сегмента $[\alpha, \beta]$ принимает значение C. Поэтому в точке N кривой L с координатами $\varphi_1(\xi)$, $\varphi_2(\xi), \ldots, \varphi_m(\xi)$ получим $f(N)=C$. Теорема доказана.

5°. Ограниченность функции, непрерывной на замкнутом ограниченном множестве.

Теорема 12.6. (первая теорема Вейерштрасса). *Если функция $u=f(M)$ непрерывна на замкнутом ограниченном множестве $\{M\}$, то она ограничена на этом множестве.*

Для того чтобы убедиться в справедливости этой теоремы, выделим (как и в доказательстве аналогичной теоремы 4.14) последовательность $\{M_n\}$ точек множества $\{M\}$, для которых $|f(M_n)|>n$. В силу теоремы Больцано—Вейерштрасса (см. п. 2 § 2) из $\{M_n\}$ можно выделить сходящуюся подпоследовательность $\{M_{k_n}\}$, предел M которой в силу замечания к теореме Больцано-Вейерштрасса принадлежит множеству $\{M\}$. Очевидно, последовательность $\{f(M_{k_n})\}$ бесконечно большая. С другой стороны, в силу непрерывности функции в точке M эта последовательность $\{f(M_{k_n})\}$ должна сходиться к $f(M)$. Полученное противоречие доказывает теорему.

6°. Достижение функцией, непрерывной на замкнутом ограниченном множестве, своих точных граней. *Точной верхней гранью* функции $f(M)$ на множестве $\{M\}$ называется такое число $\bar{u}$, которое удовлетворяет двум требованиям: 1) $f(M)\leqslant\bar{u}$ для всех точек M множества $\{M\}$;

2) для любого $\varepsilon>0$ найдется хотя бы одна точка M множества $\{M\}$, для которой $f(M)>\bar{u}-\varepsilon$.

Аналогично определяется *точная нижняя грань* $\underline{u}$ функции $f(M)$ на множестве $\{M\}$.

Для обозначения точных граней функции $f(M)$ на множестве используют следующую символику:

$$\bar{u}=\sup_{\{M\}} f(M), \quad \underline{u}=\inf_{\{M\}} f(M).$$

Теорема 12.7 (вторая теорема Вейерштрасса). *Если функция $u=f(M)$ непрерывна на замкнутом ограниченном множестве $\{M\}$, то она достигает на этом множестве своих точных верхней и нижней граней.*

Доказательство этой теоремы почти дословно повторяет доказательство теоремы 4.15 (т. е. второй теоремы Вейерштрасса для функции одной переменной).

Читатель без труда проведет это доказательство сам.

7°. Понятие равномерной непрерывности функции нескольких переменных. *Функция* $u=f(M)$ *называется* ***равномерно непрерывной*** *на множестве* $\{M\}$* *евклидова пространства* E^m, *если для любого положительного числа* ε *можно указать такое положительное* δ, *зависящее только от* ε, *что для любых двух точек* M' *и* M'' *множества, удовлетворяющих условию* $\rho(M', M'')<\delta$, *выполняется неравенство* $|f(M')-f(M'')|<\varepsilon$.

Имеет место следующая теорема.

Теорема 12.8. (теорема о равномерной непрерывности). *Непрерывная на замкнутом ограниченном множестве функция равномерно непрерывна на этом множестве.*

Доказательство этой теоремы совершенно аналогично доказательству теоремы 4.16 и получается из него путем замены термина «сегмент $[a, b]$» термином «множество $\{M\}$», замены буквы x на букву M и замены выражений типа $|x'-x''|$ на символ $\rho(M', M'')$.

Замечание. Назовем **диаметром** ограниченного множества M точную верхнюю грань чисел $\rho(M', M'')$, где M' и M'' — всевозможные точки множества $\{M\}$. Используя понятие диаметра множества, отметим следующее свойство непрерывных на замкнутых ограниченных множествах функций. *Пусть функция* $u=f(M)$ *непрерывна на замкнутом ограниченном множестве* $\{M\}$. *Тогда для любого положительного числа* ε *можно указать такое* $\delta>0$, *что на каждом принадлежащем множеству* $\{M\}$ *замкнутом подмножестве* $\{N\}$, *диаметр которого меньше* δ, *колебание* ω** *функции* $f(M)$ *меньше* ε.

Доказательство этого свойства совершенно аналогично доказательству следствия из теоремы 4.16.

Замечание. Множество $\{M\}$ m-мерного евклидова пространства называется **компактом**, если из любой системы открытых множеств, покрывающих множество $\{M\}$, можно выделить конечную подсистему, также покрывающую это множество.

Точно так же, как и для пространства всех вещественных чисел E^1 (см. п. 3 § 7 гл. 4), доказывается, что *множество m-мерного евклидова пространства E^m является компактом тогда и только тогда, когда оно является замкнутым и ограниченным.*

Таким образом, первая и вторая теоремы Вейерштрасса и теорема о равномерной непрерывности справедливы для функции, непрерывной на компакте.

---

* При этом предполагается, что множество $\{M\}$ плотно в себе, т. е. в любой δ-окрестности каждой точки M этого множества имеются отличные от M точки множества $\{M\}$.

** Колебанием ω функции $f(M)$ на множестве $\{M\}$ называется разность между точной верхней и точной нижней гранями функции $f(M)$ на этом множестве.

§ 4. ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИАЛЫ ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

1. Частные производные функции нескольких переменных. Пусть $M(x_1, x_2, ..., x_m)$ — внутренняя точка области задания функции $u=f(x_1, x_2, ..., x_m)$. Рассмотрим в данной фиксированной точке $M(x_1, x_2, ..., x_m)$ отношение частного приращения $\Delta_{x_k}u$ (см. п. 1 § 3, формулы (12.8) и (12.9)) к соответствующему приращению Δx_k аргумента x_k:

$$\frac{\Delta_{x_k}u}{\Delta x_k}=\frac{f(x_1, \ldots, x_{k-1}, x_k+\Delta x_k, x_{k+1}, \ldots, x_m)-f(x_1, x_2, \ldots, x_m)}{\Delta x_k} \tag{12.12}$$

Отношение (12.12) представляет собой функцию от Δx_k, определенную для всех отличных от нуля значений Δx_k, для которых точка $M(x_1, x_2, ..., x_{k-1}, x_k+\Delta x_k, x_{k+1}, ..., x_m)$ принадлежит области задания функции $u=f(x_1, x_2, ..., x_m)$.

Определение. *Если существует предел отношения* (12.12) *частного приращения $\Delta_{x_k}u$ функции в точке $M(x_1, x_2, ..., x_m)$ к соответствующему приращению Δx_k аргумента x_k при $\Delta x_k\to 0$, то этот предел называется частной производной функции $u=f(x_1, x_2, ..., x_m)$ в точке M по аргументу x_k и обозначается одним из следующих символов*:

$$\frac{\partial u}{\partial x_k},\ \frac{\partial f}{\partial x_k},\ u'_{x_k},\ f'_{x_k}.$$

Таким образом,

$$\frac{\partial u}{\partial x_k}=\lim_{\Delta x_k\to 0}\frac{\Delta_{x_k}u}{\Delta x_k}. \tag{12.13}$$

Отметим, что частная производная функция $u=f(x_1, x_2, ..., x_m)$ по аргументу x_k представляет собой обыкновенную производную функции одной переменной x_k при фиксированных значениях остальных переменных. Поэтому вычисление частных производных производится по обычным правилам вычисления производных функций одной переменной.

Примеры

1) $u=\operatorname{arctg}(xy^2)$, $\quad \dfrac{\partial u}{\partial x}=\dfrac{y^2}{1+x^2y^4}$, $\quad \dfrac{\partial u}{\partial y}=\dfrac{2xy}{1+x^2y^4}$.

2) $u=e^{\frac{x^2+y^2}{z}}$, $\quad \dfrac{\partial u}{\partial x}=\dfrac{2x}{z}e^{\frac{x^2+y^2}{z}}$,

$$\frac{\partial u}{\partial y}=\frac{2y}{z}e^{\frac{x^2+y^2}{z}},\quad \frac{\partial u}{\partial z}=-\frac{x^2+y^2}{z^2}e^{\frac{x^2+y^2}{z}},\quad z\neq 0 .$$

Замечание 1. Из существования у функции в данной точке всех частных производных, вообще говоря, не вытекает непрерывность функции в этой точке. Мы уже убедились, что функция (12.10)

$$u = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{при } x^2+y^2 \neq 0, \\ 0 & \text{при } x^2+y^2 = 0 \end{cases}$$

не является непрерывной в точке $O(0, 0)$ (см. пример 1° п. 2 § 3). Однако в этой точке указанная функция имеет частные производные по x и y. Это следует из того, что $f(x, 0) \equiv 0$, $f(0, y) \equiv 0$, и поэтому

$$\left.\frac{\partial f}{\partial x}\right|_{(0,0)} = 0, \quad \left.\frac{\partial f}{\partial y}\right|_{(0,0)} = 0.$$

З а м е ч а н и е 2. Подчеркнем, что данное нами определение понятия частных производных всегда пригодно для внутренних точек области задания функции, но для граничных точек этой области, вообще говоря, непригодно.

Это связано, в частности, с тем, что в граничных точках области задания функции не всегда можно вычислить частные приращения функции (так, например, обстоит дело с граничной точкой M_0 области, изображенной на рис. 12.2).

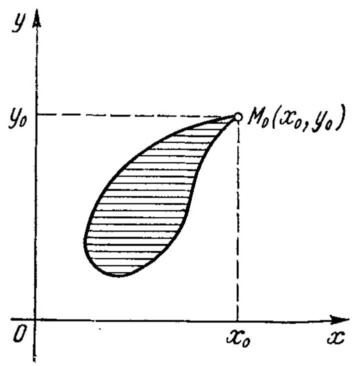


Рис. 12.2

В связи со сказанным принято определять частные производные в граничных точках как пределы этих производных при стремлении точек к границе.

2. Дифференцируемость функции нескольких переменных. Напомним, что приращением (или полным приращением) функции $u = f(x_1, x_2, ..., x_m)$ в точке $M(x_1, x_2, ..., x_m)$ соответствующим приращениям $\Delta x_1, \Delta x_2, ..., \Delta x_m$ аргументов, называется выражение*

$$\Delta u = f(x_1 + \Delta x_1,\ x_2 + \Delta x_2, ...,\ x_m + \Delta x_m) - f(x_1, x_2, ..., x_m).$$

О п р е д е л е н и е. *Функция* $u = f(x_1, x_2, ..., x_m)$ *называется* *д и ф ф е р е н ц и р у е м о й в данной точке* $M(x_1, x_2, ..., x_m)$, *если* ее полное приращение в этой точке может быть представлено в виде

$$\Delta u = A_1 \Delta x_1 + A_2 \Delta x_2 + ... + A_m \Delta x_m + \alpha_1 \Delta x_1 + \alpha_2 \Delta x_2 + ... + \alpha_m \Delta x_m, \tag{12.14}$$

* См. п. 1 § 3, формулы (12.5) и (12.6).

где $A_1, A_2, \ldots, A_m$ — некоторые не зависящие от $\Delta x_1, \Delta x_2, \ldots, \Delta x_m$ числа, а $\alpha_1, \alpha_2, \ldots, \alpha_m$ — бесконечно малые при $\Delta x_1 \to 0, \Delta x_2 \to 0, \ldots$ $\ldots, \Delta x_m \to 0$ функции, равные нулю при $\Delta x_1 = \Delta x_2 = \ldots \Delta x_m = 0$.

Соотношение (12.14) называется условием дифференцируемости функции в данной точке M. Условие (12.14) дифференцируемости функции можно записать также в иной форме. Для этого рассмотрим бесконечно малую при $\Delta x_1 \to 0$, $\Delta x_2 \to$ $\to 0, \ldots, \Delta x_m \to 0$ функцию $\rho = \sqrt{\Delta x_1^2 + x_2^2 + \ldots + \Delta x_m^2}$* и отметим, что эта функция обращается в нуль лишь при $\Delta x_1 = \Delta x_2 = \ldots$ $\ldots = \Delta x_m = 0$. Убедимся теперь, что входящая в правую часть соотношения (12.14) сумма $\alpha_1 \Delta x_1 + \alpha_2 \Delta x_2 + \ldots + \alpha_m \Delta x_m$ представляет собой бесконечно малую более высокого порядка функцию по сравнению с ρ. Иными словами, убедимся, что эта сумма представляет собой выражение $o(\rho)$. В самом деле, при $\rho \neq 0$ справедливо неравенство $\frac{|\Delta x_i|}{\rho} \leqslant 1$, и поэтому

$$|\alpha_1 \Delta x_1 + \alpha_2 \Delta x_2 + \ldots + \alpha_m \Delta x_m| \leqslant$$

$$\leqslant \left\{ |\alpha_1| \frac{|\Delta x_1|}{\rho} + |\alpha_2| \frac{|\Delta x_2|}{\rho} + \ldots + |\alpha_m| \frac{|\Delta x_m|}{\rho} \right\} \rho \leqslant$$

$$\leqslant \{ |\alpha_1| + |\alpha_2| + \ldots + |\alpha_m| \} \rho = o(\rho).$$

Таким образом, условие (12.14) дифференцируемости функции может быть записано в следующей форме:

$$\Delta u = A_1 \Delta x_1 + A_2 \Delta x_2 + \ldots + A_m \Delta x_m + o(\rho). \tag{12.15}$$

При этом величину $o(\rho)$ мы считаем равной нулю при $\rho = 0$.

Чтобы доказать, что условие (12.15) эквивалентно условию (12.14), нужно убедиться, что из представления (12.15), в свою очередь, вытекает представление (12.14). Для этой цели, считая, что не все $\Delta x_1, \Delta x_2, \ldots, \Delta x_m$ равны нулю**, представим $o(\rho)$ в виде

$$o(\rho) = \frac{o(\rho)}{\rho} \frac{\rho^2}{\rho} = \frac{o(\rho)}{\rho} \frac{\Delta x_1^2 + \Delta x_2^2 + \ldots + \Delta x_m^2}{\rho} =$$

$$= \left[\frac{o(\rho)}{\rho} \frac{\Delta x_1}{\rho} \right] \Delta x_1 + \left[\frac{o(\rho)}{\rho} \frac{\Delta x_2}{\rho} \right] \Delta x_2 + \ldots + \left[\frac{c(\rho)}{\rho} \frac{\Delta x_m}{\rho} \right] \Delta x_m.$$

Полагая $\frac{o(\rho)}{\rho} \cdot \frac{\Delta x_i}{\rho} = \alpha_i$ и учитывая, что α_i является бесконечно малой при $\rho \to 0$ (а значит, и при $\Delta x_1 \to 0, \Delta x_2 \to 0, \ldots, \Delta x_m \to 0$) функцией, мы придем к представлению (12.14).

---

* Геометрически эта функция представляет собой расстояние между точками $M(x_1, x_2, \ldots, x_m)$ и $M'(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots, x_m + \Delta x_m)$.

** Если все Δx_i равны нулю, то все члены в первой части формул (12.14) и (12.15) равиы нулю.

Итак, условие дифференцируемости функции можно записать как в виде (12.14), так и в виде (12.15).

Если хотя бы одно из чисел $A_1, A_2, \ldots, A_m$ отлично от нуля, то сумма $A_1\Delta x_1+A_2\Delta x_2+\ldots+A_m\Delta x_m$ представляет собой главную, линейную относительно приращений аргументов часть приращения дифференцируемой функции. Отметим, что при определении понятия дифференцируемости функции мы не исключали возможности обращения всех чисел $A_1, A_2, \ldots, A_m$ в нуль, и поэтому если приращение функции может быть представлено в виде (12.14) или (12.15) при $A_1=0$, $A_2=0, \ldots, A_m=0$, то функция дифференцируема в данной точке.

Справедлива следующая теорема.

Теорема 12.9. *Если функция $u=f(x_1, x_2, \ldots, x_m)$ дифференцируема в точке $M(x_1, x_2, \ldots, x_m)$, то в этой точке существуют частные производные по всем аргументам, причем $\frac{\partial u}{\partial x_i}=A_i$, где A_i определяются из условия (12.14) или (12.15) дифференцируемости функции.*

Доказательство. Из условия (12.14) дифференцируемости функции в точке $M(x_1, x_2, \ldots, x_m)$ вытекает, что ее частное приращение $\Delta_{x_i}u$ в этой точке равно $\Delta_{x_i}u=A_i\Delta x_i+\alpha_i\Delta x_i$. Отсюда вытекает, что $\frac{\Delta_{x_i}u}{\Delta x_i}=A_i+\alpha_i$, и поэтому, так как $\alpha_i\to 0$ при $\Delta x_i\to 0$,

$$\lim_{\Delta x_i\to 0}\frac{\Delta_{x_i}u}{\Delta x_i}=\frac{\partial u}{\partial x_i}=A_i.$$

Следствие 1. *Условие (12.15) дифференцируемости функции в данной точке M можно записать в следующей форме:*

$$\Delta u=\frac{\partial u}{\partial x_1}\Delta x_1+\frac{\partial u}{\partial x_2}\Delta x_2+\ldots+\frac{\partial u}{\partial x_m}\Delta x_m+o(\rho). \qquad (12.16)$$

Следствие 2. *Если функция $u=f(x_1, x_2, \ldots, x_m)$ дифференцируема в точке $M(x_1, x_2, \ldots, x_m)$, то представление ее приращения Δu в форме (12.14) или (12.15) единственно.*

В самом деле, коэффициенты A_i этих представлений равны частным производным $\frac{\partial u}{\partial x_i}$ в данной точке M и поэтому определяются единственным образом.

Убедимся в справедливости следующего важного свойства дифференцируемых функций.

Если функция $u=f(x_1, x_2, \ldots, x_m)$ дифференцируема в точке $M(x_1, x_2, \ldots, x_m)$, то она и непрерывна в этой точке.

В самом деле, из условия (12.14) дифференцируемости функции в точке вытекает, что $\lim\limits_{\Delta x_1\to 0,\ldots,\Delta x_m\to 0}\Delta u=0$, а это означает, что функция непрерывна в точке M (см. п. 1 § 3, формула (12.7)).

3. Геометрический смысл условия дифференцируемости функции двух переменных. В случае функции $u=f(x, y)$ двух переменных условие дифференцируемости может быть иллюстрировано геометрически. Введем понятие касательной плоскости к поверхности в точке.

Плоскость П, *проходящая через точку* N_0 *поверхности, называется* ***касательной*** *плоскостью в этой точке, если угол между этой плоскостью и секущей, проходящей через точку* N_0 *и любую точку* N_1 *поверхности, стремится к нулю, когда точка* N_1 *стремится к* N_0 (рис. 12.3).

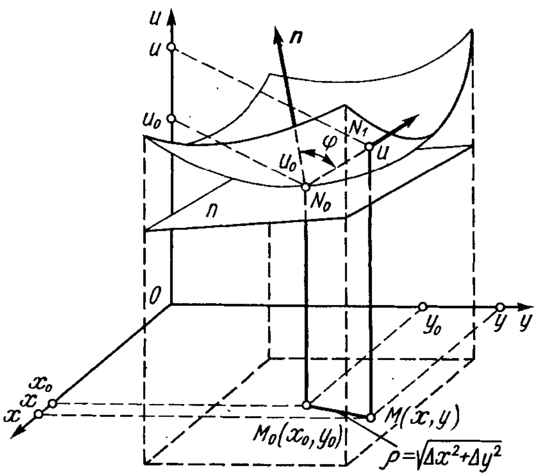


Рис. 12.3

Если в точке N_0 существует касательная плоскость, то очевидно, что касательная в точке N_0 к любой кривой расположенной на поверхности и проходящей через N_0, лежит в указанной плоскости.

Убедимся, что из условия дифференцируемости функции $u=f(x, y)$ в данной точке $M_0(x_0, y_0)$ вытекает существование касательной плоскости к графику S этой функции в точке $N_0(x_0, y_0, u_0)$. Положим $\Delta x=x-x_0$, $\Delta y=y-y_0$, $\Delta u=u-u_0$, где $u_0=f(x_0, y_0)$, $u=f(x, y)$. Очевидно, условие (12.14) дифференцируемости в рассматриваемом случае можно записать следующим образом:

$$u-u_0=A(x-x_0)+B(y-y_0)+\alpha\Delta x+\beta\Delta y=A(x-x_0)+$$
$$+B(y-y_0)+o(\rho),$$

где A и B — постоянные, равные частным производным $\frac{\partial u}{\partial x}$ и $\frac{\partial u}{\partial y}$ в точке M_0, а α и β — бесконечно малые при $\Delta x \to 0$ и $\Delta y \to 0$ функции $\rho=\sqrt{\Delta x^2+\Delta y^2}$.

Рассмотрим следующее уравнение:

$$U-u_0=A(x-x_0)+B(y-y_0).$$

Из аналитической геометрии известно, что это уравнение определяет в декартовой системе координат (x, y, U) некоторую плоскость П, проходящую через точку $N_0(x_0, y_0, u_0)$ и имеющую нормальный вектор $\mathbf{n}=\{A, B, -1\}$ *.

Докажем, что эта плоскость П является касательной плоскостью в точке N_0 поверхности S. Для этого достаточно убедиться, что: 1) плоскость П проходит через точку N_0 поверхности S и 2) угол φ между нормалью $\mathbf{n}$ этой плоскости и любой секущей N_0N_1 стремится к $\pi/2$, когда точка N_1 поверхности S стремится к точке N_0. Утверждение 1) очевидно. Перейдем к доказательству утверждения 2). Вычислим косинус угла φ, воспользовавшись известной формулой для косинуса угла между двумя векторами. Так как координаты вектора $\mathbf{n}$ равны A, B, -1, а координаты вектора $\overline{N_0N_1}$ секущей равны $x-x_0$, $y-y_0$, $u-u_0$ (см. рис. 12.3), то

$$\cos\varphi=\frac{A(x-x_0)+B(y-y_0)-(u-u_0)}{\sqrt{A^2+B^2+1}\,\sqrt{(x-x_0)^2+(y-y_0)^2+(u-u_0)^2}}.$$

Из условия дифференцируемости функции $u=f(x, y)$ вытекает, что

$$A(x-x_0)+B(y-y_0)-(u-u_0)=o(\rho).$$

Поэтому $|\cos\varphi|\leqslant\dfrac{|o(\rho)|}{\sqrt{(x-x_0)^2+(y-y_0)^2}}=\dfrac{|o(\rho)|}{\rho}\to 0$, когда $\rho\to 0$, т. е. $\lim\varphi=\dfrac{\pi}{2}$. Утверждение 2) доказано.

Таким образом, дифференцируемость функции $u=f(x, y)$ в точке $M_0(x_0, y_0)$ с геометрической точки зрения означает наличие касательной плоскости к графику функции $u=f(x, y)$ в точке $N_0(x_0, y_0, u_0)$.

Так как коэффициенты A и B равны соответственно частным производным, вычисленным в точке $M_0(x_0, y_0)$, то уравнение касательной плоскости может быть записано в виде

$$U-u_0=\frac{\partial u}{\partial x}(x-x_0)+\frac{\partial u}{\partial y}(y-y_0).$$

Нормальный вектор $\mathbf{n}=\left\{\dfrac{\partial u}{\partial x}, \dfrac{\partial u}{\partial y}, -1\right\}$ касательной плоскости принято называть нормалью к поверхности $u=f(x, y)$ в точке $N_0(x_0, y_0, z_0)$.

4. Достаточные условия дифференцируемости. Займемся выяснением достаточных условий дифференцируемости функции m переменных. Докажем следующую теорему.

Теорема 12.10. *Если функция $u=f(x_1, x_2, \ldots, x_m)$ имеет частные производные по всем аргументам в некоторой окрестности*

* Нормальным вектором плоскости называется любой ненулевой вектор $\mathbf{n}$, перпендикулярный к этой плоскости.

точки $M_0(x_1^0, x_2^0, \ldots, x_m^0)$, причем все эти частные производные непрерывны в самой точке M_0, то указанная функция дифференцируема в точке M_0.

Доказательство. Для сокращения записи проведем доказательство для функции двух переменных $u=f(x, y)$. Итак, пусть обе частные производные f_x' и f_y' существуют в окрестности точки $M_0(x_0, y_0)$ и непрерывны в этой точке. Дадим аргументам x и y столь малые приращения Δx и Δy, чтобы точка $M(x_0+\Delta x, y_0+\Delta y)$ не выходила за пределы указанной окрестности точки M_0. Полное приращение $\Delta u=f(x_0+\Delta x, y_0+\Delta y_0)-f(x_0, y_0)$ можно записать в виде

$$\Delta u=[f(x_0+\Delta x, y_0+\Delta y)-f(x_0, y_0+\Delta y)]+$$
$$+[f(x_0, y_0+\Delta y)-f(x_0, y_0)].$$

Выражение $[f(x_0+\Delta x_0, y_0+\Delta y_0)-f(x_0, y_0+\Delta y)]$ можно рассматривать как приращение функции $f(x, y_0+\Delta y)$ одной переменной x на сегменте $[x_0, x_0+\Delta x]$. Поскольку функция $u=f(x, y)$ имеет частные производные, указанная функция $f(x, y_0+\Delta y)$ дифференцируема и ее производная по x представляет собой частную производную f_x'. Применяя к указанному приращению формулу Лагранжа, найдем такое θ_1 из интервала $0<\theta_1<1$, что

$$[f(x_0+\Delta x, y_0+\Delta y)-f(x_0, y_0+\Delta y)]=f_x'(x_0+\theta_1\Delta x, y_0+\Delta y)\Delta x.$$

Рассуждая совершенно аналогично, получим

$$[f(x_0, y_0+\Delta y)-f(x_0, y_0)]=f_y'(x_0, y_0+\theta_2\Delta y)\Delta y, \quad 0<\theta_2<1.$$

Так как производные f_x' и f_y' непрерывны в точке M_0, то

$$f_x'(x_0+\theta_1\Delta x, y_0+\Delta y)=f_x'(x_0, y_0)+\alpha,$$
$$f_y'(x_0; y_0+\theta_2\Delta y)=f_y'(x_0, y_0)+\beta,$$

где α и β — бесконечно малые при $\Delta x\to 0$ и $\Delta y\to 0$ функции. Отсюда, учитывая приведенные выражения для

$$[f(x_0+\Delta x, y_0+\Delta y)-f(x_0, y_0+\Delta y_0)] \quad \text{и} \quad [f(x_0, y_0+\Delta y)-f(x_0, y_0)]$$

и выражение для Δu, найдем

$$\Delta u=f_x'(x_0, y_0)\Delta x+f_y'(x_0, y_0)\Delta y+\alpha\Delta x+\beta\Delta y.$$

Следовательно, функция $u=f(x, y)$ дифференцируема в точке M_0.

Для функции m переменных $u=f(x_1, x_2, \ldots, x_m)$ рассуждения аналогичны, нужно только полное приращение Δu такой функции представить в виде

$$\Delta u = f(x_1^0+\Delta x_1, x_2^0+\Delta x_2, \ldots, x_m^0+\Delta x_m)-f(x_1^0, x_2^0, \ldots, x_m^0) =$$
$$=\sum_{i=1}^{m}[f(x_1^0, \ldots, x_{i-1}^0, x_i^0+\Delta x_i, x_{i+1}^0+\Delta x_{i+1}, \ldots, x_m^0+\Delta x_m)-$$
$$-f(x_1^0, \ldots, x_{i-1}^0, x_i^0, x_{i+1}^0+\Delta x_{i+1}, \ldots, x_m^0+\Delta x_m)].$$

Теорема доказана.

5. Дифференциал функции нескольких переменных.

Определение. *Дифференциалом du дифференцируемой в точке $M(x_1, x_2, \ldots, x_m)$ функции $u=f(x_1, x_2, \ldots, x_m)$ называется главная линейная относительно приращений аргументов часть приращения этой функции в точке M. Если все коэффициенты A_i в представлении (12.14) приращения дифференцируемой функции равны нулю, то дифференциал du функции в точке M считается равным нулю.*

Таким образом, дифференциалом du дифференцируемой в точке M функции $u=f(x_1, x_2, \ldots, x_m)$ называется выражение

$$du=A_1\Delta x_1+A_2\Delta x_2+\ldots+A_m\Delta x_m. \tag{12.18}$$

Используя теорему 12.9, мы можем, очевидно, переписать выражение (12.18) для дифференциала du следующим образом:

$$du=\frac{\partial u}{\partial x_1}\Delta x_1+\frac{\partial u}{\partial x_2}\Delta x_2+\ldots+\frac{\partial u}{\partial x_m}\Delta x_m. \tag{12.19}$$

Введем понятие дифференциала dx_i независимой переменной x_i. Под дифференциалом dx_i независимой переменной x_i можно понимать любое (не зависящее от $x_1, x_2, \ldots, x_m$) число. Договоримся в дальнейшем брать это число равным приращению Δx_i независимой переменной x_i. Эта договоренность позволяет нам переписать формулу (12.19) в виде

$$du=\frac{\partial u}{\partial x_1}dx_1+\frac{\partial u}{\partial x_2}dx_2+\ldots+\frac{\partial u}{\partial x_m}dx_m. \tag{12.20}$$

Подчеркнем, что формула (12.20) установлена нами лишь для случая, когда аргументы $x_1, x_2, \ldots, x_m$ являются независимыми переменными. Однако ниже, в п. 7 этого параграфа, мы докажем, что формула (12.20) остается справедливой и для случая, когда аргументы $x_1, x_2, \ldots, x_m$ не являются независимыми переменными, а сами представляют собой дифференцируемые функции некоторых новых переменных.

6. Дифференцирование сложной функции. В этом пункте мы рассмотрим вопрос о дифференцировании сложной функции вида $u=f(x_1, x_2, \ldots, x_m)$, где

$$\begin{cases} x_1=\varphi_1(t_1, t_2, \ldots, t_k), \\ x_2=\varphi_2(t_1, t_2, \ldots, t_k), \\ \cdots\cdots\cdots\cdots \\ x_m=\varphi_m(t_1, t_2, \ldots, t_k). \end{cases} \tag{12.21}$$

Мы докажем, что при определенных условиях эта сложная функция является дифференцируемой функцией своих аргументов $t_1, t_2, \ldots, t_k$. При этом частные производные указанной сложной функции по аргументам $t_1, t_2, \ldots, t_k$ выражаются через частные производные функции $u=f(x_1, x_2, ..., x_m)$ и через частные производные функций (12.21) по следующим формулам:

$$\begin{aligned}\frac{\partial u}{\partial t_1} &= \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial t_1} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial t_1} + \ldots + \frac{\partial u}{\partial x_m}\frac{\partial x_m}{\partial t_1},\\ \frac{\partial u}{\partial t_2} &= \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial t_2} + \ldots + \frac{\partial u}{\partial x_m}\frac{\partial x_m}{\partial t_2},\\ \frac{\partial u}{\partial t_k} &= \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial t_k} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial t_k} + \ldots + \frac{\partial u}{\partial x_m}\frac{\partial x_m}{\partial t_k}.\end{aligned} \tag{12.22}$$

Докажем следующую основную теорему.

Теорема 12.11. *Пусть функции* (12.21) *дифференцируемы в некоторой точке* $M(t_1^0, t_2^0, ..., t_k^0)$, *а функция* $u=f(x_1, x_2, ..., x_m)$ *дифференцируема в соответствующей точке* $N(x_1^0, x_2^0, ..., x_m^0)$, *где* $x_i^0=\varphi_i(t_1^0, t_2^0, ..., t_k^0)$, $i=1, 2, ..., m$. *Тогда сложная функция* $u=f(x_1, x_2, ..., x_m)$, *где* $x_1, x_2, ..., x_m$ *определяются соотношениями* (12.21), *дифференцируема в точке M. При этом частные производные этой сложной функции в точке M определяются формулами* (12.22), *в которых все частные производные* $\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_m}$ *берутся в точке N, а все частные производные* $\frac{\partial x_i}{\partial t_j}$ *функций* (12.21) *по аргументам* $t_1, t_2, ..., t_k$ *берутся в точке M.*

Доказательство. Придадим аргументам $t_1, t_2, ..., t_k$ в точке $M(t_1^0, t_2^0, ..., t_k^0)$ произвольные приращения $\Delta t_1, \Delta t_2, ..., \Delta t_k$, не равные одновременно нулю. Этим приращениям соответствуют приращения $\Delta x_1, \Delta x_2, ..., \Delta x_m$ функций (12.21) в точке M. Приращениям $\Delta x_1, \Delta x_2, ..., \Delta x_m$, в свою очередь, соответствует приращение Δu функции $u=f(x_1, x_2, ..., x_m)$ в точке N. Поскольку функция $u=f(x_1, x_2, ..., x_m)$ предполагается дифференцируемой в точке N, указанное приращение Δu этой функции может быть записано в виде

$$\Delta u = \frac{\partial u}{\partial x_1}\Delta x_1 + \frac{\partial u}{\partial x_2}\Delta x_2 + \ldots + \frac{\partial u}{\partial x_m}\Delta x_m + \alpha_1\Delta x_1 + \alpha_2\Delta x_2 + \ldots$$
$$\ldots + \alpha_m\Delta x_m, \tag{12.23}$$

где частные производные $\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_m}$ берутся в точке N, а $\alpha_1, \alpha_2, ..., \alpha_m$ — бесконечно малые при $\Delta x_1 \to 0, \Delta x_2 \to 0, ..., \Delta x_m \to 0$ функции, равные нулю при $\Delta x_1=\Delta x_2= ... =\Delta x_m=0$. Подчеркнем, что в соотношении (12.23) $\Delta x_1, \Delta x_2, ..., \Delta x_m$ представляют собой приращения функций (12.21), отвечающие выбранным приращениям $\Delta t_1, \Delta t_2, ..., \Delta t_k$ аргументов этих функций. В силу дифференцируемости функций (12.21) в точке $M(t_1^0, t_2^0, ..., t_k^0)$ указанные приращения Δx_i можно записать в следующей форме:

$$\Delta x_i = \frac{\partial x_i}{\partial t_1}\Delta t_1 + \frac{\partial x_i}{\partial t_2}\Delta t_2 + \ldots + \frac{\partial x_i}{\partial t_k}\Delta t_k + o(\rho), \qquad (12.24)$$
$$i = 1, 2, \ldots, m,$$

где частные производные $\frac{\partial x_i}{\partial t_1}, \frac{\partial x_i}{\partial t_2}, \ldots, \frac{\partial x_i}{\partial t_k}$ берутся в точке M, а $\rho = \sqrt{(\Delta t_1)^2 + (\Delta t_2)^2 + \ldots + (\Delta t_k)^2}$.

Мы должны убедиться в том, что после подстановки в правую часть (12.23) выражений (12.24) приращение Δu может быть приведено к виду

$$\Delta u = A_1\Delta t_1 + A_2\Delta t_2 + \ldots + A_k\Delta t_k + o(\rho), \qquad (12.25)$$

где

$$A_i = \frac{\partial u}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial u}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \ldots + \frac{\partial u}{\partial x_m}\frac{\partial x_m}{\partial t_i}, \qquad (12.26)$$

Тем самым доказательство теоремы будет завершено, ибо формула (12.25) устанавливает факт дифференцируемости сложной функции, а выражение (12.26) представляет собой частную производную указанной сложной функции по переменной t_i (см. теорему 12.9).

При подстановке в правую часть (12.23) выражений (12.24) кроме группы слагаемых

$$A_1\Delta t_1 + A_2\Delta t_2 + \ldots + A_k\Delta t_k$$

мы получим и другие группы слагаемых. Нам нужно убедиться в том, что все другие группы слагаемых представляют собой величину $o(\rho)$.

Действительно, подставляя выражения (12.24) в формулу (12.23), получим

$$\Delta u = \sum_{i=1}^{m}\frac{\partial u}{\partial x_i}\Delta x_i + \sum_{i=1}^{m}\alpha_i\Delta x_i =$$
$$= \sum_{i=1}^{m}\frac{\partial u}{\partial x_i}\left[\sum_{j=1}^{k}\frac{\partial x_i}{\partial t_j}\Delta t_j + o(\rho)\right] + \sum_{i=1}^{m}\alpha_i\Delta x_i =$$
$$= \sum_{j=1}^{k}\left(\sum_{i=1}^{m}\frac{\partial u}{\partial x_i}\frac{\partial x_i}{\partial t_j}\right)\Delta t_j + \sum_{i=1}^{m}\frac{\partial u}{\partial x_i}o(\rho) + \sum_{i=1}^{m}\alpha_i\Delta x_i =$$
$$= \sum_{j=1}^{k}A_j\Delta t_j + \sum_{i=1}^{m}\frac{\partial u}{\partial x_i}o(\rho) + \sum_{i=1}^{m}\alpha_i\Delta x_i.$$

Последние две суммы написанной выше формулы представляют собой величину $o(\rho)$. В самом деле, величины $\frac{\partial u}{\partial x_i}$ берутся в точке N и поэтому представляют собой постоянные не зависящие от ρ числа. Следовательно, $\sum_{i=1}^{m} \frac{\partial u}{\partial x_i} o(\rho) = o(\rho)$. Далее, величины Δx_i для $i=1, 2, \dots, m$ удовлетворяют в силу формулы (12.24) неравенству $|\Delta x_i| \leqslant \text{const}\cdot\rho$, а величины $\alpha_i = o(1)$ для $i=1, 2, \dots, m$, ибо все α_i являются бесконечно малыми при $\Delta x_1 \to 0$, $\Delta x_2 \to 0, \dots$ $\dots, \Delta x_m \to 0$, а из дифференцируемости и вытекающей из нее непрерывности в точке M функций (12.21) следует, что Δx_1, $\Delta x_2, \dots, \Delta x_m$ стремятся к нулю при $\rho \to 0$. Поэтому $\sum_{i=1}^{m} \alpha_i \Delta x_i = o(\rho)$. Теорема доказана.

Замечание. Рассмотрим важный частный случай, когда функции (12.21) зависят от одного аргумента t. Тогда мы имеем сложную функцию одной переменной t: $u = f(x_1, x_2, \dots, x_m)$, где $x_i = \varphi_i(t)$. Производная $\frac{du}{dt}$ этой сложной функции определяется формулой

$$\frac{du}{dt} = \frac{\partial u}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial u}{\partial x_2}\frac{dx_2}{dt} + \dots + \frac{\partial u}{\partial x_m}\frac{dx_m}{dt}. \tag{12.27}$$

Применим формулу (12.27) для доказательства теоремы Эйлера об однородных функциях.

Функция $u = f(x_1, x_2, \dots, x_m)$ заданная на множестве $\{M\}$, называется однородной функцией степени p на этом множестве, если для каждой точки M $(x_1, x_2, \dots, x_m)$ множества $\{M\}$ и для каждого числа t, для которого точка $N(tx_1, tx_2, \dots, tx_m)$ принадлежит множеству M, выполняется равенство

$$f(tx_1, tx_2, \dots, tx_m) = t^p f(x_1, x_2, \dots, x_m). \tag{12.28}$$

Теорема 12.12 (теорема Эйлера об однородных функциях). *Если $u = f(x_1, x_2, \dots, x_m)$ является в некоторой области $\{M\}$ дифференцируемой однородной функцией степени p, то в каждой точке $M(x_1, x_2, \dots, x_m)$ области $\{M\}$ справедливо равенство*

$$\frac{\partial u}{\partial x_1} x_1 + \frac{\partial u}{\partial x_2} x_2 + \dots + \frac{\partial u}{\partial x_m} x_m = pu. \tag{12.29}$$

Доказательство. Пусть $M_0(x_1^0, x_2^0, \dots, x_m^0)$ — произвольная точка области $\{M\}$. Рассмотрим сложную функцию $u = f(x_1, x_2, \dots, x_m)$, где $x_i = tx_i^0$ $(i = 1, 2, \dots, m)$, т. е. функцию $u = f(tx_1^0, tx_2^0, \dots, tx_m^0)$. Так как при $t = 1$ функции $x_i = tx_i^0$ дифференцируемы и функция $u = f(x_1, x_2, \dots, x_m)$ дифференцируема в соответствующей точке M_0, то согласно теореме 12.11 и замечанию к этой теореме мы можем вычислить производную $\frac{du}{dt}$ указанной сложной функции в точке $t=1$ по формуле (12.27). Так как $\frac{dx_i}{dt}=x_i^0$, то

$$\left.\frac{du}{dt}\right|_{t=1}=\frac{\partial u}{\partial x_1}x_1^0+\frac{\partial u}{\partial x_2}x_2^0+\ldots+\frac{\partial u}{\partial x_m}x_m^0, \tag{12.30}$$

где производные $\frac{\partial u}{\partial x_i}$ берутся в точке M_0. С другой стороны, в силу (12.28) рассматриваемая сложная функция может быть представлена следующим образом:

$$u=f(tx_1^0,\ tx_2^0,\ldots,tx_m^0)=t^pf(x_1^0,\ x_2^0,\ldots,x_m^0). \tag{12.31}$$

Из (12.31) вытекает, что $\frac{du}{dt}=pt^{p-1}f(x_1^0,x_2^0,\ldots,x_m^0)$, т. е.

$$\left.\frac{du}{dt}\right|_{t=1}=pf(x_1^0,x_2^0,\ldots,x_m^0)=pu. \tag{12.32}$$

Сравнивая (12.30) и (12.32), мы и получим соотношение (12.29) для точки M_0. Так как точка M_0 — произвольная точка области $\{M\}$, то теорема доказана.

7. Инвариантность формы первого дифференциала. В п. 5 мы ввели понятие первого дифференциала du функции нескольких переменных и установили, что когда аргументы $x_1,\ x_2,\ldots,x_m$ являются *независимыми переменными*, то дифференциал du можно представить в виде

$$du=\frac{\partial u}{\partial x_1}dx_1+\frac{\partial u}{\partial x_2}dx_2+\ldots+\frac{\partial u}{\partial x_m}dx_m. \tag{12.20}$$

В этом пункте мы докажем, что формула (12.20) является универсальной и справедлива и в том случае, когда аргументы $x_1,\ x_2,\ldots,x_m$ сами являются дифференцируемыми функциями некоторых новых переменных $t_1,\ t_2,\ldots,t_k$, которые мы можем считать независимыми. Указанное свойство первого дифференциала обычно называют свойством *инвариантности его формы*.

Итак, пусть аргументы $x_1,\ x_2,\ldots,x_m$ функции $u=f(x_1,\ x_2,\ldots,x_m)$ представляют собой дифференцируемые в точке $A(t_1^0,\ t_2^0,\ldots,t_k^0)$ функции $x_i=\varphi_i(t_1,\ t_2,\ldots,t_k)$, а сама функция $u=f(x_1,\ x_2,\ldots,x_m)$ дифференцируема в точке $B(x_1^0,\ x_2^0,\ldots,x_m^0)$, где $x_i^0=\varphi_i(t_1^0,\ t_2^0,\ldots,t_k^0)$. В таком случае мы можем рассматривать u как сложную функцию независимых переменных $t_1,\ t_2,\ldots,t_k$, которая в силу теоремы 12.11 является дифференцируемой в точке A. Поэтому дифференциал du этой сложной функции можно представить в виде

$$du=\frac{\partial u}{\partial t_1}dt_1+\frac{\partial u}{\partial t_2}dt_2+\ldots+\frac{\partial u}{\partial t_k}dt_k, \tag{12.33}$$

где $\frac{\partial u}{\partial t_i}$ определяются из соотношений (12.22). Подставляя $\frac{\partial u}{\partial t_i}$

из (12.22) в (12.33) и собирая коэффициенты при $\partial u/\partial x_i$, получим

$$du = \frac{\partial u}{\partial x_1}\left(\frac{\partial x_1}{\partial t_1}dt_j + \frac{\partial x_1}{\partial t_2}dt_2 + \ldots + \frac{\partial x_1}{\partial t_k}dt_k\right) + \ldots$$

$$\ldots + \frac{\partial u}{\partial x_m}\left(\frac{\partial x_m}{\partial t_1}dt_1 + \frac{\partial x_m}{\partial t_2}dt_2 + \ldots + \frac{\partial x_m}{\partial t_k}dt_k\right).$$

Остается заметить, что в последнем соотношении коэффициент при $\partial u/\partial x_i$ равен дифференциалу dx_i функции $x_i=\varphi_i(t_1, t_2, ..., t_k)$. Мы получим для дифференциала du сложной функции формулу (12.20), в которой дифференциалы dx_i будут дифференциалами функций $x_i=\varphi_i(t_1, t_2, ..., t_k)$. Инвариантность формы первого дифференциала установлена.

Свойство инвариантности формы первого дифференциала позволяет установить следующие *правила дифференцирования.* Пусть u и v — дифференцируемые функции каких-либо переменных. Тогда

$$d(c\cdot u) = c\cdot du \quad (c=\text{const}),$$

$$d(u\pm v) = du\pm dv,$$

$$d(uv) = udv+vdu,$$

$$d\left(\frac{u}{v}\right) = \frac{v\,du - u\,dv}{v^2}.$$

Докажем, например, справедливость третьей из указанных формул. Рассмотрим функцию $w=uv$ двух переменных u и v. Дифференциал этой функции dw равен

$$dw = \frac{\partial w}{\partial u}du + \frac{\partial w}{\partial v}dv.$$

Так как $\frac{\partial w}{\partial u}=v$ и $\frac{\partial w}{\partial v}=u$, то $dw=udv+vdu$. В силу инвариантности формы первого дифференциала выражение $udv+vdu$ будет дифференциалом функции uv и в случае, когда u и v сами являются дифференцируемыми функциями каких-либо переменных.

8. **Производная по направлению. Градиент.** Начнем с рассмотрения функции трех независимых переменных $u=f(x, y, z)$. Предположим, что эта функция определена в некоторой окрестности точки $M_0(x^0, y^0, z^0)$ пространства E^3 и дифференцируема в точке M_0.

Рассмотрим всевозможные лучи, выходящие из точки M_0. Каждый такой луч задается единичным вектором **e** с координатами $(\cos\alpha, \cos\beta, \cos\gamma)$ * и определяет некоторое направление.

* Из курса аналитической геометрии известно, что если единичный вектор **e** составляет с осями координат углы, соответственно равные α, β, γ, то коор-

Фиксируем некоторый луч, выходящий из точки M_0 и определяемый единичным вектором $\mathbf{e}$ с координатами $\cos\alpha$, $\cos\beta$, $\cos\gamma$.

Взяв на прямой, содержащей этот луч, произвольную отличную от M_0 точку M, рассмотрим вектор или направленный отрезок $\overline{M_0M}$ и обозначим через l величину этого направленного отрезка на оси, определяемой единичным вектором $\mathbf{e}=(\cos\alpha, \cos\beta, \cos\gamma)$ *.

Ясно, что вектор $\overline{M_0M}$ имеет координаты

$$(l\cos\alpha,\ l\cos\beta,\ l\cos\gamma).$$

С другой стороны, если координаты точки M равны (x, y, z), то вектор $\overline{M_0M}$ имеет координаты, равные $(x-x^0, y-y^0, z-z^0)$.

Сопоставляя два полученных нами соотношения для координат вектора $\overline{M_0M}$, мы приходим к равенствам

$$x=x^0+l\cos\alpha,\qquad y=y^0+l\cos\beta,\qquad z=z^0+l\cos\gamma. \qquad (12.34)$$

Равенства (12.34) показывают, что на прямой, проходящей через точку M_0 и определяемой единичным вектором $\mathbf{e}=(\cos\alpha, \cos\beta, \cos\gamma)$, функция $u=f(x, y, z)$ представляет собой сложную функцию одной независимой переменной l вида $u=f(x^0+l\cos\alpha, y^0+l\cos\beta, z_0+l\cos\gamma)$.

Определение 1. *Производную указанной сложной функции по переменной l, взятую в точке $l=0$, назовем* ***производной функции*** $u=f(x, y, z)$ ***в точке*** M_0 ***по направлению****, определяемому единичным вектором* $\mathbf{e}$, *и будем обозначать символом* $\partial u/\partial e$.

Итак, по определению

$$\frac{\partial u}{\partial e}=\frac{\partial u}{\partial x_1}(M_0)\cos\alpha+\frac{\partial u}{\partial y}(M_0)\cos\beta+\frac{\partial u}{\partial z}(M_0)\cos\gamma. \qquad (12.35)$$

Введем понятие градиента дифференцируемой в данной точке $M_0(x_0, y_0, z_0)$ функции $u=f(x, y, z)$.

Определение 2. ***Градиентом*** *функции $u=f(x, y, z)$ в данной точке $M_0(x^0, y^0, z^0)$ называется вектор, координаты которого имеют вид*

$$\frac{\partial u}{\partial x}(M_0),\quad \frac{\partial u}{\partial y}(M_0),\quad \frac{\partial u}{\partial z}(M_0).$$

Для обозначения градиента функции $u=f(x, y, z)$ обычно используют символ

$$\operatorname{grad} u.$$

динаты этого вектора равны $\cos\alpha$, $\cos\beta$, $\cos\gamma$. Указанные три косинуса принято называть направляющими косинусами вектора $\mathbf{e}$.

* Величиной l направленного отрезка M_0M оси, определяемой единичным вектором $\mathbf{e}$, называется число, равное длине отрезка $\overline{M_0M}$, взятой со знаком «плюс», если векторы $\overline{M_0M}$ и $\mathbf{e}$ направлены в одну сторону, и со знаком «минус», если эти векторы направлены в разные стороны.

Итак, по определению

$$\operatorname{grad} u\,(M_0) = \left(\frac{\partial u}{\partial x}(M_0),\ \frac{\partial u}{\partial y}(M_0),\ \frac{\partial u}{\partial z}(M_0)\right). \tag{12.36}$$

Так как скалярное произведение двух векторов равно сумме произведений одноименных координат этих векторов, то выражение (12.35) для производной по направлению, определяемому вектором $\mathbf{e}$, можно рассматривать как скалярное произведение вектора (12.36) и $\mathbf{e} = (\cos\alpha,\ \cos\beta,\ \cos\gamma)$.

Итак, мы получаем, что

$$\frac{\partial u}{\partial e} = (\mathbf{e},\ \operatorname{grad} u). \tag{12.37}$$

С помощью равенства (12.37) убедимся, что *градиент функции* $u = f(x,\ y,\ z)$ *в точке* M_0 *характеризует направление и величину максимального роста этой функции в точке* M_0.

Точнее, докажем два утверждения:

1) *производная функции* $u = f(x,\ y,\ z)$ *в точке* M_0 *по направлению, определяемому градиентом этой функции в указанной точке, имеет максимальное значение по сравнению с производной в этой точке по любому другому направлению;*

2) *значение производной функции* $u = f(x,\ y,\ z)$ *по направлению, определенному градиентом этой функции в данной точке, равно* $|\operatorname{grad} u|$, *т. е. равно длине вектора* $\operatorname{grad} u$ *в данной точке.*

Для доказательства указанных двух утверждений заметим, что скалярное произведение двух векторов равно произведению длин этих векторов на косинус угла между ними *. Поэтому выражение (12.37) можно переписать в виде

$$\frac{\partial u}{\partial e} = |\mathbf{e}|\,|\operatorname{grad} u|\,\cos\varphi,$$

где φ — угол между векторами $\mathbf{e}$ и $\operatorname{grad} u$.

Учитывая, что $|\mathbf{e}| = 1$, мы получим, что

$$\frac{\partial u}{\partial e} = |\operatorname{grad} u|\cos\varphi.$$

Из последнего равенства вытекают оба утверждения 1) и 2). В самом деле, максимальное значение производной $\frac{\partial u}{\partial e}$ получится при $\cos\varphi = 1$, т. е. при совпадении направления $\mathbf{e}$ с направлением $\operatorname{grad} u$, причем производная $\partial u/\partial e$ в этом направлении равна $|\operatorname{grad} u|$.

Доказанные два утверждения позволяют утверждать, *что градиент не зависит от выбора системы координат* (ибо и направле-

* Этот факт устанавливается в аналитической геометрии.

ние, и длина вектора $\operatorname{grad} u$ в каждой данной точке инвариантны относительно выбора системы координат).

Для выяснения геометрического смысла вектора $\operatorname{grad} u$ целесообразно ввести понятие **поверхностей уровня** функции $u=f(x, y, z)$, понимая под этим термином те поверхности, на которых функция $u=f(x, y, z)$ сохраняет постоянное значение, т. е. удовлетворяет соотношению $f(x, y, z)=c=\text{const}$.

Если в каждой точке $M_0(x^0, y^0, z^0)$ поверхности уровня $f(x, y, z)=c$ построить касательную плоскость, то легко убедиться в том, что нормальным вектором такой плоскости будет являться вектор (12.36), т. е. $\operatorname{grad} u$*. Отсюда следует, что вектор $\operatorname{grad} u$ в каждой точке M поверхности уровня $f(x, y, z)=c$ ортогонален к этой поверхности.

Совершенно аналогично определяются производная по направлению и градиент для дифференцируемой в данной точке $M_0(x_1^0, x_2^0, \dots, x^0_m)$ функции m переменных.

Для такой функции производная $\dfrac{\partial u}{\partial e}$ в данной точке $M_0(x_1^0, x_2^0, \dots, x^0_m)$ по направлению, определяемому единичным вектором $\mathbf{e}=(\cos\alpha_1, \cos\alpha_2, \dots, \cos\alpha_m)$ **, вводится как обычная производная по переменной l сложной функции $u=f(x_1^0+l\cos\alpha_1, x_2^0+l\cos\alpha_2, \dots, x^0_m+l\cos\alpha_m)$ взятая в точке $l=0$.

Для любой дифференцируемой в данной точке $M_0(x_1^0, x_2^0, \dots, x^0_m)$ функции $u=f(x_1, x_2, \dots, x_m)$ производная в этой точке по направлению, определяемому вектором $\mathbf{e}=(\cos\alpha_1, \cos\alpha_2, \dots, \cos\alpha_m)$, равна

$$\frac{\partial u}{\partial e}=\frac{\partial u}{\partial x_1}(M_0)\cos\alpha_1+\frac{\partial u}{\partial x_2}(M_0)\cos\alpha_2+\dots+\frac{\partial u}{\partial x_m}(M_0)\cos\alpha_m. \tag{12.35*}$$

Градиентом дифференцируемой в данной точке $M_0(x_1^0, x_2^0, \dots, x^0_m)$ функции $u=f(x_1, x_2, \dots, x_m)$ называется вектор, обозна-

---

* Напомним, что нормальным вектором плоскости называется ненулевой вектор, перпендикулярный к этой плоскости. В п. 3 настоящего параграфа было установлено, что для поверхности, определяемой уравнением $f(x, y)=z$, нормальный вектор касательной плоскости в каждой точке этой поверхности имеет вид $\left(\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}, -1\right)$. Аналогично устанавливается, что для поверхности, определяемой уравнением $f(x, y, z)=c$, нормальный вектор касательной плоскости в каждой точке этой поверхности равен $\left(\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}, \dfrac{\partial f}{\partial z}\right)=\operatorname{grad} u$.

** Так как все координаты единичного вектора по модулю не превосходят единицы, то для каждой из этих координат найдется угол α_i такой, что соответствующая i-я координата равна $\cos\alpha_i$ (при этом можно считать, что $0\leqslant\alpha_i\leqslant\pi$). Таким образом, единичный вектор $\mathbf{e}$ можно записать в виде $(\cos\alpha_1, \cos\alpha_2, \dots, \cos\alpha_m)$. Величины $\cos\alpha_1, \cos\alpha_2, \dots, \cos\alpha_m$ принято называть направляющими косинусами.

чаемый символом $\operatorname{grad} u$ имеющий координаты

$$\left(\frac{\partial u}{\partial x_1}(M_0),\ \frac{\partial u}{\partial x_2}(M_0),\ \ldots,\ \frac{\partial u}{\partial x_m}(M_0)\right).$$

Так как скалярное произведение двух векторов пространства E^m равно сумме произведений одноименных координат этих векторов, то равенство (12.35 *) позволяет утверждать, что

$$\frac{\partial u}{\partial e} = (\mathbf{e}, \operatorname{grad} u).$$

Из неравенства Коши—Буняковского, справедливого для двух любых векторов пространства E^m *, вытекает, что

$$\left|\frac{\partial u}{\partial e}\right| \leqslant |\mathbf{e}|\ |\operatorname{grad} u| = |\operatorname{grad} u|,$$

причем знак $\leqslant$ переходит в знак $=$ только в случае, когда векторы $\mathbf{e}$ и $\operatorname{grad} u$ коллинеарны.

Отсюда следует, что вектор $\operatorname{grad} u$ и в пространстве E^m обладает теми же двумя свойствами 1) и 2), что и в пространстве E^3.

Заметим, что в случае функции $u=f(x, y)$ двух переменных x и y единичный вектор $\mathbf{e}$, определяющий направление в точке M_0, имеет координаты $\cos\alpha$ и $\sin\alpha$. Поэтому в указанном случае формула (12.35) принимает вид

$$\frac{\partial u}{\partial e} = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\sin\alpha.$$

§ 5. ЧАСТНЫЕ ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИАЛЫ ВЫСШИХ ПОРЯДКОВ

1. Частные производные высших порядков. Пусть частная производная $\frac{\partial u}{\partial x_i}$ по аргументу x_i функции $u=f(x_1, x_2, \ldots, x_m)$, определенной в области $\{M\}$, существует в каждой точке области $\{M\}$. В этом случае указанная частная производная представляет со-

---

* Для любых двух векторов $\mathbf{a}$ и $\mathbf{b}$ пространства E^m справедливо неравенство Коши—Буняковского $(\mathbf{a}, \mathbf{b})^2 \leqslant (\mathbf{a}, \mathbf{a})(\mathbf{b}, \mathbf{b})$ или, что то же самое, $|(\mathbf{a}, \mathbf{b})| \leqslant |\mathbf{a}|\,|\mathbf{b}|$, причем в этом неравенстве знак $\leqslant$ переходит в знак $=$ только в случае, когда векторы $\mathbf{a}$ и $\mathbf{b}$ коллинеарны.

В самом деле, при любом вещественном λ справедливо неравенство $(\lambda\mathbf{a}-\mathbf{b}, \lambda\mathbf{a}-\mathbf{b}) = \lambda^2(\mathbf{a}, \mathbf{a}) - 2\lambda(\mathbf{a}, \mathbf{b}) + (\mathbf{b}, \mathbf{b}) \geqslant 0$, причем в этом неравенстве знак $\geqslant$ переходит в знак $=$ только в случае, когда вектор $\lambda\mathbf{a}-\mathbf{b}$ является нулевым, т. е. когда $\mathbf{a}$ и $\mathbf{b}$ коллинеарны. Исключая случай нулевого вектора $\mathbf{a}$, когда доказываемое неравенство заведомо справедливо со знаком $=$, мы можем утверждать, что дискриминант квадратного трехчлена $\lambda^2(\mathbf{a}, \mathbf{a}) - 2\lambda(\mathbf{a}, \mathbf{b}) + (\mathbf{b}, \mathbf{b})$ удовлетворяет неравенству $4[(\mathbf{a}, \mathbf{b})^2 - (\mathbf{a}, \mathbf{a})(\mathbf{b}, \mathbf{b})] \leqslant 0$, причем в этом неравенстве знак $\leqslant$ переходит в знак $=$ только в случае, когда векторы $\mathbf{a}$ и $\mathbf{b}$ коллинеарны.

бой функцию переменных $x_1, x_2, \ldots, x_m$, также определенную в области $\{M\}$.

Может случиться, что эта функция $\frac{\partial u}{\partial x_i}$ имеет частную производную по аргументу x_k в некоторой точке M области $\{M\}$. Тогда указанную частную производную по аргументу x_k называют второй частной производной или частной производной второго порядка функции $u=f(x_1, x_2, \ldots, x_m)$ в точке M сначала по аргументу x_i, а затем по аргументу x_k и обозначают одним из следующих символом:

$$\frac{\partial^2 u}{\partial x_k \partial x_i}, \quad f^{(2)}_{x_i x_k}, \quad u^{(2)}_{x_i x_k}.$$

При этом если $i \neq k$, то частная производная $\frac{\partial^2 u}{\partial x_k \partial x_i}$ называется смешанной частной производной второго порядка. После того как введено понятие второй частной производной, можно последовательно ввести понятие третьей частной производной, затем четвертой и т. д.

Если предположить, что нами уже введено понятие $(n-1)$-й частной производной функции $u=f(x_1, x_2, \ldots, x_m)$ по аргументам $x_{i_1}, x_{i_2}, \ldots, x_{i_{n-1}}$ (отдельные или даже все номера которых могут совпадать) и что эта $(n-1)$-я частная производная имеет в точке M частную производную по аргументу x_{i_n}, то указанную частную производную называют n-й частной производной (или частной производной n-го порядка) функции $u=f(x_1, x_2, \ldots, x_m)$ в точке M по аргументам $x_{i_1}, x_{i_2}, \ldots, x_{i_n}$.

Таким образом, мы вводим понятие n-й частной производной *индуктивно*, переходя от первой частной производной к последующим. Соотношение, определяющее n-ю частную производную по аргументам $x_{i_1}, x_{i_2}, \ldots, x_{i_{n-1}} x_{i_n}$ имеет вид

$$\frac{\partial^n u}{\partial x_{i_n} \partial x_{i_{n-1}} \ldots \partial x_{i_2} \partial x_{i_1}} = \frac{\partial}{\partial x_{i_n}} \left(\frac{\partial^{n-1} u}{\partial x_{i_{n-1}} \ldots \partial x_{i_2} \partial x_{i_1}} \right).$$

Если не все индексы $i_1, i_2, \ldots, i_n$ совпадают между собой, то частная производная $\frac{\partial^n u}{\partial x_{i_n} \ldots \partial x_{i_2} \partial x_{i_1}}$ называется смешанной частной производной n-го порядка.

Так как частная производная функции по аргументу x_i определяется как обыкновенная производная функции одной переменной x_i при фиксированных значениях остальных переменных, то методика вычисления частных производных высших порядков предполагает умение вычислять только обыкновенные производные первого порядка. В качестве примера вычислим частные производные второго порядка функции $u = \operatorname{arctg} \frac{x}{y}$.

Имеем

$$\frac{\partial u}{\partial x}=\frac{y}{x^2+y^2},\quad \frac{\partial u}{\partial y}=-\frac{x}{x^2+y^2},$$

$$\frac{\partial^2 u}{\partial x^2}=-\frac{2xy}{(x^2+y^2)^2},\quad \frac{\partial^2 u}{\partial x\partial y}=\frac{x^2-y^2}{(x^2+y^2)^2},$$

$$\frac{\partial^2 u}{\partial y\partial x}=\frac{x^2-y^2}{(x^2+y^2)^2},\quad \frac{\partial^2 u}{\partial y^2}=\frac{2xy}{(x^2+y^2)^2}.$$

В рассмотренном примере смешанные частные производные $\frac{\partial^2 u}{\partial x\partial y}$ и $\frac{\partial^2 u}{\partial y\partial x}$ равны друг другу. Вообще говоря, значения смешанных производных зависят от порядка, в котором производятся последовательные дифференцирования. Убедимся, например, что смешанные частные производные $\frac{\partial^2 u}{\partial y\partial x}$ и $\frac{\partial^2 u}{\partial x\partial y}$ функции

$$u=\begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & \text{при } x^2+y^2\neq 0,\\ 0 \text{ при } x^2+y^2=0 \end{cases}$$

в точке (0, 0) существуют, но не равны друг другу. Действительно,

$$\frac{\partial u}{\partial x}=\begin{cases} \dfrac{y(x^4y^4+4x^2y^2)}{(x^2+y^2)^2} & \text{при } x^2+y^2\neq 0,\\ 0 & \text{при } x^2+y^2=0. \end{cases}$$

Поэтому

$$\left.\frac{\partial^2 u}{\partial y\partial x}\right|_{x=0,\,y=0}=\lim_{y\to 0}\frac{\left.\dfrac{\partial u}{\partial x}\right|_{x=0,\,y\neq 0}-\left.\dfrac{\partial u}{\partial x}\right|_{x=0,\,y=0}}{y}=-1.$$

Проводя аналогичные вычисления, получим $\left.\frac{\partial^2 u}{\partial x\partial y}\right|_{x=0,\,y=0}=1.$

Таким образом, в точке (0, 0) $\frac{\partial^2 u}{\partial x\partial y}\neq\frac{\partial^2 u}{\partial y\partial x}$.

Выясним достаточные условия независимости значений смешанных производных от порядка, в котором производятся последовательные дифференцирования.

Введем важное понятие функции m переменных, n раз дифференцируемой в данной точке.

Определение. *Функция $u=f(x_1, x_2, \ldots, x_m)$ называется n раз дифференцируемой в точке $M_0(x_1^0, x_2^0, \ldots, x_m^0)$, если все ее частные производные порядка $n-1$ являются дифференцируемыми в этой точке функциями.*

Из этого определения вытекает, что если функция $u=f(x_1, x_2, \ldots$ $\ldots, x_m)$ n раз дифференцируема в точке M_0, то при $n>1$ любая ее частная производная первого порядка $n-1$ раз дифференцируема в точке M_0, при $n>2$ любая ее частная производная второго порядка $n-2$ раз дифференцируема в точке M_0 и т. д.

Укажем достаточное условие n-кратной дифференцируемости функции в данной точке.

Для того чтобы функция $u=f(x_1, x_2, \ldots, x_m)$ была n раз дифференцируемой в точке $M_0(x_1^0, x_2^0, \ldots, x_m^0)$, достаточно, чтобы все ее частные производные n-го порядка были непрерывными точке M_0.

Справедливость этого утверждения вытекает из определения дифференцируемости функции и теоремы 12.10 о достаточных условиях дифференцируемости.

Теорема 12.13. *Пусть функция $u=f(x, y)$ дважды дифференцируема в точке $M_0(x_0, y_0)$. Тогда в этой точке частные производные $f_{xy}^{(2)}$ и $f_{yx}^{(2)}$ равны.*

Доказательство. Так как функция $u=f(x, y)$ дважды дифференцируема в точке $M_0(x_0, y_0)$, то частные производные f_x' и f_y' определены в некоторой δ-окрестности точки M_0 и представляют собой дифференцируемые функции в этой точке.

Рассмотрим выражение

$$\Phi=f(x_0+h, y_0+h)-f(x_0+h, y_0)-f(x_0, y_0+h)+f(x_0, y_0), \tag{12.38}$$

где h — любое столь малое число, что точка $M(x_0+h, y_0+h)$ находится в указанной γ-окрестности точки M_0. Выражение Φ можно рассматривать как приращение $\Delta\varphi=\varphi(x_0+h)-\varphi(x_0)$ дифференцируемой на сегменте $[x_0, x_0+h]$ функции $\varphi(x)=f(x, y_0+h)-f(x, y_0)$ одной переменной x. Поэтому по формуле Лагранжа, обозначая через θ некоторое число из интервала $0<\theta<1$, можно записать:

$$\begin{gathered}\Phi=\Delta\varphi=\varphi'(x_0+\theta h)h=\\ =[f_x'(x_0+\theta h, y_0+h)-f_x'(x_0+\theta h, y_0)]h=\\ =\{[f_x'(x_0+\theta h, y_0+h)-f_x'(x_0, y_0)]-[f_x'(x_0+\theta h, y_0)-f_x'(x_0, y_0)]\}h.\end{gathered} \tag{12.39}$$

Так как частная производная f_x' является дифференцируемой в точке M_0 функцией, то

$$\begin{gathered}[f_x'(x_0+\theta h, y_0+h)-f_x'(x_0, y_0)]=\\ =f^{(2)}_{xx}(x_0, y_0)\theta h+f^{(2)}_{xy}(x_0, y_0)\, h+\alpha_1\theta h+\beta_1 h,\\ [f_x'(x_0+\theta h, y_0)-f_x'(x_0, y_0)]=f^{(2)}_{xx}(x_0, y_0)\theta h+\alpha_2\theta h,\end{gathered}$$

где α_1, β_1 и α_2 — бесконечно малые при $h\to 0$ функции. Подставляя найденные выражения для $[f_x'(x_0+\theta h, y_0+h)-f_x'(x_0, y_0)]$ и $[f_x'(x_0+\theta h, y_0)-f_x'(x_0, y_0)]$ в формулу (12.39), получим

$$\Phi=[f^{(2)}_{xy}(x_0, y_0)+\alpha]h^2, \tag{12.40}$$

где $\alpha=\alpha_1\theta+\beta_1-\alpha_2\theta$ — бесконечно малая при $h\to 0$ функция. С другой стороны, выражение Φ, определяемое соотношением (12.38), можно рассматривать как приращение $\Delta\psi=\psi(y_0+h)-\psi(y_0)$ дифференцируемой на сегменте $[y_0, y_0+h]$ функции $\psi(y)=f(x_0+h, y)-f(x_0, y)$. Применяя формулу Лагранжа и учитывая дифференцируемость частной производной f_y' в точке M_0, мы получим совершенно аналогично предыдущему следующее выражение для Φ:

$$\Phi=[f^{(2)}{}_{yx}(x_0, y_0)+\beta]h^2, \tag{12.41}$$

где β — бесконечно малая при $h\to 0$ функция. Приравнивая правые части соотношений (12.40) и (12.41) и сокращая обе части полученного равенства на h^2, найдем, что $f^{(2)}{}_{xy}(x_0, y_0)+\alpha=f^{(2)}{}_{yx}(x_0,y_0)+\beta$. Так как α и β — бесконечно малые при $h\to 0$ функции, то из последнего равенства следует, что $f^{(2)}{}_{xy}(x_0, y_0)=f^{(2)}{}_{yx}(x_0, y_0)$. Теорема доказана.

Теорема 12.13 утверждает, что в данной точке $M_0(x_0, y_0)$ имеет место равенство $f^{(2)}{}_{xy}=f^{(2)}{}_{yx}$, если в этой точке дифференцируемы f_x' и f_y'. Из дифференцируемости f_x' и f_y' в точке M_0 вытекает существование в *этой точке всех* частных производных второго порядка. Однако равенство $f^{(2)}{}_{xy}$ и $f^{(2)}{}_{yx}$ имеет место и при условии существования лишь производных $f^{(2)}{}_{xy}$ и $f^{(2)}{}_{yx}$, но при дополнительном требовании непрерывности этих производных в рассматриваемой точке. Именно справедлива следующая теорема.

Теорема 12.13*. *Пусть в некоторой окрестности точки $M_0(x_0, y_0)$ функция $u=f(x, y)$ имеет частные производные f_x', f_y', $f^{(2)}{}_{xy}$, $f^{(2)}{}_{yx}$. Пусть, кроме того, производные $f^{(2)}{}_{xy}$ и $f^{(2)}{}_{yx}$ непрерывны в точке M_0. Тогда в этой точке $f^{(2)}{}_{xy}=f^{(2)}{}_{yx}$.*

Для доказательства воспользуемся выражением Φ, определенным соотношением (12.38). Из (12.39) вытекает, что Φ представляет собой умноженную на h разность значений функции $f_x'(x, y)$ в точках $(x_0+\theta h, y_0+h)$ и $(x_0+\theta h, y_0)$. Применяя к этой разности формулу Лагранжа конечных приращений по переменной y на сегменте $[y_0, y_0+h]$, получим

$$\Phi=f^{(2)}{}_{xy}(x_0+\theta h, y_0+\theta_1 h)h^2, \quad \text{где} \quad 0<\theta_1<1.$$

В силу непрерывности $f^{(2)}{}_{xy}$ в точке $M_0(x_0, y_0)$ из последнего равенства получаем

$$\Phi=[f^{(2)}{}_{xy}(x_0, y_0)+\alpha(h)]h^2,$$

где $\alpha(h)\to 0$ при $h\to 0$.

С другой стороны, эта же величина Φ представляет собой умноженную на h разность значений функции $f_y'(x, y)$ в точках $(x_0+h, y_0+\theta_2 h)$ и $(x_0, y_0+\theta_2 h)$. Применяя к этой разности формулу Лагранжа конечных приращений по переменной x на сегменте $[x_0, x_0+h]$ и учитывая непрерывность $f^{(2)}{}_{xy}$ в точке $M_0(x_0, y_0)$, получим

$$\Phi=[f^{(2)}{}_{yx}(x_0, y_0)+\beta(h)]h^2,$$

где $\beta(h)\to 0$ при $h\to 0$.

Приравнивая последние два выражения для Φ и рассуждая так же, как и в конце доказательства теоремы 12.13, мы убедимся в справедливости нужного нам равенства

$$f^{(2)}{}_{xy}(x_0, y_0) = f^{(2)}{}_{yx}(x_0, y_0).$$

Докажем теперь теорему о независимости значения любой смешанной частной производной n-го порядка от порядка, в котором производятся последовательные дифференцирования.

Теорема 12.14. *Пусть функция $u=f(x_1, x_2, \ldots, x_m)$ n раз дифференцируема в точке $M_0(x_1^0, x_2^0, \ldots, x_m^0)$. Тогда в этой точке значение любой смешанной частной производной n-го порядка не зависит от порядка, в котором производятся последовательные дифференцирования.*

Доказательство. Очевидно, достаточно доказать независимость значения любой n-й смешанной производной от порядка проведения двух последовательных дифференцирований. Иными словами, достаточно доказать равенство

$$\frac{\partial^n u}{\partial x_{i_n} \ldots \partial x_{i_{k+1}} \partial x_{i_k} \ldots \partial x_{i_1}} = \frac{\partial^n u}{\partial x_{i_n} \ldots \partial x_{i_k} \partial x_{i_{k+1}} \ldots \partial x_{i_1}}. \qquad (12.42)$$

Рассмотрим функцию $\dfrac{\partial^{k-1} u}{\partial x_{i_{k-1}} \ldots \partial x_{i_1}}$ Эта функция представляет собой дважды дифференцируемую функцию переменных x_{i_k} и $x_{i_{k+1}}$. Поэтому в силу теоремы 12.13

$$\frac{\partial^{k+1} u}{\partial x_{i_{k+1}} \partial x_{i_k} \partial x_{i_{k-1}} \ldots \partial x_{i_1}} = \frac{\partial^{k+1} u}{\partial x_{i_k} \partial x_{i_{k+1}} \partial x_{i_{k-1}} \ldots \partial x_{i_1}}.$$

Отсюда и вытекает справедливость равенства (12.42). Теорема доказана.

Отметим, что в случае n раз дифференцируемой в точке M_0 функции $u=f(x_1, x_2, \ldots, x_m)$ любую ее частную производную n-го порядка можно записать в виде

$$\frac{\partial^n u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_m^{\alpha_m}},$$

где $\alpha_1, \alpha_2, \ldots, \alpha_m$ — целые числа, удовлетворяющие условиям:

$$0 \leqslant \alpha_i \leqslant n, \quad \alpha_1 + \alpha_2 + \ldots + \alpha_m = n.$$

2. Дифференциалы высших порядков. Выше мы использовали для обозначения дифференциалов аргументов функции $u = f(x_1, x_2, \ldots, x_m)$ и для обозначения дифференциала самой этой функции символы $dx_1, dx_2, \ldots, dx_m$ и du соответственно.

Теперь нам придется использовать для обозначения дифференциалов аргументов указанной функции и дифференциала самой этой функции и другие символы. В частности, мы будем обозначать дифференциалы аргументов функции $u=f(x_1, x_2, \ldots, x_m)$ и дифференциал самой этой функции символами $\delta x_1, \delta x_2, \ldots, \delta x_m$ и δu соответственно. В этих обозначениях инвариантное по форме выражение для первого дифференциала этой функции (12.20) (см. п. 7, § 3), будет иметь вид

$$\delta u = \frac{\partial u}{\partial x_1}\delta x_1 + \frac{\partial u}{\partial x_2}\delta x_2 + \ldots + \frac{\partial u}{\partial x_m}\delta x_m.$$

Возвращаясь к прежним обозначениям, рассмотрим выражение (12.20) для первого дифференциала дифференцируемой в данной точке $M(x_1, x_2, \ldots, x_m)$ функции $u=f(x_1, x_2, \ldots, x_m)$:

$$du = \frac{\partial u}{\partial x_1}dx_1 + \frac{\partial u}{\partial x_2}dx_2 + \ldots + \frac{\partial u}{\partial x_m}dx_m. \tag{12.20}$$

Предположим, что величина, стоящая в правой части (12.20), представляет собой функцию аргументов $x_1, x_2, \ldots, x_m$, дифференцируемую в данной точке $M(x_1, x_2, \ldots, x_m)$. Для этого достаточно потребовать, чтобы функция $u=f(x_1, x_2, \ldots, x_m)$ была два раза дифференцируема в данной точке $M(x_1, x_2, \ldots, x_m)$, а аргументы $x_1, x_2, \ldots, x_m$ являлись либо независимыми переменными, либо два раза дифференцируемыми функциями некоторых независимых переменных $t_1, t_2, \ldots, t_k$.

При этих предположениях мы можем рассмотреть дифференциал

$$\delta(du) = \delta\left[\sum_{k=1}^{m}\frac{\partial u}{\partial x_k}dx_k\right]$$

от величины (12.20).

Определение 1. *Значение $\delta(du)$ дифференциала от первого дифференциала* (12.20), *взятое при $\delta x_1=dx_1, \delta x_2=dx_2, \ldots, \delta x_m=dx_m$, называется* ***вторым дифференциалом*** *функции $u=f(x_1, x_2, \ldots, x_m)$ (в данной точке $M(x_1, x_2, \ldots, x_m)$) и обозначается символом d^2u.*

Итак, по определению*

$$d^2u = \delta(du)\Bigg|_{\substack{\delta x_1=dx_1,\\ \delta x_2=dx_2,\\ \ldots\ldots\\ \delta x_m=dx_m}} = \left\{\delta\left[\sum_{k=1}^{m}\frac{\partial u}{\partial x_k}dx_k\right]\right\}\Bigg|_{\substack{\delta x_1=dx_1,\\ \delta x_2=dx_2,\\ \ldots\ldots\\ \delta x_m=dx_m}}.$$

Дифференциал d^nu любого порядка n введем по индукции.

---

* Символ $\{\ \}\Big|_{\substack{\delta x_1=dx_1,\\ \delta x_2=dx_2,\\ \ldots\ldots\\ \delta x_m=dx_m}}$ обозначает, что в выражении, заключенном в фигурные скобки, следует положить $\delta x_1=dx_1, \delta x_2=dx_2, \ldots, \delta x_m=dx_m$.

Предположим, что уже введен дифференциал $d^{n-1}u$ порядка $n-1$ и что функция $u=f(x_1, x_2, \dots, x_m)$ n раз дифференцируема в данной точке $M(x_1, x_2, \dots, x_m)$, а ее аргументы $x_1, x_2, \dots, x_m$ являются либо независимыми переменными, либо n раз дифференцируемыми функциями некоторых независимых переменных $t_1, t_2, \dots$ $\dots, t_k$.

Определение 2. *Значение $\delta(d^{n-1}u)$ дифференциала от $(n-1)$-го дифференциала $d^{n-1}u$, взятое при $\delta x_1=dx_1$, $\delta x_2=dx_2, \dots$ $\dots, \delta x_m=dx_m$, называется n-м дифференциалом функции $u=f(x_1, x_2, \dots, x_m)$ (в данной точке $M(x_1, x_2, \dots, x_m)$) и обозначается символом d^nu.*

Итак, по определению

$$d^nu = \delta(d^{n-1}u)\Big|_{\substack{\delta x_1=dx_1,\\ \delta x_2=dx_2,\\ \dots\dots\\ \delta x_m=dx_m}}.$$

При вычислении второго и последующих дифференциалов du приходится существенно различать два случая: 1) случай, когда аргументы $x_1, x_2, \dots, x_m$ являются независимыми переменными, 2) случай, когда аргументы $x_1, x_2, \dots, x_m$ являются соответствующее число раз дифференцируемыми функциями некоторых независимых переменных $t_1, t_2, \dots, t_k$.

Рассмотрим сначала первый случай. Если $x_1, x_2, \dots, x_m$ являются независимыми переменными, то мы имеем право считать, что $dx_1, dx_2, \dots, dx_m$ не зависят от $x_1, x_2, \dots, x_m$.

Каждый дифференциал dx_k мы можем взять равным одному и тому же приращению Δx_k для всех точек $M(x_1, x_2, \dots, x_m)$. При этом мы получим, что

$$\delta(dx_k) = \sum_{i=1}^{m} \frac{\partial(dx_k)}{\partial x_i}\delta x_i = 0.$$

Последнее соотношение и правила дифференцирования, установленные в конце п. 7 § 3, позволяют нам записать для два раза дифференцируемой в данной точке M функции $u=f(x_1, x_2, \dots, x_m)$, следующую цепочку равенства:

$$d^2u = \delta(du)\Big|_{\substack{\delta x_1=dx_1,\\ \dots\dots\\ \delta x_m=dx_m}} = \delta\left[\sum_{k=1}^{m}\frac{\partial u}{\partial x_k}dx_k\right]\Big|_{\substack{\delta x_1=dx_1,\\ \dots\dots\\ \delta x_m=dx_m}} =$$

$$= \sum_{k=1}^{m}\delta\left[\frac{\partial u}{\partial x_k}dx_k\right]\Big|_{\substack{\delta x_1=dx_1,\\ \dots\dots\\ \delta x_m=dx_m}} = \sum_{k=1}^{m}\left\{dx_k\delta\left(\frac{\partial u}{\partial x_k}\right)+\frac{\partial u}{\partial x_k}\delta(dx_k)\right\}\Big|_{\substack{\delta x_1=dx_1,\\ \dots\dots\\ \delta x_m=dx_m}} =$$

$$= \sum_{k=1}^{m} \left\{ dx_k \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left(\frac{\partial u}{\partial x_k} \right) \delta x_i \right\} \Bigg|_{\substack{\delta x_1 = dx_1, \\ \cdots \\ \delta x_m = dx_m}} =$$

$$= \sum_{k=1}^{m} \sum_{i=1}^{m} \frac{\partial^2 u}{\partial x_i \partial x_k} \delta x_i dx_k \Bigg] \Bigg|_{\substack{\delta x_1 = dx_1, \\ \cdots \\ \delta x_m = dx_m}} = \sum_{k=1}^{m} \sum_{i=1}^{m} \frac{\partial^2 u}{\partial x_i \partial x_k} dx_i dx_k. \quad (12.43)$$

(Мы воспользовались еще и тем, что для два раза дифференцируемой функции смешанные производные второго порядка не зависят от того, в какой последовательности производится дифференцирование).

Итак, мы получаем, что в случае, когда аргументы $x_1, x_2, \ldots, x_m$ являются независимыми переменными, для второго дифференциала два раза дифференцируемой в данной точке функции $u = f(x_1, x_2, \ldots, x_m)$ справедливо представление

$$d^2 u = \sum_{i=1}^{m} \sum_{k=1}^{m} \frac{\partial^2 u}{\partial x_i \partial x_k} dx_i dx_k. \quad (12.44)$$

Замечание 1. Функция m переменных $t_1, t_2, \ldots, t_k$ вида $\Phi = \sum_{i=1}^{m} \sum_{k=1}^{m} a_{ik} t_i t_k$, где a_{ik} — постоянные вещественные числа, называется **квадратичной формой** от переменных $t_1, t_2, \ldots, t_k$, а числа a_{ik} — ее коэффициентами.

Квадратичная форма называется **симметричной**, если ее коэффициенты удовлетворяют условию $a_{ik} = a_{ki}$ (для всех $i = 1, 2, \ldots, m$; $k = 1, 2, \ldots, m$).

Полученное нами выражение (12.44) позволяет утверждать, что для случая, когда аргументы $x_1, x_2, \ldots, x_m$ являются независимыми переменными, второй дифференциал два раза дифференцируемой в данной точке M функции $u = f(x_1, x_2, \ldots, x_m)$ представляет собой симметричную *, квадратичную форму от переменных $dx_1, dx_2, \ldots$ $\ldots, dx_m$, коэффициенты которой равны соответствующим частным производным второго порядка функции $u = f(x_1, x_2, \ldots, x_m)$, взятым в данной точке M.

Отметим, что полученное нами выражение для дифференциала второго порядка (12.44) можно переписать и в другом виде, используя формальный символ

$$d = dx_1 \frac{\partial}{\partial x_1} + dx_2 \frac{\partial}{\partial x_2} + \ldots + dx_m \frac{\partial}{\partial x_m}. \quad (12.45)$$

---

* Симметричность этой квадратичной формы вытекает из равенства $\frac{\partial^2 u}{\partial x_i \partial x_k}(M) = \frac{\partial^2 u}{\partial x_k \partial x_i}(M)$.

С помощью этого символа выражение (12.44) может быть переписано в виде

$$d^2u = \left(dx_1 \frac{\partial}{\partial x_1} + dx_2 \frac{\partial}{\partial x_2} + \ldots + dx_m \frac{\partial}{\partial x_m}\right)^2 u. \qquad (12.46)$$

По индукции легко убедиться в том, что в случае, когда аргументы $x_1, x_2, \ldots, x_m$ n раз дифференцируемой в данной точке $M(x_1, x_2, \ldots, x_m)$ функции $u=f(x_1, x_2, \ldots, x_m)$ являются независимыми переменными, для n-го дифференциала этой функции справедливо представление

$$d^n u = \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \cdots \sum_{i_n=1}^{m} \frac{\partial^n u}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_n}} dx_{i_1} dx_{i_2} \ldots dx_{i_m}.$$

Это представление с помощью формального символа (12.45) может быть переписано в виде

$$d^n u = \left(dx_1 \frac{\partial}{\partial x_2} + dx_2 \frac{\partial}{\partial x_2} + \ldots + dx_m \frac{\partial}{\partial x_m}\right)^n u. \qquad (12.47)$$

Совершенно другой вид имеют представления для второго и последующего дифференциалов в случае, когда аргументы $x_1, x_2, \ldots, x_m$ функции $u=f(x_1, x_2, \ldots, x_m)$ являются соответствующее число раз дифференцируемыми функциями некоторых независимых переменных $t_1, t_2, \ldots, t_k$.

Обращаясь к этому случаю, установим выражение для второго дифференциала два раза дифференцируемой в данной точке $M(x_1, x_2, \ldots, x_m)$ функции $u=f(x_1, x_2, \ldots, x_m)$, аргументы $x_1, x_2, \ldots$ $\ldots, x_m$ которой являются два раза дифференцируемыми функциями некоторых независимых переменных $t_1, t_2, \ldots, t_k$.

Повторяя рассуждения из цепочки (12.43), мы на этот раз получим

$$d^2u = \delta(du)\Big|_{\substack{\delta x_1 = dx_1, \\ \ldots \\ \delta x_m = dx_m}} = \sum_{k=1}^{m} \left\{ dx_k \delta\left(\frac{\partial u}{\partial x_k}\right) + \frac{\partial u}{\partial x_k} \delta(dx_k) \right\}\Big|_{\substack{\delta x_1 = dx_1, \\ \ldots \\ \delta x_m = dx_m}} =$$

$$= \sum_{k=1}^{m} \left\{ dx_k \sum_{i=1}^{m} \frac{\partial}{\partial x_i}\left(\frac{\partial u}{\partial x_k}\right) \delta x_i \right\}\Big|_{\substack{\delta x_1 = dx_1, \\ \ldots \\ \delta x_m = dx_m}} + \sum_{k=1}^{m} \left\{ \frac{\partial u}{\partial x_k} \delta(dx_k) \right\}\Big|_{\substack{\delta x_1 = dx_1, \\ \ldots \\ \delta x_m = dx_m}}.$$

Заметим, что в силу определения второго дифференциала функции $u=x_k$ (где k — любой из номеров $1, 2, \ldots, m$)

$$[\delta(dx_k)]\Big|_{\substack{\delta x_1 = dx_1, \\ \ldots \\ \delta x_m = dx_m}} = d^2 x_k.$$

Учитывая это соотношение, мы приходим к следующему представлению для второго дифференциала:

$$d^2u=\sum_{k=1}^{m}\sum_{i=1}^{m}\frac{\partial^2 u}{\partial x_i \partial x_k}dx_i dx_k+\sum_{k=1}^{m}\frac{\partial u}{\partial x_k}d^2x_k,$$

или (с использованием символа **(12.45)**)

$$d^2u=\left(dx_1\frac{\partial}{\partial x_1}+dx_2\frac{\partial}{\partial x_2}+\ldots+dx_m\frac{\partial}{\partial x_m}\right)^2u+$$
$$+\left(\frac{\partial u}{\partial x_1}d^2x_1+\frac{\partial u}{\partial x_2}d^2x_2+\ldots+\frac{\partial u}{\partial x_m}d^2x_m\right). \qquad (12.48)$$

Сравнивая полученное нами представление (12.48) с представлением (12.46), мы убедимся в том, что (в отличие от первого дифференциала) второй дифференциал уже не обладает свойством инвариантности формы.

Тем более не обладают свойством инвариантности формы все последующие дифференциалы.

Замечание 2. Укажем важный частный случай, когда второй и последующие дифференциалы функции m переменных $u=f(x_1, x_2, \ldots, x_m)$ все же обладают инвариантностью формы и определяются той самой формулой (12.47), что и для случая независимых переменных $x_1, x_2, \ldots, x_m$.

Будем говорить, что переменные $x_1, x_2, \ldots, x_m$ являются линейными функциями независимых переменных $t_1, t_2, \ldots, t_k$, если они определяются равенствами

$$x_i=a_{i0}+a_{i1}t_1+a_{i2}t_2+\ldots+a_{ik}t_k \quad (i=1, 2, \ldots, m),$$

в которых через $a_{i0}, a_{i1}, \ldots, a_{ik}$ обозначены некоторые постоянные.

Заметим, что *если функция $u=f(x_1, x_2, \ldots, x_m)$ является n раз дифференцируемой в данной точке $M(x_1, x_2, \ldots, x_m)$, а ее аргументы $x_1, x_2, \ldots, x_m$ являются линейными функциями независимых переменных $t_1, t_2, \ldots, t_k$, то n-й дифференциал функции $u=f(x_1, x_2, \ldots$ $\ldots, x_m)$ определяется той же самой формулой* (12.47), *что и для случая независимых переменных $x_1, x_2, \ldots, x_m$.*

Для того чтобы убедиться в этом, заметим, что поскольку $t_1, t_2, \ldots, t_k$ являются независимыми переменными, то n-й дифференциал x_i как функции аргументов $t_1, t_2, \ldots, t_k$ определяется равенством типа (12.47), а точнее, равенством

$$d^nx_i=\left(dt_1\frac{\partial}{\partial t_1}+dt_2\frac{\partial}{\partial t_2}+\ldots+dt_k\frac{\partial}{\partial t_k}\right)^nx_i.$$

Но любая частная производная выше первого порядка от линейной функции x_i равна нулю.

Значит, $d^2x_i=0$, $d^3x_i=0$, ..., $d^nx_i=0$.

Равенство $d^2x_i=0$ (при всех $i=1, 2, \ldots, m$) и представление (12.48) дают право заключить, что d^2u определяется равенством (12.46). Совершенно аналогично, используя соотношения $d^3x_i=0, \ldots$ $\ldots, d^nx_i=0$, мы по индукции докажем, что $d^3u, d^4u, \ldots, d^nu$ определяются равенством (12.47).

Замечание 3. При проведении вычислений иногда требуется расшифровать равенство (12.47) и, учитывая, что в этом равенстве имеются совпадающие члены, выписать все различные члены этого равенства со стоящими перед ними коэффициентами.

Для этой цели может быть использована формула полинома Ньютона, имеющая вид

$$(a_1+a_2+\ldots+a_m)^n = \sum_{\substack{\alpha_1+\alpha_2+\ldots+\alpha_m=n \\ 0\leqslant\alpha_i\leqslant n}} \frac{(\alpha_1+\alpha_2+\ldots+\alpha_m)!}{\alpha_1!\,\alpha_2!\ldots\alpha_m!} a_1^{\alpha_1}a_2^{\alpha_2}\ldots a_m^{\alpha_m}. \qquad (12.49)$$

(Суммирование в правой части этой формулы идет по всем целочисленным индексам $\alpha_1, \alpha_2, \ldots, \alpha_m$, каждый из которых удовлетворяет неравенствам $0\leqslant\alpha_i\leqslant n$ при условии, что сумма всех этих индексов $\alpha_1+\alpha_2+\ldots+\alpha_m$ равна n.)

Формулу (12.49) нетрудно установить по индукции. В самом деле, при $m=2$ и любом целом n эта формула заведомо справедлива, ибо она переходит в известную формулу бинома Ньютона.

Предположим, что эта формула справедлива для некоторого номера $m\geqslant 2$ и любого целого n, и проверим, что в таком случае она справедлива и для номера $m+1$ и любого целого n.

Представив $(a_1+a_2+\ldots+a_m+a_{m+1})^n$ в виде

$$(a_1+a_2+\ldots+a_{m+1})^n=[(a_1+a_2+\ldots+a_m)+a_{m+1}]^n,$$

подсчитаем с помощью бинома Ньютона коэффициент при $a_1^{\alpha_1}a_2^{\alpha_2}\ldots a_m^{\alpha_m}a_{m+1}^{\alpha_{m+1}}$. В силу равенства $\alpha_1+\alpha_2+\ldots+\alpha_{m+1}=n$, формулы бинома Ньютона и предположения о справедливости формулы (12.49) для номера m и любого целого n этот коэффициент равен *

$$C_n^{\alpha_{m+1}} \frac{(\alpha_1+\alpha_2+\ldots+\alpha_m)!}{\alpha_1!\,\alpha_2!\ldots\alpha_m!} =$$

---

* Мы учитываем, что $C_n^k = \dfrac{n!}{k!\,(n-k)!}$, и потому

$$C_n^{\alpha_{m+1}} = C_{\alpha_1+\alpha_2+\ldots+\alpha_{m+1}}^{\alpha_{m+1}} = \frac{(\alpha_1+\alpha_2+\ldots+\alpha_m+\alpha_{m+1})!}{\alpha_{m+1}!\,(\alpha_1+\alpha_2+\ldots+\alpha_m)!}.$$

$$= \frac{(\alpha_1+\alpha_2+\ldots+\alpha_m+\alpha_{m+1})!}{\alpha_{m+1}!\,(\alpha_1+\alpha_2+\ldots+\alpha_m)!}\,\frac{(\alpha_1+\alpha_2+\ldots+\alpha_m)!}{\alpha_1!\,\alpha_2!\ldots\alpha_m!} =$$
$$= \frac{(\alpha_1+\alpha_2+\ldots+\alpha_m+\alpha_{m+1})!}{\alpha_1!\,\alpha_2!\ldots\alpha_m!\,\alpha_{m+1}!}.$$

Полученное нами выражение для коэффициента при $a_1^{\alpha_1}a_2^{\alpha_2}\ldots a_m^{\alpha_m}a_{m+1}^{\alpha_{m+1}}$ в точности совпадает с тем выражением, которое получится из формулы (12.49), если в этой формуле заменить номер m на $m+1$.

Индукция завершена, и формула (12.49) доказана.

Формула (12.49) дает нам право переписать выражение (12.47) для n-го дифференциала в следующем виде:

$$d^n u(M) = \sum_{\substack{\alpha_1+\alpha_2+\ldots+\alpha_m=n\\ 0\leqslant\alpha_i\leqslant n}} \frac{\partial^n u}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\ldots\partial x_m^{\alpha_m}}(M)(dx_1^{\alpha_1})(dx_2)^{\alpha_2}\ldots(dx_m)^{\alpha_m}.$$

3. Формула Тейлора с остаточным членом в форме Лагранжа и в интегральной форме. Договоримся обозначать k-й дифференциал функции m переменных $u=f(x_1, x_2, \ldots, x_m)$ в точке M пространства E^m символом $d^k u|_M$.

Докажем следующую важную теорему.

Теорема 12.15. *Пусть $n\geqslant 0$ — целое число, функция $u=f(x_1, x_2, \ldots, x_m)$ задана в некоторой ε-окрестности* точки $M_0(x_1, x_2, \ldots, x_m)$ и $n+1$ раз дифференцируема в указанной окрестности. Тогда полное приращение $\Delta u=f(M)-f(M_0)$ этой функции в точке M может быть представлено в следующей форме*:

$$\Delta u = du\Big|_{M_0} + \frac{1}{2!}d^2u\Big|_{M_0} + \ldots + \frac{1}{n!}d^n u\Big|_{M_0} + \frac{1}{(n+1)!}d^{n+1}u\Big|_{N}, \quad (12.50)$$

при этом N — некоторая точка указанной окрестности, зависящая, вообще говоря, от $M(x_1, x_2, \ldots, x_m)$, а дифференциалы dx_i переменных x_i, входящие в выражения $d^k u|_{M_0}$ и $d^{n+1}u|_N$, равны $\Delta x_i = x_i - x_i^0$. Формула (12.50) *называется формулой Тейлора для функции $u=f(M)$ с центром разложения в точке M_0, а последний член формулы* (12.50) *называется остаточным членом, записанным в форме Лагранжа.*

Доказательство. Для сокращения записи проведем рассуждения для функции $u=f(x, y)$ двух переменных x и y. Предварительно запишем в специальной форме формулу Тейлора для $n+1$ раз дифференцируемой в некоторой окрестности точки t_0 функции $u=F(t)$ одной переменной t. Напомним, что формула Тей-

* Вместо ε-окрестности точки M_0 можно взять любую звездную окрестность этой точки (т. е. любое открытое множество, содержащее точку M_0, и обладающее следующим свойством: если точка M принадлежит этому множеству, то и весь отрезок прямой M_0M ему принадлежит).

лора с центром разложения в t_0 для функции $u=F(t)$ одной переменной имеет следующий вид:

$$F(t)=F(t_0)+F'(t_0)(t-t_0)+\frac{1}{2!}F^{(2)}(t_0)(t-t_0)^2+\ldots$$
$$\ldots+\frac{1}{n!}F^{(n)}(t_0)(t-t_0)^n+\frac{1}{(n+1)!}F^{(n+1)}(t_0+\theta(t-t_0))(t-t_0)^{n+1},$$
$$0<\theta<1. \qquad (12.51)$$

Так как аргумент t является независимой переменной, то приращение $\Delta t=t-t_0$ представляет собой дифференциал dt независимой переменной t. Поэтому

$$F^{(k)}(t_0)(t-t_0)^k=F^{(k)}(t_0)\,dt^k=d^kF(t_0)=d^ku|_{t_0}$$
и
$$F^{(n+1)}(t_0+\theta(t-t_0))(t-t_0)^{n+1}=d^{n+1}u|_{t_0+\theta(t-t_0)}. \qquad (12.52)$$

Если мы обозначим разность $F(t)-F(t_0)$ через Δu, то согласно (12.52) формулу Тейлора (12.51) можно записать в следующей специальной форме:

$$\Delta u=du\Big|_{t_0}+\frac{1}{2!}d^2u\Big|_{t_0}+\ldots+\frac{1}{n!}d^nu\Big|_{t_0}+\frac{1}{(n+1)!}d^{(n+1)}u|_{t_0+\theta(t-t_0)}. \qquad (12.53)$$

Рассмотрим теперь в ε-окрестности точки $M_0(x_0, y_0)$ произвольную точку $M(x_0+\Delta x, y_0+\Delta y)$ и соединим точки M_0 и M прямой линией. Очевидно, координаты x и y точек указанной прямой представляют собой следующие *линейные функции* новой переменной t:

$$x=x_0+t\Delta x, \qquad y=y_0+t\Delta y; \qquad (12.54)$$

при этом координаты точек отрезка M_0M соответствуют значениям переменной t из сегмента $[0, 1]$. Отметим, что значению $t=0$ отвечает точка M_0, а значению $t=1$ — точка M. Так как по условию функция $u=f(x, y)$ двух переменных x и y $n+1$ раз дифференцируема в рассматриваемой окрестности точки M_0, то из формул (12.54) вытекает, что на прямой M_0M эта функция является сложной функцией переменной t из сегмента $[0, 1]$. Обозначим эту сложную функцию через $F(t)$ и запишем для нее формулу Тейлора с центром разложения в точке $t_0=0$ в специальной форме (12.53) при

$$\Delta u=F(1)-F(0)=f(M)-f(M_0).$$

Фигурирующие в формуле (12.53) дифференциалы различных порядков представляют собой дифференциалы сложной функции $u=f(x, y)$, где x и y являются линейными функциями (12.54). Согласно замечанию 2 предыдущего пункта при этих условиях дифференциалы любого порядка функции $u=f(x, y)$ могут быть записаны в форме (12.47). Поэтому

$$d^k u\,|_{t_0=0} = \left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y}\right)^k u\,|_{M_0(x_0, y_0)} d^k u\,|_{M_0}, \qquad (12.55)$$

$$d^{n+1}u\,|_{t_0+\theta(t-t_0)} \left(dx\frac{\partial}{\partial x} + dy\frac{\partial}{\partial y}\right)^{n+1} u\,|_{N(x_0+\theta\Delta x,\ y_0+\theta\Delta y)} = d^{n+1}u\,|_N,$$

причем в формулах (12.55) dx и dy находятся из соотношений (12.54) при $dt=\Delta t=1-0=1$. Таким образом, в формулах (12.55)

$$dx=dt\Delta x=\Delta x \text{ и } dy=dt\Delta y=\Delta y. \qquad (12.56)$$

Подставляя $d^k u|_{t_0}$ и $d^{n+1}u|_{t_0+\theta(t-t_0)}$ из (12.55) в формулу (12.53) и учитывая соотношения (12.56), мы получим формулу Тейлора (12.50).

Приведем развернутое выражение формулы Тейлора (12.50) для функции $u=f(x_1, x_2, \ldots, x_m)$:

$$f(x_1, x_2, \ldots, x_m) = f(x_1^0, x_2^0, \ldots, x_m^0) +$$

$$+\sum_{k=1}^{n}\frac{1}{k!}\left[(x_1-x_1^0)\frac{\partial}{\partial x_1} + (x_2-x_2^0)\frac{\partial}{\partial x_2} + \ldots + (x_m-x_m^0)\frac{\partial}{\partial x_m}\right]^k \times$$

$$\times f(x_1, x_2, \ldots x_m) + \frac{1}{(n+1)!}\left[(x_1-x_1^0)\frac{\partial}{\partial x_1} + (x_2-x_2^0)\frac{\partial}{\partial x_2} + \ldots\right.$$

$$\left.\ldots + (x_m-x_m^0)\frac{\partial}{\partial x_m}\right]^{n+1} f(x_1^0 +$$

$$+\theta(x_1-x_1^0),\ x_2^0+\theta(x_2-x_2^0), \ldots, x_m^0+\theta(x_m-x_m^0)). \qquad (12.57)$$

Следствие. *Если функция $u=f(x_1, x_2, \ldots, x_m)$ удовлетворяет тем же условиям, что и в теореме* 12.15, *и, сверх того, все частные производные этой функции порядка $n+1$ непрерывны в рассматриваемой ε-окрестности точки M_0, то остаточный член, т. е. последний член в формулах* (12.50) *и* (12.57), *может быть записан в виде*

$$\frac{1}{n!}\int_0^1 (1-t)^n \left[\sum_{j=1}^{m}(x_j-x_j^0)\frac{\partial}{\partial x_j}\right]^{n+1} \times$$

$$\times f(x_1^0+t(x_1-x_1^0),\ x_2^0+t(x_2-x_2^0), \ldots, x_m^0+t(x_m-x_m^0))\,dt.$$

Такую форму остаточного члена естественно назвать интегральной. Для получения формулы Тейлора с остаточным членом в интегральной форме следует записать в интегральной форме остаточный член в формуле Тейлора для функции одной переменной $F(t)$, рассмотренной при доказательстве теоремы 12.15, т. е. воспользоваться результатами п. 4 § 4 гл. 9. В рассматриваемом случае указанный остаточный член имеет вид

$$\frac{1}{n!}\int_0^1 F^{(n+1)}(t)(1-t)^n dt.$$

Это и приводит нас к написанному выше выражению остаточного члена для функции $u=f(x_1, x_2, \dots, x_m)$.

4. Формула Тейлора с остаточным членом в форме Пеано.

Теорема 12.15*. *Пусть $n \geqslant 1$ — целое число, функция $u=f(M)=f(x_1, x_2, \dots, x_m)$ задана и $n-1$ раз дифференцируема в* ε-окрестности точки $M_0(x_1^0, x_2^0, \dots, x_m^0)$ *и n раз дифференцируема в самой точке M_0**.

Тогда для любой точки M из указанной ε-окрестности справедлива следующая формула:

$$f(M)=f(M_0)+\frac{1}{1!}du\Big|_{M_0}+\frac{1}{2!}d^2u\Big|_{M_0}+\dots+\frac{1}{n!}d^nu\Big|_{M_0}+o(\rho^n), \tag{12.58}$$

в которой через ρ обозначено расстояние $\rho(M_0, M)$, а символ $o(\rho^n)$ обозначает бесконечно малую при $\rho\to 0$ (или при $M\to M_0$) функцию более высокого порядка малости, чем ρ^n.

Формула (12.58) называется формулой Тейлора (с центром в точке M_0) с остаточным членом в форме Пеано.

Замечание. В более подробной записи формула Тейлора (12.58) имеет вид

$$f(x_1, x_2, \dots, x_m)=f(x_1^0, x_2^0, \dots, x_m^0)+$$
$$+\sum_{k=1}^{n}\frac{1}{k!}\left[(x_1-x_1^0)\frac{\partial}{\partial x_1}+\dots+(x_m-x_m^0)\frac{\partial}{\partial x_m}\right]^k\times$$
$$\times f(x_1^0, x_2^0, \dots, x_m^0)+o(\rho^n). \tag{12.59}$$

Заметим, что в правой части (12.59) стоит сумма многочлена степени n от m переменных $x_1, x_2, \dots, x_m$ и остаточного члена $o(\rho^n)$.

Обозначим через $g_n(M)$ разность между $f(M)$ и указанным многочленом, т. е. положим

$$g_n(M)=f(M)-f(M_0)-$$
$$-\sum_{k=1}^{n}\frac{1}{k!}\left[(x_1-x_1^0)\frac{\partial}{\partial x_1}+\dots+(x_m-x_m^0)\frac{\partial}{\partial x_m}\right]^k f(M_0). \tag{12.60}$$

Теорема 12.15* будет доказана, если мы установим, что при выполнении условий этой теоремы $g_n(M)=o(\rho^n)$.

* При $n=1$ следует требовать, чтобы функция $u=f(M)$ была только задана в ε-окрестности точки M_0 и дифференцируема в самой точке M_0.

Доказательству теоремы 12.15* предпошлем две леммы.

Лемма 1. *Если функция $f(M)=f(x_1, x_2, \dots, x_m)$ n раз дифференцируема в точке $M_0(x_1^0, x_2^0, \dots, x_m^0)$, то как сама функция $g_n(M)$, определяемая равенством* (12.60), *так и все ее частные производные по любым переменным $x_1, x_2, \dots, x_m$ до порядка n включительно обращаются в нуль в точке M_0.*

Доказательство. При $n=1$ функция (12.60) принимает вид

$$g_1(M)=f(M)-f(M_0)-(x_1-x_1^0)\frac{\partial f}{\partial x_1}(M_0)-\dots-(x_m-x_m^0)\frac{\partial f}{\partial x_m}(M_0),$$

и равенства $g_1(M_0)=0$, $\frac{\partial g_1}{\partial x_i}(M_0)=0$ при всех $i=1, 2, \dots, m$ проверяются элементарно.

Для проведения индукции предположим, что лемма справедлива для некоторого номера $n\geqslant 1$, и докажем, что в таком случае она справедлива и для номера $n+1$.

Пусть функция $f(M)$ $n+1$ раз дифференцируема в точке M_0 и

$$g_{n+1}(M)=$$
$$=f(M)-f(M_0)-\sum_{k=1}^{n+1}\frac{1}{k!}\left[(x_1-x_1^0)\frac{\partial}{\partial x_1}+\dots+(x_m-x_m^0)\frac{\partial}{\partial x_m}\right]^k f(M_0). \tag{12.61}$$

Равенство $g_{n+1}(M)=0$ проверяется элементарно (достаточно учесть, что каждая круглая скобка $(x_i-x_i^0)$ в (12.61) обращается в нуль в точке M_0).

Нам остается доказать, что для любого $i=1, 2, \dots, m$ сама функция $\frac{\partial g_{n+1}}{\partial x_i}(M)$ и все частные производные этой функции до порядка n включительно обращаются в нуль в точке M_0, а для этого в силу сделанного нами предположения о справедливости леммы для номера n достаточно доказать, что функция $\frac{\partial g_{n+1}}{\partial x_i}(M)$ определяется равенством типа (12.60), а точнее, равенством

$$\frac{\partial g_{n+1}}{\partial x_i}(M)=\frac{\partial f}{\partial x_i}(M)-\frac{\partial f}{\partial x_i}(M_0)-$$
$$-\sum_{k=1}^{n}\frac{1}{k!}\left[(x_1-x_1^0)\frac{\partial}{\partial x_1}+\dots+(x_m-x_m^0)\frac{\partial}{\partial x_m}\right]^k\frac{\partial f}{\partial x_i}(M_0). \tag{12.62}$$

Так как все переменные x_i ($i=1, 2, \dots, m$) равноправны и входят в выражение для $g_{n+1}(M)$ симметрично, то достаточно доказать равенство (12.62) для $i=1$, т. е. доказать равенство

$$\frac{\partial g_{n+1}}{\partial x_1}(M) = \frac{\partial f}{\partial x_1}(M) - \frac{\partial f}{\partial x_1}(M_0) -$$

$$-\sum_{k=1}^{n} \frac{1}{k!}\left[(x_1 - x_1^0)\frac{\partial}{\partial x_1} + \ldots + (x_m - x_m^0)\frac{\partial}{\partial x_m}\right]^k \frac{\partial f}{\partial x_1}(M_0). \quad (12.63)$$

Из (12.61) очевидно, что для доказательства (12.63) достаточно убедиться, что для каждого номера $k=1, 2, \ldots, n+1$ при фиксированных $x_2, x_3, \ldots, x_m$

$$\frac{d}{dx_1}\left[(x_1 - x_1^0)\frac{\partial}{\partial x_1} + (x_2 - x_2^0)\frac{\partial}{\partial x_2} + \ldots + (x_m - x_m^0)\frac{\partial}{\partial x_m}\right]^k f(M_0) =$$

$$= k\left[(x_1 - x_1^0)\frac{\partial}{\partial x_1} + (x_2 - x_2^0)\frac{\partial}{\partial x_2} + \ldots + (x_m - x_m^0)\frac{\partial}{\partial x_m}\right]^{k-1} \frac{\partial f}{\partial x_1}(M_0).$$

$$(12.64)$$

Так как при дифференцировании по x_1 переменные $x_2, x_3, \ldots$ $\ldots, x_m$ фиксированы, то величину

$$D = (x_2 - x_2^0)\frac{\partial}{\partial x_2} + \ldots + (x_m - x_m^0)\frac{\partial}{\partial x_m}$$

при дифференцировании по x_1 можно рассматривать как постоянную. К этому следует добавить, что поскольку символы $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_m}$ используются для образования частных производных функции f в **фиксированной** точке M_0, то при дифференцировании по x_1 указанные символы нужно рассматривать как постоянные величины.

В силу сказанного для доказательства равенства (12.64) достаточно убедиться в справедливости равенства

$$\frac{d}{dx_1}\left[(x_1 - x_1^0)\frac{\partial}{\partial x_1} + D\right]^k = k\frac{\partial}{\partial x_1}\left[(x_1 - x_1^0)\frac{\partial}{\partial x_1} + D\right]^{k-1}. \quad (12.65)$$

Дифференцируя функцию $\left[(x_1 + x_1^0)\frac{\partial}{\partial x_1} + D\right]^k$ по x_1 как сложную и учитывая отмеченную выше независимость от x_1 символов D и $\frac{\partial}{\partial x_1}$, мы получим равенство (12.65). Индукция завершена.

Лемма 1 доказана.

Лемма 2. *Пусть* $g(M) = g(x_1, x_2, \ldots, x_m)$ — *произвольная функция, удовлетворяющая двум требованиям:*

1) $g(M)$ *n раз дифференцируема в точке* $M_0(x_1^0, x_2^0, \ldots, x_m^0)$;

2) *сама функция* $g(M)$ *и все ее частные производные по любым переменным* $x_1, x_2, \ldots, x_m$ *до порядка n включительно обращаются*

в нуль в указанной точке M_0. Тогда для функции $g(M)$ справедлива оценка

$$g(M)=o(\rho^n), \tag{12.66}$$

в которой через ρ обозначено расстояние $\rho(M, M_0)$ между точками M и M_0.

Доказательство. При $n=1$ утверждение леммы вытекает из условия дифференцируемости* функции $g(M)$ в точке M_0, которое имеет вид $g(M)-g(M_0)=\sum\limits_{i=1}^{n}\frac{\partial g}{\partial x_i}(M_0)(x_i-x_i^0)+o(\rho)$. Учитывая, что $g(M_0)=0$, $\frac{\partial g}{\partial x_i}(M_0)=0$ для всех $i=1, 2, \ldots, m$ мы и получим, что $g(M)=o(\rho)$.

Для проведения индукции предположим, что лемма 2 справедлива для некоторого номера $n\geqslant 1$, и докажем, что в таком случае она справедлива и для номера $n+1$.

Пусть функция $g(M)$ удовлетворяет двум требованиям леммы 2 для номера $n+1$. Тогда, очевидно, любая частная производная этой функции первого порядка $\frac{\partial g}{\partial x_i}(M)$, $i=1, 2, \ldots, m$, будет удовлетворять двум требованиям леммы 2 для номера n, а потому (в силу сделанного нами предположения о справедливости леммы 2 для номера n) будет справедлива оценка

$$\frac{\partial g}{\partial x_i}(M)=o(\rho^n). \tag{12.66*}$$

Заметим теперь, что поскольку $n\geqslant 1$, то $n+1\geqslant 2$ и функция $g(M)$, удовлетворяющая двум требованиям леммы 2 для номера $n+1$, во всяком случае, один раз дифференцируема в окрестности точки M_0. Поэтому для этой функции $g(M)$ выполнены условия теоремы 12.15 для номера $n=0$. Согласно указанной теореме для любой точки M из достаточно малой ε-окрестности точки M_0 на отрезке** M_0M найдется точка N такая, что справедлива формула

$$g(M)=g(M_0)+\frac{1}{1!}\sum_{i=1}^{m}(x_i-x_i^0)\frac{\partial g}{\partial x_i}(N). \tag{12.67}$$

Заметим теперь, что поскольку точка N лежит между точками M_0 и M, а ρ — это расстояние между точками M_0 и M, то $\rho(N, M_0)\leqslant\rho$, и потому из (12.66*) вытекает, что

$$\frac{\partial g}{\partial x_i}(N)=o(\rho^n).$$

* См. соотношение (12.16) из п. 2 § 4 настоящей главы.

** Т. е. на множестве точек вида $M_0+t(M-M_0)$, где t — любое число из сегмента $0\leqslant t\leqslant 1$.

Подставляя последнюю оценку в (12.67) и учитывая, что $g(M_0) = 0$, мы получим

$$g(M) = o(\rho^n) \sum_{i=1}^{m} |x_i - x_i^0|.$$

Так как $|x_i - x_i^0| \leqslant \sqrt{\sum_{i=1}^{m} (x_i - x_i^0)^2} = \rho$, то мы окончательно получим, что $g(M) = o(\rho^{n+1})$.

Индукция завершена. Лемма 2 доказана.

Доказательство теоремы 12.15* легко проводится с помощью леммы 1 и 2.

В самом деле, выше уже отмечалось, что для доказательства теоремы 12.15 достаточно установить, что при выполнении условий этой теоремы для функции (12.60) справедлива оценка

$$g_n(M) = o(\rho^n).$$

В силу леммы 1 сама функция (12.60) и все ее частные производные по любым переменным $x_1, x_2, \ldots, x_m$ до порядка n включительно обращаются в нуль в точке M_0. Но тогда в силу леммы 2 для функции (12.60) справедлива оценка $g_n(M) = o(\rho^n)$. Теорема 12.15* доказана.

§ 6. ЛОКАЛЬНЫЙ ЭКСТРЕМУМ ФУНКЦИИ m ПЕРЕМЕННЫХ

1. Понятие экстремума функции m переменных. Необходимые условия экстремума. Пусть функция m переменных $u = f(M) = f(x_1, x_2, \ldots, x_m)$ определена в некоторой окрестности точки $M_0(x_1^0, x_2^0, \ldots, x_m^0)$ пространства E^m.

Определение 1. *Будем говорить, что функция $u = f(M)$ имеет в точке M_0 локальный максимум [локальный минимум], если найдется такая δ-окрестность точки M_0, в пределах которой значение $f(M_0)$ является наибольшим [наименьшим] среди всех значений $f(M)$ этой функции.*

Определение 2. *Будем говорить, что функция $u = f(M)$ имеет в точке M_0 локальный экстремум, если она имеет в этой точке либо локальный максимум, либо локальный минимум.*

Установим необходимые условия локального экстремума функции $u = f(M)$, обладающей в данной точке M_0 частными производными первого порядка по всем переменным.

Докажем следующее

Утверждение. *Если функция $u = f(M) = f(x_1, x_2, \ldots, x_m)$ обладает в точке $M_0(x_1^0, x_2^0, \ldots, x_m^0)$ частными производными первого порядка по всем переменным $x_1, x_2, \ldots, x_m$ и имеет в этой* *точке локальный экстремум, то все частные производные первого порядка обращаются в точке M_0 в нуль, т. е.*

$$\frac{\partial u}{\partial x_1}(M_0)=0,\ \frac{\partial u}{\partial x_2}(M_0)=0,\ \ldots,\ \frac{\partial u}{\partial x_m}(M_0)=0. \qquad (12.68)$$

Доказательство. Установим справедливость первого равенства (12.68). Фиксируем у функции $u=f(x_1, x_2, \ldots, x_m)$ аргументы $x_2, x_3, \ldots, x_m$, положив их равными соответствующим координатам точки M_0, т. е. положив $x_2=x_2^0, x_3=x_3^0, \ldots, x_m=x_m^0$. При этом мы получим функцию $u=f(x_1, x_2^0, \ldots, x_m^0)$ одной переменной x_1. Производная этой функции одной переменной в точке $x_1=x_1^0$ совпадает с частной производной $\frac{\partial u}{\partial x_1}(M_0)$.

Так как функция m переменных $u=f(M)$ имеет локальный экстремум в точке M_0, то указанная функция одной переменной $u=f(x_1, x_2^0, \ldots, x_m^0)$ имеет локальный экстремум в точке $x_1=x_1^0$, и поэтому (в силу результатов п. 2 § 1 гл. 7) производная этой функции одной переменной в точке $x_1=x_1^0$, совпадающая с частной производной $\frac{\partial u}{\partial x_1}(M_0)$, равна нулю.

Первое равенство (12.68) доказано. Остальные равенства (12.68) доказываются аналогично.

Подчеркнем, что равенства (12.68) (т .е. обращение в нуль в данной точке M_0 всех частных производных первого порядка) являются лишь необходимыми и не являются достаточными условиями локального экстремума функции $u=f(M)$ в точке M_0.

Например, у функции двух переменных $u=xy$ обе частные производные $\frac{\partial u}{\partial x}$ и $\frac{\partial u}{\partial y}$ обращаются в нуль в точке $M_0(0, 0)$, но никакого экстремума в этой точке $M_0(0, 0)$ указанная функция не имеет, ибо эта функция $u=xy$ равна нулю в самой точке $M_0(0, 0)$, а в как угодно малой δ-окрестности этой точки принимает как положительные, так и отрицательные значения.

Точки, в которых обращаются в нуль все частные производные первого порядка функции $u=f(M)$, называются ***стационарными точками*** *этой функции.*

В каждой стационарной точке у функции $u=f(M)$ возможен локальный экстремум, однако наличие этого экстремума можно установить лишь с помощью достаточных условий локального экстремума, выяснению которых посвящен следующий пункт.

Из доказанного выше утверждения вытекает и другая форма необходимых условий локального экстремума: *если функция $u=f(M)$ дифференцируема в точке M_0 и имеет в этой точке локальный экстремум, то дифференциал $du|_{M^0}$ этой функции в точке M_0 равен нулю тождественно относительно дифференциалов независимых переменных $dx_1, dx_2, \ldots, dx_m$.*

В самом деле, поскольку

$$du|_{M_0} = \frac{\partial u}{\partial x_1}(M_0)\,dx_1 + \frac{\partial u}{\partial x_2}(M_0)\,dx_2 + \ldots + \frac{\partial u}{\partial x_m}(M_0)\,dx_m,$$

то из равенств (12.68) вытекает, что при любых $dx_1, dx_2, \ldots, dx_m$ справедливо равенство $du|_{M_0} = 0$.

2. Достаточные условия локального экстремума функции m переменных. При формулировке достаточных условий локального экстремума функции m переменных $u=f(M)$ важную роль будет играть второй дифференциал этой функции в обследуемой точке M_0.

В п. 2 § 5 настоящей главы мы убедились в том, что для случая, когда аргументы $x_1, x_2, \ldots, x_m$ два раза дифференцируемой функции $u=f(x_1, x_2, \ldots, x_m)$ являются либо независимыми переменными, либо линейными функциями некоторых независимых переменных, второй дифференциал этой функции в данной точке M_0 представляет собой квадратичную форму относительно дифференциалов аргументов $dx_1, dx_2, \ldots, dx_m$ следующего вида:

$$d^2u|_{M_0} = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}\,dx_i\,dx_k, \tag{12.69}$$

где

$$a_{ik} = a_{ki} = \frac{\partial^2 u}{\partial x_i \partial x_k}(M_0). \tag{12.70}$$

Для формулировки достаточных условий локального экстремума нам понадобятся некоторые сведения из теории квадратичных форм, которые мы для удобства читателя приводим ниже*.

Квадратичная форма относительно переменных $h_1, h_2, \ldots, h_m$

$$\Phi(h_1, h_2, \ldots, h_m) = \sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}h_ih_k \tag{12.71}$$

называется *положительно определенной* [*отрицательно определенной*], если для любых значений $h_1, h_2, \ldots, h_m$, одновременно не равных нулю, эта форма принимает строго положительные [строго отрицательные] значения.

Квадратичная форма (12.71) называется *знакоопределенной*, если она является либо положительно определенной, либо отрицательно определенной.

Квадратичная форма (12.71) называется *знакопеременной*, если она принимает как строго положительные, так и строго отрицательные значения.

---

* Все приводимые здесь определения и утверждения можно найти, например, в книге В. А. Ильина и Э. Г. Позняка «Линейная алгебра», (М., Наука, 1974).

Квадратичная форма (12.71) называется **квазизнакоопределенной**, если она принимает либо только неотрицательные, либо только неположительные значения, но при этом обращается в нуль для некоторых значений $h_1, h_2, \ldots, h_m$, одновременно не равных нулю.

Сформулируем так называемый **критерий Сильвестра** знакоопределенности квадратичной формы *.

Назовем **матрицей квадратичной формы** (12.71) следующую матрицу:

$$A=\begin{pmatrix} a_{11}\,a_{12} & \ldots & a_{1m} \\ a_{21}\,a_{22} & \ldots & a_{2m} \\ \cdots & \cdots & \cdots \\ a_{m1}a_{m2} & \ldots & a_{mm} \end{pmatrix}. \tag{12.72}$$

Если все элементы матрицы A удовлетворяют условию $a_{ik}=a_{ki}$ $(i=1, 2, \ldots, m;\ k=1, 2, \ldots, m)$, то указанная матрица называется **симметричной**.

Назовем **главными минорами** симметричной матрицы (12.72) следующие определители:

$$A_1=a_{11},\quad A_2=\begin{vmatrix} a_{11}a_{12} \\ a_{21}a_{22} \end{vmatrix},\quad A_3=\begin{vmatrix} a_{11}a_{12}a_{13} \\ a_{21}a_{22}a_{23} \\ a_{31}a_{32}a_{33} \end{vmatrix},\ \ldots,\ A_m=\begin{vmatrix} a_{11}\,a_{12}\ldots a_{1m} \\ a_{21}\,a_{22}\ldots a_{2m} \\ \cdots\cdots\cdots \\ a_{m1}a_{m2}\ldots a_{mm} \end{vmatrix}.$$

Критерий Сильвестра формулируется в виде следующих двух утверждений.

1°. *Для того чтобы квадратичная форма* (12.71) *с симметричной матрицей* (12.72) *являлась положительно определенной, необходимо и достаточно, чтобы все главные миноры матрицы* (12.72) *были положительны, т. е. чтобы были справедливы неравенства*

$$A_1>0,\ A_2>0,\ \ldots,\ A_m>0.$$

2°. *Для того чтобы квадратичная форма* (12.71) *с симметричной матрицей* (12.72) *являлась отрицательно определенной, необходимо и достаточно, чтобы знаки главных миноров матрицы* (12.72) *чередовались, причем знак A_1 был отрицателен, т. е. чтобы были справедливы неравенства*

$$A_1<0,\ A_2>0,\ A_3<0,\ A_4>0,\ \ldots .$$

Теперь мы подготовлены для того, чтобы сформулировать и доказать теорему, устанавливающую достаточные условия локального экстремума.

Теорема 12.16. *Пусть функция m переменных $u=f(M)=f(x_1, x_2, \ldots, x_m)$ один раз дифференцируема в некоторой окрестности точки $M_0(x_1^0, x_2^0, \ldots, x_m^0)$ и два раза дифференцируема в самой точке M_0. Пусть, кроме того, точка M_0 является стационар-*

* Дж. Сильвестр — английский математик (1814—1897).

ной точкой функции $u=f(M)$, т. е. $du|_{M_0}=0$. Тогда если второй дифференциал (12.69), (12.70) представляет собой положительно определенную [отрицательно определенную] квадратичную форму от переменных $dx_1, dx_2, \dots, dx_m$, то функция $u=f(M)$ имеет в точке M_0 локальный минимум [локальный максимум]. Если же второй дифференциал (12.69), (12.70) представляет собой знакопеременную квадратичную форму, то функция $u=f(M)$ не имеет локального экстремума в точке M_0.

Доказательство. Докажем сначала первую часть теоремы, предполагая ради определенности, что второй дифференциал (12.69), (12.70) представляет собой положительно определенную квадратичную форму от переменных $dx_1, dx_2, \dots, dx_m$. Докажем, что в этом случае функция $u=f(M)$ имеет в точке M_0 локальный минимум.

Разложим функцию $u=f(M)$ в окрестности точки M_0 по формуле Тейлора с остаточным членом в форме Пеано, беря в этой формуле $n=2$ *. Мы получим при этом, что

$$f(M)-f(M_0)=du|_{M_0}+\frac{1}{2}d^2u|_{M_0}+o(\rho^2), \tag{12.73}$$

причем в равенстве (12.73) дифференциалы dx_i переменных x_i, входящие в выражение для $du|_{M_0}$ и $d^2u|_{M_0}$, равны соответствующим приращениям $x_i-x_i^0$ эти переменных, а величина ρ равна

$$\begin{aligned}\rho&=\sqrt{(dx_1)^2+(dx_2)^2+\dots+(dx_m)^2}=\\&=\sqrt{(x_1-x_1^0)^2+(x_2-x_2^0)^2+\dots+(x_m-x_m^0)^2}.\end{aligned} \tag{12.74}$$

По условию теоремы точка M_0 является стационарной. Поэтому на основании результатов предыдущего пункта $du|_{M_0}=0$. Учитывая это равенство и полагая в выражениях (12.69), (12.70) для второго дифференциала $dx_i=x_i-x_i^0$, мы придадим формуле Тейлора (12.73) следующий вид:

$$f(M)-f(M_0)=\frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m}a_{ik}(x_i-x_i^0)(x_k-x_k^0)+o(\rho^2). \tag{12.75}$$

Достаточно доказать, что для всех достаточно малых ρ правая часть (12.75) положительна. (Это и будет означать, что в достаточной малой окрестности точки M_0 разность $f(M)-f(M_0)$ положительна, т. е. функция $u=f(M)$ имеет в точке M_0 локальный минимум.)

---

* Для функции $u=f(M)$ выполнены при $n=2$ все условия теоремы 12.15* (см. п. 4 § 5 настоящей главы).

Положим $h_i = \dfrac{x_i - x_i^0}{\rho}$, где $i=1, 2, \ldots$. Тогда из выражения (12.74) для ρ вытекают следующие соотношения:

$$|h_i| \leqslant 1, \quad h_1^2 + h_2^2 + \ldots + h_m^2 = 1. \qquad (12.76)$$

С помощью введенных обозначений равенство (12.75) может быть переписано в виде

$$f(M) - f(M_0) = \frac{\rho^2}{2} \sum_{i=1}^{m} \sum_{k=1}^{m} a_{ik} h_i h_k + o(\rho^2). \qquad (12.75^*)$$

Отношение $o(\rho^2)/\rho^2$ представляет собой бесконечно малую при $\rho \to 0$ (или при $M \to M_0$) функцию, которую мы обозначим через $\alpha(\rho)$. Введение этой функции позволяет нам записать равенство $o(\rho^2) = \rho^2 \alpha(\rho)$, с помощью которого мы придадим соотношению (12.75*) вид

$$f(M) - f(M_0) = \rho^2 \left[\frac{1}{2} \sum_{i=1}^{m} \sum_{k=1}^{m} a_{ik} h_i h_k + \alpha(\rho) \right]. \qquad (12.75^{**})$$

Теперь уже нетрудно доказать, что правая часть (12.75**) является положительной для всех достаточно малых ρ. Квадратичная форма $\Phi = \sum\limits_{i=1}^{m} \sum\limits_{k=1}^{m} a_{ik} h_i h_k$ представляет собой функцию, определенную и непрерывную на поверхности единичной сферы (12.76), представляющей собой замкнутое и ограниченное множество. По второй теореме Вейерштрасса (см. теорему 12.7 из п. 3 § 3) эта функция достигает на указанном множестве своей точной нижней грани μ, причем из положительной определенности квадратичной формы (12.71) и из того, что $h_1, h_2, \ldots, h_m$, удовлетворяющие соотношению (12.76), не равны одновременно нулю, вытекает, что указанная точная нижняя грань μ строго положительна.

Так как бесконечно малая при $\rho \to 0$ функция $\alpha(\rho)$ при всех достаточно малых ρ удовлетворяет неравенству $|\alpha(\rho)| < \mu$, то вся правая часть (12.75**) является положительной при всех достаточно малых ρ, т. е. при всех M, достаточно близких к M_0.

Это и означает, что функция $u = f(M)$ имеет в точке M_0 локальный минимум.

Совершенно аналогично доказывается, что в случае, когда второй дифференциал (12.69), (12.70) представляет собой отрицательно определенную квадратичную форму, функция $u = f(M)$ имеет в точке M_0 локальный максимум.

Докажем теперь вторую часть теоремы, т. е. докажем, что в случае, когда второй дифференциал (12.69), (12.70) представляет собой знакопеременную квадратичную форму, функция $u=f(M)$ не имеет локального экстремума в точке M_0.

Прежде всего установим следующее свойство знакопеременной квадратичной формы (12.71):

Если квадратичная форма $\Phi(h_1, h_2, \ldots, h_m)$ *знакопеременна, то найдутся две совокупности переменных* $(h_1', h_2', \ldots, h_m')$ *и* $(h_1'', h_2'', \ldots, h_m'')$ *такие, что*

$$(h_1')^2+(h_2')^2+\ldots+(h_m')^2=1, \quad (h_1'')^2+(h_2'')^2+\ldots+(h_m'')^2=1, \tag{12.77}$$

причем

$$\Phi(h_1' h_2', \ldots, h_m')>0, \quad \Phi(h_1'', h_2'', \ldots, h_m'')<0. \tag{12.78}$$

В самом деле, в силу определения знакопеременной квадратичной формы найдутся две совокупности аргументов $t_1', t_2', \ldots$ $\ldots, t_m'$ и $t_1'', t_2'', \ldots, t_m''$, состоящие из чисел, одновременно не равных нулю, и такие, что

$$\Phi(t_1', t_2', \ldots, t_m')>0, \quad \Phi(t_1'', t_2'', \ldots, t_m'')<0. \tag{12.79}$$

Положив

$$h_i'=\frac{t_i'}{\sqrt{(t_1')^2+(t_2')^2+\ldots+(t_m')^2}}, \quad h_i''=\frac{t_i''}{\sqrt{(t_1'')^2+(t_2'')^2+\ldots+(t_m'')^2}} \tag{12.80}$$

и учитывая, что из определения (12.71) квадратичной формы сразу же вытекает, что

$$\Phi(h_1', h_2', \ldots, h_m')=\frac{1}{(t_1')^2+(t_2')^2+\ldots+(t_m')^2}\Phi(t_1', t_2', \ldots, t_m'),$$

$$\Phi(h_1'', h_2'', \ldots, h_m'')=\frac{1}{(t_1'')^2+(t_2'')^2+\ldots+(t_m'')^2}\Phi(t_1'', t_2'', \ldots, t_m''),$$

мы получим (в силу (12.79)) неравенства (12.78), причем из соотношений (12.80) сразу же вытекают равенства (12.77).

Зафиксируем две совокупности переменных $h_1', h_2', \ldots, h_m'$ и $h_1'', h_2'', \ldots, h_m''$, удовлетворяющие соотношениям (12.77) и (12.78), и докажем, что для любого $\rho>0$ найдутся две точки $M'(x_1', x_2', \ldots$ $\ldots, x_m')$ и $M''(x_1'', x_2'', \ldots, x_m'')$ пространства E^m такие, что $\rho(M', M_0)=\rho(M'', M_0)=\rho$, причем

$$\frac{x_i'-x_i^0}{\rho}=h_i', \quad \frac{x_i''-x_i^0}{\rho}=h_i'' \quad \text{для всех } i=1, 2, \ldots, m. \tag{12.81}$$

В самом деле, положив для любого $\rho>0$ и для каждого номера i

$$x_i'=x_i^0+\rho h_i', \qquad x_i''=x_i^0+\rho h_i'',$$

мы удовлетворим соотношениям (12.81), причем в силу равенств (12.77) будут справедливы равенства

$$\rho(M', M_0) = \sqrt{(x_1'-x_1^0)^2+(x_2'-x_2^0)^2+\ldots+(x_m'-x_m^0)^2} =$$

$$= \rho\sqrt{(h_1')^2+(h_2')^2+\ldots+(h_m')^2} = \rho,$$

$$\rho(M'', M_0) = \sqrt{(x_1''-x_1^0)^2+(x_2''-x_2^0)^2+\ldots+(x_m''-x_m^0)^2} =$$

$$= \rho\sqrt{(h_1'')^2+(h_2'')^2+\ldots+(h_m'')^2} = \rho.$$

Теперь уже нетрудно убедиться в том, что для случая, когда второй дифференциал (12.69), (12.70) представляет собой знакопеременную квадратичную форму, функция $u=f(M)$ не имеет экстремума в точке M_0.

Записывая для функции $u=f(M)$ разложение в окрестности точки M_0 по формуле Тейлора с остаточным членом в форме Пеано и беря это разложение в указанных выше точках M' и M'', мы получим вместо (12.75) следующие два разложения:

$$f(M')-f(M_0) = \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}(x_i'-x_i^0)(x_k'-x_k^0)+o(\rho^2), \quad (12.82)$$

$$f(M'')-f(M_0) = \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}(x_i''-x_i^0)(x_k''-x_k^0)+o(\rho^2), \quad (12.83)$$

справедливые для всех достаточно малых $\rho>0$.

Подставляя в эти разложения значения $(x_i'-x_i^0)$ и $(x_i''-x_i^0)$ из равенств (12.81) и учитывая, что $o(\rho^2)=\rho^2\alpha(\rho)$, где $\alpha(\rho)\to 0$ при $\rho\to 0$, мы придадим разложениям (12.82) и (12.83) следующий вид:

$$f(M')-f(M_0) = \rho^2\left[\frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}h_i'h_k'+\alpha(\rho)\right],$$

$$f(M'')-f(M_0) = \rho^2\left[\frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} a_{ik}h_i''h_k''+\alpha(\rho)\right].$$

Последние два соотношения можно также переписать в виде

$$f(M')-f(M_0) = \rho^2\left[\frac{1}{2}\Phi(h_1', h_2', \ldots, h_m')+\alpha(\rho)\right], \quad (12.82^*)$$

$$f(M'')-f(M_0) = \rho^2\left[\frac{1}{2}\Phi(h_1'', h_2'', \ldots, h_m'')+\alpha(\rho)\right]. \quad (12.83^*)$$

Учитывая соотношения (12.78) и тот факт, что величины $\Phi(h_1', h_2', \ldots, h_m')>0$ и $\Phi(h_1'', h_2'', \ldots, h_m'')<0$ не зависят от ρ, и вспоминая, что $\rho=\rho(M', M_0)=\rho(M'', M_0)$, мы получим из соотношений (12.82*) и (12.83*), что для достаточно малого $\rho>0$ справедливы неравенства $f(M')>f(M_0)$ и $f(M'')<f(M_0)$, которые и доказывают отсутствие экстремума в точке M_0.

Теорема 12.16 полностью доказана.

Замечание. Если второй дифференциал два раза дифференцируемой в данной стационарной точке M_0 функции $u=f(M)$ представляет собой в этой точке квазизнакоопределенную квадратичную форму, то нельзя сказать ничего определенного о наличии или отсутствии в этой точке локального экстремума.

Так, например, у каждой из двух функций $u_1=x^3+y^3$ и $u_2=$ $=x^4+y^4$ второй дифференциал в стационарной точке $M_0(0, 0)$ тождественно равен нулю (т. е. представляет собой квазизнакоопределенную квадратичную форму), но только одна вторая из указанных двух функций имеет в этой точке локальный экстремум.

Для решения вопроса о локальном экстремуме для случая, когда второй дифференциал представляет собой квазизнакоопределенную квадратичную форму, следует привлечь дифференциалы более высоких порядков, но это выходит за рамки нашего курса.

Пример. Найти точки локального экстремума функции трех переменных

$$u=\lambda x^2+y^2+z^2+2y+2z, \tag{12.84}$$

где λ — вещественное число, отличное от нуля.

Так как $\dfrac{\partial u}{\partial x}=2\lambda x$, $\dfrac{\partial u}{\partial y}=2y+2$, $\dfrac{\partial u}{\partial z}=2z+2$, то единственной стационарной точкой является точка $M_0(0, -1, -1)$.

Далее очевидно, что второй дифференциал в этой точке имеет вид

$$d^2u|_{M_0}=2\lambda(dx)^2+2(dy)^2+2(dz)^2.$$

При $\lambda>0$ второй дифференциал в точке M_0 представляет собой положительно определенную квадратичную форму*, и потому функция (12.84) имеет в точке $M_0(0, -1, -1)$ локальный минимум.

При $\lambda<0$ указанный второй дифференциал представляет собой знакопеременную квадратичную форму**, и потому функция (12.84) не имеет локального экстремума в точке M_0.

3. Случай функции двух переменных. На практике часто встречается задача об экстремуме функции двух переменных $u=f(x,y)$.

---

* Ибо этот второй дифференциал принимает строго положительные значения при dx, dy и dz, одновременно не равных нулю.

** Ибо при $\lambda<0$ этот второй дифференциал положителен при $dx=0$, $dy=1$, $dz=1$ и отрицателен при $dx=1$, $dy=0$, $dz=0$.

В этом пункте мы приведем результаты, относящиеся к этому случаю.

Пусть частные производные $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x \partial y}$, $\frac{\partial^2 u}{\partial y^2}$ в некоторой точке $M_0(x_0, y_0)$ обозначены символами a_{11}, a_{12}, a_{21} соответственно. Справедливо следующее

Утверждение. *Пусть функция двух переменных $u=f(x, y)$ один раз дифференцируема в окрестности точки M_0 и два раза дифференцируема в самой точке M_0, и пусть M_0 является стационарной точкой. Тогда если в точке M_0 выполнено условие $a_{11}a_{22}-a_{12}^2>0$, то функция $u=f(x, y)$ имеет в точке M_0 локальный экстремум (максимум при $a_{11}<0$ и минимум при $a_{11}>0$).*

Если же в точке M_0 $a_{11}a_{22}-a_{12}^2<0$, то функция $u=f(x, y)$ не имеет в этой точке локального экстремума.*

Доказательство. Справедливость первой части утверждения непосредственно вытекает из теоремы 12.16 и критерия Сильвестра знакоопределенности квадратичной формы, ибо

$$A_1=a_{11}, \qquad A_2=a_{11}a_{22}-a_{12}^2.$$

Докажем вторую часть утверждения. Итак, пусть в точке M_0 справедливо неравенство $a_{11}a_{22}-a_{12}^2<0$. Докажем, что в этом случае второй дифференциал d^2u в точке M_0 представляет собой знакопеременную форму. Рассмотрим сначала случай $a_{11}\neq 0$. Используя введенные выше обозначения

$$\rho=\sqrt{(x-x_0)^2+(y-y_0)^2}, \quad h_1=\frac{x-x_0}{\rho}, \quad h_2=\frac{y-y_0}{\rho},$$

получим следующее выражение для второго дифференциала:

$$d^2u\,|_{M_0}=\rho^2\sum_{i,k=1}^{2} a_{ik}h_ih_k=\rho^2\left(a_{11}h_1^2+2a_{12}h_1h_2+a_{22}h_2^2\right)=$$

$$=\frac{\rho^2}{a_{11}}\left[(a_{11}h_1+a_{12}h_2)^2+(a_{11}a_{22}-a_{12}^2)\,h_2^2\right].$$

Легко проверить, что при $h_1=1$, $h_2=0$ и при $h_1=\frac{a_{12}}{\sqrt{a_{11}^2+a_{12}^2}}$, $h_2=\frac{-a_{11}}{\sqrt{a_{11}^2+a_{12}^2}}$** дифференциал $d^2u\,|_{M_0}$ имеет разные знаки, т. е. является знакопеременной формой, и поэтому согласно теореме 12.16 функция не имеет в точке M_0 локального экстремума.

---

* Случай $a_{11}a_{22}-a_{12}^2=0$ требует дополнительного исследования.

** При этом ρ может быть как угодно малой величиной. Условие $h_1^2+h_2^2=1$ выполнено.

Рассмотрим теперь случай $a_{11}=0$. Тогда из условия $a_{11}a_2-a_{12}^2<$ <0 вытекает, что $a_{12}\neq 0$. Следовательно, так же как и выше, имеем

$$d^2u\,|_{M_0}=\rho^2h_2(2a_{12}h_1+a_{22}h_2). \tag{12.85}$$

Пусть $h_1\neq 0$ и величина h_2 столь мала (из условия $h_1^2+h_2^2=1$ следует, что такой выбор h_1 и h_2 возможен), что выражение $(2a_{12}h_1+a_{22}h_2)$ сохраняет знак величины $2a_{12}h_1$. Тогда из формулы (12.85) вытекает, что $d^2u\,|_{M_0}$ имеет разные знаки при $h_2>0$ и $h_2<0$, т. е. функция $u=f(x, y)$ не имеет локального экстремума в точке M_0. Утверждение полностью доказано.

Замечание. *Требование* $d^2f(M_0)\geqslant 0$ [*соответственно* $d^2f(M_0)\leqslant$ $\leqslant 0$] *является необходимым условием локального минимума* [*максимума*] *в точке* M_0 *дважды дифференцируемой в этой точке функции* $f(M)$.

В самом деле, пусть ради определенности $f(M)$ имеет в точке M_0 локальный минимум, но условие $d^2f(M_0)\geqslant 0$ не выполнено. Тогда найдутся $h_1, h_2, \ldots, h_m$ такие, что

$$d^2f(M_0)=\sum_{i=1}^{m}\sum_{k=1}^{m}\frac{\partial^2 f}{\partial x_i\partial x_k}(M_0)\,h_ih_k<0.$$

Рассмотрим функцию $F(t)=f(x_1^0+th_1, \ldots, x_m^0+th_m)$, заведомо определенную при всех t, достаточно малых по модулю. Функция $F(t)$ обязана иметь локальный минимум в точке $t=0$, чему противоречит условие

$$F''(0)=d^2f(M_0)<0.$$

ДОПОЛНЕНИЕ 1

Градиентный метод поиска экстремума сильно выпуклой функции

В этом дополнении излагается теория широко применяемого на практике градиентного метода поиска экстремума сильно выпуклой функции.

Идея этого метода чрезвычайно проста. Для приближенного отыскания точки минимума функции m переменных используется тот факт, что градиент этой функции имеет направление, совпадающее с направлением *наибольшего возрастания* этой функции. Значит, вектор $-\operatorname{grad} f(x_0)$ в каждой точке $x_0=$ $=(x_1^0, x_2^0, \ldots, x_m^0)$ направлен в сторону *наибольшего убывания* функции $f(x)$. Это дает основание ожидать, что если, отправляясь от некоторого нулевого приближения $x_0=(x_1^0, x_2^0, \ldots, x^0_m)$ мы построим k-е приближение $x_k=(x_1^k, x_2^k, \ldots, x_m^k)$ по рекуррентной формуле

$$x_{k+1}=x_k-\alpha\operatorname{grad} f(x_k),$$

то при достаточно малом положительном α последовательность точек $\{x_k\}$ сойдется к точке минимума функции $f(x)$.

Строгой реализации этой простой идеи и посвящено настоящее дополнение.

1. Выпуклые множества и выпуклые функции. Пусть $x_1 = (x_1^1, x_2^1, \ldots, x_m^1)$ и $x_2 = (x_1^2, x_2^2, \ldots, x_m^2)$ — две точки m-мерного евклидова пространства E^m, которые мы можем рассматривать как векторы в E^m с соответствующими координатами.

Назовем отрезком или сегментом, соединяющим точки x_1 и x_2, множество точек пространства E^m вида $x_1 + t(x_2 - x_1)$, где t — любое число из сегмента $0 \leqslant t \leqslant 1$.

Будем обозначать отрезок, соединяющий точки x_1 и x_2, символом $\overline{x_1 x_2}$.

Определение 1. *Множество Q точек пространства E^m называется выпуклым, если оно обладает следующим свойством: каковы бы ни были две точки x_1 и x_2, принадлежащие множеству Q, отрезок $\overline{x_1 x_2}$, их соединяющий, также принадлежит этому множеству.*

Примером выпуклого множества в пространстве E^m может служить m-мерный шар (безразлично, открытый или замкнутый) или полупространство $x_m \geqslant 0$ (т. е. множество всех точек $(x_1, x_2, \ldots, x_m)$ пространства E^m, m-я координата которых удовлетворяет условию $x_m \geqslant 0$).

Примером множества Q, не являющегося выпуклым, может служить дополнение m-мерного шара или m-мерный открытый шар, из которого удалена хотя бы одна точка.

Пусть Q — некоторое множество точек пространства E^m, а x — любая фиксированная точка этого пространства.

Назовем расстоянием от точки x до множества Q точную нижнюю грань расстояний от точки x до всевозможных точек этого множества.

Будем обозначать расстояние от точки x до множества Q символом $\rho(x, Q)$. Итак, по определению

$$\rho(x, Q) = \inf_{y \in Q} \rho(x, y).$$

Для любого множества Q пространства E^m и любой точки x этого пространства существует расстояние $\rho(x, Q)$ *. В частности, если точка x принадлежит множеству Q, то $\rho(x, Q) = 0$.

Однако у множества Q не всегда существует точка y такая, что $\rho(x, y) = \rho(x, Q)$.

Так, например, если множество Q представляет открытый m-мерный шар, а x — точка E^m, лежащая вне этого шара, то у такого множества Q не существует точки y такой, что $\rho(x, y) =$

* Ибо множество $\rho(x, y)$ для всевозможных y, принадлежащих Q, всегда ограничено снизу (например, числом нуль).

$=\rho(x, Q)$ (ибо для всех точек y открытого шара справедливо неравенство $\rho(x, y)>\rho(x, Q)$).

Если все же у множества Q существует точка y такая, что $\rho(x, y)=\rho(x, Q)$, то эта точка y называется **проекцией точки** x **на множество** Q.

Проекцию точки x на множество Q будем обозначать символом $P_Q(x)$.

Подчеркнем, что если точка x принадлежит множеству Q то $P_Q(x)=x$.

Итак, проекция $P_Q(x)$ точки x на множество Q определяется соотношением

$$\rho(x, P_Q(x))=\rho(x, Q)=\inf_{y\in Q}\rho(x, y).$$

Полезно отметить, что может существовать несколько проекций точки x на множество Q. Так, например, если Q—m-мерная сфера с центром в точке x, то любая точка Q является проекцией точки x на множество Q.

Справедлива, однако, следующая лемма:

Лемма 1. *Если множество Q пространства E^m является выпуклым и замкнутым, а x — любая точка E^m, то существует и притом единственная проекция точки x на множество Q.*

Доказательство. Сначала докажем существование *хотя бы одной* проекции точки x на множество Q. Обозначим через $\rho(x, Q)$ расстояние от точки x до множества Q. По определению $\rho(x, Q)$ как точной нижней грани $\inf\limits_{y\in Q}\rho(x, y)$ найдется последовательность $\{y_n\}$ точек множества Q такая, что $\rho(x, y_n)\to\rho(x, Q)$.

По определению предела числовой последовательности для любого $\varepsilon>0$ все элементы y_n, начиная с некоторого номера, удовлетворяют соотношению

$$\rho(x, Q)-\varepsilon<\rho(x, y_n)<\rho(x, Q)+\varepsilon.$$

Отсюда следует, что последовательность $\{y_n\}$ точек пространства E^m во всяком случае является ограниченной и потому в силу теоремы Больцано — Вейерштрасса (см. п. 2 § 2 гл. 12) из этой последовательности можно выделить сходящуюся подпоследовательность y_{k_n}, где $n=1, 2, 3, \ldots$. Обозначим через y предел подпоследовательности $\{y_{k_n}\}$. В силу замкнутости множества Q точка y принадлежит этому множеству. Остается доказать, что

$$\rho(x, y)=\rho(x, Q)=\lim_{n\to\infty}\rho(x, y_{k_n}).$$

Для доказательства этого заметим, что в силу неравенств треугольника

$$\rho(x, y_{k_n}) \leqslant \rho(x, y) + \rho(y, y_{k_n}) \text{ и } \rho(x, y) \leqslant \rho(x, y_{k_n}) + \rho(y_{k_n}, y)$$

справедливо соотношение

$$|\rho(x, y_{k_n}) - \rho(x, y)| \leqslant \rho(y, y_{k_n}).$$

Из этого соотношения и их сходимости подпоследовательности $\{y_{k_n}\}$ к y вытекает, что $\lim\limits_{n\to\infty} \rho(x, y_{k_n}) = \rho(x, y)$, т. е. $\rho(x, Q) = \rho(x, y)$.

Тем самым доказательство *существования* хотя бы одной проекции точки x на множество Q завершено.

Докажем теперь, что существует *только одна* проекция точки x на множество Q. Предположим, что существуют две различные проекции y_1 и y_2 точки x на множество Q. Так как множество Q является выпуклым, то весь отрезок $\overline{y_1y_2}$, соединяющий точки y_1 и y_2, принадлежит множеству Q. В частности, множеству Q принадлежит середина $\frac{y_1+y_2}{2}$ указанного отрезка. Убедимся в том, что *расстояние* $\rho\left(x, \frac{y_1+y_2}{2}\right)$ *от точки x до указанной середины отрезка $\overline{y_1y_2}$ строго меньше расстояния* $\rho(x, y_1) = \rho(x, y_2)$.

Исключим из рассмотрения тривиальный случай, когда $\frac{y_1+y_2}{2} = x$. В этом случае $\rho\left(x, \frac{y_1+y_2}{2}\right) = 0$, в то время как $\rho(x, y_1) = \rho(x, y_2) > 0$, ибо в противном случае обе точки y_1 и y_2 совпадали бы с x и не могли бы быть различными. Итак, в тривиальном случае $\frac{y_1+y_2}{2} = x$ неравенство

$$\rho\left(x, \frac{y_1+y_2}{2}\right) < \rho(x, y_1) = \rho(x, y_2) \tag{12.1.1}$$

очевидно.

Докажем теперь неравенство (12.1.1) в случае, когда $\frac{y_1+y_2}{2} \neq x$.

Используя свойства скалярного произведения двух векторов пространства E^m (см., например, замечание 2 из п. 1 § 1 гл. 12), мы получим соотношение

$$\rho^2\left(x, \frac{y_1+y_2}{2}\right) = \left(\frac{y_1+y_2}{2} - x, \frac{y_1+y_2}{2} - x\right) =$$

$$= \left(\frac{y_1-x}{2} + \frac{y_2-x}{2}, \frac{y_1-x}{2} + \frac{y_2-x}{2}\right) =$$

$$= \frac{1}{4}[(y_1-x, y_1-x) + 2(y_1-x, y_2-x) + (y_2-x, y_2-x)]. \tag{12.1.2}$$

Убедимся теперь в справедливости строгого неравенства

$$|(y_1-x,\ y_2-x)|<\sqrt{(y_1-x,\ y_1-x)}\ \sqrt{(y_2-x,\ y_2-x)}. \qquad (12.1.3)$$

В сноске на с. 485 доказано, что для любых векторов* a и b пространства E^m, не коллинеарных друг другу (т. е. таких, что $a\neq\lambda b$ ни для одного вещественного λ), справедливо строгое неравенство Коши — Буняковского

$$|(a,\ b)|<\sqrt{(a,\ a)(b,\ b)}.$$

Это означает, что для доказательства неравенства (12.1.3) нам достаточно убедиться в том, что векторы y_1-x и y_2-x не коллинеарны, т. е. убедиться в том, что ни для одного вещественного λ не может быть справедливо равенство

$$y_1-x=\lambda(y_2-x). \qquad (12.1.4)$$

Если бы равенство (12.1.4) было справедливо для такого λ, для которого $|\lambda|\neq 1$, то было бы невозможно равенство $\rho(y_1,\ x)=\rho(y_2,\ x)$.

Справедливость равенства (12.1.4) для $\lambda=1$ противоречила бы тому, что точки y_1 и y_2 являются различными.

Наконец, справедливость равенства (12.1.4) для $\lambda=-1$ означала бы, что $\frac{y_1+y_2}{2}=x$, а этот случай мы исключили.

Итак, равенство (12.1.4) несправедливо ни для одного вещественного λ, а потому доказательство неравенства (12.1.3) завершено.

Сопоставляя равенство (12.1.2) с неравенством (12.1.3), получим, что

$$\rho^2\left(x,\ \frac{y_1+y_2}{2}\right)<\frac{1}{4}[(y_1-x,\ y_1-x)+$$

$$+2\sqrt{(y_1-x,\ y_1-x)}\ \sqrt{(y_2-x,\ y_2-x)}+(y_2-x,\ y_2-x)]=$$

$$=\frac{1}{4}\left[\sqrt{(y_1-x,\ y_1-x)}+\sqrt{(y_2-x,\ y_2-x)}\right]^2=$$

$$=\frac{1}{4}[\rho(x,\ y_1)+\rho(x,\ y_2)]^2=\rho^2(x,\ y_1)=\rho^2(x,\ y_2).$$

Тем самым доказательство неравенства (12.1.1) завершено. Но это неравенство означает, что у множества Q нашлась точка $\frac{y_1+y_2}{2}$, более близкая к x, чем точки y_1 и y_2, а это противоречит тому, что каждая из точек y_1 и y_2 является проекцией точки x на множество Q, т. е. является точной нижней гранью расстояния $\rho(x,\ y)$ для всевозможных y, принадлежащих Q.

* Векторы в данном дополнении не будем выделять жирным шрифтом.

Полученное противоречие показывает, что наше предположение о том, что существуют две различные проекции y_1 и y_2 точки x на множество Q, является ошибочным.

Доказательство леммы 1 полностью завершено.

Перейдем теперь к определению выпуклой функции.

Определение 2. *Функция $f(x)$, заданная на выпуклом множестве Q пространства E^m, называется выпуклой вниз или просто выпуклой на этом множестве, если для любых двух точек x_1 и x_2 множества Q и для любого вещественного числа t из сегмента $0 \leqslant t \leqslant 1$ справедливо неравенство*

$$f[x_1+t(x_2-x_1)] \leqslant f(x_1)+t[f(x_2)-f(x_1)]. \qquad (12.1.5)$$

Определение 3. *Функция $f(x)$, заданная на выпуклом множестве Q пространства E^m, называется строго выпуклой на этом множестве, если для любых двух точек x_1 и x_2 множества Q и для любого вещественного числа t из интервала $0<t<1$ справедливо строгое неравенство*

$$f[x_1+t(x_2-x_1)] < f(x_1)+t[f(x_2)-f(x_1)]. \qquad (12.1.6)$$

Ясно, что всякая строго выпуклая на множестве Q функция $f(x)$ является выпуклой на этом множестве.

Легко установить достаточное условие выпуклости [соответственно строгой выпуклости] дважды дифференцируемой на выпуклом множестве Q функции $f(x)$.

Лемма 2. *Пусть функция $f(x)$ задана и два раза дифференцируема на выпуклом множестве Q. Тогда, для того чтобы эта функция являлась выпуклой [строго выпуклой], на множестве Q достаточно, чтобы второй дифференциал d^2f этой функции во всех точках Q являлся квазиположительно определенной [строго положительно определенной] квадратичной формой.*

Доказательство. Пусть x_1 и x_2 — любые две фиксированные точки множества Q. Рассмотрим на сегменте $0 \leqslant t \leqslant 1$ следующую функцию одной независимой переменной t:

$$F(t)=f[x_1+t(x_2-x_1)]-f(x_1)-t[f(x_2)-f(x_1)]. \qquad (12.1.7)$$

Напомним, что второй дифференциал d^2f функции $f(x) = f(x_1, x_2, \ldots, x_m)$ независимых переменных $(x_1, x_2, \ldots, x_m)$ в данной точке $x=(x_1, x_2, \ldots, x_m)$ равен *

$$d^2f(x)=\sum_{i=1}^{m}\sum_{k=1}^{m}\frac{\partial^2 f}{\partial x_i \partial x_k}(x)\,\Delta x_i \Delta x_k. \qquad (12.1.8)$$

Дифференцируя функцию (12.1.7) два раза по t по правилу дифференцирования сложной функции, получим

* См. п. 2 § 5 гл. 12.

$$F''(t) = \sum_{i=1}^{m} \sum_{k=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_k} [x_1 + t(x_2 - x_1)](x_i^2 - x_i^1)(x_k^2 - x_k^1), \quad (12.1.9)$$

где $(x_1^1, x_2^1, \ldots, x_m^1)$ и $(x_1^2, x_2^2, \ldots, x_m^2)$ — координаты точек x_1 и x_2 соответственно.

Сопоставляя соотношения (12.1.8) и (12.1.9), мы убедимся в справедливости равенства

$$F''(t) = d^2f[x_1 + t(x_2 - x_1)], \quad (12.1.10)$$

в котором в выражении для d^2f приращения Δx_i взяты равными $x_i^2 - x_i^1$.

Дальнейшие рассуждения ради определенности проведем для случая, когда второй дифференциал d^2f во всех точках Q является *квазиположительно определенной квадратичной формой*.

В этом случае для всех t из сегмента $0 \leqslant t \leqslant 1$ правая (а значит, и левая) часть (12.1.10) неотрицательна, т. е. для всех t из сегмента $0 \leqslant t \leqslant 1$

$$F''(t) \geqslant 0. \quad (12.1.11)$$

В силу определения 2 и соотношения (12.1.7) нам достаточно доказать, что для всех t из сегмента $0 \leqslant t \leqslant 1$ справедливо неравенство

$$F(t) \leqslant 0. \quad (12.1.12)$$

Для доказательства неравенства (12.1.12) используем соотношение (12.1.11) и легко проверяемые равенства

$$F(0) = 0, \qquad F(1) = 0. \quad (12.1.13)$$

Предположим, что внутри сегмента $0 \leqslant t \leqslant 1$ существует хотя бы одна точка t, в которой $F(t) > 0$. Тогда функция $F(t)$ достигает своего максимального на сегменте $0 \leqslant t \leqslant 1$ значения в некоторой внутренней точке t_0 этого сегмента, причем $F(t_0) > 0$. В этой точке t_0 функция $F(t)$ имеет локальный максимум, а потому $F'(t_0) = 0$. Но из неравенства (12.1.11) вытекает, что производная $F'(t)$ не убывает на всем сегменте $0 \leqslant t \leqslant 1$, а потому и на сегменте $t_0 \leqslant t \leqslant 1$. Отсюда и из условия $F'(t_0) = 0$ следует, что производная $F'(t)$ неотрицательна всюду на сегменте $t_0 \leqslant t \leqslant 1$, а поэтому функция $F(t)$ не убывает на этом сегменте. Это приводит нас к неравенству

$$F(1) \geqslant F(t_0) > 0,$$

противоречащему второму соотношению (12.1.13).

Полученное противоречие доказывает, что наше предположение о том, что на сегменте $0 \leqslant t \leqslant 1$ существует хотя бы одна точка t, в которой $F(t) > 0$, является ошибочным, т. е. доказывает справедливость всюду на сегменте $0 \leqslant t \leqslant 1$ неравенства (12.1.12).

Тем самым первая часть леммы (о выпуклости $f(x)$ при условии, что d^2f является квазиположительно определенной квадратичной формой) доказана.

Вторая часть леммы (о строгой выпуклости $f(x)$ при условии, что d^2f является *строго положительно определенной квадратичной формой*) доказывается аналогично. Исходя из неравенства (12.1.11), справедливого на этот раз со знаком $>$, и из равенств (12.1.13) и предположив, что внутри сегмента $0\leqslant t\leqslant 1$ существует хотя бы одна точка t, в которой $F(t)\geqslant 0$, мы придем к выводу, что $F(t)$ имеет внутри сегмента $0\leqslant t\leqslant 1$ точку локального максимума t_0, причем $F(t_0)\geqslant 0$. Но тогда, поскольку $F'(t_0)=0$, мы получим из (12.1.11), что $F'(t)>0$ всюду на полуинтервале $t_0<t\leqslant 1$, а это означает, что $F(1)>F(t_0)\geqslant 0$.

Мы снова получаем противоречие со вторым соотношением (12.1.13), которое доказывает, что $F(t)<0$ всюду на интервале $0<t<1$, т. е. доказывает строгую выпуклость $f(x)$ на множестве Q.

Лемма 2 полностью доказана.

Доказанная лемма естественно наводит на мысль о рассмотрении следующего еще более узкого класса выпуклых на выпуклом множестве Q и два раза дифференцируемых на этом множестве функций.

Определение 4. *Два раза дифференцируемая на выпуклом множестве Q функция $f(x)$ называется сильно выпуклой на этом множестве, если существуют такие две положительные постоянные k_1 и k_2, что второй дифференциал d^2f этой функции, определяемый соотношением* (12.1.8), *во всех точках x множества Q удовлетворяет неравенствам*

$$k_1(\Delta x)^2\leqslant d^2f\leqslant k_2(\Delta x)^2. \tag{12.1.14}$$

(В этих неравенствах через Δx обозначен вектор с координатами $(\Delta x_1, \Delta x_2, \dots, \Delta x_m)$, а символ $(\Delta x)^2$ обозначает скалярный квадрат этого вектора, т. е. скалярное произведение $(\Delta x, \Delta x)$.)

Из левого неравенства (12.1.14) сразу же вытекает, что второй дифференциал сильно выпуклой функции представляет собой строго положительно определенную во всех точках множества Q функцию, а потому (в силу леммы 2) сильно выпуклая на множестве Q функция заведомо является строго выпуклой на этом множестве.

Вместе с тем класс сильно выпуклых функций достаточно широк и важен в прикладных задачах, и мы ограничимся этим классом при изложении теории градиентного метода поиска минимума.

Начнем с выяснения вопроса о существовании и единственности минимума.

2. Существование минимума у сильно выпуклой функции и единственность минимума у строго выпуклой функции. Пусть функция $f(x)$ определена на выпуклом множестве Q.

Будем говорить, что эта функция имеет в точке x_0 множества Q **локальный минимум**, если существует такая δ-окрестность этой точки C_δ, что значение $f(x_0)$ не больше значений $f(x)$ этой функции во всех точках пересечения δ-окрестности C_δ и множества Q. При таком определении понятие локального минимума включает в себя и точки краевого минимума функции $f(x)$ на границе множества Q.

Таким образом, при данном нами определении можно подразделить точки минимума на точки **внутреннего локального минимума** (для случая, когда эти точки являются внутренними точками Q) и точки **краевого локального минимума** (для случая, когда эти точки являются граничными точками Q).

Для изучения вопроса о существовании и единственности точки локального минимума нам понадобится следующая вспомогательная лемма.

Лемма 3. *Пусть на выпуклом множестве Q задана дифференцируемая выпуклая функция $f(x)$. Для того чтобы эта функция имела локальный минимум в точке x_0 множества Q, необходимо и достаточно, чтобы для любого вектора Δx, для которого точка $x+\Delta x$ принадлежит множеству Q, было справедливо неравенство**

$$(\operatorname{grad} f(x_0),\ \Delta x) \geqslant 0. \tag{12.1.15}$$

Доказательство. *Необходимость.* В силу утверждения, доказанного в п. 8 § 4 гл. 12, левая часть (12.1.15) равна произведению производной функции $f(x)$ в точке x_0 **по направлению** вектора Δx на длину $|\Delta x|$ этого вектора:

$$(\operatorname{grad} f(x_0),\ \Delta x) = \frac{\partial f}{\partial e}(x_0)\,|\Delta x|, \tag{12.1.16}$$

где $e=\Delta x/|\Delta x|$ — единичный вектор в направлении Δx.

Так как x_0 является точкой локального минимума функции $f(x)$, то производная $\partial f/\partial e(x_0)$ по любому направлению $e=\Delta x/|\Delta x|$ неотрицательна (точнее, равна нулю в случае, если x_0 — точка внутреннего локального экстремума, и неотрицательна в случае, если x_0 — точка краевого локального экстремума).

Итак, правая часть (12.1.16) (а потому и левая часть (12.1.15)) неотрицательна. Необходимость доказана.

Достаточность. Пусть для любого вектора Δx, для которого точка $x_0+\Delta x$ принадлежит Q, справедливо неравенство (12.1.15). Докажем, что точка x_0 является точкой локального минимума функции $f(x)$.

* В неравенстве (12.1.15) берется скалярное произведение векторов $\operatorname{grad} f(x_0)$ и Δx. Определение $\operatorname{grad} f(x)$ см. в п. 8 § 4 гл. 12.

Так как функция $f(x)$ по условию является выпуклой на множестве Q, то для любых двух точек x_1 и x_2 этого множества и любого числа t из сегмента $0 \leqslant t \leqslant 1$ справедливо неравенство (12.1.5). Полагая в этом неравенстве $x_1 = x_0$, $x_2 = x_0 + \Delta x$, мы можем переписать это неравенство в виде

$$f(x_0 + \Delta x) - f(x_0) \geqslant \frac{f(x_0 + t\Delta x) - f(x_0)}{t}. \tag{12.1.17}$$

Считая x_0 и Δx фиксированными, перейдем в неравенстве (12.1.17) к пределу при $t \to 0+0$. По определению производной по направлению (см. п. 8 § 4 гл. 12) предел при $t \to 0+0$ правой части (12.1.17) в точности равен произведению, стоящему в правой части (12.1.16). Поэтому в силу соотношений (12.1.15) и (12.1.16) этот предел неотрицателен. Учитывая, что левая часть (12.1.17) не зависит от t, мы получим в пределе при $t \to 0+0$ из неравенства (12.1.17), что

$$f(x_0 + \Delta x) - f(x_0) \geqslant 0.$$

Последнее неравенство, справедливое для любого вектора Δx, для которого точка $x_0 + \Delta x$ принадлежит Q, доказывает, что функция $f(x)$ имеет в точке x_0 локальный минимум. Достаточность доказана.

Лемма 3 полностью доказана.

Замечание 1. Из приведенного нами доказательства очевидно, что для случая, когда точка x_0 является внутренней точкой множества Q, т. е. когда речь идет о внутреннем локальном минимуме, в формулировке леммы 3 знак $\geqslant$ в неравенстве (12.1.15) можно заменить на знак $=$.

Замечание 2. При доказательстве необходимости леммы 3 мы не использовали требования выпуклости функции $f(x)$. Поэтому доказательство необходимости проходит без требования выпуклости функции $f(x)$. Иными словами, справедливо следующее

Утверждение. *Если функция $f(x)$ дифференцируема на выпуклом множестве Q и имеет локальный минимум во внутренней [в граничной] точке x_0 этого множества, то для любого вектора Δx для которого точка $x_0 + \Delta x$ принадлежит Q, справедливо неравенство*

$$(\operatorname{grad} f(x_0), \Delta x) = 0 \quad [(\operatorname{grad} f(x_0), \Delta x) \geqslant 0].$$

Перейдем к вопросу об единственности и о существовании точки локального минимума.

Теорема (об единственности локального минимума у строго выпуклой функции). *Если функция $f(x)$ дифференцируема и строго выпукла на выпуклом множестве Q, то она может иметь локальный минимум только в одной точке этого множества.*

Доказательство. Предположим, что функция $f(x)$ имеет локальный минимум в двух различных точка x_1 и x_2 множества Q. Тогда условие выпуклости (12.1.5) для точек x_1 и x_2 можно записать в виде

$$f(x_2)-f(x_1) \geqslant \frac{f[x_1+t(x_2-x_1)]-f(x_1)}{t} \qquad (12.1.18)$$

(здесь t — любое число из интервала $0<t<1$).

Меняя в соотношении (12.1.8) точки x_1 и x_2 ролями, мы получим неравенство

$$f(x_1)-f(x_2) \geqslant \frac{f[x_2+t(x_1-x_2)]-f(x_2)}{t}. \qquad (12.1.19)$$

В пределе при $t\to 0+0$ правая часть (12.1.18) [соответственно правая часть (12.1.19)] дает производную функции $f(x)$ по направлению вектора x_2-x_1 [соответственно вектора x_1-x_2], взятую в точке x_1 [соответственно в точке x_2], умноженную на $|x_2-x_1|$. Так как обе точки x_1 и x_2 являются точками локального минимума, то обе указанные производные по направлению неотрицательны, т. е. пределы правых частей (12.1.18) и (12.1.19) при $t\to 0+0$ оба неотрицательны.

Таким образом, из неравенств (12.1.18) и (12.1.19) в пределе при $t\to 0+0$ мы получим

$$f(x_2)-f(x_1)\geqslant 0,$$

$$f(x_1)-f(x_2)\geqslant 0.$$

Сопоставление двух последних неравенств приводит к заключению о том, что $f(x_1)=f(x_2)$.

Используя равенство $f(x_1)=f(x_2)$, мы получим из условия строгой выпуклости (12.1.6), что

$$f[x_1+t(x_2-x_1)]<f(x_1) \qquad (12.1.20)$$

для *всех* t из интервала $0<t<1$.

Неравенство (12.1.20) противоречит тому, что функция $f(x)$ имеет локальный минимум в точке x_1 (в точке $x_1+t(x_2-x_1)$, как угодно близкой при малом t к точке x_1, функция $f(x)$ имеет значение, меньшее значения $f(x_1)$).

Полученное противоречие доказывает, что наше предположение о том, что функция $f(x)$ имеет локальный минимум в двух различных точках множества Q, является ошибочным. Теорема доказана.

Существование локального минимума докажем при более сильных ограничениях, чем единственность.

Теорема (о существовании локального минимума у сильно выпуклой функции). *Если функция $f(x)$ сильно выпукла на замкнутом выпуклом множестве Q, то* *у этой функции существует на множестве Q точка x_0 локального минимума* *.

Доказательство. Сначала отметим, что теорема заведомо справедлива для случая, когда выпуклое замкнутое множество Q является, кроме того, ограниченным. В этом случае по второй теореме Вейерштрасса (см. теорему 12.7) функция $f(x)$, будучи во всяком случае непрерывной на множестве Q, достигает в некоторой точке x_0 этого множества своего минимального на Q значения. Указанная точка x_0 и является точкой локального минимума.

Остается доказать теорему в случае, когда выпуклое замкнутое множество Q не является ограниченным. В этом случае мы фиксируем некоторую внутреннюю точку x_1 множества Q и разложим функцию $f(x)$ по формуле Тейлора с центром в точке x_1, взяв остаточный член $R_2(x)$ в форме Лагранжа ** (см. п. 3 § 5 гл. 12). Указанное разложение будет иметь вид

$$f(x)=f(x_1)+df(x_1)+\frac{1}{2}d^2f[x_1+\theta(x-x_1)], \qquad (12.1.21)$$

где θ — число из интервала $0<\theta<1$, так что точка $x_1+\theta(x-x_1)$ принадлежит отрезку, соединяющему точки x_1 и x_2 ***.

Если обозначить через Δx вектор $x-x_1$, то для $df(x_1)$ будет справедливо равенство

$$df(x_1)=(\operatorname{grad} f(x_1),\ \Delta x).$$

Из этого равенства вытекает, что

$$|df(x_1)|\leqslant|\operatorname{grad} f(x_1)|\,|\Delta x|. \qquad (12.1.22)$$

Далее, используя левое неравенство в определении сильной выпуклости (12.1.14), мы придем к неравенству

$$d^2f[x_1+\theta(x-x_1)]\geqslant k_1(\Delta x)^2. \qquad (12.1.23)$$

Из соотношений (12.1.21)—(12.1.23) заключаем, что

$$f(x)-f(x_1)\geqslant-|df(x_1)|+\frac{1}{2}d^2f[x_1+\theta(x-x_1)]\geqslant$$

$$\geqslant-|\operatorname{grad} f(x_1)|\,|\Delta x|+\frac{k_1}{2}|\Delta x|^2,$$

* Так как сильно выпуклая на выпуклом множестве Q функция $f(x)$ является строго выпуклой на этом множестве, то по предыдущей теореме точка x_0 будет единственной точкой локального минимума.

** Мы учитываем, что сильно выпуклая на множестве Q функция $f(x)$ два раза дифференцируема на этом множестве.

*** Какова бы ни была точка x множества Q, отрезок, соединяющий точки x_1 и x, принадлежит множеству Q в силу выпуклости этого множества. В сноске к теореме Тейлора 12.15 отмечалось, что в качестве окрестности центра разложения можно брать любую звездную окрестность этого центра, т. е. можно брать все множество Q.

так что

$$f(x)-f(x_1)\geqslant|\Delta x|\left[\frac{k_1}{2}|\Delta x|-|\operatorname{grad} f(x_1)|\right]. \qquad (12.1.24)$$

Учитывая, что точка x_1 фиксирована и величина $|\operatorname{grad} f(x_1)|$ представляет собой некоторое фиксированное число, мы заведомо можем выбрать положительное число R настолько большим, чтобы при $|\Delta x|>R$ выражение в квадратных скобках в (12.1.24) было положительным.

Это означает, что при $|\Delta x|>R$ справедливо неравенство $f(x)>f(x_1)$, т. е. всюду вне замкнутого шара C_R радиуса R с центром в точке x_1 значения $f(x)$ превосходят значение $f(x_1)$ (в центре указанного шара).

Обозначим через Q_R пересечение множества Q с указанным шаром C_R. Так как оба множества Q и C_R являются выпуклыми и замкнутыми, то и их пересечение Q_R также является выпуклым и замкнутым. Так как, кроме того, множество Q_R является ограниченным, то по доказанному выше функция $f(x)$ имеет на множестве Q_R единственную точку x_0 локального минимума.

Поскольку мы доказали, что во всех точках Q, лежащих за пределами Q_R, значения $f(x)$ превосходят $f(x_1)$, то эти значения тем более превосходят $f(x_0)$, т. е. точка x_0 является точкой локального минимума $f(x)$ и на всем множестве Q.

Теорема полностью доказана.

3. Поиск минимума сильно выпуклой функции. Мы доказали, что сильно выпуклая функция $f(x)$, заданная на замкнутом выпуклом множестве Q, имеет на этом множестве единственную точку x_0 локального минимума.

Обратимся к построению и обоснованию алгоритма, с помощью которого отыскивается эта точка x_0.

Фиксируем произвольную точку x_1 множества Q и произвольное число α, удовлетворяющее неравенствам

$$0<\alpha<\frac{2}{k_2}, \qquad (12.1.25)$$

где k_2 — постоянная из неравенства (12.1.14), определяющего сильную выпуклость функции $f(x)$.

Отправляясь от x_1 как от первого приближения, составим итерационную последовательность $\{x_k\}$ с помощью рекуррентного соотношения

$$x_{k+1}=P_Q(x_k-\alpha \operatorname{grad} f(x_k)), \quad k=1,\ 2,\ 3,\ \ldots . \qquad (12.1.26)$$

В настоящем пункте мы докажем следующее утверждение.

Основная теорема. *Пусть функция $f(x)$ является сильно выпуклой на замкнутом множестве Q, и пусть x_1 — произвольная точка множества Q. Тогда итерационная последовательность $\{x_k\}$, определяемая рекуррентным соотношением* (12.1.26), *при любом α, удовлетворяющем неравенствам* (12.1.25), *сходится к точке x_0 локального минимума функции $f(x)$.*

Подчеркнем, что эта теорема дает алгорит отыскания любого (внутреннего или краевого) локального минимума функции $f(x)$, являющейся сильно выпуклой на произвольном (не обязательно ограниченном) замкнутом выпуклом множестве.

Доказательству основной теоремы предпошлем четыре леммы.

Лемма 4. *Если Q — выпуклое замкнутое множество E^m, x — произвольная фиксированная точка E^m, а y — произвольная точка множества Q*, то

$$(x-P_Q(x),\ y-P_Q(x)) \leqslant 0. \tag{12.1.27}$$

Доказательство. Предположим, что неравенство (12.1.27) несправедливо. Тогда существует точка y множества Q такая, что

$$(x-P_Q(x),\ y-P_Q(x)) > 0. \tag{12.1.28}$$

Из (12.1.28) сразу же вытекает, что точка y не совпадает с $P_Q(x)$.

В силу выпуклости множества Q любая точка $z=P_Q(x)+ +t(y-P_Q(x))$ отрезка, соединяющего точки $P_Q(x)$ и y, принадлежит множеству Q. Вычислим расстояние между любой такой точкой z и точкой x:

$$\rho^2(z,\ x) = (x-P_Q(x)-t(y-P_Q(x)),\ x-P_Q(x)-t(y-P_Q(x))) = \\ =\rho^2(x,\ P_Q(x))-2t(x-P_Q(x),\ y-P_Q(x))+t^2\rho^2(y,\ P_Q(x)). \tag{12.1.29}$$

Так как x и y фиксированы, а t — любое число из сегмента $0\leqslant t\leqslant 1$, то в силу неравенства (12.1.28) мы можем взять t удовлетворяющим неравенству

$$0 < t < \frac{2(x-P_Q(x),\ y-P_Q(x))}{\rho^2(y,\ P_Q(x))}.$$

При таком выборе t

$$-2t(x-P_Q(x),\ y-P_Q(x))+t^2\rho^2(y,\ P_Q(x)) < 0,$$

и мы получим из (12.1.29), что

$$\rho^2(z,\ x) < \rho^2(x,\ P_Q(x)).$$

Последнее неравенство противоречит тому, что точка $P_Q(x)$ является проекцией точки x на множество Q: у множества Q нашлась точка z, удаленная от x меньше, чем $P_Q(x)$ от x. Полученное противоречие завершает доказательство леммы.

Лемма 5. *Пусть $f(x)$ дифференцируема и выпукла на замкнутом выпуклом множестве Q. Если при некотором положительном α проекция $P_Q(x_0-\alpha\,\mathrm{grad}\,f(x_0))$ точки $x_0-\alpha\,\mathrm{grad}\,f(x_0)$ на* *множество Q совпадает с точкой x_0 этого множества, то функция $f(x)$ имеет в точке x_0 локальный минимум.*

Доказательство. Используя лемму 4, запишем неравенство (12.1.27) для точек $x = x_0 - \alpha \operatorname{grad} f(x_0)$ и $y = x_0 + \Delta x$, где Δx — любой вектор, для которого точка $y = x_0 + \Delta x$ принадлежит Q. В результате получим

$$(x_0 - \alpha \operatorname{grad} f(x_0) - P_Q(x_0 - \alpha \operatorname{grad} f(x_0)),$$
$$x_0 + \Delta x - P_Q(x_0 - \alpha \operatorname{grad} f(x_0))) \leqslant 0.$$

Учитывая, что $P_Q(x_0 - \alpha \operatorname{grad} f(x_0)) = x_0$, получим из последнего неравенства следующее соотношение:

$$(\operatorname{grad} f(x_0),\ \Delta x) \geqslant 0.$$

Это соотношение, справедливое для любого вектора Δx, для которого точка $x_0 + \Delta x$ принадлежит Q, в силу леммы 3 устанавливает, что функция $f(x)$ имеет в точке x_0 локальный минимум. Лемма 5 доказана.

Предположим, что функция $f(x)$ является сильно выпуклой на ограниченном замкнутом выпуклом множестве Q. Обозначим через m минимальное значение $f(x)$ на множестве Q, а через μ число, строго большее m, так что

$$\mu > m = \min_{x \in Q} f(x).$$

Фиксируем число ν, строго большее μ, и обозначим через $\hat{Q}$ — подмножество тех точек x множества Q, для которых

$$\mu \leqslant f(x) \leqslant \nu. \tag{12.1.30}$$

Множество $\hat{Q}$ как подмножество ограниченного множества Q само является ограниченным.

Убедимся в том, что множество $\hat{Q}$ является *замкнутым*. Пусть $\{x_k\}$ — произвольная сходящаяся последовательность точек множества $\hat{Q}$. Требуется доказать, что предел x_0 * этой последовательности также принадлежит множеству $\hat{Q}$. Так как каждая точка x_k принадлежит множеству $\hat{Q}$, то для каждого номера k

$$\mu \leqslant f(x_k) \leqslant \nu. \tag{12.1.31}$$

Сильно выпуклая функция $f(x)$ во всяком случае непрерывна на Q, а поэтому из сходимости последовательности $\{x_k\}$ к x_0 в силу определения непрерывности функции по Гейне вытекает сходимость последовательности $f(x_k)$ к числу $f(x_0)$. Так как все элементы сходящейся числовой последовательности $f(x_k)$ удовлетворяют неравенствам (12.1.31), то и предел $f(x_0)$ этой последовательности удовлетворяет неравенствам $\mu \leqslant f(x_0) \leqslant \nu$

* Так как исходное множество Q является замкнутым, то предел x_0 во всяком случае принадлежит Q.

(см. гл. 3, следствие 2 из теоремы 3.13), а это и означает, что точка x_0 принадлежит множеству $\hat{Q}$. Доказательство замкнутости множества $\hat{Q}$ завершено.

Итак, множество $\hat{Q}$ всех точек x из множества Q, для которых справедливы неравенства (12.1.30), является замкнутым и ограниченным.

Докажем теперь следующую лемму.

Лемма 6. *Пусть функция $f(x)$ сильно выпукла на выпуклом замкнутом множестве Q, x — любая точка Q, α — любое положительное число, символ Δx обозначает разность*

$$\Delta x = P_Q(x - \alpha \operatorname{grad} f(x)) - x. \qquad (12.1.32)$$

Тогда справедливо неравенство

$$(\operatorname{grad} f(x), \Delta x) \leqslant -\frac{1}{\alpha}|\Delta x|. \qquad (12.1.33)$$

Если же множество Q, кроме того, ограничено и точка x принадлежит подмножеству $\hat{Q}$ тех точек Q, для которых справедливы неравенства (12.1.30) *при* $\mu > \min_{x\in Q} f(x)$, *то найдется строго положительное число γ такое, что справедливо неравенство*

$$|\Delta x| \geqslant \gamma. \qquad (12.1.34)$$

Доказательство. Докажем сначала неравенство (12.1.33). Фиксируем произвольную точку x множества Q и, привлекая лемму 4, запишем неравенство (12.1.27), взяв в нем вместо x точку $x - \alpha \operatorname{grad} f(x)$, а вместо y точку x. При этом получим неравенство

$$(x - \alpha \operatorname{grad} f(x) - P_Q(x - \alpha \operatorname{grad} f(x)),\ x - P_Q(x - \alpha \operatorname{grad} f(x))) \leqslant 0,$$

которое с учетом обозначения $\Delta x = P_Q(x - \alpha \operatorname{grad} f(x)) - x$ перепишется в виде

$$(-\alpha \operatorname{grad} f(x) - \Delta x,\ -\Delta x) \leqslant 0.$$

Из последнего неравенства и из свойств скалярного произведения вытекает, что

$$\alpha(\operatorname{grad} f(x), \Delta x) + |\Delta x^2| \leqslant 0,$$

а это и приводит к неравенству (12.1.33).

Остается доказать, что при дополнительном предположении о том, что Q ограничено и что x принадлежит подмножеству $\hat{Q}$, существует $\gamma > 0$ такое, что справедливо неравенство (12.1.34). Рассмотрим неотрицательную функцию точки x вида

$$|\Delta x| = |P_Q(x - \alpha \operatorname{grad} f(x)) - x|. \qquad (12.1.35)$$

Убедимся в том, что эта функция является непрерывной на множестве Q функцией точки x.

Докажем сначала, что векторная функция $P_Q(x)$ является непрерывной функцией точки x. Для этого достаточно доказать неравенство

$$|P_Q(x+\Delta x)-P_Q(x)|\leqslant|\Delta x|, \qquad (12.1.36)$$

справедливое для любых векторов x и Δx.

В силу леммы 4 справедливы неравенства

$$(x-P_Q(x),\ P_Q(x+\Delta x)-P_Q(x))\leqslant 0,$$
$$(x+\Delta x-P_Q(x+\Delta x),\ P_Q(x)-P_Q(x+\Delta x))\leqslant 0.$$

Используя эти неравенства и неравенство Коши — Буняковского получим цепочку соотношений

$$|P_Q(x+\Delta x)-P_Q(x)|^2=(P_Q(x+\Delta x)-P_Q(x),$$
$$P_Q(x+\Delta x)-P_Q(x))=(P_Q(x+\Delta x)-x,\ P_Q(x+\Delta x)-P_Q(x))+$$
$$+(x-P_Q(x),\ P_Q(x+\Delta x)-P_Q(x))\leqslant$$
$$\leqslant(P_Q(x+\Delta x)-x,\ P_Q(x+\Delta x)-P_Q(x))=P_Q(x+\Delta x)-x-\Delta x,$$
$$P_Q(x+\Delta x)-P_Q(x))+(\Delta x,\ P_Q(x+\Delta x)-P_Q(x))\leqslant$$
$$\leqslant(\Delta x,\ P_Q(x+\Delta x)-P_Q(x))\leqslant|\Delta x|\,|P_Q(x+\Delta x)-P_Q(x)|,$$

из которой и вытекает неравенство (12.1.36).

Итак, доказано, что $P_Q(x)$ является непрерывной векторной функцией точки x. Из сильной выпуклости $f(x)$ на Q вытекает, что функция $\alpha\,\mathrm{grad}\,f(x)$ также является непрерывной на Q векторной функцией точки x. Но тогда из теоремы о непрерывности сложной функции и непрерывности разности непрерывных функций вытекает, что и функция

$$P_Q(x-\alpha\,\mathrm{grad}\,f(x))-x$$

является непрерывной на множестве Q векторной функцией точки x.

Модуль указанной векторной функции, т. е. скалярная функция (12.1.35) тем более является непрерывной на множестве Q, а потому и на его подмножестве $\hat{Q}$.

Итак, функция (12.1.35) непрерывна и неотрицательна всюду на замкнутом ограниченном множестве $\hat{Q}$. В таком случае, по второй теореме Вейерштрасса (см. теорему 12.7) эта функция достигает на множестве $\hat{Q}$ своего неотрицательного минимального значения γ. Указанное минимальное значение γ заведомо строго положительно, ибо если бы γ равнялась нулю, то на множестве $\hat{Q}$ нашлась бы точка x_0 такая, что $P_Q(x_0-\alpha\,\mathrm{grad}\,f(x_0))-x_0=0$, а это означало бы в силу леммы 5, что в этой точке x_0 множества $\hat{Q}$ функция $f(x)$ имеет единственный на множестве Q локальный минимум (в то время как этот минимум по определению $\hat{Q}$ лежит вне Q). Итак, $\gamma>0$, и неравенство (12.1.34) доказано.

Лемма 6 полностью доказана.

Лемма 7. *Пусть функция $f(x)$ сильно выпукла на выпуклом замкнутом множестве Q, x — любая точка Q, α — любое число, удовлетворяющее неравенствам* (12.1.25), Δx — *разность вида* (12.1.32). *Тогда при переходе из точки x в точку $x^* = P_Q(x - \alpha\,\mathrm{grad}\, f(x))$ значение функции $f(x)$ не возрастает, причем* *

$$f(x) - f(x^*) \geqslant \left(\frac{1}{\alpha} - \frac{k_2}{2}\right)|\Delta x|^2. \qquad (12.1.37)$$

Если же множество Q, кроме того, ограничено и точка x принадлежит подмножеству $\hat{Q}$ тех точек Q, для которых справедливо неравенство (12.1.30) *при* $\mu > \min\limits_{x\in Q} f(x)$, *то неравенство* (12.1.37) *переходит в неравенство*

$$f(x) - f(x^*) \geqslant \left(\frac{1}{\alpha} - \frac{k_2}{2}\right)\gamma^2, \qquad (12.1.38)$$

где $\gamma > 0$ — постоянная из леммы 6.

Доказательство. Достаточно для любой точки x множества Q установить неравенство (12.1.37), ибо из этого неравенства и из неравенства (12.1.34) сразу вытекает и неравенство (12.1.38) (для точек x, принадлежащих $\hat{Q}$, при условии, что Q ограничено).

Сначала докажем неравенство (12.1.37) для случая, когда точка x является внутренней точкой множества Q. Имея в виду, что точка $x^* = P_Q(x - \alpha\,\mathrm{grad}\, f(x))$ принадлежит множеству Q, на котором функция $f(x)$ сильно выпукла, выразим значение $f(x^*)$ по формуле Тейлора с центром в точке x, взяв остаточный член $R_2(x^*)$ в форме Лагранжа. При этом получим

$$f(x^*) = f(x) + (\mathrm{grad}\, f(x), \Delta x) + \frac{1}{2} d^2 f(x + \theta\Delta x), \qquad (12.1.39)$$

где $\Delta x = x^* - x = P_Q(x - \alpha\,\mathrm{grad}\, f(x)) - x$, $0 < \theta < 1$.

Используя неравенство (12.1.33) и правое неравенство (12.1.14), мы получим из формулы Тейлора (12.1.39)

$$f(x^*) - f(x) \leqslant -\frac{1}{\alpha}|\Delta x|^2 + \frac{k_2}{2}|\Delta x|^2,$$

так что для случая внутренней точки x неравенство (12.1.37) доказано.

* Из (13.1.25) вытекает, что $\dfrac{1}{\alpha} - \dfrac{k_2}{2} > 0$.

Пусть теперь x является граничной точкой множества Q. По определению граничной точки* найдется последовательность $\{x_n\}$ внутренних точек множества Q, сходящихся к x.

Для каждой точки x_n по формуле Тейлора с центром в этой точке мы получим

$$f(x^*)=f(x_n)+(\operatorname{grad} f(x_n),\ (x^*-x_n))+ \\ +\frac{1}{2}\,d^2f[x_n+\theta_n(x^*-x_n)], \tag{12.1.40}$$

где $0<\theta_n<1$.

Учитывая, что правое неравенство (12.1.14) справедливо для d^2f в любой точке множества Q и что $\operatorname{grad} f(x)$ является непрерывной векторной функцией точки x на множестве Q, мы получим, что в пределе при $n\to+\infty$ из соотношения (12.1.40) вытекает справедливость неравенства (12.1.37) для граничной точки x множества Q.

Лемма 7 доказана.

Перейдем теперь непосредственно к доказательству основной теоремы. Сначала докажем основную теорему при дополнительном предположении о том, что замкнутое выпуклое множество Q является также ограниченным.

Возьмем произвольную точку x_1 множества Q и составим итерационную последовательность $\{x_k\}$ точек, определяемых рекуррентным соотношением (12.1.26), при условии, что число α удовлетворяет неравенствам (12.1.25).

Из леммы 7 (а точнее, из неравенства (12.1.37)) сразу же вытекает, что

$$f(x_k)-f(x_{k+1})\geqslant\left(\frac{1}{\alpha}-\frac{k_2}{2}\right)|x_k-x_{k+1}|^2\geqslant 0.$$

Таким образом, последовательность $\{f(x_k)\}$ является невозрастающей. Так как, кроме того, эта последовательность ограничена снизу (минимальным значением m функции $f(x)$ на множестве Q), то она является сходящейся (см. теорему 3.15 из гл. 3). Обозначим предел последовательности $\{f(x_k)\}$ через μ. Ясно, что $\mu\geqslant m$, где m — минимальное значение $f(x)$ на множестве Q.

Кроме того, поскольку все члены невозрастающей сходящейся последовательности не меньше ее предела (см. замечание 3 к теореме 3.15, гл. 3), то для всех номеров k справедливо неравенство

$$f(x_k)\geqslant\mu. \tag{12.1.41}$$

Докажем, что для предела μ справедливо равенство

$$\mu=m=\min_{x\in Q} f(x).$$

* Так как множество Q выпукло, то оно не может иметь изолированных граничных точек.

Предположим, что это равенство несправедливо, т. е. предположим, что $\mu > m$. Тогда если обозначить через ν максимальное значение $f(x)$ на множестве Q, а через $\hat{Q}$ подмножество тех точек Q, для которых справедливы неравенства (12.1.30), то в силу леммы 7 найдется строго положительная постоянная γ такая, что справедливо неравенство (12.1.38), которое приводит к следующему неравенству:

$$f(x_k) - f(x_{k+1}) \geqslant \left(\frac{1}{\alpha} - \frac{k_2}{2}\right)\gamma^2 > 0, \tag{12.1.42}$$

справедливому для любого номера k.

Суммируя неравенства (12.1.42), записанное для номеров k, равных 1, 2, ..., $n-1$, мы получим, что для любого номера n

$$f(x_1) - f(x_n) \geqslant (n-1)\left(\frac{1}{\alpha} - \frac{k_2}{2}\right)\gamma^2$$

или, что то же самое,

$$f(x_n) \leqslant f(x_1) - (n-1)\left(\frac{1}{\alpha} - \frac{k_2}{2}\right)\gamma^2. \tag{12.1.43}$$

Неравенства (12.1.43), справедливые для любого номера n, противоречат неравенству (12.1.41), ибо величина, стоящая в правой части (12.1.43), при достаточно большом номере n становится меньше числа μ.

Полученное противоречие доказывает ошибочность предположения $\mu > m$, т. е. доказывает что $\mu = m$. Итак, доказано, что последовательность $f(x_k)$ сходится к минимальному значению m функции $f(x)$ на множестве Q.

Остается доказать, что сама итерационная последовательность $\{x_k\}$ сходится к той точке x_0, в которой это минимальное значение достигается*.

Фиксируем произвольное положительное число ε и обозначим через C_ε открытый m-мерный шар радиуса ε с центром в точке x_0. Далее обозначим через $\hat{Q}_\varepsilon$ ту часть множества Q, которая не содержит точек шара C_ε. Ясно, что $\hat{Q}_\varepsilon$ — замкнутое ограниченное множество, так что функция $f(x)$ достигает (в силу второй теоремы Вейерштрасса) своего минимального на этом множестве значения, которое мы обозначим через m_ε.

Можно утверждать, что $m_\varepsilon > m$, ибо в противном случае нарушалось бы условие существования у функции $f(x)$ на множестве Q единственной точки локального минимума.

Далее, можно утверждать, что на множестве $\hat{Q}_\varepsilon$ имеется лишь конечное число точек последовательности $\{x_k\}$ (ибо последовательность $\{f(x_k)\}$, у которой имеется бесконечное чис-

* Нами уже доказано, что минимальное значение функции $f(x)$ на множестве Q достигается в единственной точке этого множества.

ло элементов, удовлетворяющих неравенству $f(x_k) \geqslant m_\varepsilon > m$, не может сходиться к числу m).

Значит, мы доказали, что для любого $\varepsilon > 0$ найдется номер N, начиная с которого все элементы последовательности $\{x_k\}$ лежат в шаре C_ε радиуса ε с центром в точке x_0. Это и означает, что последовательность $\{x_k\}$ сходится к точке x_0.

Тем самым для случая ограниченного замкнутого выпуклого множества Q основная теорема доказана.

Пусть теперь Q — неограниченное замкнутое выпуклое множество. Снова фиксируем произвольную точку x_1 этого множества и составим итерационную последовательность (12.1.26) при условии, что число α удовлетворяет неравенствам (12.1.25).

При доказательстве теоремы о существовании локального минимума у сильно выпуклой функции (см. п. 2) мы установили, что точка x_0 локального минимума сильно выпуклой функции $f(x)$ на неограниченном замкнутом выпуклом множестве Q лежит в той части Q_R множества Q, которая содержится в шаре C_R с центром в точке x_1, радиус R которого выбран из условия

$$\frac{k_1}{2} R - |\operatorname{grad} f(x_1)| > 0.$$

Там же установлено, что подмножество Q_R множества Q является ограниченным выпуклым замкнутым множеством и что всюду вне Q_R значения $f(x)$ превосходят $f(x_1)$.

Так как в силу леммы 7 (а точнее, в силу неравенства (12.1.37)) последовательность $\{f(x_k)\}$ является невозрастающей, а за пределами Q_R все значения $f(x)$ превосходят $f(x_1)$, то все точки итерационной последовательности $\{x_k\}$ лежат в Q_R, а потому для любого номера k

$$P_Q(x_k - \alpha \operatorname{grad} f(x_k)) = P_{Q_R}(x_k - \alpha \operatorname{grad} f(x_k)).$$

Значит, итерационную последовательность (12.1.26) можно заменить на

$$x_{k+1} = P_{Q_R}(x_k - \alpha \operatorname{grad} f(x_k)),$$

после чего все дальнейшие рассуждения сведутся к ограниченному замкнутому выпуклому множеству Q_R, т. е. к уже рассмотренному выше случаю.

Основная теорема полностью доказана.

Замечание 1. Особенно просто выглядит последовательность (12.1.26) для случая, когда множество Q совпадает со всем пространством E^m. В этом случае для любой точки x справедливо равенство $P_Q(x) = x$, и потому рекуррентная формула (12.1.26) принимает вид

$$x_{k+1} = x_k - \alpha \operatorname{grad} f(x_k).$$

ДОПОЛНЕНИЕ 2

Метрические, нормированные пространства

В этом дополнении будут изложены важные понятия и факты общей топологии, которые часто употребляются в различных разделах математики. Читатель без труда обнаружит, что эти понятия и факты являются естественным обобщением ряда определений и утверждений, содержащихся в предыдущих главах.

Метрические пространства

1. Определение метрического пространства. Примеры. Выше мы уже подчеркивали, что фундаментальную роль в анализе играет понятие предела. В основе этого понятия лежит определение расстояния между числами, т. е. абсолютная величина разности этих чисел. Поэтому представляется естественным ввести понятие расстояния уже не между двумя числами, а между двумя произвольными элементами некоторого абстрактного множества X. Ясно, что при этом это расстояние должно обобщать свойства расстояния между числами числовой оси. В связи с вышесказанным дадим следующее определение.

Определение 1. *На множестве X определена структура метрического пространства, если задана функция $\rho(x, y)$ двух произвольных элементов этого множества x и y, удовлетворяющая аксиомам:*

1) $\rho(x, y)=0$ *тогда и только тогда, когда* $x=y$;

2) $\rho(x, y)=\rho(y, x)$ *(аксиома симметрии)*;

3) $\rho(x, y)\leqslant\rho(y, z)+\rho(z, x)$ *(неравенство треугольника).*

Функция $\rho(x, y)$ называется метрикой или функцией расстояния, число $\rho(x, y)$ называется расстоянием между точками x и y множества X.

Таким образом, метрическое пространство образуют множество X и функция расстояния ρ. Поэтому обозначается метрическое пространство обычно так: (X, ρ) или просто X, если ясно, о какой метрике идет речь.

Если в аксиоме 3) положить $x=y$, то, учитывая 1), получается, что $0\leqslant\rho(y, z)$, т. е. функция расстояния — неотрицательная функция своих аргументов.

Приведем примеры наиболее часто встречающихся метрических пространств.

Примеры.

1) Множество вещественных чисел превращается в метрическое пространство, если для любых чисел x, y положить $\rho(x, y)=|x-y|$. Это метрическое пространство обычно обозначается E^1.

2) Координатное n-мерное пространство $X=A^n$, точки которого (или элементы множества X) — упорядоченные наборы n чисел, $x=(x_1, \dots, x_n)$, превращается в метрическое пространство, если положить $\rho(x, y)=\left(\sum_{i=1}^{n}|x_i-y_i|^2\right)^{1/2}$, где $x=(x_1, \dots, x_n)$, $y=(y_1, \dots, y_n)$; обозначается оно* через $E^n=(X, \rho)$. Аксиомы 1) и 2) определения метрического пространства, как легко видеть, выполняются. Справедливость аксиомы 3) вытекает из неравенства Коши — Буняковского для сумм (см. § 5 гл. 9):

$$\rho^2(x, y)=\sum_{i=1}^{n}|x_i-y_i|^2=\sum_{i=1}^{n}[(x_i-z_i)+(z_i-y_i)]^2=$$

$$=\sum_{i=1}^{n}(x_i-z_i)^2+2\sum_{i=1}^{n}(x_i-z_i)(z_i-y_i)+\sum_{i=1}^{n}(z_i-y_i)^2\leqslant$$

$$\leqslant\sum_{i=1}^{n}(x_i-z_i)^2+2\sum_{i=1}^{n}|x_i-z_i||z_i-y_i|+\sum_{i=1}^{n}(z_i-y_i)^2\leqslant$$

$$\leqslant\sum_{i=1}^{n}(x_i-z_i)^2+2\left(\sum_{i=1}^{n}|x_i-z_i|^2\right)^{1/2}\left(\sum_{i=1}^{n}|z_i-y_i|^2\right)^{1/2}+\sum_{i=1}^{n}(z_i-y_i)^2=$$

$$=\rho^2(x, z)+2\rho(x, z)\rho(z, y)+\rho^2(z, y)=[\rho(x, z)+\rho(z, y)]^2.$$

Таким образом, $\rho(x, y)\leqslant\rho(x, z)+\rho(z, y)$, и неравенство треугольника в этом случае установлено. Заметим, что выше мы применили неравенство Коши — Буняковского к сумме $\sum_{i=1}^{n}|x_i-z_i||z_i-y_i|=\sum_{i=1}^{n}a_ib_i$, где положено

$$a_i=|x_i-z_i|, \quad b_i=|z_i-y_i|, \quad i=1, \dots, n.$$

На множестве X, элементами которого являются упорядоченные наборы n чисел $x=(x_1, \dots, x_n)$, можно вводить и другие функции расстояния, например положить, что:

$$\rho_0(x, y)=\frac{\rho(x, y)}{1+\rho(x, y)}, \quad x, y\in X, \qquad \text{а)}$$

где функция $\rho(x, y)$ введена выше, в примере 2); или положить

$$\rho_1(x, y)=\max_{1\leqslant i\leqslant n}|x_i-y_i|; \qquad \text{б)}$$

---

* Этим же символом E^n мы обозначим и евклидово пространство, состоящее из элементов произвольной природы и имеющее размерность n.

$$\rho_2(x, y) = \sum_{i=1}^{n} |x_i - y_i|; \qquad \text{в)}$$

$$\rho_3(x, y) = \begin{cases} 1, & \text{если } x \neq y, \\ 0, & \text{если } x = y; \end{cases} \qquad \text{г)}$$

$$\rho_4(x, y) = \begin{cases} \rho(x, y), & \text{если } \rho(x, y) < 1, \\ 1, & \text{если } \rho(x, y) \geqslant 1, \end{cases} \qquad \text{д)}$$

где функция $\rho(x, y)$ определяется, как указано выше.

Естественно, что при этом одно и то же множество X превращается в разные метрические пространства (X, ρ_i), где $i=0, 1, 2, 3, 4$.

3) Пусть Y — множество непрерывных функций, заданных на сегменте $[a, b]$. Введем метрику, полагая что $\rho(x, y) = \max\limits_{a \leqslant t \leqslant b} |x(t) - y(t)|$. Получившееся пространство представляет собой метрическое пространство. Оно обозначается через

$$C[a, b] = (Y, \rho).$$

Точно так же множество Z n раз непрерывно дифференцируемых функций на сегменте $[a, b]$, $n \geqslant 1$, становится метрическим пространством, если ввести метрику по правилу:

$$\rho(x, y) = \max_{0 \leqslant i \leqslant n} \max_{a \leqslant t < b} |x^{(i)}(t) - y^{(i)}(t)|,$$

где $x^{(0)}(t) \equiv x(t)$, $y^{(0)}(t) \equiv y(t)$. Это пространство обозначается обычно так:

$$C_n[a, b] = (Z, \rho), \qquad n \geqslant 1.$$

Пространство $C[a, b]$ иногда обозначается символом $C_0[a, b]$.

4) Пусть V — множество всех ограниченных последовательностей $x = (x_1, \ldots, x_n, \ldots)$ действительных чисел. Положим $\rho(x, y) = \sup\limits_i |x_i - y_i|$. Справедливость аксиом 1)—3) метрического пространства очевидна. Это пространство обозначается через $m = (V, \rho)$.

Заметим, что каждое подмножество X_0 метрического пространства (X, ρ), в свою очередь, является метрическим пространством с той же самой функцией расстояния ρ. Действительно, если аксиомы, определяющие метрику ρ, выполнены для любых $x, y, z \in X$, то они выполнены и для x, y, z, принадлежащих X_0. Таким образом, каждое подмножество E^n — метрическое пространство с функцией $\rho(x, y) = \left(\sum\limits_{i=1}^{n} |x_i - y_i|^2\right)^{1/2}$. Точно так же любое подмножество $C_n[a, b]$ — метрическое пространство, $n = 0, 1, 2, \ldots$.

Пара (X_0, ρ) *называется* ***подпространством*** *пространства* (X, ρ).

Заметим также, что если $\rho(x, y)$ — функция расстояния в некотором метрическом пространстве, то по формулам а) или д) примера 2) мы можем получить новую функцию расстояния.

2. Открытые и замкнутые множества. *Шаром* $O(a, r)$ *в метрическом пространстве* X (*замкнутым шаром* $K(a, r)$) *с центром в точке* a *и радиусом* r *называется совокупность всех точек* $x \in X$ *таких, что* $\rho(x, a) < r$ $(\rho(x, a) \leqslant r)$.

Определение 2. *Множество* $\Sigma \subset X$ *называется* ***открытым*** *в пространстве* $X = (X, \rho)$, *если вместе с каждой своей точкой* x *оно содержит и некоторый шар* $O(x, r)$. $\varnothing$ *также назовем открытым.*

Определение 3. ***Окрестностью точки*** $x \in X$ *называется любое открытое множество, содержащее* x. ***Окрестностью некоторого подмножества*** X *(быть может, самого* X*) называется любое открытое множество, содержащее данное подмножество*. Окрестность точки* x *будем обозначать через* Σ_x.

Определение 4. *Пусть множество* $Y \subset X$, *тогда точка* $x \in X$ *называется* ***предельной точкой*** *множества* Y, *если каждая окрестность точки* x *содержит по крайней мере одну точку множества* Y, *отличную от* x.

Точка $y \in Y$ *называется* ***изолированной точкой*** *множества* Y, *если существует окрестность точки* y, *в которой нет точек* Y, *отличных от* y.

Определение 5. *Точка* y, *принадлежащая множеству* Y — *подмножеству* X, — *называется* ***внутренней***, *если она содержится в* Y *вместе с некоторой своей окрестностью. Точки, внутренние для дополнения* Y *в* X, *называются* ***внешними*** *по отношению к* Y. *Если точка не является ни внутренней, ни внешней по отношению к* Y, *то она называется* ***граничной*** *для* Y. *Множество граничных точек для* Y *обозначается через* ∂Y.

Определение 6. *Множество в метрическом пространстве называется* ***замкнутым***, *если его дополнение открыто.*

Справедливо следующее утверждение.

Лемма 1. *Сумма любого числа открытых множеств, пересечение любого конечного числа открытых множеств есть мно-*

---

* Шар $O(a, r)$ в метрическом пространстве, очевидно, является открытым множеством, а поэтому его также можно считать окрестностью точки a. То, что шар $O(a, r)$ — открытое множество, следует из того, что этот шар содержит любую свою точку x вместе с некоторым шаром $O(x, \varepsilon)$, где $0 < \varepsilon < r - \rho(a, x)$. Действительно, если $y \in O(x, \varepsilon)$, то $\rho(a, y) \leqslant \rho(a, x) + \rho(x, y) < \rho(a, x) + r - \rho(a, x) = r$, т. е. $y \in O(a, r)$, поэтому любая точка $y \in O(x, \varepsilon)$ принадлежит также шару $O(a, r)$. Следовательно, шар $O(x, \varepsilon) \subset O(a, r)$, и поэтому исходный шар $O(a, r)$ — открытое множество (см. также пример 3) в конце настоящего раздела).

жество открытое, $\varnothing$ и все метрическое пространство X открыты.

Пересечение любого числа замкнутых множеств замкнуто, сумма любого конечного числа замкнутых множеств замкнута, $\varnothing$ и X замкнуты.

Доказательство. Пусть $\{\Sigma_\alpha\}$ — семейство открытых в X множеств. Если $x \in \bigcup_\alpha \Sigma_\alpha$, то существует α_0 такое, что $x \in \Sigma_{\alpha_0}$ следовательно, существует число $r>0$ такое, что $O(x,r) \subset \Sigma_{\alpha_0}$, т. е. $O(x,r) \subset \bigcup_\alpha \Sigma_\alpha$. Следовательно, $\bigcup_\alpha \Sigma_\alpha$ — открытое множество.

Далее, если $\Sigma_1, \ldots, \Sigma_n$ открыты в X, то из того, что $x \in \bigcap_{i=1}^n \Sigma_i$, следует, что для любого $i=1, \ldots, n$ $x \in \Sigma_i$, т. е. для любого $i=1, \ldots, n$ существует такое $r_i>0$, что $O(x, r_i) \subset \Sigma_i$. Взяв $r = \min_{1 \leqslant i \leqslant n} r_i$, получаем, что для любого $i=1, \ldots, n$

$$O(x,r) \subset O(x,r_i), \quad O(x,r) \subset \bigcap_{i=1}^n \Sigma_i,$$

т. е. пересечение множеств Σ_i, $i=1, \ldots, n$ — открытое множество.

Второе утверждение непосредственно следует из первого, если воспользоваться принципом двойственности для множеств.

Например, пусть $\{F_\alpha\}$ — семейство замкнутых в X множеств. Для каждого α введем в рассмотрение открытое множество $\Sigma_\alpha = F_\alpha'$ *. Тогда $(\bigcap_\alpha F_\alpha)' = \bigcup_\alpha F_\alpha' = \bigcup_\alpha \Sigma_\alpha$, т. е. $(\bigcap_\alpha F_\alpha)'$ — открытое множество, следовательно, $\bigcap_\alpha F_\alpha$ замкнуто. То, что $\varnothing$ и X одновременно открыты и замкнуты, очевидно **.

Определение 7. *Замыканием $\overline{Y}$ множества Y называется пересечение всех замкнутых множеств, содержащих Y.*

Очевидно, что $\overline{Y}$ содержится в каждом замкнутом множестве, содержащем Y. Следовательно, замыкание множества Y есть наименьшее из всех замкнутых множеств, содержащих Y.

Справедлива следующая лемма.

Лемма 2. *Операция замыкания в метрическом пространстве удовлетворяет следующим свойствам:*

$$1)\ \overline{Y} \supset Y, \quad 2)\ \overline{\overline{Y}} = \overline{Y}, \quad 3)\ \overline{Y \cup Z} = \overline{Y} \cup \overline{Z}, \quad 4)\ \overline{\varnothing} = \varnothing, \quad \overline{X} = X.$$

Доказательство. Свойство 1) очевидно: если $x \in Y$, то x принадлежит любому замкнутому множеству, содержащему Y, следовательно, x принадлежит и пересечению этих замкнутых множеств, т. е. $x \in \overline{Y}$.

* Напомним, что A' обозначает дополнение множества A.

** Мы *по определению* полагаем $\varnothing$ открытым. Так как X, очевидно, открыто, то $\varnothing = X'$ будет являться по определению 6 и замкнутым множеством. Аналогично $X = \varnothing'$ замкнуто.

Свойство 2) вытекает из того, что $\overline{Y}$, будучи пересечением замкнутых множеств, является в силу леммы 1 замкнутым множеством.

Докажем свойство 3), так как $Y \subset Y \cup Z$, то каждое замкнутое множество, содержащее $Y \cup Z$, содержит и Y, следовательно, пересечение этих замкнутых множеств также содержит Y, т. е. $Y \subset \overline{Y \cup Z}$, но $\overline{Y}$ принадлежит любому замкнутому множеству, содержащему Y, следовательно, $\overline{Y} \subset \overline{Y \cup Z}$. Аналогично $\overline{Z} \subset \overline{Y \cup Z}$. Таким образом, $\overline{Y} \cup \overline{Z} \subset \overline{Y \cup Z}$; обратно, $\overline{Y} \cup \overline{Z}$ в силу леммы 1 замкнуто, следовательно, $\overline{Y} \cup \overline{Z} \supset \overline{Y \cup Z}$.

Утверждение 4) означает, что $\varnothing$ и X — замкнутые множества.

Лемма полностью доказана.

Попутно мы показали, что если множество $A \subset B$, то $\overline{A} \subset \overline{B}$.

Дадим следующее определение.

Определение 8. *Пространство (X, ρ) называется связным, если его нельзя представить в виде суммы двух непустых открытых непересекающихся подмножеств.*

Очевидно, что пространство связно тогда и только тогда, когда его нельзя представить в виде суммы двух непустых замкнутых непересекающихся множеств.

Множество Y, принадлежащее метрическому пространству X, связно, если Y связно как подпространство в X.

3. Прямое произведение метрических пространств. Если (X_1, ρ_1), (X_2, ρ_2) — два метрических пространства, то можно определить прямое произведение этих метрических пространств. Пусть $X = X_1 \times X_2$ — прямое произведение множеств X_1 и X_2, т. е. множество всевозможных пар (x_1, x_2), где $x_1 \in X_1$, $x_2 \in X_2$. Метрику на множестве X можно ввести, например, по следующему правилу:

$$\rho(x, y) = [\rho_1^2(x_1, y_1) + \rho_2^2(x_2, y_2)]^{1/2},$$

где $x = (x_1, x_2)$, $y = (y_1, y_2)$, $x_i \in X_i$, $y_i \in Y_i$, $i = 1, 2$. Пара (X, ρ), $X = X_1 \times X_2$, ρ — функция расстояния, введенная выше, является метрическим пространством и, по определению, называется прямым произведением метрических пространств (X_1, ρ_1) и (X_2, ρ_2).

Если задана счетная последовательность метрических пространств (X_1, ρ_1), (X_2, ρ_2), ... и через X обозначено прямое произведение множеств X_i, $i = 1, 2, \ldots$, $X = \prod_{i=1}^{\infty} X_i$, т. е. X состоит из элементов $x = (x_1, x_2, \ldots)$, $x_i \in X_i$, то можно дать, например, следующее определение.

Определение 9. *Прямым произведением метрических пространств (X_1, ρ_1), (X_2, ρ_2), ... называется пара (X, ρ), где*

$$X=\prod_{i=1}^{\infty} X_i, \quad \rho(x, y)=\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{2^k}\frac{\rho_k(x_k, y_k)}{1+\rho_k(x_k, y_k)},$$

$$x, y \in X, \; x_k, y_k \in X_k, \; x=(x_1, x_2, \ldots), \; y=y_1, y_2, \ldots),$$

$$k=1, 2, \ldots .$$

Обозначается оно так: $(X, \rho)=\prod\limits_{i=1}^{\infty}(X_i, \rho_i)$.

4. Всюду плотные и совершенные множества.

Определение 10. *Пусть A и B — два множества в метрическом пространстве (X, ρ). Множество A называется плотным в B, если $\bar{A}\supset B$. Множество A называется всюду плотным в пространстве X, если $\bar{A}=X$.*

Пространства, в которых имеются счетные всюду плотные множества, называются сепарабельными.

Рассмотренные выше примеры метрических пространств E^1, E^n, $C_n[a, b]$, $n\geqslant 0$, являются примерами *сепарабельных* метрических пространств. Так, в E^n счетным всюду плотным множеством является множество точек, у которых все координаты — рациональные числа. В пространствах $C_n[a, b]$, $n\geqslant 0$, такими множествами являются множества многочленов с рациональными коэффициентами.

Однако *пространство m есть пример несепарабельного пространства*. Если рассмотреть множество E_0 последовательностей, состоящих только из нулей и единиц, то мощность такого множества есть континуум.*

Убедимся в этом. Рассмотрим какое-либо число $x\in$ $\in[0, 1]$. Разобьем сегмент $[0, 1]$ пополам. Возьмем в качестве первого элемента последовательности число 0, если x принадлежит левой половине, т. е. сегменту $[0, 1/2]$, и число 1 — в противном случае. Тот из двух рассматриваемых сегментов, которому принадлежит x, обозначим $[a_1, b_1]$. В качестве второго элемента последовательности возьмем 0, если x принадлежит левой половине сегмента $[a_1, b_1]$, и 1 — в противном случае. Продолжая эти рассуждения далее, мы поставим в соответствие рассматриваемому числу x вполне определенную последовательность из нулей и единиц. Если $x\neq y$, то в результате описанного процесса эти точки на некотором этапе станут принадлежать разным отрезкам, а поэтому и последовательности, им отвечающие, будут разные. Следовательно, множество всевозможных последовательностей из нулей и единиц есть множе-

* См. пример 4 из п. 1.

ство мощности, не меньшей, чем континуум. Для наших целей этого и достаточно *.

Взаимные расстояния в пространстве m между двумя различными элементами p и q множества E_0 есть величина, равная единице. Следовательно, приблизить сколь угодно близко каждый из этих элементов элементами счетного множества нельзя, так как множество шаров радиуса 1/3 с центрами в точках множества E_0 имеет мощность не менее континуума и эти шары не пересекаются.

Поскольку $E_0 \subset m$, где $m=(V, \rho)$, то и все пространство m несепарабельно.

Множество A называется ***нигде не плотным*** *в метрическом пространстве X, если любое открытое множество этого пространства содержит другое открытое множество, целиком свободное от точек множества A.*

Например, в пространстве $C[0, 1]$ множество A функций вида $y=nx^2$ (n — целые числа) нигде не плотно. Другой пример нигде не плотного множества на сегменте $[0, 1]$ (рассматриваемом как метрическое пространство) дает так называемое канторово совершенное множество.

Множество A точек метрического пространства называется ***совершенным****, если оно замкнуто и если каждая точка множества A является его предельной точкой.*

Канторово совершенное множество на сегменте $I=[0, 1]$ строится следующим образом. Из сегмента $[0, 1]$ удаляется интервал $(1/3, 2/3)$ и оставшееся множество — объединение двух сегментов $[0, 1/3]$, $[2/3, 1]$ — обозначается через I_1. Из этих двух сегментов, в свою очередь, удаляются их трети — интервалы $(1/9, 2/9)$, $(7/9, 8/9)$. Объединение оставшихся сегментов обозначим через I_2. Продолжим этот процесс неограниченно. Очевидно, $I \supset I_1 \supset \ldots \supset I_n \supset \ldots$ и I_n есть объединение 2^n сегментов, длина каждого из которых 3^{-n}.

Множество $K=\bigcap\limits_{n=1}^{\infty} I_n$ и называется множеством Кантора.

Покажем, что K совершенно. То, что оно замкнуто, следует из построения и леммы 1; остается показать, что K не содержит изолированных точек. Пусть $x \in K$, и пусть Σ_x — произвольная окрестность точки x, т. е. открытое множество. Тогда по определению открытого множества найдется интервал σ_x (шар с центром в точке x), содержащий точку x и принадлежащий Σ_x. Пусть A_n — тот сегмент множества I_n, который содержит точку x. Если n достаточно большое, то $A_n \subset \sigma_x$. Обозначим через a_n тот конец сегмента A_n, который не совпадает

* На самом деле множество всевозможных последовательностей из нулей и единиц есть множество мощности континуум.

с x. Очевидно, это следует из построения множества K, $a_n \in K$. Следовательно, произвольная окрестность точки x — множество Σ_x — содержит точку $a_n \neq x : a_n \in A_n \subset \sigma_x \subset \Sigma_x$, т. е. точка x — предельная для множества K, и K совершенно.

Докажем теперь, что множество K нигде не плотно на сегменте $[0, 1]$, рассматриваемом как метрическое пространство с обычным евклидовым расстоянием. Поскольку любое открытое множество на сегменте содержит внутри себя интервал, то достаточно показать, что любой интервал (шар) содержит внутри себя другой интервал, свободный от точек, принадлежащих K. Пусть σ — произвольный интервал сегмента $[0, 1]$. Если он не содержит точек K, то построение в этом случае закончено. Если же имеется точка $x \in K$ и $x \in \sigma$, то мы можем выбрать столь большое m, что $x \in A_m \subset I_m$ и $A_m \subset \sigma$, m — натуральное. Найдем интервал длины $3^{-(m+1)}$ с центром в середине A_m. Этот интервал не содержит точек K и содержится в σ.

Таким образом, нигде не плотность множества K на сегменте $[0, 1]$ доказана.

5. Сходимость. Непрерывные отображения.

Определение 11. *Последовательность $\{a_n\}$ точек метрического пространства называется сходящейся к точке a этого пространства, если любая окрестность точки a содержит все элементы этой последовательности, за исключением конечного их числа. Если последовательность $\{a_n\}$ сходится к a, то пишут $a_n \to a$ при $n \to \infty$ или $\lim\limits_{n\to\infty} a_n = a$.*

Непосредственно из данного определения следует, что $a_n \to a$, если $\rho(a_n, a) \to 0$, $n \to \infty$.

Лемма 3. *Точка a метрического пространства X принадлежит замыканию $\bar{A}$ некоторого множества A тогда и только тогда, когда существует последовательность $\{a_n\}$ точек множества A, сходящаяся к a.*

Доказательство. Пусть $a \in \bar{A}$, тогда a принадлежит каждому замкнутому множеству, содержащему A. Возьмем в качестве окрестности точки a шар $O(a, 1/n)$, n — натуральное. В этом шаре имеется по крайней мере одна точка $a_n \in A$. Если бы это было не так, то $a \in A$ и существовала бы окрестность точки a, свободная от точек множества A. Дополнение до этой окрестности было бы замкнутым множеством, содержащим множество A, и точка a не принадлежала бы этому замкнутому множеству. Таким образом, получилось противоречие условию $a \in \bar{A}$, следовательно, мы построили последовательность точек $a_n \to a$ при $n \to \infty$, $a_n \in A$. Заметим, что данная последовательность $\{a_n\}$ может оказаться стационарной, т. е. такой, что все $a_n \equiv a$.

Верно и обратное: если $a_n \to a$ и $a_n \in A$, то $a \in \bar{A}$.

Для этого заметим, что если любая окрестность точки a пересекается с A, то $a \in \bar{A}$. Действительно, если $a \in A$, то $a \in \bar{A}$. Если $a \in A'$, то, допустив, что $a \in \bar{A}$, и взяв дополнение множества $\bar{A}$, получим окрестность точки a — множество $(\bar{A})'$, не пересекающееся с A, т. е. получим противоречие.

Доказательство леммы теперь завершается так. Нам дано, что $a_n \to a$ и $a_n \in A$. Следовательно, любая окрестность точки a содержит точки a_n множества A. По сказанному $a \in \bar{A}$, что и требовалось.

Одновременно нами доказано следующее утверждение.

Утверждение. *Точка a принадлежит замыканию множества A в том и только том случае, если каждая окрестность Σ_a точки a пересекается с A.*

В гл. 4 было подробно изучено понятие непрерывности функции числового аргумента. Оказывается, что это понятие допускает естественное обобщение на случай, когда задана уже не обычная функция числового аргумента, а отображение одного метрического пространства в другое. Введем понятие непрерывного отображения.

Определение 12. *Отображение g одного метрического пространства (X, ρ) в другое (X_0, ρ_0) называется **непрерывным в точке** x, если для каждой окрестности $\Sigma_{g(x)}$ точки $g(x)$ найдется такая окрестность Σ_x точки x, что $g(\Sigma_x) \subset \Sigma_{g(x)}$. Если g непрерывно в каждой точке пространства X, то такое отображение называется **непрерывным на** X.*

Справедлива следующая лемма.

Лемма 4. *Отображение $g: X \to X_0$ одного метрического пространства X в другое X_0 непрерывно на X тогда и только тогда, когда прообраз любого открытого множества открыт.*

Доказательство. *Необходимость.* Пусть $g: X \to X_0$ — непрерывное отображение X в X_0 и Σ_0 — открытое множество в X_0. Если $g^{-1}(\Sigma_0) = \{x \in X : g(x) \in \Sigma_0\} = \varnothing$, то открытость $g^{-1}(\Sigma_0)$ очевидна, так как $\varnothing$ открыто.

Пусть $g^{-1}(\Sigma_0) \neq \varnothing$ и $x \in \Sigma = g^{-1}(\Sigma_0)$. Так как $x \in g^{-1}(\Sigma_0)$, то $g(x) \in \Sigma_0$, следовательно, Σ_0 можно рассматривать как окрестность точки $g(x)$. Ввиду того, что g — непрерывное отображение на X (а значит, и в точке x), найдется окрестность Σ_x точки x такая, что $g(\Sigma_x) \subset \Sigma_0$, т. е. $\Sigma_x \subset g^{-1}(\Sigma_0)$. Итак, для любой точки $x \in g^{-1}(\Sigma_0)$ найдется окрестность Σ_x такая, что $\Sigma_x \subset g^{-1}(\Sigma_0)$.

Поскольку $\Sigma = \bigcup_{x \in \Sigma} \Sigma_x$, то множество Σ открыто как объединение открытых множеств Σ_x. Таким образом, прообраз любого открытого множества открыт.

Достаточность. Пусть дано, что при отображении $g: X \to X_0$ прообраз любого открытого множества открыт. Возьмем любую точку $x \in X$ и произвольную окрестность $\Sigma_{g(x)}$ ее образа в X_0. Тогда $\Sigma_x = g^{-1}(\Sigma_{g(x)})$ по условию — открытое множество в X, т. е. является окрестностью точки x, причем образ Σ_x при отображении g содержится в $\Sigma_{g(x)}$. Следовательно, отображение g по определению непрерывно в точке x. Поскольку эти рассуждения применимы для любой точки $x \in X$, то отображение g непрерывно на X и лемма доказана.

Учитывая, что шар является открытым множеством и всякое открытое множество содержит любую свою точку вместе с некоторым шаром, можно определение 12 непрерывности отображения $g: X \to X_0$ в точке $x \in X$ переформулировать следующим образом.

Определение 13. *Отображение g одного метрического пространства (X, ρ) в другое (X_0, ρ_0) называется **непрерывным в точке** x, если для любого числа $\varepsilon > 0$ существует такое число $\delta > 0$, что если точка y принадлежит открытому шару $O(x, \delta)$ с центром в точке x радиуса δ, то точка $g(y)$ принадлежит открытому шару $O(g(x), \varepsilon)$ с центром в точке $g(x)$ радиуса ε, лежащему в пространстве (X_0, ρ_0).*

Последнее определение можно переформулировать также следующим образом: *отображение $g: X \to X_0$ непрерывно в точке x, если* $\lim\limits_{y \to x} g(y) = g(x)$.

Из неравенства треугольника легко заключаем, что функция расстояния $\rho(x, y)$ непрерывна в точке x при фиксированном y. На самом деле она является непрерывной функцией и по двум переменным *.

В случае, если метрическое пространство X есть числовая ось с обычным расстоянием между числами, т. е. пространство E^1, а отображение g — обычная скалярная функция на E^1, данное определение 13 иепрерывности, очевидно, совпадает с определениями гл. 4.

Введем понятие гомеоморфизма.

Определение 14. *Отображение $g: X \to X_0$ метрического пространства X в метрическое пространство X_0 называется **гомеоморфизмом**, если g отображает X на X_0 взаимно однозначно и g непрерывно вместе с обратным отображением g^{-1}.*

6. Компактность. Покрытием множества A в метрическом пространстве называется любое семейство открытых множеств, объединение которых содержит A.

*Определение 15. Метрическое пространство X называется **компактным** или **компактом**, если любое его по-*

---

* Функция $\rho(x, y)$ определена на произведении $X \times X$ и отображает $X \times X$ в E^1. Функция $\rho(x, y)$ непрерывна на произведении метрических пространств $X \times X$. Ее непрерывность следует из неравенства четырехугольника: $|\rho(x, z) - \rho(y, u)| \leqslant |\rho(x, y) + \rho(z, u)|$, которое получается из двух неравенств:

$$\rho(x, z) \leqslant \rho(x, y) + \rho(y, z) \leqslant \rho(x, y) + \rho(y, u) + \rho(u, z),$$

$$\rho(y, u) \leqslant \rho(y, x) + \rho(u, x) \leqslant \rho(y, x) + \rho(x, z) + \rho(z, u).$$

крытие содержит конечное подпокрытие, т. е. из всякого покрытия пространства X открытыми множествами можно выделить конечную систему, содержащую все пространство (подпокрытие).

Сегмент $[a, b]$ числовой оси является компактом, что вытекает из того факта, что из всякого покрытия этого сегмента открытыми множествами можно выделить конечное подпокрытие.

Можно показать, что *метрическое пространство компактно тогда и только тогда, когда из всякой последовательности его точек можно выделить подпоследовательность, сходящуюся к некоторой точке из этого пространства.*

Определение 16. *Система подмножеств $\{A_\alpha\}$ множества называется* центрированной, *если любое конечное подсемейство этой системы имеет непустое пересечение.*

Справедливы следующие леммы.

Лемма 5. *Для того чтобы метрическое пространство X было компактным, необходимо и достаточно, чтобы каждая центрированная система замкнутых его подмножеств имела непустое пересечение.*

Доказательство. Пусть X компактно, и пусть $\{F_\alpha\}$ — центрированная система замкнутых подмножеств. Множества $G_n = X \setminus F_\alpha$ открыты, и никакая конечная система из этих множеств G_n, $1 \leqslant n \leqslant N < \infty$ не покрывает X. Значит, поскольку X компактно, $\{G_\alpha\}$ не могут служить покрытием компактного пространства X. В противном случае мы могли бы выбрать конечное подпокрытие X из системы $\{G_\alpha\}$. Но если $\{G_\alpha\}$ не покрывает X, то $\bigcap\limits_\alpha F_\alpha$ не пусто, так как $\bigcap\limits_\alpha F_\alpha = \bigcap\limits_\alpha (X \setminus G_\alpha) = X \setminus \bigcup\limits_\alpha G_\alpha$.

Обратно, пусть любая центрированная система замкнутых подмножеств из X имеет непустое пересечение. Пусть $\{G_\alpha\}$ — открытое покрытие X. Выберем из этого покрытия конечное подпокрытие. Положим $F_\alpha = X \setminus G_\alpha$ и заметим, что так как $\{G_\alpha\}$ покрывает все X, то $\bigcap\limits_\alpha F_\alpha = \varnothing$. Следовательно, $\{F_\alpha\}$ не является центрированной, т. е. существуют такие $F_1, F_2, \ldots, F_M$, $M < \infty$, что $\bigcap\limits_{i=1}^{M} F_i = \varnothing$, но тогда $\{G_i\}^M{}_{i=1} = \{X \setminus F_i\}^M{}_{i=1}$ — конечное подпокрытие покрытия $\{G_\alpha\}$, что и требовалось доказать.

Лемма 6. *Замкнутое подмножество компактного метрического пространства компактно*.*

Доказательство. Пусть F — замкнутое подмножество компактного метрического пространства X и $\{\Sigma_\alpha\}$ — некоторая система открытых множеств — покрытие F. К системе $\{\Sigma_\alpha\}$ при-

* Т. е. из всякого покрытия этого подмножества открытыми в данном пространстве множествами можно выделить конечное подпокрытие.

соединим открытое множество $G=X\setminus F$, и полученное покрытие всего пространства обозначим через $\{\Sigma_\alpha^1\}=\{\Sigma_\alpha\}\cup G$. Выберем в силу компактности X из системы $\{\Sigma_\alpha^1\}$ конечное покрытие всего пространства — систему $\{\Sigma_i^1\}^N_{i=1}$. Выбрасывая, если это необходимо, из системы $\{\Sigma_i^1\}^N_{i=1}$ множество G, мы получим конечное покрытие множества F, выбранное из системы $\{\Sigma_\alpha\}$. Лемма доказана.

Лемма 7. *Образ компактного пространства X при непрерывном отображении — компактное множество.*

Доказательство. Пусть g — непрерывное отображение X в X_0. Пусть $\{\Sigma_\alpha\}$ — покрытие $g(X)\subset X_0$ открытыми множествами, а $\Psi_\alpha=g^{-1}\{\Sigma_\alpha\}$. Множества Ψ_α открыты (см. лемму 4), и $\{\Psi_\alpha\}$ — покрытие $g(X)$. Выберем из этого покрытия в силу компактности X конечное подпокрытие: $\{\Psi_i\}^M_{i=1}$; тогда $\{\Sigma_i\}^M_{i=1}$, $M<\infty$, — покрытие X_0, $\Sigma_i=g(\Psi_i)$, $i=1, \ldots, M$, что и требовалось доказать.

Лемма 8. *Компактное подмножество метрического пространства X замкнуто*.*

Доказательство. Пусть F — компактное подмножество, и пусть $a\in X\setminus F$; для любого $x\in F$ существуют окрестности Σ_a и Σ_x точек a и x соответственно такие, что $\Sigma_a\cap\Sigma_x=\varnothing$. В качестве таких окрестностей можно, например, взять шары $O(a, r)$ и $O(x, r)$, $r=\frac{1}{3}\rho(a,x)$. Множество $G=\bigcup\limits_{x\in F}\Sigma_x$ — покрытие множества F. В силу компактности F выберем из этого покрытия конечное подпокрытие: $\{\Sigma_{x_i}\}_{i=1}^N$. Рассмотрим соответствующие Σ_{x_i} окрестности $\Sigma_a{}^i$, которые по построению не пересекаются с Σ_{x_i}, и $\bigcap\limits_{i=1}^{N}\Sigma_a^i=\Sigma$ является окрестностью точки a. Очевидно, что $\Sigma\cap\Sigma_{x_i}=\varnothing$, $i=1,\ldots,N$, и поэтому $\Sigma\cap F=\varnothing$. Следовательно, $\Sigma\subset X\setminus F$, т. е. множество $X\setminus F$ открыто, а F замкнуто.

Лемма 9. *Пусть $g: X\to E^1$, где E — действительная числовая ось. Если g — непрерывное отображение, а X — компакт, то g ограничено и достигает своих точных верхней и нижней граней.*

Доказательство. Пусть $g(X)$ — непрерывный образ компакта (компактного пространства). По лемме 7 подмножество $g(X)$ метрического пространства E^1 компактно, следовательно, ограниченно и замкнуто (см. гл. 4, § 7, п. 3). Осталось заметить, что непустое ограниченное замкнутое множество чис-

* Заметим, что в отличие от множеств числовой оси (или E^n) в случае общего метрического пространства ограниченность и замкнутость множества еще недостаточна для его компактности (ср. с гл. 4, § 7, п. 3).

ловой оси содержит точные верхнюю и нижнюю грани *. Лемма доказана.

7. Базис пространства. Введем понятие базиса.

Определение 17. *Система открытых множеств $\{\Sigma_\alpha\}$ метрического пространства X называется базисом этого пространства, если всякое непустое открытое множество пространства X может быть получено как объединение некоторых Σ_α из множеств системы.*

Справедлива следующая лемма.

Лемма 10. *Для того чтобы система $\{\Sigma_\alpha\}$ открытых множеств была базисом пространства X, необходимо и достаточно, чтобы для всякого открытого множества G и всякой точки $a\in$ $\in G$ нашлось бы такое множество Σ_{α_0} из данной системы, что $a\in\Sigma_{\alpha_0}\subset G$.*

Доказательство. Допустим, что система $\{\Sigma_\alpha\}$ — базис пространства. Тогда любое открытое множество G есть сумма некоторых из множеств $\{\Sigma_\alpha\}$. Поэтому для всякой точки $a\in G$ существует такое множество Σ_{α_0}, что $a\in\Sigma_{\alpha_0}\subset G$ и Σ_{α_0} принадлежит данной системе $\{\Sigma_\alpha\}$.

Обратно, пусть выполнены условия, сформулированные в лемме. Тогда всякое открытое множество G представимо в виде

$$G=\bigcup_{a\in G}\Sigma_{\alpha_0},$$

т. е. система $\{\Sigma_{\alpha_0}\}$ — базис пространства. Лемма доказана.

Определение 18. *Метрическое пространство X называется пространством со счетным базисом, если в нем существует хотя бы один базис, состоящий не более чем из счетного числа множеств. Пространства со счетным базисом называются также пространствами со второй аксиомой счетности.*

Лемма 11. *Метрическое пространство X является пространством со счетным базисом тогда, когда в нем имеется счетное, всюду плотное множество. Обратно, если в пространстве имеется счетный базис, то в нем есть и счетное всюду плотное множество.*

Доказательство. Пусть $A=\{a_n\}^\infty{}_{n=1}$ — счетное, всюду плотное множество в X. Всевозможные шары $O(a_n,\ 1/m)$, где n, m — всевозможные натуральные числа, образуют базис пространства, причем счетный.

Обратно, если в X имеется счетный базис $\{\Sigma_n\}^\infty{}_{n=1}$, то, выбрав по точке $a_n\in\Sigma_n$, мы получаем множество $A=\{a_n\}^\infty{}_{n=1}$, которое всюду плотно. Действительно, если бы $\bar A\neq X$, то открытое множество $G=X\setminus\bar A$ было бы не пустым и не содержало бы ни одной точки из $A=\{a_n\}^\infty{}_{n=1}$, что невозможно, так как G —

* См. также гл. 4, § 7, п. 3.

открытое множество и оно есть объединение некоторых из множеств системы $\{\Sigma_n\}$, а $a_n \in \Sigma_n$.

Примеры. 1) Легко построить метрическое пространство (X, ρ) и замкнутые шары $K_1(x_1, r_1)$ и $K_2(x_2, r_2)$ такие, что $K_1 \subset K_2$, $r_1 > r_2$.

Действительно, пусть (X, ρ) — метрическое пространство, состоящее из всех точек $\{x, y\}$ замкнутого круга на плоскости $xy: X \equiv \{x, y: x^2+y^2 \leqslant 9\}$ с обычной евклидовой метрикой ρ. Шар K_2 определим так: $K_2 \equiv (X, \rho)$. Пусть шар $K_1 \equiv K_2 \cap \{x, y: (x-2)^2+y^2 \leqslant 16\}$. Тогда $K_1 \subset K_2$, $r_1=4$, $r_2=3$, $r_1 > r_2$.

2) Множество A метрического пространства замкнуто тогда и только тогда, когда оно совпадает со своим замыканием $\bar{A}$, т. е. $A=\bar{A}$.

Действительно, если $A=\bar{A}$, то, так как замыкание любого множества замкнуто (как пересечение замкнутых), A — замкнуто.

Обратно, $\bar{A}$ принадлежит любому замкнутому множеству, содержащему A; одним из таких замкнутых множеств является в данном случае само A, т. е. $\bar{A} \subset A$. С другой стороны, по определению замыкания $A \subset \bar{A}$, т. е. если A замкнуто, то $A=\bar{A}$.

3) Шар $O(x_0, r)$, как уже говорилось выше, в метрическом пространстве X — открытое множество. (Если $x \in O(x, r)$, т. е. $\rho(x, x_0) < r$, то шар $O(x, \varepsilon)$ при $0 < \varepsilon < r - \rho(x_0, x)$ будет принадлежать исходному множеству: $O(x, \varepsilon) \subset O(x_0, r)$. Действительно, если $y \in O(x, \varepsilon)$, т. е. $\rho(x, y) < \varepsilon$, то

$$\rho(x_0, y) \leqslant \rho(x_0, x) + \rho(x, y) < \rho(x_0, x) + r - \rho(x_0, x) = r.)$$

Аналогично замкнутый шар — множество $K(x, r) = \{y: \rho(x, y) \leqslant r\}$ — замкнутое множество в X. Это следует из того, что его дополнение есть множество открытое.

4) Замыкание $\bar{A}$ множества A метрического пространства состоит из всех точек, которые являются либо предельными точками множества A, либо элементами A *. Действительно, всякое множество содержится в своем замыкании: $A \subset \bar{A}$. Докажем, что замыкание $\bar{A}$ содержит и предельные точки для A. В самом деле, замыкание множества — множество замкнутое, а если множество замкнуто, то оно содержит все свои предельные точки: если точка не принадлежит замкнутому множеству, то она принадлежит его дополнению, которое и является окрестностью этой точки, свободной от точек замкнутого множества, т. е. точка не является предельной для замкнутого множества. Иными словами, $\bar{A}$ содержит все свои предельные

* Это утверждение остается справедливым, а доказательство полностью сохраняется и для топологических пространств, которые будут рассмотрены в следующем разделе. Заметим также, что предельная точка множества A может оказаться элементом множества A.

точки; так как $A \subset \bar{A}$, то $\bar{A}$ содержит и все предельные точки множества A. Таким образом, замыкание A содержит точки множества A и точки, предельные для A. Обратно, пусть $x \in \bar{A}$, покажем, что тогда либо x — предельная для A, либо $x \in A$. Действительно, если $x \in \bar{A}$ и $x \in A$, то все доказано. Если же $x \in \bar{A}$, но $x \,\overline{\in}\, A$, то x — предельная для A. В самом деле, допустим противное, тогда существует окрестность точки x — Σ_x, свободная от точек множества A. Ее дополнение — замкнутое множество, содержащее A и не содержащее x, т. е. точка $x \,\overline{\in}\, \bar{A}$, что невозможно.

5) В метрическом пространстве (X, ρ) можно построить открытый шар $O(x, r) = \{y : \rho(x, y) < r\}$ и замкнутый шар $K(x, r) = \{y : \rho(x, y) \leqslant r\}$ с общим центром и равными радиусами такие, что $\bar{O}(x, r) \neq K(x, r)$. Действительно, пусть X — множество, состоящее более чем из одной точки, и пусть

$$\rho(x, y) = \begin{cases} 1, & \text{если } x \neq y, \\ 0, & \text{если } x = y, \end{cases}$$

тогда $K(x, 1) = X$, $O(x, 1) = x$, $\bar{O}(x, 1) = x \neq K(x, 1)$.

Свойства метрических пространств

В предыдущем разделе были изложены основные свойства метрических пространств, базирующиеся на понятии открытого и замкнутого множеств. Следует подчеркнуть, что фактически все утверждения предыдущего раздела используют только свойства открытых множеств — то, что сумма любого числа и пересечение конечного числа открытых множеств есть множество открытое, все пространство и пустое множество открыты, и не используют такие понятия, как шар, расстояние. Все это наводит на мысль, что можно строить пространства, в которых открытые (замкнутые) множества определяются аксиоматически с сохранением свойств, сформулированных в лемме 1. Тогда надобность в определениях 1 и 2 отпадает, а утверждение леммы 1 станет аксиомой. Именно так мы и поступим в следующем разделе, когда будем изучать топологические пространства. Ясно, что при этом будет достигнуто полное единообразие в изучении основных свойств метрических и топологических пространств.

Такой подход к изучению свойств метрических пространств (формулировки и доказательства, использующие только свойства открытых (замкнутых) множеств) обладает, конечно, рядом преимуществ — например, допускает обобщения на классы более общих пространств. Вместе с тем, поскольку в структуру метрического пространства введена функция расстояния, эти пространства должны обладать и своими, присущими только им свойствами. Более того, чаще всего именно эти свойства и изучаются при рассмотрении метрических пространств.

В настоящем разделе мы и изложим эти фундаментальные свойства метрических пространств. Все они используют понятие полноты пространства.

Определение 1. *Последовательность* $\{x_n\}$ *элементов метрического пространства* (X, ρ) *называется* ***фундаментальной***, *если для любого числа* $\varepsilon>0$ *существует такое натуральное число* N, *что для любых* $n, m>N$ *выполняется неравенство* $\rho(x_n, x_m)<\varepsilon$.

Определение 2. *Метрическое пространство* (X, ρ) *называется* ***полным***, *если в нем всякая фундаментальная последовательность сходится к некоторому пределу, являющемуся элементом этого пространства.*

Заметим, что, как мы уже говорили, если последовательность $\{x_n\}$ сходится к элементу x, то $\rho(x, x_n)\to 0$ при $n\to\infty$ *.

Теорема (принцип вложенных шаров). *Для того чтобы метрическое пространство было полным, необходимо и достаточно, чтобы в нем всякая последовательность замкнутых вложенных друг в друга шаров, радиусы которых стремятся к нулю, имела непустое пересечение.*

Доказательство. *Необходимость.* Пусть (X, ρ) — полное метрическое пространство, а $K_1(x_1, r_1)\supset K_2(x_2, r_2)\supset\ldots$ — вложенные друг в друга замкнутые шары. Последовательность их центров фундаментальна, так как $\rho(x_n, x_m)<r_n$, $m>n$, а $r_n\to 0$, $n\to\infty$.

Поскольку пространство X полное, то существует элемент x пространства X такой, что $x=\lim\limits_{n\to\infty} x_n$. Очевидно, что $x\in\bigcap\limits_{n=1}^{\infty} K_n$. Действительно, x — предельная точка для каждого K_n (см. определение 4 п. 1) ввиду того, что при $m\geqslant n$ $x_m\in K_m\subset K_n$, а так как K_n — замкнутое множество, то $x\in K_n$ (см. пример 4 с. 549).

Достаточность. Пусть выполнены условия теоремы относительно шаров. Докажем, что пространство X полное. Мы должны показать, что если $\{x_n\}$ — фундаментальная последовательность, то она имеет предел $x\in X$. Выберем точку x_{n_1} такую, что $\rho(x_n, x_{n_1})<1/2$, $n>n_1$. Примем x_{n_1} за центр замкнутого шара радиуса 1, который обозначим через $K(x_{n_1}, 1)$. Выберем, далее, точку x_{n_2} из последовательности $\{x_n\}$, удовлетворяющую следующим условиям: $\rho(x_{n_1}, x_{n_2})<1/2^2$ для любого $n>n_2$, $n_2>n_1$. Примем точку x_{n_2} за центр шара $K(x_{n_2}, 1/2)$ радиуса 1/2. Пусть $x_{n_1}, x_{n_2}, \ldots, x_{n_k}$ $(n_1<n_2,\ldots<n_k)$ уже выбраны,

* Очевидно, что любая сходящаяся последовательность является фундаментальной.

тогда $x_{n_{k+1}}$ выберем так, чтобы выполнялись условия: $\rho(x_n, x_{n_{k+1}}) < 1/2^{k+1}$ для любого $n > n_{k+1}$, $n_{k+1} > n_k$. Как и выше, примем $x_{n_{k+1}}$ за центр замкнутого шара $K(x_{n_{k+1}}, 1/2^k)$ радиуса $1/2^k$ и т. д. Мы получили последовательность вложенных друг в друга замкнутых шаров, радиусы которых стремятся к нулю. По предположению существует точка x, общая для всех шаров. Ясно, что $\rho(x_{n_k}, x) \to 0, n_k \to \infty$, т. е. фундаментальная последовательность $\{x_n\}$ содержит последовательность $\{x_{n_k}\}$, сходящуюся к некоторой точке x пространства. Тогда и сама последовательность сходится к этому же пределу. Действительно, применяя свойство треугольника для функции расстояния (см. определение 1 п. 1), имеем

$$\rho(x_n, x) \leqslant \rho(x_{n_k}, x) + \rho(x_n, x_{n_k}) \to 0, \quad n_k, n \to \infty.$$

Таким образом, $\lim\limits_{n\to\infty} x_n = x \in X$. Теорема доказана.

Замечание. В условиях теоремы все условия являются существенными: полнота пространства, замкнутость шаров, условие того, что они вложены, стремление их радиусов к нулю. Наименее очевидным является необходимость последнего условия. Вот пример полного метрического пространства и последовательности вложенных друг в друга замкнутых шаров, имеющих пустое пересечение.

Пусть $X = (N, \rho)$, где N — множество натуральных чисел, а

$$\rho(m, n) = \begin{cases} 1 + \dfrac{1}{m+n}, & \text{если } n \neq m, \\ 0, & \text{если } n = m. \end{cases}$$

Пусть

$$K\left(n, 1 + \frac{1}{2n}\right) = \{m : \rho(m, n) \leqslant 1 + 1/2n\}, \quad n = 1, 2, \ldots .$$

Тогда очевидно, что $K\left(n, 1 + \dfrac{1}{2n}\right)$ замкнуты и вложены друг в друга, а пространство полно, поскольку каждая фундаментальная последовательность сходится в пространстве (она является, как говорят, «стационарной»). Однако пересечение указанных шаров, очевидно, пусто.

Определение 3. *Множество M, расположенное в метрическом пространстве X, называется множеством* ***первой категории***, *если его можно представить в виде объединения счетного числа нигде не плотных в X множеств.*

Все остальные множества называются множествами ***второй категории***.

Теорема (теорема о категориях). *Пусть (X, ρ) — непустое полное метрическое пространство, тогда X — множест-*

во второй категории, т. е. X нельзя представить в виде объединения счетного числа нигде не плотных множеств.

Доказательство. Предположим противное, что $X = \bigcup_{n=1}^{\infty} A_n$ и каждое из множеств A_n, $n=1, 2, \ldots$, нигде не плотно в X. Пусть K_0 — некоторый замкнутый шар радиуса единица. Поскольку множество A_1 нигде не плотно, то существует замкнутый шар K_1 радиуса меньше $1/2$ такой, что $K_1 \subset K_0$ и $K_1 \cap A_1 = \varnothing$. Поскольку множество A_2 нигде не плотно, то точно так же существует замкнутый шар K_2 радиуса меньше $1/2^2$, принадлежащий K_1, для которого $K_2 \cap A_2 = \varnothing$ и т. д. В результате мы получаем последовательность вложенных друг в друга замкнутых шаров $\{K_n\}^{\infty}_{n=1}$, радиусы которых стремятся к нулю. По предыдущей теореме существует $x \in K_n$ для любого n, $x \in X$. Так как по построению $K_n \cap A_n = \varnothing$, то $x \overline{\in} A_n$ для любого n. Следовательно, $x \overline{\in} \bigcup_{n=1}^{\infty} A_n = X$. Это противоречит предположению: $X = \bigcup_{n=1}^{\infty} A_n$. Теорема доказана.

Определение 4. *Отображение g метрического пространства X в это же пространство называется* сжимающим, *если существует такое число $0<\alpha<1$, что*

$$\rho(g(x), g(y)) \leqslant \alpha\rho(x, y).^*$$

Теорема (принцип сжимающих отображений. *Всякое сжимающее отображение полного метрического пространства (X, ρ) в это же пространство имеет и при том только одну неподвижную точку $x \in X$, т. е. такую точку $x \in X$, что $g(x) = x$.*

Доказательство. Пусть x_0 — некоторая точка из X. Определим последовательность $\{x_n\}^{\infty}_{n=1}$ по правилу: $x_1 = g(x_0)$, $x_2 = g(x_1), \ldots$.

Последовательность $\{x_n\}$ фундаментальна в X. Действительно, если $m > n$, то

$$\begin{gathered}\rho(x_n, x_m) = \rho(g(x_{n-1}), g(x_{m-1})) \leqslant \alpha\rho(x_{n-1}, x_{m-1}) \leqslant \\ \leqslant \ldots \alpha^n\rho(x_0, x_{m-n}) \leqslant \alpha^n\{\rho(x_0, x_1) + \rho(x_1, x_2) + \\ + \ldots + \rho(x_{m-n-1}, x_{m-n})\} \leqslant \alpha^n\rho(x_0, x_1)\{1 + \alpha + \ldots + \alpha^{m-n-1}\} < \\ < \alpha^n\rho(x_0, x_1)\cdot(1-\alpha)^{-1},\end{gathered}$$

где $\alpha < 1$.

* Отсюда следует, что сжимающее отображение g непрерывно, т. е. $\lim\limits_{y\to x} g(y) = g(x)$.

Таким образом, $\rho(x_n, x_m) \to 0$, $m, n \to \infty$. В силу полноты пространства X существует $\lim\limits_{n\to\infty} x_n$. Пусть $x = \lim\limits_{n\to\infty} x_n$. Тогда в силу непрерывности отображения g имеем

$$g(x) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

Итак, неподвижная точка существует. Докажем ее единственность. Если $g(x) = x$ и $g(y) = y$, то $\rho(x, y) < \alpha\rho(x, y)$, т. е. $\rho(x, y) = 0$ и $x = y$.

Замечание. Если отображение g метрического пространства X в себя обладает лишь тем свойством, что $\rho(g(x), g(y)) < \rho(x, y)$ при $x \neq y$, то неподвижной точки может и не быть. Вот соответствующий пример: пусть $X = (P, \rho)$, где $P = [1, \infty)$, а ρ — обычная евклидова метрика. Пусть $g(x) = x + \frac{1}{x}$. Тогда $\rho(g(x), g(y)) = \left|x + \frac{1}{x} - y - \frac{1}{y}\right| < |x - y|$. Однако неподвижной точки нет: $g(x) = x + \frac{1}{x} \neq x$ ни для какого $x \in X$.

Определение 5. Будем говорить, что два отображения g и g_1 метрического пространства (X, ρ) в это же пространство коммутируют, если для любого $x \in X$ справедливо равенство $g(g_1(x)) = g_1(g(x))$.

Теорема (обобщение принципа сжатых отображений). *Пусть g_1 и g_2 — отображения полного метрического пространства (X, ρ) в это же пространство. Тогда, если отображение g_1 сжимающее и отображения g и g_1 коммутируют, то уравнение $g(x) = x$ имеет решение $x \in X$.*

Доказательство. По предыдущей теореме существует и притом только одна точка $x \in X$ такая, что $g_1(x) = x$. Применим к обеим частям равенства отображение g. Воспользовавшись тем, что отображения коммутируют, получим

$$g(g_1(x)) = g(x), \quad g_1(g(x)) = g(x), \quad g_1(y) = y,$$

где $y = g(x)$. Учитывая, что отображение g_1 сжимающее и неподвижная точка у этого отображения одна, получим, что $x = y = g(x)$. Следовательно, и у отображения g существует неподвижная точка, а именно найденная выше точка x.

Замечание. Степенью n отображения g называется отображение g^n, полученное в результате n последовательных применений отображения g:

$$g^n(x) = g(g(\ldots g(x)) \ldots), \quad x \in X.$$

Из предыдущей теоремы в силу того, что отображения g и g^n коммутируют, следует, что если отображение g таково, что некоторая степень n его — сжимающее отображение, то уравнение $g(x)=x$ имеет единственное решение*.

Примеры. 1) *Нахождение корней функции.* Пусть $\varphi(t)$, $a \leqslant t \leqslant b$ — действительная функция, удовлетворяющая условию Липшица:

$$|\varphi(t_1)-\varphi(t_2)| \leqslant \theta |t_1-t_2|, \quad 0<\theta<1,$$

и отображающая сегмент $[a, b]$ в себя. Если ввести метрическое пространство (X, ρ), где $X=[a, b]$, а ρ — обычная евклидова метрика на сегменте, то отображение φ в X сжимающее, и потому числовая последовательность t_0, $t_1=\varphi(t_0)$, $t_2=\varphi(t_1), \ldots$ сходится к единственному корню уравнения $\varphi(t)=t$ для любого $t_0 \in [a, b]$. Отображение φ сжимающее, если, например,

$$|\varphi'(t)| \leqslant \theta < 1, \; t \in [a, b].$$

Пусть нам надо решить уравнение вида $F(t)=0$, причем $F(t)$ действительная, определенная на $[a, b]$, функция, $F(a)<0$, $F(b)>0$, $0<\theta_1 \leqslant F'(t) \leqslant \theta_2$ при $t \in [a, b]$. Тогда, если рассмотреть функцию $\varphi(t)=t-\lambda F(t)$, $\lambda \in E^1$ и найти корень уравнения $\varphi(t)=t$, то мы решим исходную задачу. К этому последнему уравнению можно применить предыдущие рассуждения, если, например $|\varphi'(t)| \leqslant \theta < 1$. Имеем $1-\lambda\theta_2 \leqslant \varphi'(t) \leqslant 1-\lambda\theta_1$. Нетрудно подобрать действительное число $\lambda>0$, чтобы было выполнено условие: $0 \leqslant \varphi'(t) \leqslant \theta < 1$. Тогда, как нетрудно убедиться, функция $\varphi(t)$ отображает сегмент $[a, b]$ в себя. Действительно, так как $\varphi'(t) \geqslant 0$, то $\varphi(t)$ не убывает, следовательно, $\varphi(a) \leqslant \varphi(t) \leqslant \varphi(b)$ для любого t из сегмента $[a, b]$. Но $\varphi(a)=a-\lambda F(a)>a$, $\varphi(b)=b-\lambda F(b)<b$, т. е. $a<\varphi(t)<b$ для любого $t \in [a, b]$.

2) *Нахождение решений систем уравнений вида* $y=Ax+b$.

Пусть $y_i=\sum\limits_{j=1}^{n} a_{ij}x_j+b$, $i=1, \ldots, n$ — отображение n-мерного пространства в себя: набор $(x_1, x_2, \ldots, x_n)$ переходит в набор $(y_1, y_2, \ldots, y_n) \in X$, т. е. отображение $g(x)=Ax+b$ переводит набор из n чисел в набор из n чисел. Здесь x, b, $g(x)=y$ — векторы, A — матрица, $A=\{a_{ij}\}^n_{i,j=1}$.

Если отображение $g(x)$ будет в некотором метрическом пространстве и при некоторых условиях сжимающим, то векторное уравнение $g(x)=x$ будет иметь по предыдущему одно и только одно решение. Найдем такие условия на отображении g и введем метрику на множестве X, т. е. образуем соответствующие метрические пространства:

* Единственность решения следует из того факта, что всякая точка, неподвижная относительно отображения g, будет неподвижной и относительно отображения g^n.

А) Пусть $\rho_1(x, y) = \max\limits_{1 \leqslant i \leqslant n} |x_i - y_i|$ и $x, y \in X$. Тогда

$$\rho_1(y', y'') \leqslant \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}| \rho_1(x', x''),$$

$$x', x'', y', y'' \in X, \; y' = Ax' + b, \; y'' = Ax'' + b,$$

и условие того, что отображение g сжимающее, будет выполнено, если

$$\sum_{j=1}^{n} |a_{ij}| \leqslant \alpha < 1, \quad i = 1, 2, \ldots, n.$$

Б) Если ввести метрику на X по правилу

$$\rho_2(x, y) = \sum_{i=1}^{n} |x_i - y_i|,$$

то, как нетрудно вычислить, условие того, что отображение сжимающее, будет иметь вид

$$\sum_{i=1}^{n} |a_{ij}| \leqslant \alpha < 1, \quad j = 1, 2, \ldots, n.$$

В) Наконец, если метрика задана так:

$$\rho(x, y) = \left[\sum_{i=1}^{n} |x_i - y_i|^2\right]^{1/2},$$

то отображение g сжимающее, если

$$\sum_{i,j=1}^{n} |a_{ij}|^2 \leqslant \alpha < 1.$$

Выписанные условия *достаточны*, для того чтобы уравнение $g(x) = x$ имело и притом единственное решение или, что то же самое, чтобы система $x_i = \sum\limits_{j=1}^{n} a_{ij} x_j + b_i$, $i = 1, 2, \ldots, n$, имела и притом одно решение.

3) *Существование и единственность решения обыкновенного дифференциального уравнения первого порядка.* Рассмотрим так называемую задачу Коши. Надлежит найти такую дифференцируемую функцию $y(t)$, которая удовлетворяла бы уравнению $y' = f(t, y)$ и при $t = t_0$ имела заданное значение $y(t_0) = y_0$, где y_0 — некоторое число; при этом надлежит доказать, что при определенных условиях такое решение $y(t)$ только одно.

Предположим, что функция $f(t, y)$ непрерывна на множестве: $a \leqslant t \leqslant b$, $-\infty < y < \infty$ и удовлетворяет с константой K условию Липшица по y, т. е. для всех точек $t \in [a, b]$

$$|f(t, y_1) - f(t, y_2)| \leqslant K|y_1 - y_2|,$$

и пусть t_0 — внутренняя точка сегмента $[a, b]$. Легко убедиться, что решение задачи Коши эквивалентно решению интегрального уравнения $y(t) = y_0 + \int_{t_0}^{t} f(\xi, y(\xi))\, d\xi$. Значит, задано отображение множества функций $\{y\}$ по правилу: $g(y) = y_0 + \int_{t_0}^{t} f(\xi, y(\xi))\, d\xi$. Введем пространство $C[a, b]$. Тогда отображение g определено на этом пространстве и отображает его в себя, а задача о нахождении решения интегрального уравнения сводится к нахождению неподвижной точки отображения g — нахождению функции y такой, что $g(y) = y$. Для того чтобы такая точка существовала и была одна, достаточно, чтобы отображение g было сжимающее.

Поскольку из условия Липшица следует, что

$$|f(t, y_1) - f(t, y_2)| \leqslant K|y_1 - y_2| \leqslant K\rho(y_1, y_2),$$

то

$$\rho(g(y), g(z)) \leqslant \max_{a \leqslant t \leqslant b} \int_{t_0}^{t} K\rho(y, z)\, d\xi \leqslant K(b-a)\rho(y, z).$$

Здесь ρ — метрика в $C[a, b]$. Следовательно, отображение сжимающее, если сегмент $[a, b]$ достаточно мал, так что

$$K(b-a) = \theta < 1.$$

При этих условиях мы получаем теорему существования и единственности решения задачи Коши на сегменте $[a, b]$, содержащем точку t_0.

4) *Оператор Вольтерра, свойства его n-й степени.* Покажем, что некоторая степень отображения, задаваемого интегральным оператором Вольтерра, имеющим вид

$$g(f) = \lambda \int_{a}^{t} K(t, \xi) f(\xi)\, d\xi + \varphi(t),$$

где λ — некоторое число, $K(t, \xi)$, $\varphi(t)$ — непрерывные функции своих аргументов, представляет собой сжимающее отображение в $C[a, b]$, $a \leqslant t \leqslant b$, $a \leqslant \xi \leqslant b$.

Пусть

$$M = \max_{a \leqslant t, \xi \leqslant b} |K(t, \xi)|, \quad \rho(f_1, f_2) = \max_{a \leqslant t \leqslant b} |f_1(t) - f_2(t)|.$$

Тогда

$$|g(f_1)-g(f_2)| = |\lambda| \left| \int_a^b K(t, \xi)(f_1(\xi)-f_2(\xi))\,d\xi \right| \leqslant$$

$$\leqslant |\lambda| M(b-a)\rho(f_1, f_2).$$

Отсюда

$$|g^n(f_1)-g^n(f_2)| \leqslant |\lambda|^n M^n \frac{(b-a)^n}{n!}\rho(f_1, f_2).$$

Всегда можно выбрать такое n, что $\frac{|\lambda|^n M^n (b-a)^n}{n!} < 1$, и, следовательно, при этом n отображение g^n будет сжимающим. Согласно замечанию после предыдущей теоремы интегральное уравнение вида $g(f)=f$ имеет при любом λ решение, и притом единственное.

Топологические пространства

В настоящем разделе будут рассмотрены основные свойства топологических пространств. Материал данного раздела вполне аналогичен изложенному в разделе о метрических пространствах, и поэтому мы повторим лишь основные определения, а доказательства некоторых теорем, поскольку они являются дословным повторением соответствующих доказательств в разделе о метрических пространствах, опустим. Подробнее мы остановимся лишь на специфических особенностях топологических пространств.

1. Определение топологического пространства. Хаусдорфово топологическое пространство. Примеры.

Определение 1. *Говорят, что на множестве X определена* *структура топологического пространства, если задана система $\{\Sigma\}$ его подмножеств, обладающая свойствами:*

1) *само множество X и пустое множество $\varnothing$ принадлежат $\{\Sigma\}$;*

2) *сумма любого числа множеств системы $\{\Sigma\}$ и пересечение любого конечного числа множеств системы $\{\Sigma\}$ принадлежат $\{\Sigma\}$.*

Система $\{\Sigma\}$, удовлетворяющая условиям 1)—2), *называется* *топологией на множестве X, а составляющие ее множества* — *открытыми в этой топологии* *.

Таким образом, пара, состоящая из множества X и топологии $\{\Sigma\}$, является топологическим пространством, которое иногда удобно обозначать через (X, Σ).

* Множество, являющееся дополнением к открытому, называется замкнутым в топологическом пространстве.

Определение 1 выделяет весьма общий класс пространств. Обычно этот класс несколько сужают, добавляя к свойствам 1) и 2) так называемые аксиомы отделимости. Из этих аксиом мы рассмотрим наиболее часто используемые.

Аксиома T_2 (Хаусдорфа): *для любых различных точек x и y, принадлежащих множеству X, существуют такое множество Σ_x, содержащее точку x, и множество Σ_y, содержащее точку y, такие что они оба принадлежат системе $\{\Sigma\}$ и не пересекаются, т. е. $\Sigma_x \cap \Sigma_y = \varnothing$.*

Топологические пространства, удовлетворяющие аксиоме T_2 (аксиоме Хаусдорфа), называются хаусдорфовыми.

Аксиома T_1: *для любых двух различных точек x и y, принадлежащих множеству X, существует множество Σ_x, принадлежащее системе $\{\Sigma\}$, содержащее точку x и не содержащее точку y, а также существует множество Σ_y из системы $\{\Sigma\}$, содержащее точку y и не содержащее точку x.*

Топологические пространства, удовлетворяющие аксиоме T_1, называются T_1-пространствами.

Ясно, что если выполнена аксиома T_2, то аксиома T_1 выполнена, т. е. класс топологических пространств, удовлетворяющих аксиомам 1), 2), T_2, более узкий, чем класс топологических пространств, удовлетворяющих аксиомам 1), 2), T_1.

Примером пространства, удовлетворяющего аксиомам 1), 2), T_1 и не удовлетворяющего аксиомам 1), 2), T_2, является следующее топологическое пространство. Множество X состоит из точек отрезка $[0, 1]$, а открытыми считаются следующие множества: X, $\varnothing$, $\Sigma_{\alpha_n} = [0, 1] \setminus \{\alpha_n\}$, где $\{\alpha_n\}$ — произвольное, не более чем счетное множество отрезка $[0, 1]$. Очевидно, что аксиомы 1), 2), T_1 выполнены. Однако аксиома T_2 не выполняется.

Не всякое топологическое пространство удовлетворяет аксиоме T_1. Вот традиционный пример. Множество $X = \{a, b\}$ состоит из двух точек. Топологию зададим открытыми множествами, к которым отнесем все X, пустое множество $\varnothing$ и точку b. Аксиомы 1) и 2) выполнены, а аксиома T_1 — нет.

Приведем наиболее часто встречающиеся примеры топологических пространств.

Примеры.

1) Рассмотрим произвольное метрическое пространство (X, ρ). Открытые множества в силу леммы 1 раздела о метрических пространствах удовлетворяют свойствам 1) и 2) определения 1 топологического пространства. Аксиома T_2 (Хаусдорфа) также выполняется в метрическом пространстве: если $x \neq y$, то $\rho(x, y) = a > 0$ и шары $O\left(x, \frac{a}{3}\right)$, $O\left(y, \frac{a}{3}\right)$ — открытые множества в (X, ρ), такие что $O\left(x, \frac{a}{3}\right) \cap O\left(y, \frac{a}{3}\right) = \varnothing$.

Таким образом, всякое метрическое пространство (X, ρ) является и хаудорфовым топологическим пространством (X, Σ), где $\{\Sigma\}$ — система открытых множеств в (X, ρ).

2) Рассмотрим множество X произвольной природы. Отнесем к системе $\{\Sigma\}$ только все множество X и пустое множество $\varnothing$. Аксиомы 1), 2), очевидно, выполнены. Однако аксиомы T_1 и T_2 не выполнены. Такая топология называется антидискретной.

3) Пусть X — произвольное множество. Отнесем к системе $\{\Sigma\}$ все подмножества множества X. Легко проверить, что $\{\Sigma\}$ — хаусдорфова топология. Такая топология называется дискретной.

Дадим следующее определение.

Определение 2. *Окрестностью точки x, принадлежащей топологическому пространству (X, Σ), называется любое открытое множество, содержащее точку x. Окрестностью некоторого подмножества X (быть может, самого X) называется любое открытое множество, содержащее данное подмножество (или X). Окрестность точки x будем обозначать Σ_x.*

Предположим, что для каждой точки x, принадлежащей топологическому пространству (X, Σ) среди всех окрестностей этой точки выделены некоторые, причем так, что какова бы ни была точка x и ее произвольная окрестность Σ_x, существует окрестность Σ_x^1 точки x из выделенной системы, что $x \in \Sigma_x^1 \subset \subset \Sigma_x$.

Определение 3. *Система выделенных окрестностей $\{\Sigma_x^1\}$ называется определяющей системой окрестностей данного топологического пространства*.*

Справедлива следующая лемма, которая дает удобный способ задания топологии.

Лемма 1. *Пусть X — произвольное множество. Для каждой точки x определим некоторые подмножества Σ_x, называемые «окрестностями» точки x и удовлетворяющие условиям:*

а) *каждая точка имеет хотя бы одну свою «окрестность» и принадлежит любой своей «окрестности»;*

б) *пересечение двух «окрестностей» точки содержит некоторую «окрестность» этой же точки;*

в) *какова бы ни была «окрестность» Σ_x точки $x \in X$ и точка $y \in \Sigma_x$, существует «окрестность» Σ_y точки y такая, что $\Sigma_y \subset \Sigma_x$;*

Тогда если отнести к системе $\{\Sigma\}$ всевозможные «окрестности» Σ_x точек $x \in X$, их всевозможные объединения и пустое множество, то будет задана топология на множестве X и (X, Σ) — топологическое пространство, в котором система всех

---

* Если x фиксировано, то система $\{\Sigma_x^1\}$ называется определяющей системой окрестностей данной точки x.

«окрестностей» является определяющей системой. Обратно, всякое топологическое пространство может быть получено таким способом.

Доказательство. Проверим выполнение аксиом 1)—2) топологического пространства. То, что все X принадлежит $\{\Sigma\}$, очевидно, $\varnothing$ отнесено к $\{\Sigma\}$ по условию.

Аксиома 1) выполнена.

Для проверки аксиомы 2) надо убедиться лишь в том, что $\Sigma_1 \cap \Sigma_2 \in \{\Sigma\}$, если $\Sigma_1 \in \{\Sigma\}$, $\Sigma_2 \in \{\Sigma\}$. Следовательно, надо установить, что $\Sigma_1 \cap \Sigma_2$ может быть получено как объединение некоторых «окрестностей», т. е. надо убедиться, что для любой точки $x \in \Sigma_1 \cap \Sigma_2$ существует «окрестность» $\Sigma_x \subset \Sigma_1 \cap \Sigma_2$. Но Σ_1 и Σ_2 принадлежат $\{\Sigma\}$, поэтому имеются «окрестности» $\Sigma_x^1 \subset \Sigma_1$ и $\Sigma_x^2 \subset \subset \Sigma_2$ — их пересечение содержит по условию б) некоторую «окрестность» Σ_x точки x, которая содержится, очевидно, в $\Sigma_1 \cap \Sigma_2$.

Обратно, если задано топологическое пространство (X, Σ), то в качестве «окрестности» точки x, удовлетворяющей условиям а)—в), можно взять произвольное множество из системы $\{\Sigma\}$, содержащее точку x.

Используя эту лемму, приведем примеры еще двух хаусдорфовых топологических пространств.

Примеры.

1) В качестве X возьмем двумерную плоскость E^2. Окрестность любой точки $x \in X$ получим, если из любого открытого круга с центром в x удалим все отличные от самой точки x точки, лежащие на вертикальном диаметре этого круга. Полученное топологическое пространство является хаусдорфовым.

2) Рассмотрим в качестве X отрезок $[0, 1]$, окрестности всех точек, кроме точки 0, определим обычным образом, а окрестностями точки 0 будем считать всевозможные полуинтервалы $[0, \alpha)$, $\alpha > 0$, из которых выкинуты точки $\frac{1}{n}$, где n — натуральное число. Это, как легко видеть, пример хаусдорфова топологического пространства.

Пусть (X, Σ) — топологическое пространство, а Y — подмножество X. Тогда на подмножестве X можно рассмотреть след системы $\{\Sigma\}$, т. е. множества вида $\{\Sigma_Y\} = \{Y \cap \Sigma_\alpha\}$, $\Sigma_\alpha \in \{\Sigma\}$. Легко видеть, что тем самым на Y задана топология, поэтому Y само превращается в топологическое пространство и (Y, Σ_Y) называется **подпространством** пространства (X, Σ). Топология, задаваемая системой $\{\Sigma_Y\} = \{Y \cap \Sigma_\alpha\}$, $\Sigma_\alpha \in \{\Sigma\}$ называется **индуцированной топологией**.

Так же, как и в случае метрических пространств, пространство (X, Σ) называется **связным**, если его нельзя представить в виде суммы двух непустых открытых непересекающихся подмножеств. Множество Y в топологическом пространстве (X, Σ) связно, если Y связно, как подпространство в (X, Σ).

2. Замечание о топологических пространствах. После того как введены открытые множества для топологических пространств, можно ввести все понятия, введенные для метрических пространств. Так, дословно сохраняются определения предельной точки множества (см. определение 4 раздела о метрических пространствах), определение внутренней точки* (см. определение 5 раздела о метрических пространствах), определение замкнутого множества (см. определение 6 раздела о метрических пространствах), определение замыкания множества (см. определение 7 раздела о метрических пространствах), определение плотного и всюду плотного множества (см. определение 10 раздела о метрических пространствах), полностью сохраняется определение понятия нигде не плотного, совершенного множества, данные для метрических пространств. Точно так же, как и в случае метрических пространств, в случае топологических пространств определяется важное понятие непрерывного отображения (определение 12 раздела о метрических пространствах), понятие гомеоморфного отображения (определение 14 раздела о метрических пространствах), определение компактного топологического пространства, или компакта (см. определение 15 раздела о метрических пространствах). Так же, как и для метрических пространств, для топологических пространств вводится понятие центрированной системы (определение 16 раздела о метрических пространствах), определение базы топологии топологического пространства, топологического пространства со счетной базой. Топологические пространства со счетной базой называются топологическими пространствами *со второй аксиомой счетности*.

Читатель без труда сформулирует эти определения для случая топологических пространств, для этого в соответствующих определениях раздела о метрических пространствах выражение «метрическое пространство» следует заменить на выражение «топологическое пространство».

Согласно этим определениям в случае топологического пространства остаются справедливыми основные утверждения, доказанные для метрических пространств. Это вполне естественно, поскольку доказательства этих утверждений в основном используют понятие открытого и замкнутого множества и непосредственно не зависят от введенной там метрики.

Так, лемма 1, утверждающая, что объединение произвольного числа открытых множеств и пересечения конечного их чис-

* Совокупность всех внутренних точек множества Y называется внутренностью множества Y и обозначается через $\overset{\circ}{Y}$. Внутренность $\overset{\circ}{Y}$, очевидно, есть объединение всех открытых множеств, принадлежащих Y.

ла является множеством открытым, есть соответствующая аксиома топологического пространства; утверждение леммы 2, в том числе и доказательства свойств операции замыкания, полностью сохраняется.

На топологические пространства переносится также и понятие сходящейся последовательности (определение 11 раздела о метрических пространствах), а именно: последовательность $\{a_n\}$ точек топологического пространства называется **сходящейся к точке** a этого пространства, если любая окрестность точки a содержит все точки последовательности $\{a_n\}$ за исключением конечного числа. Если последовательность $\{a_n\}$ сходится к точке a, то пишут, что $a_n \to a$ при $n \to \infty$ или $\lim\limits_{n \to \infty} a_n = a$.

Однако в топологических пространствах это понятие не играет столь большой роли, как в метрических пространствах. В самом деле, лемма 3 утверждала, что в метрическом пространстве точка a принадлежит замыканию $\bar{A}$ некоторого множества A тогда и только тогда, когда существует последовательность $\{a_n\}$ точек множества A, сходящаяся к a. В топологическом пространстве этот факт может быть несправедлив. (Вспомним, что при доказательстве этого утверждения в метрических пространствах мы строили последовательность шаров $O\left(a, \frac{1}{n}\right)$, вложенных друг в друга для любого натурального n.)

Можно выделить класс топологических пространств, обладающих аналогичным свойством.

Назовем топологическое пространство **пространством с первой аксиомой счетности**, если для любой его точки a существует счетная система ее окрестностей $\{\Sigma_a{}^n\}$ такая, что для любого открытого множества Σ_a, содержащего точку a, найдется окрестность $\Sigma_a^{n_0}$, обладающая свойством: $\Sigma_a^{n_0} \subset \Sigma_a$. Такая система окрестностей называется **счетной определяющей системой окрестностей точки** a.

В метрическом пространстве первая аксиома счетности, очевидно, выполнена.

В топологическом пространстве (X, Σ) с **первой аксиомой счетности** справедливо утверждение: *точка $a \in X$ принадлежит замыканию $\bar{A}$ некоторого множества A тогда и только тогда, когда существует последовательность $\{a_n\}$ точек множества A, сходящаяся к a.*

Доказательство этого утверждения аналогично доказательству соответствующего утверждения для случая метрических пространств. Последовательность шаров $O\left(a, \frac{1}{n}\right)$ следует заменить на последовательность окрестностей из системы $\{\Sigma_a{}^n\}$, причем всегда можно считать, что $\Sigma_a{}^{n+1}\subset\Sigma_a{}^n$. В противном случае $\Sigma_a{}^n$ надо заменить на $\bigcap\limits_{k=1}^{n}\Sigma_a^k$.

Утверждение на с. 544 полностью сохраняется. Сохраняется также и утверждение леммы 4 — критерий непрерывности отображения.

На случай топологических пространств полностью переносится критерий компактности в терминах центрированной системы замкнутых подмножеств (лемма 5), утверждения лемм 6, 7, 8, 9 о свойствах компакта и непрерывных функций на нем.

Заметим, что топологическое пространство может не быть пространством со счетной базой топологии даже тогда, когда оно является пространством с первой аксиомой счетности и в нем имеется счетное всюду плотное множество. Однако если в топологическом пространстве есть счетная база топологии, то топологическое пространство сепарабельно и удовлетворяет первой аксиоме счетности (ср. с леммой 11). Точно так же, как и в случае метрических пространств (см. определение 18), топологическое пространство называется пространством со второй аксиомой счетности, если в нем существует хотя бы одна база топологии, состоящая не более чем из счетного числа множеств.

Линейные нормированные пространства, линейные операторы

Понятие линейного пространства играет фундаментальную роль в анализе. Линейное пространство и линейные операторы в таких пространствах играют также важную роль во многих других разделах математики.

1. Определение линейного пространства. Примеры.

Определение 1. *Множество элементов L, содержащее хотя бы один элемент, называется* л и н е й н ы м *или* в е к т о р н ы м п р о с т р а н с т в о м, *если выполнены следующие аксиомы.*

1. *Для любых двух элементов x и y множества L однозначно определен третий элемент z этого множества, называемый их* с у м м о й *и обозначаемый символом $x+y$, причем справедливы следующие свойства:*

а) $x+y=y+x$ *(коммутативность)*;

б) $(x+y)+z=x+(y+z)$ *(ассоциативность)*;

в) *существует такой элемент O, что $x+O=x$ для любого элемента $x\in L$; элемент O называется* н у л е в ы м *или* н у л е м *пространства L*;

г) *для любого элемента $x\in L$ существует такой элемент x',* что $x+x'=0$; *элемент x' называется* п р о т и в о п о л о ж н ы м *к элементу x и обозначается обычно так: $-x$.*

2. *Для любого числа* α *и любого элемента* x *пространства* L *определен элемент* y *пространства* L, *называемый* ***произведением*** α ***на*** x *и обозначаемый символом* αx, *причем справедливы следующие свойства*:

а) $\alpha(\beta x)=(\alpha\beta)x$ (*для любых чисел* α *и* β *и любого элемента* x);

б) $(\alpha+\beta)x=\alpha x+\beta x$ (*для любых чисел* α *и* β *и любого элемента* x);

в) $\alpha(x+y)=\alpha x+\alpha y$ (*для любого числа* α *и любых элементов* x *и* y);

г) $1\cdot x=x$ *для любого элемента* x.

Если в аксиоме 2 используются только действительные числа, пространство L называется **действительным пространством**. Если же рассмотрения ведутся с использованием комплексных чисел, то и пространство L называется **комплексным**.

Элементы линейного пространства называют также **векторами**.

Рассмотрим некоторые **примеры** линейных пространств. При изучении метрических пространств были рассмотрены следующие множества (см. раздел о метрических пространствах, примеры): множество вещественных чисел, координатное n-мерное пространство, множество непрерывных на сегменте функций и совокупность ограниченных последовательностей. Все эти множества представляют собой и примеры линейных пространств, если ввести операции сложения и умножения по следующим правилам.

1) В совокупности A^1 вещественных чисел — обычные арифметические операции сложения и умножения.

2) В n-мерном координатном пространстве A^n — по формулам

$$x+y=(x_1,\ x_2,\ldots,x_n)+(y_1,\ y_2,\ldots,y_n)=(x_1+y_1,\ x_2+y_2,\ldots,x_n+y_n),$$
$$\alpha x=\alpha(x_1,\ x_2,\ldots,x_n)=(\alpha x_1,\ \alpha x_2,\ldots,\alpha x_n).$$

3) В пространстве непрерывных функций на отрезке — обычные операции сложения двух функций, т. е. $f+g=f(x)+g(x)$, и умножение функции на число, т. е. $\alpha f=\alpha f(x)$.

В множестве ограниченных последовательностей введем операции сложения и умножения по формулам

$$(x+y)=(x_1,\ x_2,\ldots)+(y_1,\ y_2,\ldots)=(x_1+y_1,\ x_2+y_2,\ldots),$$
$$\alpha x=\alpha(x_1,\ x_2,\ldots)=(\alpha x_1,\ \alpha x_2,\ldots).$$

Нетрудно проверить во всех перечисленных выше примерах выполнимость аксиом, определяющих линейное пространство. За этими линейными пространствами мы сохраним те же обозначения, что и в случае, когда рассматривались метрические пространства, т. е. соответственно в случае 1) — обозначение E^1, в случае 2) — E^n, в случае 3) — обозначение $C[a, b]$ и в случае 4) — обозначение m.

Определение 2. *Линейные пространства L и L^* называются* **изоморфными**, *если между их элементами можно установить взаимно однозначное соответствие так, что если элемент $x \in L$ соответствует элементу $x^* \in L^*$, а элемент $y \in L$ соответствует элементу $y^* \in L^*$, то элемент $x+y$ соответствует элементу x^*+y^*, а элемент αx соответствует элементу αx^* (α — любое число).*

Изоморфные пространства можно рассматривать как различные реализации одного и того же пространства, и такие пространства можно не различать.

2. Нормированные пространства. Банаховы пространства. Примеры.

Определение 3. *Функция $f(x) = \|x\|$, ставящая каждому элементу x линейного пространства L в соответствие вещественное число $\|x\|$, называется* **нормой** *в линейном пространстве L, если она удовлетворяет следующим аксиомам:*

1) $f(x) = \|x\| = 0$ *тогда и только тогда, когда* $x = 0$;

2) $f(\alpha x) = \|\alpha x\| = |\alpha| \, \|x\| = |\alpha| f(x)$ *для любого числа* α;

3) $f(x_1 + x_2) = \|x_1 + x_2\| \leqslant \|x_1\| + \|x_2\|$ *для любых x_1 и x_2, принадлежащих L.*

Определение 4. *Линейное пространство L с введенной функцией $f(x) = \|x\|$ называется* **линейным нормированным пространством.**

Чтобы подчеркнуть, что пространство L нормированное, обозначают его обычно буквой N.

Значение нормы на элементе x пространства N называется **нормой вектора** x или **длиной** или **модулем** этого вектора.

Норма вектора всегда неотрицательна и равна нулю только для нулевого вектора. Действительно, полагая в аксиоме 3) $x_1 = -x_2$ с учетом 1) и 2), получаем, что

$$0 = \|0\| = \|x_1 - x_1\| \leqslant \|x_1\| + \|-x_1\| = \|x_1\| + |-1| \, \|x_1\| = 2\|x_1\|,$$

т. е.

$$f(x_1) = \|x_1\| \geqslant 0.$$

В любом нормированном пространстве может быть введена естественная метрика по правилу $\rho(x, y) = \|x - y\|$. Из аксиом, задающих норму, вытекает, что функция $\rho(x, y)$ действительно задает расстояние в пространстве, т. е. удовлетворяет аксиомам расстояния.

Кроме того, поскольку N — линейное пространство, метрика $\rho(x, y)$ **инвариантна относительно сдвигов**, т. е.

$$\rho(x_1 + x, x_2 + x) = \|(x_1 + x) - (x_2 + x)\| = \|x_1 - x_2\| = \rho(x_1, x_2),$$

и **положительно однородна**, т. е.

$$\rho(\alpha x, \alpha y) = \|\alpha x - \alpha y\| = \|\alpha(x-y)\| = |\alpha|\,\|x-y\| = |\alpha|\rho(x, y).$$

С появлением естественной метрики в нормированном пространстве могут быть введены все те понятия, которые мы рассматривали в метрических пространствах, например полнота и т. п.

Следующее определение выделяет из всех нормированных пространств важнейший класс пространств, называемых банаховыми *.

Определение 5. *Банаховым пространством называется полное линейное нормированное пространство.*

Приведем примеры банаховых пространств.

Примеры. 1) Пространство E^1 — действительная числовая ось с обычными арифметическими операциями, как мы знаем, является линейным пространством. Оно превращается в нормированное пространство, если норму числа x положить равной его абсолютной величине: $\|x\| = |x|$. Из свойств абсолютной величины вытекает, что выполнены все свойства нормы. Это нормированное пространство E^1 является полным, т. е. банаховым. Его полнота, т. е. полнота E^1 как метрического пространства с обычным расстоянием между действительными числами ($\rho(x, y) = \|x-y\| = |x-y|$), была установлена в гл. 3.

2) Линейное пространство E^n, рассмотренное нами выше, также является нормированным пространством, если норму вектора $x = (x_1, x_2, \ldots, x_n)$ ввести по правилу

$$\|x\| = (x_1^2 + x_2^2 + \ldots + x_n^2)^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}.$$

Аксиомы 1) и 2) нормы, очевидно, выполнены, а неравенство треугольника превращается в неравенство

$$\|x+y\| = \left(\sum_{i=1}^{n} |x_i + y_i|^2\right)^{1/2} \leqslant \left(\sum_{i=1}^{n} (|x_i| + |y_i|)^2\right)^{1/2} \leqslant$$

$$\leqslant \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} + \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2} = \|x\| + \|y\|,$$

которое является частным случаем (при $p=2$) неравенства Минковского для сумм (см. § 5 гл. 9). Это пространство E^n является полным, так как сходимость x_k к x в E^n означает, что все координаты x_k^i вектора x_k сходятся к соответствующим координатам вектора x. Осталось применить утверждение примера 1).

* Стефан Банах — известный польский математик (1892—1945).

3) Пространство $C[a, b]$ непрерывных функций с обычными операциями сложения функций и умножения функций на число, как мы уже говорили выше, является линейным пространством. Оно превращается в нормированное пространство, если положить $\|f\| = \max\limits_{x\in[ab]} |f(x)|$, где $f(x)$ — непрерывная на сегменте $[a, b]$ функция. Используя свойства абсолютной величины и функции max, легко убедиться, что все аксиомы нормы здесь выполнены *.

4) Пространство m-ограниченных последовательностей также можно сделать нормированным, положив $\|x\| = \sup\limits_k |x_k|$ для $x = (x_1, x_2, \ldots)$. Аксиомы нормы здесь также проверяются без труда. Можно убедиться также, что оно является полным нормированным пространством, т. е. банаховым пространством.

3. Операторы в линейных и нормированных пространствах. Пусть L_1 и L_2 — два линейных пространства и A — отображение пространства L_1 в L_2, т. е. $A: L_1 \to L_2$.

Определение 6. *Отображение $A: L_1 \to L_2$ одного линейного пространства L_1 в другое L_2 называется линейным отображением или линейным оператором, если:*

а) $A(x_1 + x_2) = Ax_1 + Ax_2$ *для любых векторов x_1 и x_2 пространства L_1;*

б) $A(\alpha x) = \alpha Ax$ *для любого вектора $x \in L_1$ и любого числа α.*

Отображение $A: L_1 \times L_2 \times \ldots \times L_n \to L$ прямого произведения линейных пространств $L_1, L_2, \ldots, L_n$** в линейное пространство L называется полилинейным, если линейно каждое отображение

$$A_i: L_i \to L, \quad i = 1, 2, \ldots, n,$$

получаемое из $A(x, y, \ldots, z)$ фиксированием всех переменных, кроме переменной, стоящей на i-м месте. Здесь $x \in L_1$, $y \in L_2, \ldots$ $\ldots, z \in L_n$.

Если задано линейное отображение $A: L_1 \to L_2$ линейного пространства L_1 в линейное пространство L_2 и пространство L_2 является действительной числовой осью (или комплексной плоскостью), то оператор A называют функционалом.

Рассмотрим некоторые примеры.

Примеры. 1) Пусть L — произвольное линейное пространство. Оператор E ставит в соответствие каждому элементу $x \in L$ тот же элемент x этого же пространства, т. е. $E: L \to L$ и $Ex = x$ для любого $x \in L$. Такой оператор называется единичным или единицей.

* Можно убедиться, что это пространство является и полным, т. е. банаховым.

** Линейные операции в $L_1 \times L_2 \times \ldots \times L_n$ определены равенствами $[x_1, \ldots, x_n] + [y_1, \ldots, y_n] = [x_1 + y_1, \ldots, x_n + y_n]$, $\alpha[x_1, \ldots, x_n] = [\alpha x_1, \ldots, \alpha x_n]$, где $[x_1, \ldots, x_n]$, $[y_1, \ldots, y_n] \in L_1 \times \ldots \times L_n$, $x_i, y_i \in L_i$, $i = 1, 2, \ldots, n$; α — число.

2) В трехмерном пространстве E^3 зададим оператор A как линейное преобразование, состоящее в проектировании каждого вектора на ось Ox (т. е. каждому вектору в соответствие ставится его координата на оси Ox). Оператор можно задать матрицей. Например, в базисе из единичных векторов e_1, e_2, e_3, направленных по осям координат Ox, Oy, Oz соответственно, указанный оператор A можно задать так:

$$Ax=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},\quad x=(x_1,\ x_2,\ x_3)\in E^3.$$

3) В пространстве $C[a,\ b]$ линейный оператор A задан по правилу: $Af=\int\limits_a^b f(x)\,dx$. Оператор A отображает $C[a,\ b]$ в числовую ось и является функционалом.

4) Модуль непрерывности $\omega(f,\ \delta)$ (см. гл. 4) и дельта-функция $\delta(x-a)$, т. е. оператор $\delta(x-a)$, действующий по правилу $\delta(x-a)[f(x)]=f(a)$, также являются функционалами на $C[a,\ b]$.

4. Пространство операторов. Пусть L_1 и L_2 — два линейных пространства. Рассмотрим совокупность $\{A\}$ всех линейных операторов, отображающих L_1 в L_2. В множестве $\{A\}$, элементами которого являются линейные операторы, отображающие L_1 в L_2, можно ввести алгебраические операции. Пусть A_1 и A_2 — такие операторы. Определим *сложение* этих операторов посредством равенства

$$(A_1+A_2)x=A_1x+A_2x,\quad x\in L_1.$$

Умножение линейного оператора на число определим формулой

$$(\alpha A)x=\alpha Ax.$$

Очевидно, что при таких определениях все необходимые аксиомы, задающие линейное пространство, будут выполнены и рассматриваемое множество $\{A\}$ линейных операторов будет линейным пространством. Нулем этого пространства будет нулевой оператор O, т. е. оператор, переводящий любой вектор x в нулевой вектор: $Ox=0$. Пространство линейных операторов, которое мы ввели выше, обычно обозначается так: $(L_1\to L_2)$. Если пространства L_1 и L_2, кроме того, обладают некоторой топологией, например нормированы, то и пространство операторов $(L_1\to L_2)$ будет обладать определенной топологией.

5. Норма оператора. Пусть N_1 и N_2 — два линейных нормированных пространства и оператор A отображает N_1 в N_2.

Определение 7. *Линейный оператор* $A: N_1 \to N_2$, *отображающий линейное нормированное пространство* N_1 *в линейное нормированное пространство* N_2, *называется* ***ограниченным***, *если существует такая постоянная* M, что $\|Ax\|_{N_2} \leq M\|x\|_{N_1}$ *для всех* $x \in N_1$. Индекс внизу символа нормы означает то пространство, в котором вычисляется норма вектора. Если это не будет вызывать недоразумений, эти индексы мы будем опускать.

Справедливо следующее утверждение.

Теорема. *Для того чтобы линейный оператор* $A: N_1 \to N_2$, *отображающий линейное нормированное пространство* N_1 *в* N_2, *был непрерывен, необходимо и достаточно, чтобы он был ограничен.*

(Заметим, что непрерывность оператора понимается как непрерывность соответствующего отображения.)

Доказательство. Пусть A — непрерывный оператор. Если бы он был неограничен, то нашлась бы последовательность элементов $\{x_n\}$ такая, что

$$\|Ax_n\| > n\|x_n\|.$$

Пусть $y_n = \dfrac{x_n}{n\|x_n\|}$. Тогда $y_n \to 0$, так как $\|y_n\| = \dfrac{1}{n} \to 0$ при $n \to \infty$. С другой стороны,

$$\|Ay_n\| = \frac{\|Ax_n\|}{n\|x_n\|} > 1.$$

Заметим, что (в силу линейности оператора A) $A\cdot 0 = 0$; действительно, $A\cdot 0 = A(x-x) = Ax - Ax = 0$. Итак, Ay_n не стремится к $A\cdot 0 = 0$, т. е. оператор A не является непрерывным в нулевой точке, что противоречит условию теоремы.

Обратно, если A ограничен, то $\|Ax\| \leq M\|x\|$. Пусть $x_n \to x$, т. е. $\|x_n - x\| \to 0$ при $n \to \infty$. Тогда $\|Ax_n - Ax\| = \|A(x_n - x)\| \leq M\|x_n - x\| \to 0$ при $n \to \infty$. Следовательно, $Ax_n \to Ax$ и оператор A непрерывен в точке x.

Определение 8. *Пусть A — линейный ограниченный оператор. Наименьшая из постоянных M, удовлетворяющих условию*

$$\|Ax\| \leq M\|x\|,$$

называется ***нормой оператора*** A *и обозначается символом* $\|A\|$.

Таким образом, согласно определению 8 норма оператора обладает следующими свойствами:

а) для любого $x \in N_1$ справедливо неравенство

$$\|Ax\| \leq \|A\|\|x\|;$$

б) для любого числа $\varepsilon>0$ найдется такой элемент x_ε, что

$$\|Ax_\varepsilon\|>(\|A\|-\varepsilon)\cdot\|x_\varepsilon\|.$$

Покажем, что $\|A\|=\sup\limits_{\|x\|\leqslant 1}\|Ax\|$ *или, что то же самое,* $\|A\|=\sup\limits_{\|x\|\neq 0}\frac{\|Ax\|}{\|x\|}$. Действительно, если $\|x\|\leqslant 1$, то $\|Ax\|\leqslant \leqslant\|A\|\cdot\|x\|\leqslant\|A\|$. Значит, и $\sup\limits_{\|x\|\leqslant 1}\|Ax\|\leqslant\|A\|$. С другой стороны, для любого $\varepsilon>0$ существует элемент x_ε такой, что

$$\|Ax_\varepsilon\|>(\|A\|-\varepsilon)\|x_\varepsilon\|.$$

Пусть $\xi_\varepsilon=\frac{x_\varepsilon}{\|x_\varepsilon\|}$. Тогда $\|A\xi_\varepsilon\|=\frac{\|Ax_\varepsilon\|}{\|x_\varepsilon\|}>\frac{1}{\|x_\varepsilon\|}(\|A\|-\varepsilon)\|x_\varepsilon\|= =\|A\|-\varepsilon$. Так как $\xi_\varepsilon=1$, то

$$\sup_{\|x\|\leqslant 1}\|Ax\|\geqslant\|A\xi_\varepsilon\|>\|A\|-\varepsilon.$$

Следовательно, $\sup\limits_{\|x\|\leqslant 1}\|Ax\|\geqslant\|A\|$, и потому $\|A\|=\sup\limits_{\|x\|\leqslant 1}\|Ax\|$.

Замечание. Из проведенных рассуждений следует, что

$$\|A\|=\sup_{\|x\|=1}\|Ax\|.$$

Выше (см. п. 4) нами было введено пространство $(L_1\to L_2)$ операторов, отображающих линейное пространство L_1 в линейное пространство L_2. Это пространство играет важную роль в различных разделах анализа, и мы сейчас продолжим его изучение.

Предположим теперь, что указанные выше линейные пространства L_1 и L_2 являются нормированными. Переобозначим их через N_1 и N_2 соответственно, а соответствующее линейное пространство, элементами которого являются линейные ограниченные операторы, через $(N_1\to N_2)$. В пространстве $(N_1\to N_2)$ можно ввести норму. Для этого норму элемента A пространства $(N_1\to N_2)$ введем по правилу: $\|A\|=\sup\limits_{\|x\|=1}\|Ax\|$. Легко видеть, что эта норма удовлетворяет аксиомам определения нормы. Таким образом, линейное пространство $(N_1\to N_2)$, элементами которого являются линейные ограниченные операторы, есть линейное нормированное пространство. Возникает естественный вопрос: когда это пространство является полным, т. е. банаховым?

Ответ на этот вопрос содержится в доказываемой ниже теореме.

Теорема. *Если линейное нормированное пространство $N_2=B_2$ банахово, то пространство $(N_1\to B_2)$ линейных ограниченных операторов также является банаховым.*

Доказательство. Пусть $\{A_n\}$ — фундаментальная последовательность пространства операторов $(N_1 \to B_2)$, т. е. $\|A_n - A_m\| \to 0$, когда $n, m \to \infty$. Для любого x получим, что $\|A_n x - A_m x\| \leqslant \|A_n - A_m\| \|x\| \to 0$ при $n, m \to \infty$. Следовательно, если $x \in N_1$ фиксировано, то последовательность элементов $\{A_n x\}$ фундаментальна в B_2, т. е. в силу полноты B_2 эта последовательность сходится. Обозначим $y = \lim\limits_{n\to\infty} A_n x$. Мы получили, таким образом, отображение N_1 в B_2. Оператор, осуществляющий это отображение, обозначим через A. Из свойств предела следует, что он линеен. Докажем его ограниченность. Из того, что $\|A_n - A_m\| \to 0$ при $n, m \to \infty$, следует, что $|\|A_n\| - \|A_m\|| \to 0$ при $n, m \to \infty$, т. е. числовая последовательность $\{\|A_n\|\}$ фундаментальна в E^1, а следовательно, и ограничена. Существует такая постоянная M, что $\|A_n\| \leqslant M$ для любого натурального n. Отсюда получаем, что $\|A_n x\| \leqslant \|A_n\| \|x\| \leqslant M\|x\|$, т. е. в силу того, что функция, определяющая норму (расстояние), непрерывна, имеем

$$\|Ax\| = \lim_{n\to\infty} \|A_n x\| \leqslant M\|x\|.$$

Итак, оператор A — ограничен. Оператор A был нами определен как оператор, отображающий N_1 в B_2 по указанному выше правилу. Покажем, что A является пределом последовательности $\{A_n\}$ в смысле сходимости по норме в пространстве $(N_1 \to B_2)$. Зададим произвольное $\varepsilon > 0$ и выберем n_0 так, что $\|A_{n+p}x - A_n x\| < \varepsilon$ для $n \geqslant n_0$ и любого x такого, что $\|x\| \leqslant 1$. Пусть $p \to \infty$. Тогда $\|Ax - A_n x\| \leqslant \varepsilon$ для $n \geqslant n_0$ и всех x с нормой, не превосходящей единицы. Поэтому для $n \geqslant n_0$ получим $\|A_n - A\| = \sup\limits_{\|x\| \leqslant 1} \|(A_n - A)x\| \leqslant \varepsilon$. Следовательно, $A = \lim\limits_{n\to\infty} A_n$ в смысле сходимости по норме в пространстве $(N_1 \to B_2)$, т. е. это пространство банахово, что и требовалось доказать.

6. Понятие гильбертова * пространства.

Определение 9. *Гильбертовым пространством называется множество H элементов f, g, h, ..., обладающее следующими свойствами:*

1) *H представляет собой линейное пространство, т. е. в H определены действия сложения и умножения на действительные или комплексные числа (в зависимости от этого H называется действительным или комплексным пространством).*

2) *H является метрическим пространством, причем метрика вводится с помощью скалярного произведения, т. е. числовой функции (f, g) от пары аргументов f и g, называемой их скалярным произведением и удовлетворяющей аксиомам:*

* Давид Гильберт — немецкий математик (1862—1943).

а) $(\alpha f, g)=\alpha(f, g)$ для любого числа α;

б) $(f+g, h)=(f, h)+(g, h)$;

в) $(f, g)=\overline{(g, f)}$ *;

г) $(f, f)>0$ при $f\neq 0$; $(f, f)=0$ при $f=0$.

Норма $\|f\|$ *элемента* f *определяется равенством* $\|f\|=(f, f)^{1/2}$, *а расстояние между элементами* f *и* g *полагается равным*

$$\rho(f, g)=\|f-g\|.$$

3) *H является полным пространством, как метрическое пространство с введенным выше расстоянием.* (*Конечномерное пространство всегда полно.*)

Возьмем произвольные элементы $f, g\in H$, пусть λ — действительное число. Тогда

$$0\leqslant(g+\lambda(f, g)g, f+\lambda(f, g)g)=(f, f)+2\lambda|(f, g)|^2+ +\lambda^2|(f, g)|^2(g, g)$$

и, следовательно, такой многочлен относительно λ не может иметь различных действительных корней. Отсюда вытекает, что

$$|(f, g)|^4-(f, f)|(f, g)|^2(g, g)\leqslant 0.$$

Таким образом (даже в случае $(f, g)=0$),

$$|(f, g)|^2\leqslant(f, f)(g, g) \text{ или } |(f, g)|\leqslant\|f\|\|g\|.$$

Мы получили неравенство Коши — Буняковского. Знак равенства в нем, помимо тривиального случая $f=0$ или $g=0$, достигается только тогда, когда $f=-\lambda(f, g)g$ при некотором значении λ, т. е. когда векторы f и g коллинеарны.

Используя это неравенство, легко проверить, что норма элемента $\|f\|$ и расстояние $\rho(f, g)=\|f-g\|$ удовлетворяют всем аксиомам, входящим в их определение.

Вместе с метрикой в гильбертовом пространстве появляются понятия, связанные с предельным переходом в смысле введенного расстояния.

Наличие скалярного произведения позволяет ввести в H понятия угла между векторами (если H вещественно). Угол (f, g) между векторами f и g определяется равенством

$$\cos(\widehat{f, g})=\frac{(f, g)}{\|f\|\|g\|}.$$

Это понятие, в свою очередь, позволяет назвать два вектора ортогональными, если они образуют угол в 90°. Другими словами, векторы f и g называются ортогональными, если $(f, g)=0$.

---

* $\overline{(g, f)}$ означает комплексно сопряженное к числу (g, f).

Если вектор f ортогонален векторам $g_1, \ldots, g_n$, то он ортогонален и их линейной комбинации $\sum_{i=1}^{n} \alpha_i g_i$.

Если векторы $g_1, \ldots, g_n, \ldots$ ортогональны вектору f и $g = \lim_{n\to\infty} g_n$, то вектор g ортогонален вектору f.

Из сказанного следует, что совокупность всех векторов, ортогональных векторам $f_1, \ldots, f_n$, где n фиксировано, образует замкнутое линейное многообразие, т. е. подпространство, называемое **ортогональным дополнением** к множеству $\{f_1, \ldots, f_n\}$.

В гильбертовом пространстве H можно ввести важное понятие сопряженного оператора.

Определение 10. *Оператор A^* называется* ***сопряженным*** *к линейному ограниченному оператору A, если для любых элементов $x, y \in H$ выполняется равенство $(Ax, y) = (x, A^*y)$.*

Ограниченный линейный оператор A, действующий в гильбертовом пространстве H, называется **самосопряженным**, если $(Ax, y) = (x, Ay)$ для всех x и y из H.

Примеры. 1) В n-мерном пространстве E^n, элементами которого являются наборы чисел $x = (x_1, x_2, \ldots, x_n)$, можно ввести скалярное произведение по формуле $(x, y) = \sum_{i=1}^{n} x_i y_i$. Учитывая, что конечномерное пространство E^n полное, заключаем, что E^n является гильбертовым пространством. Аксиомы скалярного произведения здесь, очевидно, выполнены.

2) Операторы E (единичный), O (нулевой) являются примерами самосопряженных операторов, действующих в гильбертовом пространстве. Для них всегда выполнены соотношения

$$(Ex, y) = (x, y) = (x, Ey),$$

$$(Ox, y) = (0, y) = 0 = (x, Oy).$$

ДОПОЛНЕНИЕ 3

Дифференциальное исчисление в линейных нормированных пространствах

В гл. 5, 6, 7 были изучены вопросы дифференциального исчисления функции одной переменной, а также были исследованы экстремальные свойства функции одной переменной. В предыдущих параграфах настоящей главы эти же вопросы изучались уже для функций многих переменных.

Подчеркнем, что всюду в гл. 5, 6, 7 под функцией мы понимали соответствие между точками множества $\{M\}$ числовой оси (или точками множества $\{M\}$ m-мерного евклидова пространства) и подмножеством числовой оси. Другими словами, такие функции отображают множество $\{M\}$ числовой оси или m-мерного евклидова пространства в подмножество числовой оси. Такие функции могут быть названы **числовыми (скалярными) функциями**, ибо множество значений таких функций есть числа (скаляры) вещественной оси.

В дополнении 2 к этой главе мы вводили понятие функции, отображающей одно абстрактное множество в другое (например, одно метрическое пространство в другое метрическое пространство, одно нормированное пространство в другое нормированное пространство и т. д.). Такие функции называются **операторами**, отображениями, функциями множеств и т. д.

В частности, можно рассматривать функции, отображающие m-мерное евклидово пространство в n-мерное евклидово пространство. Такие функции называются уже **векторными функциями**, поскольку значениями таких функций являются не числа, а векторы некоторого пространства. Например, если отображение происходит в n-мерное евклидово пространство, то значениями функции, осуществляющей это отображение, являются векторы n-мерного пространства.

Примером функции, осуществляющей отображение одного метрического пространства X в то же пространство X может быть тождественное отображение E, ставящее в соответствие каждой точке $x \in X$ ту же точку $x \in X$. Отображение $\|x\|$, ставящее в соответствие точке x нормированного пространства N число $\|x\|$ — норму элемента x, есть пример функции, заданной на нормированном пространстве N.

В этом параграфе будет построено дифференциальное исчисление функций, заданных на нормированном пространстве. В качестве элементарного следствия наших построений могут быть получены факты, относящиеся к функциям, осуществляющим отображения m-мерного евклидова пространства в n-мерное евклидово пространство (при этом натуральное число n может как совпадать, так и не совпадать с натуральным числом m).

1. Понятие дифференцируемости. Сильная и слабая дифференцируемость в линейных нормированных пространствах. Пусть N_1 и N_2 — два нормированных пространства и F — отображение (функция), действующее из N_1 в N_2 ($F: N_1 \to N_2$) и определенное на некотором открытом множестве Σ пространства N_1. Напомним, что поскольку N_1 — нормированное пространство, то оно, в частности, является и метрическим пространством; следовательно, все понятия, введенные в метрических пространствах, такие, как открытое, замкнутое, ограниченное множество, расстояние между точками и т. д. имеют смысл. Подчеркнем, что в этом дополнении будут использоваться многие понятия, введенные в дополнении 2. Введем понятие дифференцируемого отображения.

Определение. *Назовем отображение* $F: N_1 \to N_2$ *дифференцируемым** *в данной точке* x, *принадлежащей открытому множеству* $\Sigma \subset N_1$, *если существует такой ограниченный линейный оператор* $L_x \in (N_1 \to N_2)$ **, *что для любого* $\varepsilon > 0$ *существует* $\delta > 0$ *такое, что если* $h \in N_1$ *и* $\|h\|_{N_1} < \delta$, то

$$\|F(x+h) - F(x) - L_x h\|_{N_1} \leqslant \varepsilon \|h\|_{N_1}$$

или, что то же самое,

$$\|F(x+h) - F(x) - L_x h\|_{N_2} = o(h),$$

где $o(h)/\|h\| \to 0$ *при* $\|h\|_{N_1} \to 0$. *Здесь* $x+h \in \Sigma \subset N_1$.

Индексы N_2 или N_1 у знака $\|\cdot\|$ нормы означают, что норма берется соответственно в пространстве N_2 или N_1. Для простоты записи договоримся в дальнейшем о том, что там, где не будет возникать недоразумений, эти индексы опускать.

Выберем последовательность чисел $\varepsilon_n \to 0$. Согласно сформулированному определению ей отвечает последовательность чисел δ_n и для любой последовательности точек h_n такой, что $\|h_n\| < \delta_n$, $\|h_n\| \to 0$, мы получим, что

$$\|F(x+h_n) - F(x)\| \leqslant \|F(x+h_n) - F(x) - L_x h_n\| + \|L_x h_n\| \leqslant$$
$$\leqslant \varepsilon_n \|h_n\| + \|L_x\| \|h_n\| \to 0,$$

когда $n \to \infty$ (поскольку $\|h_n\| \to 0$, $\varepsilon_n \to 0$, а $\|L_x\| \leqslant C$ в силу ограниченности линейного оператора L_x).

Следовательно, $\lim\limits_{n\to\infty} F(x+h_n) = F(x)$ при $\|h_n\| \to 0$, т. е. если обозначить $x+h_n$ через x_n, то $\lim\limits_{n\to\infty} x_n = x$, $\lim\limits_{n\to\infty} F(x_n) = F(x)$. Та-

---

* Если, в частности, $F(x)$ есть скалярная функция, определенная на некотором интервале (α, β) ($F(x)$ отображает подмножество числовой оси в подмножество числовой оси $F(x): (\alpha, \beta) \to E^1$), то в случае ее дифференцируемости в точке x_0 из интервала (α, β) можно записать:

$$F(x_0+h) - F(x_0) - Lh = o(h), \quad o(h)/h \to 0, \text{ если } h \to 0,$$

где h — некоторое число — приращение аргумента функции, $L = L(x_0)$ — линейный оператор, являющийся производной отображения $F(x)$ в точке x_0, принадлежит в данном случае пространству операторов $(E^1 \to E^1)$, т. е. является просто оператором умножения на число. С другой стороны, если рассматривать $F(x)$ как обычную числовую функцию, то число L есть, очевидно, производная функции $F(x)$ в точке x_0, т. е. $L = F'(x_0)$.

** Через $(N_1 \to N_2)$ в дополнении 2 было обозначено пространство линейных ограниченных операторов, отображающих одно линейное нормированное пространство N_1 в другое линейное нормированное пространство N_2.

ким образом, *дифференцируемое в точке x отображение $F(x)$ непрерывно в этой точке.*

Выражение $L_x h$ при каждом $h \in N_1$ является элементом пространства N_2 и называется ***сильным дифференциалом*** *(или* ***дифференциалом Фреше****) отображения F в точке x (и иногда обозначается символом dF). Линейный оператор L_x называется* ***производной*** *или* ***сильной производной*** *отображения F в точке x. Будем обозначать эту производную символом $F'(x)$. Таким образом, можно записать, что сильный дифференциал отображения F по определению равен $L_x h$, т. е. $dF = L_x h$, а сильная производная $F'(x)$ равна L_x, т. е. $F'(x) = L_x$.*

Если отображение F дифференцируемо в точке x, то соответствующая производная определяется **единственным образом**. В самом деле, пусть

$$F(x+h) - F(x) - L^1{}_x h = \mathbf{o}(h)^* \quad \text{и} \quad F(x+h) - F(x) - L^2{}_x h = \mathbf{o}(h).$$

Тогда

$$L^1{}_x h - L^2{}_x h = \mathbf{o}(h).$$

Следовательно, для любого $\varepsilon > 0$ найдется такое $\delta > 0$, что из неравенства $\|h\| < \delta$ следует неравенство

$$\|L^1{}_x h - L^2{}_x h\| \leqslant \varepsilon \|h\|.$$

Разделим обе части этого неравенства на $\|h\| \neq 0$. Тогда получим, что выполнено неравенство

$$\left\| L_x^1 \frac{h}{\|h\|} - L_x^2 \frac{h}{\|h\|} \right\| \leqslant \varepsilon,$$

справедливое в силу свойства нормы и линейности операторов $L^1{}_x$ и $L^2{}_x$. Полагая $\frac{h}{\|h\|} = e$, где $\|e\| = 1$, мы получим, что

$$\|L^1{}_x e - L^2{}_x e\| \leqslant \varepsilon$$

для любого вектора e из единичной сферы (т. е. для любого вектора, с нормой, равной единице).

В силу произвольности ε отсюда следует совпадение операторов $L^1{}_x$ и $L^2{}_x$ всюду на единичной сфере. Поскольку операторы $L^1{}_x$ и $L^2{}_x$ линейные, то они, очевидно, совпадают и всюду, т. е.

$$L^1{}_x = L^2{}_x.$$

* Написанное равенство следует понимать так:

$$\|F(x+h) - F(x) - L_x{}^1 h\| = o(h).$$

В равенстве $F(x+h) - F(x) - L_x{}^1 h = \mathbf{o}(h)$, $\frac{\mathbf{o}(h)}{\|h\|} \to 0$, если $\|h\| \to \infty$, т. е. норма величины $\mathbf{o}(h)$ является числовой величиной $o(\|h\|)$.

Для того чтобы проиллюстрировать определение сильной дифференцируемости отображения одного нормированного пространства в другое на конкретном примере, рассмотрим случай отображения $F: E^m \to E^1$, т. е. отображения m-мерного евклидова пространства в числовую ось. В этом случае* отображение $F = F(x_1, x_2, \dots, x_m)$ есть обычная числовая функция от m переменных. Если обозначить приращения аргумента x_1 через h_1, аргумента x_2 через h_2, ..., аргумента x_m через h_m, то, как это следует из формулы (12.15), условие дифференцируемости функции m переменных $F(x_1, x_2, \dots, x_m)$ в точке $x = (x_1, x_2, \dots, x_m)$, принадлежащей некоторому открытому множеству $\Sigma \subset E^m$, записывается в виде

$$F(x_1+h_1, x_2+h_2, \dots, x_m+h_m) - F(x_1, x_2, \dots, x_m) =$$
$$= F(x+h) - F(x) = A_1h_1 + A_2h_2 + \dots + A_mh_m + o(\rho),$$

где $x+h = (x_1+h_1, x_2+h_2, \dots, x_m+h_m)$, $A_1 = \dfrac{\partial F(x)}{\partial x_1}$,

$$A_2 = \frac{\partial F(x)}{\partial x_2}, \dots, A_m = \frac{\partial F(x)}{\partial x_m}, \quad \rho = (h_1^2 + h_2^2 + \dots + h_m^2)^{1/2},$$

$$x = (x_1, x_2, \dots, x_m).$$

Заметим, что справедливо равенство $\rho = \|h\|$, где $\|h\|$ берется в пространстве E^m, как в нормированном пространстве.

Поэтому если функция m переменных дифференцируема, то она, очевидно, и сильно дифференцируема, как отображение нормированного пространства E^m в нормированное пространство E^1, причем сильная производная L_x определяется из условия $L_x h = \sum\limits_{i=1}^{m} \dfrac{\partial F}{\partial x_i} h_i$, т. е. $L_x h$ равно скалярному произведению вектора $F_x = \left(\dfrac{\partial F}{\partial x_1}, \dfrac{\partial F}{\partial x_2}, \dots, \dfrac{\partial F}{\partial x_m}\right)$ на вектор $h = (h_1, h_2, \dots, h_m)$. Из курса линейной алгебры мы знаем, что всякий линейный функционал из пространства $(E^n \to E^1)$ имеет вид скалярного произведения.

Таким образом, в случае отображения m-мерного евклидова пространства в числовую ось понятие сильной дифференцируемости функции $F(x_1, x_2, \dots, x_m)$, очевидно, в силу единственности производной совпадает с понятием ее дифференцируемости.

Установим теперь некоторые элементарные факты, непосредственно вытекающие из определения производной.

---

* Выбрав базис в E^m, мы всегда, независимо от природы элементов $x \in E^m$, можем считать, что x является упорядоченной совокупностью m чисел.

Свойство 1. *Если* $F(x)=F=\mathrm{const}$ *(оператор* F *от* x *не зависит — является постоянным), то* $F'\equiv O$, *где* O *— нулевой оператор.*

Доказательство этого свойства очевидно.

Свойство 2. *Производная непрерывного (т. е. ограниченного) линейного отображения* A *есть само это отображение:* $A'(x)=A$.

Действительно,

$$A(x+h)-A(x)=A(x)+A(h)-A(x)=A(h).$$

Поэтому

$$A'(x)=A.$$

Свойство 3 (производная сложной функции). *Пусть* N_1, N_2, N_3 *— три нормированных пространства,* Σ_{x_0} *— окрестность точки* $x_0\in N_1$, F *— отображение этой окрестности в* N_2, $y_0=F(x_0)$, Σ_{y_0} *— окрестность точки* $y_0\in N_2$ *и* G *— отображение этой окрестности в* N_3. *Тогда если отображение* F *дифференцируемо в точке* x_0, *а* G *дифференцируемо в точке* y_0, *то отображение* $H=GF$ *(которое определено в некоторой окрестности точки* x_0 *и отображает ее в* N_3*) дифференцируемо в точке* x_0 *и*

$$H'(x_0)=G'(y_0)\cdot F'(x_0).$$

В самом деле, согласно условиям дифференцируемости отображений F и G

$$F(x_0+\xi)=F(x_0)+F'(x_0)\xi+\mathbf{o}_1(\xi)$$

и

$$G(y_0+\eta)=G(y_0)+G'(y_0)\eta+\mathbf{o}_2(\eta),$$

где $\dfrac{\|\mathbf{o}_1(\xi)\|}{\|\xi\|}$ — величина, стремящаяся к нулю при стремлении к нулю $\|\xi\|$; $\dfrac{\|\mathbf{o}_2(\eta)\|}{\|\eta\|}$ — величина, стремящаяся к нулю при стремлении к нулю $\|\eta\|$.

Операторы $F'(x_0)$ и $G'(y_0)$ — постоянные операторы, ограниченные по норме. Поэтому $\mathbf{o}(F'(x_0)\xi)=\mathbf{o}(\xi)$ и $G(y_0)\mathbf{o}(\xi)=\mathbf{o}(\xi)$. В самом деле, $\|\mathbf{o}(F'(x_0)\xi)\|\leqslant\varepsilon\|F'(x_0)\xi\|$, если $\|F'(x_0)\xi\|<\delta$, т. е. $\|\mathbf{o}(F'(x_0)\xi)\|\leqslant\varepsilon\|F'(x_0)\xi\|\leqslant\varepsilon\|F'(x_0)\|\|\xi\|=\varepsilon_1\|\xi\|$, где $\varepsilon_1=\varepsilon\|F'(x_0)\|$, $\|F'(x_0)\|$ — норма ограниченного оператора $F'(x_0)$ *. Таким образом, для всякого $\varepsilon_1>0$ существует $\delta_1>0$ (а именно $\delta_1=\delta/\|F'(x_0)\|$, $\|F'(x_0)\|\neq 0$) такое, что $\|\mathbf{o}(F'(x_0)\xi)\|\leqslant\varepsilon_1\|\xi\|$, если $\|\xi\|\leqslant\delta_1=\delta/\|F'(x_0)\|$, т. е. если $\|F'(x_0)\|\|\xi\|<\delta$ (поскольку, если выполнено это неравенство, то тем более $\|F'(x_0)\xi\|\leqslant\|F'(x_0)\|\|\xi\|<\delta$.

* Мы воспользовались оценкой $\|Ax\|\leqslant\|A\|\|x\|$, справедливой для любого ограниченного оператора (см. дополнение 2).

Следовательно, $\mathbf{o}(F'(x_0)\xi) = \mathbf{o}(\xi)$ *. Соотношение $G(y_0)\mathbf{o}(\xi) = \mathbf{o}(\xi)$ доказывается еще проще.

Учитывая доказанные соотношения, получим

$$H(x_0+\xi) = GF(x_0+\xi) = G[F(x_0) + F'(x_0)\xi + \mathbf{o}_1(\xi)] =$$
$$= G[y_0 + F'(x_0)\xi + \mathbf{o}_1(\xi)] = G(y_0+\eta),$$

где $\eta = F'(x_0)\xi + \mathbf{o}_1(\xi)$. Далее,

$$G(y_0+\eta) = G(y_0) + G'(y_0)\eta + \mathbf{o}_2(\eta) =$$
$$= G(y_0) + G'(y_0)[F'(x_0)\xi + \mathbf{o}_1(\xi)] + \mathbf{o}_2(F'(x_0)\xi + \mathbf{o}_1(\xi)) =$$
$$= G(y_0) + G'(y_0)F'(x_0)\xi + G'(y_0)\mathbf{o}_1(\xi) + \mathbf{o}_2(F'(x_0)\xi) + \mathbf{o}_1(\xi) =$$
$$= G(y_0) + G'(y_0)F'(x_0)\xi + \mathbf{o}_3(\xi).$$

Следовательно,

$$H(x_0+\xi) - G(y_0) = H(x_0+\xi) - G(F(x_0)) = H(x_0+\xi) - H(x_0) =$$
$$= G'(y_0)F'(x_0)\xi + \mathbf{o}_3(\xi),$$

и формула для производной сложной функции полностью доказана.

Если F, G, H — числовые функции, то это формула превращается в известное нам правило дифференцирования сложной функции.

Для отображений F, G, H это правило называется еще правилом вычисления производной композиции отображений.

Свойство 4. *Пусть F и G — два непрерывных отображения, действующих из N_1 в N_2. Если F и G дифференцируемы в точке x_0, то и отображения $F+G$ и aF, где a — число, тоже дифференцируемы в этой точке, причем*

$$(F+G)'(x_0) = F'(x_0) + G'(x_0),$$
$$(aF)'(x_0) = aF'(x_0) \text{ **}.$$

В самом деле, из определения суммы операторов и произведения оператора на число *** получаем, что

$$(F+G)(x_0+h) = F(x_0+h) + G(x_0+h) =$$
$$= F(x_0) + G(x_0) + F'(x_0)h + G'(x_0)h + \mathbf{o}_1(h);$$
$$(aF)(x_0+h) = aF(x_0) + aF'(x_0)h + \mathbf{o}_2(h).$$

Из этих равенств и получаем требуемые соотношения.

Рассмотрим теперь еще одно понятие, связанное с дифференцируемостью отображения.

---

* Если $F'(x_0) = 0$, то $\|F'(x_0)\| = 0$ и соотношение $\mathbf{o}(F'(x_0)\xi) = \mathbf{o}(\xi)$ очевидно.

** Символами $(F+G)'(x_0)$ и $(aF)'(x_0)$ обозначены значения операторов $(F+G)'$ и $(aF)'$ в точке x_0.

*** Для нелинейных операторов эти определения такие же, как и в случае линейных операторов (см. дополнение 1).

Определение. *Слабым дифференциалом (или дифференциалом Гато) отображения $F: N_1 \to N_2$ называется предел*

$$\lim_{t\to 0}\frac{F(x+th)-F(x)}{t}=\frac{d}{dt}F(x+th)\Big|_{t=0},$$

где сходимость понимается по норме в пространстве N_2.

Слабый дифференциал называется еще первой вариацией отображения F в точке x и обозначается символом $DF(x, h)$.

Слабый дифференциал $DF(x, h)$ может и не быть линеен по h. *Если же оператор $DF(x, h)$ оказывается линейным, т. е. если*

$$DF(x, h)=F_c'(x)h,$$

где $F_c'(x)$ — ограниченный линейный оператор, то этот оператор называется слабой производной (или производной Гато).

2. Формула Лагранжа конечных приращений. При изучении дифференцируемых числовых функций важную роль играла формула Лагранжа конечных приращений. Выведем такую формулу в случае дифференцируемого отображения $F: N_1 \to N_2$. Пусть Σ — открытое множество в N_1, содержащее точку x_0, и пусть множество точек $x_0+t(x-x_0)$* целиком содержится в Σ, $0 \leqslant t \leqslant 1$.

Теорема. *Пусть $F: N_1 \to N_2$ — отображение, определенное на открытом множестве $\Sigma \subset N_1$, непрерывное на $[x_0, x]$ и имеющее сильную производную F' в каждой точке интервала (x_0, x). Тогда*

$$\|F(x)-F(x_0)\| \leqslant \sup_{\xi\in(x_0,x)}\|F'(\xi)\|\cdot\|\Delta x\| =$$

$$=\sup_{0<\theta<1}\|F'(x_0+\theta\Delta x)\|\cdot\|\Delta x\|, \quad \Delta x=x-x_0.$$

Доказательство. Рассмотрим сначала непрерывное отображение $f(t)$, заданное на сегменте $[0, 1]$ и отображающее его в пространство N_2, $f: [0, 1] \to N_2$ и непрерывную функцию $g(t)$, заданную на сегменте $[0, 1]$, принимающую числовые значения. Покажем, что *если $f(t)$ и $g(t)$ сильно дифференцируемы на интервале* $(0, 1)$ ($f(t)$ как отображение одного нормированного пространства, а именно интервала $(0, 1)$, в другое нормированное пространство N_2, а $g(t)$ дифференцируе-

---

* Множество точек вида $x_0+t(x-x_0)$, где $0 \leqslant t \leqslant 1$, называется отрезком или сегментом и обозначается символом $[x_0, x]$ в N_1. При $t=0$ получаем точку x_0, а при $t=1$ — точку x; если $0<t<1$, то множество точек $x_0+t(x-x_0)$ называется интервалом (x_0, x).

ма как числовая функция, определенная на интервале $(0, 1)$), *и если*

$$\|f(t)\| \leqslant g'(t), \quad 0<t<1,$$

то

$$\|f(1)-f(0)\| \leqslant g(1)-g(0). \tag{*}$$

Из этой формулы утверждение теоремы уже будет следовать легко.

Докажем эту формулу. При каждом фиксированном $\varepsilon>0$ обозначим через A_ε множество точек сегмента $[0, 1]$, в которых выполнено неравенство

$$\|f(t)-f(0)\| \leqslant g(t)-g(0)+\varepsilon t+\varepsilon. \tag{**}$$

Покажем, что при любом $\varepsilon>0$ число 1 принадлежит множеству A_ε. Тогда при $t=1$ и $\varepsilon \to 0$ получим доказываемую формулу.

В силу непрерывности отображения $f(t)$ и функции $g(t)$ множество A_ε замкнуто на сегменте $[0, 1]$. Действительно, если $t_n \in A_\varepsilon$ и $t_n \to t_0$, то в неравенстве

$$\|f(t_n)-f(0)\| \leqslant g(t_n)-g(0)+\varepsilon t_n+\varepsilon,$$

пользуясь непрерывностью функции нормы (или, что то же самое, непрерывностью функции расстояния): $\|x\|=\rho(x, 0)$, непрерывностью отображения $f(t)$ и непрерывностью функции $g(t)$, можно перейти к пределу при $t_n \to t_0$ и заключить, что

$$\|f(t_0)-f(0)\| \leqslant g(t_0)-g(0)+\varepsilon t_0+\varepsilon,$$

т. е. $t_0 \in A_\varepsilon$, а следовательно, множество A_ε замкнуто.

Множество A_ε непусто, поскольку (**), очевидно, выполнено при достаточно малых t (левая часть неравенства (**) при малых t в силу непрерывности $f(t)$ мала, а правая имеет положительный член ε в качестве слагаемого, который не зависит от t).

Пусть $a=\sup A_\varepsilon$. Поскольку множество $A_\varepsilon \subset [0, 1]$ замкнуто, то $a \in A_\varepsilon$. Покажем, что a не может быть меньше числа 1. Допустим противное, что $a<1$, т. е. $0<a<1$. В силу дифференцируемости $f(t)$ и $g(t)$ в точке a для достаточно малых положительных чисел h будем иметь

$$\|f(a+h)-f(a)\| = \|f'(a)h+o(h)\| \leqslant \|f'(a)h\| + \|o(h)\| \leqslant$$
$$\leqslant \|f'(a)\| h + \frac{\varepsilon}{2} h$$

(в данном случае $\|h\|=|h|=h$, поскольку h — положительное число, а норма элемента h сегмента $[0, 1]$ совпадает с модулем числа h).

Далее,

$$g(a+h)-g(a) = |g(a+h)-g(a)|, \quad h>0,$$

поскольку функция $g(t)$ на интервале $(0, 1)$ не убывает. Это следует из того, что $g'(t)\geqslant 0$, $0<t<1$, что, в свою очередь, следует из неравенства $0\leqslant\|f'(t)\|\leqslant g'(t)$, $0<t<1$.

Заметим теперь, что

$$|g(a+h)-g(a)| = |g'(a)h+o(h)|\geqslant$$
$$\geqslant |g'(a)h|-|o(h)|\geqslant g'(a)h-\frac{\varepsilon}{2}h$$

для достаточно малых h. Поэтому

$$\|f(a+h)-f(a)\|\leqslant\|f'(a)\|h+\frac{\varepsilon}{2}h\leqslant g'(a)h+\frac{\varepsilon}{2}h\leqslant$$
$$\leqslant g(a+h)-g(a)+\varepsilon h.$$

Поскольку

$$\|f(a)-f(0)\|\leqslant g(a)-g(0)+\varepsilon a+\varepsilon,$$

то

$$\|f(a+h)-f(0)\|=\|f(a+h)-f(a)+f(a)-f(0)\|\leqslant$$
$$\leqslant\|f(a+h)-f(a)\|+\|f(a)-f(0)\|\leqslant$$
$$\leqslant g(a+h)-g(a)+\varepsilon h+g(a)-g(0)+\varepsilon a+\varepsilon=$$
$$=g(a+h)-g(0)+\varepsilon(a+h)+\varepsilon.$$

Мы видим, что $a+h\in A_\varepsilon$, а это противоречит выбору числа a, т. е. число a не может быть меньшим единицы. Следовательно, $a=1$, и $1\in A_\varepsilon$. Формула $(*)$ доказана.

Завершим теперь доказательство теоремы. Положим

$$f(t)=F(x_0+t\Delta x),\quad g(t)=M\|\Delta x\|t,\quad 0<t<1,$$

где

$$M=\sup_{\xi\in(0,1)}\|f'(\xi)\|.$$

Отображение $f(t)$ и функция $g(t)$ удовлетворяют, очевидно, всем условиям вспомогательного утверждения, установленного нами выше. Поэтому, подставляя эти выражения для $f(t)$ и $g(t)$ в формулу $(*)$, получаем формулу Лагранжа. Теорема доказана.

Следствие. *Если $A\in(N_1\to N_2)$ — линейное непрерывное отображение нормированного пространства N_1 в нормированное пространство N_2, не зависящее от x, а $F(x)$ — отображение открытого множества $\Sigma\subset N_1$ в N_2, удовлетворяющее условиям теоремы, то*

$$\|F(x)-F(x_0)-A\Delta x\|\leqslant\sup_{\xi\in(x_0,x)}\|F'(\xi)-A\|\,\|\Delta x\|.$$

Действительно, применив теорему к отображению

$$F(x)-A\Delta x,\quad \Delta x=x-x_0,$$

получим

$$\|F(x)-A\,\Delta x-F(x_0)\|\leqslant \sup_{\xi\in(x_0,x)} \|F'(\xi)-A(x-x_0)'|_{x=\xi}\|\|\Delta x\| =$$

$$= \sup_{\xi\in(x_0x)} \|F'(\xi)-A\|\,\|\Delta x\|,$$

что и требовалось доказать.

Можно показать, что *теорема и следствие из нее верны, если всюду вместо сильной производной $F'(x)$ рассматривать слабую производную $F_c'(x)$.*

3. Связь между слабой и сильной дифференцируемостью. Сильная и слабая дифференцируемость представляет собой различные понятия даже в случае m-мерного евклидова пространства при $m\geqslant 2$.

Действительно, рассмотрим, например, функцию двух переменных

$$f(x)=f(x_1,\ x_2)=\begin{cases}\dfrac{x_1^3x_2}{x_1^4+x_2^2}, & \text{если } x_1^2+x_2^2>0,\\ 0, & \text{если } x_1=0 \text{ и } x_2=0,\end{cases}$$

$x=(x_1,\ x_2)$ — точка плоскости.

Легко проверить, что эта функция непрерывна всюду на плоскости, включая точку $(0,\ 0)$. В точке $(0,\ 0)$ имеем

$$\lim_{t\to 0}\frac{f(0+th)-f(0)}{t}=\lim_{t\to 0}\frac{t^4h_1^3h_2}{t^4h_1^4+t^2h_2^2}=\lim_{t\to 0}\frac{t^2h_1h_2}{h_2^2+t^2h_1^4},$$

где $h=(h_1,\ h_2)$ — точка плоскости, представляющая собой некоторое приращение аргумента x функции $f(x)$ в точке $(0,\ 0)$.

Таким образом, мы видим, что в точке $(0,\ 0)$ слабый дифференциал $f(x)$ существует и равен нулю.

С другой стороны, в начальной точке имеем $f'_{x_1}(0,\ 0)=$ $=f'_{x_2}(0,\ 0)=0$; это непосредственно вытекает из самого определения частных производных и из того, что $f(x_1,\ 0)=0$, $f(0,\ x_2)=0$. Поэтому, записывая приращение функции в точке $(0,\ 0)$ в виде

$$\Delta f(h_1,\ h_2)=f(h_1,\ h_2)-f(0,\ 0)=\frac{\partial f(0,\ 0)}{\partial x_1}h_1+\frac{\partial f(0,\ 0)}{\partial x_2}h_2+o(h),$$

где $h=(h_1,\ h_2)$ — приращение аргумента, $\dfrac{\|o(h)\|}{\|h\|}\to 0$ при $\|h\|=(h_1^2+h_2^2)^{1/2}\to 0$, получим, что

$$\Delta f=o(h)=\frac{h_1^3h_2}{h_1^4+h_2^2}.$$

Выбирая $h_2=h_1^2$, имеем $\mathbf{o}(h)|_{h_2=h_1^2}=h_1/2$, что приводит к противоречию. Действительно, величина $|h_1|/2$, поделенная на норму h, при $h_2=h_1^2$ стремится к 1/2 при $\|h\|\to 0$, а величина $\mathbf{o}(h)$, поделенная на $\|h\|$, при любых h_1 и h_2 стремится к нулю при $\|h\|\to 0$, что следует из определения $\mathbf{o}(h)$.

Следовательно, функция $f(x)$ не дифференцируема в точке (0, 0) в сильном смысле.

Однако *если отображение F имеет сильную производную, то оно имеет и слабую производную, причем сильная и слабая производные совпадают.* Действительно, для сильно дифференцируемого отображения имеем

$$F(x+th)-F(x)=F'(x)(th)+\mathbf{o}(th)=tF'(x)h+\mathbf{o}(th),$$

$$\lim_{t\to 0}\frac{F(x+th)-F(x)}{t}=\lim_{t\to 0}\left[F'(x)h+\frac{\mathbf{o}(th)}{t}\right]=F'(x)h,$$

что и трсбовалось.

Выясним условия, при которых из слабой дифференцируемости отображения F следует его сильная дифференцируемость.

Докажем следующую теорему.

Теорема. *Если слабая производная $F_c'(x)$ отображения $F:N_1\to N_2$ существует в некоторой окрестности Σ_{x_0} точки x_0 и представляет собой в этой окрестности функцию от x, непрерывную в точке x_0*, то в точке x_0 сильная производная $F'(x_0)$ существует и совпадает со слабой.*

Доказательство. По условию отображение F имеет слабую производную $F_c'(x_0)$ в точке x_0. Выбрав h так, что $x_0+h\in\Sigma_{x_0}$, рассмотрим выражение

$$r(x_0,h)=F(x_0+h)-F(x_0)-F_c'(x_0)\,h.$$

Элемент $r(x_0,h)$ принадлежит пространству N_2. Пусть φ — произвольный линейный ограниченный функционал на пространстве N_2. Позже мы укажем некоторые условия на его выбор. Тогда из формулы для $r(x_0,h)$ получим

$$\varphi(r(x_0,h))=\varphi(F(x_0+h)-F(x_0))-\varphi(F_c'(x_0)h).$$

Рассмотрим функцию $f(t)=\varphi(F(x_0+th)-F(x_0))$ числового аргумента t, $0\leqslant t\leqslant 1$. Эта функция дифференцируема по t и для нее

$$\frac{df}{dt}=\lim_{\Delta t\to 0}\varphi\left(\frac{F(x_0+th+\Delta th)-F(x_0+th)}{\Delta t}\right)=\varphi\left(F_c'(x_0+th)\,h\right).$$

* $F_c'(x)$ отображает множество Σ_{x_0} пространства N_1 в некоторое множество пространства операторов $(N_1\to N_2)$, т. е. $F_c'(x)$ — операторная функция. Понятие непрерывной функции было нами определено даже для отображений одного метрического пространства в другое (см. дополнение 2).

(Здесь мы воспользовались непрерывностью функционала φ, поменяв местами символы lim и φ, а также слабой дифференцируемостью $F(x)$ в окрестности Σ_{x_0}.)

Применив к функции $f(t)$ формулу конечных приращений на сегменте $[0, 1]$, получим

$$f(1)=f(0)+f'(\theta), \quad 0<\theta<1,$$

или

$$\varphi(F(x_0+h)-F(x_0))=\varphi(F_c'(x_0+\theta h)h), \quad 0<\theta<1.$$

Таким образом,

$$\varphi(r(x_0, h))=\varphi(F(x_0+h)-F(x_0))-\varphi(F_c'(x_0)h)=$$
$$=\varphi(F_c'(x_0+\theta h)h)-\varphi(F_c'(x_0)h)=$$
$$=\varphi(F_c'(x_0+\theta h)h-F_c'(x_0)h).$$

Отметим теперь следующий факт: запись $\psi(x)$, где $x \in N$ — некоторому нормированному пространству, а ψ — линейный непрерывный функционал на N, можно рассматривать с двух точек зрения. Во-первых, можно фиксировать функционал ψ и менять аргумент $x \in N$. В этом случае, например, как мы отмечали в дополнении 2, $\|\psi\| = \sup\limits_{\|x\|=1} |\psi(x)|$. Во-вторых, можно фиксировать элемент $x \in N$, а менять функционалы ψ. Линейные непрерывные функционалы ψ отображают пространство N в пространство P действительных (или комплексных) чисел и принадлежат тоже нормированному (даже банаховому в силу полноты P, см. дополнение 2) пространству $(N \to P)$. В этом случае, например, $\|x\| = \sup\limits_{\|\psi\|=1} |\psi(x)|$, где верхняя грань уже берется по всем функционалам $\psi \in (N \to P)$; $\|\psi\|=1$, т. е. по функционалам, имеющим нормы, равные единице. Поэтому, по определению верхней грани, существует такой функционал ψ, $\|\psi\|=1$, что при фиксированном $x \in N$ $|\psi(x)| \geqslant \frac{1}{2}\|x\|$. Воспользуемся этим фактом и выберем функционал φ, $\|\varphi\|=1$, таким, что

$$|\varphi(r(x_0, h))| \geqslant \frac{1}{2}\|r(x_0, h)\|,$$

где h, а следовательно, $r(x_0, h)$ фиксированы.

Отсюда и из равенства для $\varphi(r(x_0, h))$ получаем, что

$$\|r(x_0, h)\| \leqslant 2\|\varphi(F_c'(x_0+\theta h)h-F_c'(x_0, h)\| \leqslant$$
$$\leqslant 2\|F_c'(x_0+\theta h)h-F_c'(x_0)h\|\|\varphi\| \leqslant$$
$$\leqslant 2\|F_c'(x_0+\theta h)-F_c'(x_0)\|\|h\|.$$

Но по условию $F_c'(x)$ есть непрерывная в точке x_0 функция от x; поэтому

$$\lim_{h\to 0}\|F_c'(x_0+\theta h)-F_c'(x_0)\|=0,$$

так что $\|(x_0, h)\|$ есть величина выше первого порядка малости относительно $\|h\|$, т. е. $F_c'(x_0)h$, как это следует из формулы $r(x_0, h)=F(x_0+h)-F(x_0)-F_c'(x_0)h$ есть главная, линейная по h, часть разности $F(x_0+h)-F(x_0)$. Тем самым доказано существование сильной производной $F'(x_0)$ и ее совпадение со слабой производной $F_c'(x_0)$.

4. Дифференцируемость функционалов. В предыдущих пунктах нами было введено понятие дифференцируемого отображения $F:N_1\to N_2$, отображающего нормированное пространство N_1 в нормированное пространство N_2. Мы уже отмечали, что производная $F'(x)$ такого отображения представляет собой при каждом x линейный оператор, действующий из N_1 в N_2, т. е. элемент пространства операторов $(N_1\to N_2)$. В частности, если $N_2=P$, где P — числовая прямая или комплексная плоскость, то отображение F принимает числовые значения на N_1 и называется функционалом. При этом производная функционала F в точке x_0 есть линейный функционал, зависящий от x_0.

Для примера найдем главную линейную часть приращения функционала $F(x)=\|x\|^2$, заданного в действительном гильбертовом пространстве H. Имеем

$$F(x+h)-F(x)=\|x+h\|^2-\|x\|^2=(x+h,\ x+h)-(x,\ x)=$$
$$=(x,\ x)+2(x,\ h)-(x,\ x)+(h,\ h)=2(x,\ h)+(h,\ h).$$

Можно убедиться, что $F'(x)=2x$. (Для этого следует заметить, что всякий функционал в H имеет вид скалярного произведения.)

5. Интеграл от абстрактных функций. В этом пункте будет изложен материал, являющийся, с одной стороны, дополнением к изложенному в гл. 9 материалу об определенном интеграле, а с другой стороны, являющийся необходимым в теории дифференцирования в банаховых пространствах.

Предположим, что отображение F действует из банахова пространства B_1 в другое банахово пространство B_2, причем пространство B_1 есть числовая ось E^1. Таким образом, $F:E^1\to B_2$.

Отображение F, сопоставляющее числу x элемент банахова пространства B_2, назовем ***абстрактной функцией*** *на числовой оси.* Производная $F'(x)$ абстрактной функции при условии, что она существует, представляет собой при каждом x элемент пространства B_2 — касательный вектор к кривой $F(x)$. *Для абстрактной функции*, представляющей собой функцию числового аргумента x, очевидно, *слабая дифферен-*

цируемость совпадает с сильной. В этом случае, используя соотношение

$$\|F(x+th)-F(x)-F_c'(x)ht\|=o(|t|), \quad h\neq 0, \quad t\to 0,$$

вытекающее из слабой дифференцируемости, и полагая в нем $th=h_1$, получим, что

$$\|F(x+h_1)-F(x)-F_c'(x)h_1\|=o\left(\frac{|h_1|}{|h|}\right)=o(|h_1|),$$

где число $h_1\to 0$, т. е. $F(x)$ сильно дифференцируемо и сильная производная совпадает со слабой.

Построим интеграл от абстрактной функции $F(x)$, определенной на сегменте $[a, b]$. Пусть $\{x_k\}$ — разбиение сегмента $[a, b]$. Введем интегральную сумму

$$\sigma=\sum_{k=1}^{n} F(\xi_k)\,\Delta x_k, \quad \xi_k\in[x_{k-1}, x_k].$$

Пусть $d=\max\limits_{1\leqslant k\leqslant n}\Delta x_k$ — диаметр разбиения.

Абстрактную функцию $F(x)$ назовем ***интегрируемой*** *на сегменте $[a, b]$, если для этой функции на указанном сегменте существует предел I ее интегральных сумм при стремлении диаметра d разбиений $\{x_k\}$ к нулю, причем этот предел берется по норме пространства B_2.* Таким образом, абстрактная функция $F(x)$ интегрируема, если

$$\lim_{d\to 0}\left\|\sum_{k=1}^{n} F(\xi_k)\,\Delta x_k-I\right\|_{B_2}=0.$$

Предел I называется ***интегралом*** *от абстрактной функции по сегменту $[a, b]$ и обозначается символом*

$$I=\int_a^b F(x)\,dx.$$

Очевидно, что I является элементом B_2, поскольку σ является элементом B_2 и пространство B_2 полное, т. е. и предел σ при $d\to 0$ принадлежит B_2.

Следует отметить, что построение теории интеграла от абстрактной функции мало чем отличается от построения теории интеграла от числовой функции. Подчеркнем, что интеграл от абстрактной функции не является уже числом, как обычный интеграл, поэтому, например, всюду в доказательствах модуль интеграла от числовой функции надо заменять на норму интеграла от абстрактной функции и т. п.

Отметим следующие простые свойства интеграла от абстрактных функций. Их доказательства аналогичны доказательствам, приведенным в гл. 9 при построении интеграла Римана.

1. *Интеграл от абстрактной непрерывной функции $F(x)$ существует, т. е. такая функция интегрируема.*

2. *Если C — линейное отображение пространства B_2 в банахово пространство B_3 (предполагается, что C также и непрерывно)*, то

$$\int_a^b CF(x)\,dx = C\int_a^b F(x)\,dx.$$

3. *Справедливо неравенство*

$$\left\|\int_a^b F(x)\,dx\right\| \leqslant \int_a^b \|F(x)\|\,dx,$$

где справа стоит обычный интеграл Римана от числовой функции.

4. *Если $F(x)$ имеет вид $f(x)y_0$, где $f(x)$ — числовая функция, а y_0 — фиксированный элемент из B_2, то*

$$\int_a^b F(x)\,dx = y_0\int_a^b f(x)\,dx,$$

где справа также стоит интеграл Римана от функции $f(x)$ по сегменту $[a, b]$.

6. Формула Ньютона — Лейбница для абстрактных функций. В проведенных выше рассмотрениях через $(N_1 \to N_2)$ мы обозначали пространство всех линейных ограниченных (непрерывных) отображений из линейного нормированного пространства N_1 в линейное нормированное пространство N_2. Топологию* в пространстве $(N_1 \to N_2)$ задает, например, норма в этом пространстве. Заметим, что для задания топологии (системы окрестностей каждой точки) в линейном пространстве $(N_1 \to N_2)$ достаточно задать систему окрестностей начала координат, в силу линейности пространства тем самым будут заданы окрестности каждой точки.

Рассмотрим теперь *линейное пространство* $(N_1 \to N_2)_1$ *всех непрерывных ограниченных, быть может, нелинейных, отображений пространства N_1 в N_2.*

* Напомним, что для задания топологии на некотором множестве необходимо указать систему множеств, удовлетворяющих аксиомам 1, 2 (см. дополнение 2, раздел топологических пространств). Множества, удовлетворяющие этим аксиомам, называются открытыми. Открытые множества метрического или нормированного пространства удовлетворяют упомянутым аксиомам.

Нелинейное отображение $F: N_1 \to N_2$ называется ограниченным, *если для всякого ограниченного множества $A \subset N_1$ соответствующее множество $F(A)$ ограничено в N_2* (множество в нормированном пространстве ограничено, если его можно поместить внутрь некоторого шара с центром в начале координат). *Нелинейное непрерывное отображение* не обязано *быть ограниченным.*

Очевидно, что пространство $(N_1 \to N_2)$ является подпространством $(N_1 \to N_2)_1$. В линейном пространстве $(N_1 \to N_2)_1$ зададим систему окрестностей нуля (для любого натурального n и произвольного $\varepsilon > 0$):

$$\Sigma_{n,\varepsilon} = \{F : \sup_{\|x\| \leqslant n} \|F(x)\| < \varepsilon\}$$

и тем самым зададим топологию в пространстве $(N_1 \to N_2)_1$. Читателю предлагается проверить, что данной системой окрестностей в пространстве $(N_1 \to N_2)_1$ задана топология. На подпространстве $(N_1 \to N_2) \subset (N_1 \to N_2)_1$ эта топология совпадает с обычной топологией, задаваемой операторной нормой.

Рассмотрим сегмент $[x_0, x_0 + \Delta x]$ в N_1. Пусть задано непрерывное отображение $F(x)$ этого сегмента в пространство $(N_1 \to N_2)_1$, т. е. каждой точке $x \in [x_0, x_0 + \Delta x]$ сопоставлено некоторое отображение $F(x) \in (N_1 \to N_2)_1$, непрерывно зависящее от точки $x \in N_1$.

Определим интеграл *от $F(x)$* по сегменту $[x_0, x_0 + \Delta x]$, *по определению полагая*

$$\int_{x_0}^{x_0 + \Delta x} F(x)\, dx = \int_0^1 F(x_0 + t\,\Delta x)(\Delta x)\, dt.$$

Здесь оператор $F(x_0 + t\Delta x)$ применяется к элементу $\Delta x \in N_1$. При каждом $t \in [0, 1]$ величина $F(x_0 + t\Delta x](\Delta x)$ представляет собой элемент пространства N_2 — образ элемента Δx при отображении $F(x_0 + t\Delta x)$.

Согласно построениям п. 5 интеграл, стоящий в правой части формулы, существует и является элементом пространства N_2.

Используя эти замечания, докажем *формулу Ньютона — Лейбница для интеграла от абстрактной функции.*

Рассмотрим отображение F, которое действует из N_1 в N_2 и имеет на сегменте $[x_0, x_0 + \Delta x]$ сильную производную $F'(x)$, непрерывную по x.

Напомним, что отображение $F(x): N_1 \to N_2$ непрерывно в точке $x \in N_1$, если для любого шара с центром в точке $F(x)$, лежащего в N_2, найдется такой шар с центром в точке x, лежащий в N_1, образ которого при отображении $F(x)$ целиком содержится в указанном шаре с центром в точке $F(x)$.

Отображение $F(x)$ непрерывно на сегменте $[x_0, x_0+\Delta x]$ (т. е. на множестве точек вида $\{x_0+t\Delta x\}$, где $0 \leqslant t \leqslant 1$), если оно непрерывно в любой внутренней точке сегмента (т. е. когда точка отвечает параметру t такому, что $0<t<1$) и, кроме того, непрерывно в точке x_0 справа и в точке $x_0+\Delta x$ слева *.

Поскольку отображение $F'(x)$ непрерывно на сегменте $[x_0, x_0+\Delta x]$, то существует интеграл $\int\limits_{x_0}^{x_0+\Delta x} F'(x)\,dx$.

Докажем, что *имеет место равенство*

$$\int\limits_{x_0}^{x_0+\Delta x} F'(x)\,dx = F(x_0+\Delta x) - F(x_0),$$

которое и называется формулой Ньютона — Лейбница для абстрактных функций.

По определению интеграла можем записать

$$\int\limits_{x_0}^{x_0+\Delta x} F'(x)\,dx = \int\limits_0^1 F'(x_0+t\,\Delta x)(\Delta x)\,dt =$$

$$= \lim_{d\to 0} \sum_{k=0}^{n-1} F'(x_0+t_k\,\Delta x)\,\Delta x\,(t_{k+1}-t_k) = \lim_{d\to 0} \sum_{k=0}^{n-1} F'(x_k)(\Delta x_k),$$

где $x_k = x_0 + t_k\Delta x_k$, $\Delta x_k = (t_{k+1}-t_k)\Delta x$, d — диаметр разбиения сегмента $[0, 1]$.

С другой стороны, легко заметить, что

$$F(x_0+\Delta x) - F(x_0) = \sum_{k=0}^{n-1} [F(x_0+t_{k+1}\,\Delta x) - F(x_0+t_k\,\Delta x)] =$$

$$= \sum_{k=0}^{n-1} [F(x_{k+1}) - F(x_k)].$$

Рассмотрим разность $F(x_{k+1}) - F(x_k) - F'(x_k)\Delta x_k$ и применим к ней формулу из следствия к теореме п. 2 Лагранжа. При $A = F'(x_k)$ получим

$$\|F(x_{k+1}) - F(x_k) - F'(x_k)\,\Delta x_k\| \leqslant$$

$$\leqslant \sup_{0<\theta_k<1} \|F'(x_k+\theta_k\,\Delta x_k) - F'(x_k)\|\;\|\Delta x_k\|.$$

---

* Говорят, что точка $x \in [x_0, x_0+\Delta x]$ стремится к точке x_0 справа, если $x \in \{x_0+t\Delta x\}$ и $t \to 0+0$. Аналогично определяется стремление x к точке $x_0+\Delta x$ слева. Отображение $F(x)$, определенное на сегменте $[x_0, x_0+\Delta x]$, непрерывно в точке x_0 справа, если предельное значение $F(x)$ при стремлении x к точке x_0 справа равно $F(x_0)$. Аналогично определяется непрерывность $F(x)$ в точке $x_0+\Delta x$ слева.

Просуммируем эти неравенства по всем k от 0 до $n-1$ и вместо $\|\Delta x_k\|$ поставим $\|\Delta x\|\,|t_{k+1}-t_k|$. Получим

$$\left\|\sum_{k=1}^{n-1}[F(x_{k+1})-F(x_k)-F'(x_k)\Delta x_k]\right\|\leqslant$$

$$\leqslant\|\Delta x\|\sum_{k=1}^{n-1}(t_{k+1}-t_k)\sup_{0<\theta_k<1}\|F'(x_k+\theta_k\Delta x_k)-F'(x_k)\|.$$

Так как производная $F'(x)$ непрерывна на сегменте $[x_0,\ x_0+\Delta x]$, то она является и равномерно непрерывной* на этом сегменте. Поэтому правая часть неравенства, написанного выше, стремится к нулю при неограниченном измельчении разбиения сегмента $[x_0,\ x_0+\Delta x]$, откуда и вытекает, что

$$F(x_0+\Delta x)-F(x_0)=\sum_{k=0}^{n-1}[F(x_{k+1})-F(x_k)]=$$

$$=\sum_{k=0}^{n-1}[F(x_{k+1})-F(x_k)-F'(x_k)\Delta x_k]+\sum_{k=0}^{n-1}F'(x_k)\Delta x_k\to$$

$$\to\int\limits_{x_0}^{x_0+\Delta x}F'(x)\,dx,$$

если $d\to 0$. Формула Ньютона — Лейбница доказана.

7. Производные второго порядка. Рассмотрим дифференцируемое отображение $F:N_1\to N_2$. Мы уже отмечали, что производная $F'(x)$ этого отображения при каждом фиксированном $x\in N_1$ есть элемент пространства $(N_1\to N_2)$ — пространства линейных ограниченных операторов на N_1, действующих в N_2. Допустим, что отображение $F'(x)$, в свою очередь, дифференцируемо. *Производная отображения $F'(x)$ называется второй производной отображения F и обозначается символом F''.* При каждом фиксированном x отображение $F''(x)$ является, очевидно, элементом пространства $(N_1\to(N_1\to N_2))$ — пространства линейных ограниченных операторов, действующих из N_1 в $(N_1\to N_2)$.

Элементы этого пространства допускают и другую интерпретацию в виде так называемых билинейных отображений.

* Абстрактная функция $F(x):N_1\to N_2$ называется равномерно непрерывной на множестве $\{x\}=M\subset N_1$, если для всякого $\varepsilon>0$ найдется такое $\delta>0$, что $\|F(x_1)-F(x_2)\|<\varepsilon$ для любых x_1 и x_2 из множества M, если только $\|x_1-x_2\|<\delta$. Доказательство того факта, что непрерывная абстрактная функция на компакте в N_1 является равномерно непрерывной, аналогично доказательству этого факта для числовых функций на компакте (см. гл. 4, § 6, п. 3, теорему 4.16).

Определение. *Отображение $B:(N_1 \to N_2)$, отображающее нормированное пространство N_1 в нормированное пространство N_2, называется билинейным, если каждой упорядоченной паре элементов (x_1, y_1)* *) *из N_1 поставлен в соответствие элемент $B(x_1, y_1)$ из N_2 так, что выполнены следующие свойства:*

1. *Для любых x_1, y_1, x_2, y_2 из N_1 и любых чисел α и β имеют место равенства*

$$B(\alpha x_1+\beta x_2, y_1)=\alpha B(x_1, y_1)+\beta B(x_2, y_1),$$
$$B(x_1, \alpha y_1+\beta y_2)=\alpha B(x_1, y_1)+\beta B(x_1, y_2).$$

2. *Существует такое положительное число M, что*

$$\|B(x, y)\| \leqslant M\|x\|\|y\|$$

при всех x, y, из N_1.

Первое из этих условий означает, что билинейное отображение B линейно по каждому из двух своих аргументов. Второе условие означает ограниченность билинейного отображения B. Так же, как и в случае линейного отображения (см. дополнение 1), показывается, что *ограниченное билинейное отображение является непрерывным уже по совокупности своих аргументов.*

Определение. *Наименьшая из постоянных M, удовлетворяющих условию $\|B(x, y)\| \leqslant M\|x\|\|y\|$, называется нормой билинейного отображения B и обозначается символом $\|B\|$.*

Точно так же, как и над линейными операциями (операторами) (см. дополнение 1), над билинейными операциями можно определить операцию сложения двух билинейных отображений, по определению положив $(B_1+B_2)(x, y)=B_1(x, y)+B_2(x, y)$, а также операцию умножения отображения B на число α: $(\alpha B)(x, y)=\alpha B(x, y)$. Поскольку для билинейной операции B определена и его норма, то *билинейные отображения пространства N_1 в пространство N_2 сами образуют линейное пространство. Обозначим его $(N_1^2 \to N_2)$* (здесь индекс 2 над символом N_1 означает, что из пространства N_1 берется два элемента (x, y), которые переводятся отображением B в один элемент пространства N_2).

Так же, как и в случае линейных отображений, показывается (см. дополнение 1), что *если пространство N_2 банахово (полное), то и пространство $(N_1^2 \to N_2)$ будет банаховым.*

* Пара элементов x и y называется упорядоченной, если указано, какой из ее элементов является первым. Если элемент x является первым элементом пары, а элемент y — вторым, то упорядоченную пару записывают так: (x, y).

Докажем следующее

Утверждение. *Между элементами пространства $(N_1 \to (N_1 \to N_2))$ и пространства $(N_1^2 \to N_2)$ можно установить взаимно однозначное соответствие, сохраняющее линейные операции (т. е. если элементу $A_1 \in (N_1 \to (N_1 \to N_2))$, отвечает элемент $B_1 \in (N_1^2 \to N_2)$, а элементу $A_2 \in (N_1 \to (N_1 \to N_2))$ отвечает элемент $B_2 \in (N_1^2 \to N_2)$, то элементу $\alpha A_1 + \beta A_2$ отвечает элемент $\alpha B_1 + \beta B_2$, где α и β — любые числа).*

Действительно, каждому элементу $A \in (N_1 \to (N_1 \to N_2))$ поставим в соответствие элемент B из $(N_1^2 \to N_2)$ по правилу

$$B(x, y) = (Ax)y.$$

Это соответствие, очевидно, линейно.

Покажем, что это соответствие сохраняет и нормы элементов*, т. е. покажем, что $\|A\| = \|B\|$. Отсюда, в частности и будет следовать взаимная однозначность соответствия. Действительно, если бы два различных элемента при изометричном соответствии переходили бы в один элемент, то их разность соответствовала бы нулевому элементу. Норма этой разности равнялась бы норме нулевого элемента, т. е. нулю. Таким образом, элементы совпадали бы, что противоречит их выбору.

Итак, докажем, что если элементу $A \in (N_1 \to (N_1 \to N_2))$ отвечает элемент $B \in (N_1^2 \to N_2)$, то $\|A\| = \|B\|$, причем нормы элементов берутся в соответствующих пространствах.

Пусть $z = B(x, y) = (Ax)y$. Тогда

$$\|z\| \leqslant \|Ax\| \|y\| \leqslant \|A\| \|x\| \|y\|,$$

откуда

$$\|B\| \leqslant \|A\|.$$

С другой стороны, если задано билинейное отображение B, то при фиксированном $x \in N_1$ отображение $y \to (Ax)y = B(x, y)$ является линейным отображением пространства N_1 в N_2.

Таким образом, каждому $x \in N_1$ ставится в соответствие элемент Ax пространства $(N_1 \to N_2)$. Очевидно, что Ax линейно зависит от x, т. е. билинейное отображение B определяет некоторый элемент A пространства $(N_1 \to (N_1 \to N_2))$.

При этом отображение B восстанавливается по A при помощи формулы

$$B(x, y) = (Ax)y.$$

Далее,

$$\|Ax\| = \sup_{\|y\| \leqslant 1} \|(Ax)\, y\| = \sup_{\|y\| \leqslant 1} \|B(x, y)\| \leqslant \|B\| \, \|x\|.$$

---

* Соответствие между двумя упорядоченными пространствами, сохраняющее нормы отвечающих друг другу элементов, называется **изометричным**. Подчеркнем, что нормы берутся в соответствующих простраиствах.

Таким образом,

$$\|A\| \leqslant \|B\|.$$

Следовательно, из соотношений $\|B\| \leqslant \|A\|$ и $\|A\| \leqslant \|B\|$ получаем, что

$$\|A\| = \|B\|.$$

Нами, таким образом, доказано, что соответствие между пространствами $(N_1 \to (N_1 \to N_2))$ и $(N_1^2 \to N_2)$ линейно и изометрично. При этом образ пространства $(N_1 \to (N_1 \to N_2))$ при таком соответствии есть все пространство $(N_1^2 \to N_2)$. Утверждение доказано.

Из доказанного утверждения следует, что если берется какой-нибудь элемент пространства $(N_1 \to (N_1 \to N_2))$, то всегда можно указать его образ в пространстве $(N_1^2 \to N_2)$.

Более того, норма элемента в исходном пространстве и норма его образа совпадают. Сохраняются также линейные операции при построенном соответствии в пространствах $(N_1 \to \to (N_1 \to N_2))$ и $(N_1^2 \to N_2)$. Такие пространства можно не различать; называются они **изоморфными**. Эти пространства отличаются только тем, что их элементы имеют различную природу, а все остальные свойства пространств $(N_1 \to (N_1 \to \to N_2))$ и $(N_1^2 \to N_2)$ одинаковы.

В частности, мы выяснили, что вторая производная $F''(x)$ отображения $F: N_1 \to N_2$ является элементом пространства $(N_1 \to (N_1 \to N_2))$. В соответствии с только что сказанным мы можем считать $F''(x)$ элементом пространства $(N_1^2 \to N_2)$.

8. Отображение m-мерного евклидова пространства в n-мерное. Важным частным случаем введенных выше отображений является случай отображения $F(x)$ m-мерного евклидова пространства в n-мерное. Напомним, что в евклидовом пространстве $N_1 = E^m$ норма элемента $h = (h_1, h_2, \ldots, h_m)$ записывается в виде

$$\|h\|_{E^m} = \left[\sum_{i=1}^{m} h_i^2\right]^{1/2},$$

а в пространстве $N_2 = E^n$ норма элемента* $F(x) = (F_1(x), F_2(x), \ldots, F_n(x))$ записывается в виде

$$\|F(x)\|_{E^n} = \left[\sum_{i=1}^{n} F_i^2(x)\right]^{1/2},$$

* Мы, естественно, подразумеваем, что в пространствах E^m и E^n выбраны базисы, поэтому отображение $F(x): E^m \to E^n$ можно рассматривать в координатной форме. Конкретная реализация элементов пространств E^m, E^n может быть любой.

где x — фиксированная точка из E^m; $F_1(x)$, $F_2(x)$, ..., $F_n(x)$ — координаты вектора $F(x)$.

В случае отображения $F(x)=(F_1(x), F_2(x), ..., F_n(x))$, $x=(x_1, x_2, ..., x_m)$ m-мерного евклидова пространства E^m в n-мерное евклидово пространство E^n естественно считать это отображение или, что то же, эту вектор-функцию $F(x)$ * дифференцируемой в точке $x \in E^m$, если каждая компонента $F_1(x_1, x_2, ..., x_m)$, $F_2(x_1, x_2, ..., x_m)$, ..., $F_n(x_1, x_2, ..., x_m)$ дифференцируема в точке x как функция m переменных x_1, x_2, ..., x_m.

Отображение $F(x)$ можно рассматривать и с общей точки зрения (а не как векторную функцию), т. е. как отображение одного нормированного пространства $N_1=E^m$ в другое нормированное пространство $N_2=E^n$.

Определение дифференцируемого отображения $F(x): N_1 \to N_2$ в точке x в том случае будет таким же, как и в случае общих нормированных пространств, только нормы, фигурирующие в этом определении, будут определяться формулами для норм в евклидовом пространстве. А именно мы *назовем отображение* $F(x): E^m \to E^n$, *определенное на некотором открытом подмножестве* $\Sigma \subset E^m$, *дифференцируемым в точке* $x \in \Sigma$, *если существует такой ограниченный линейный оператор* $L_x \in (E^m \to E^n)$, *что для любого* $\varepsilon>0$ *можно найти* $\delta>0$, *при котором из неравенства* $\|h\|_{E^m} < \delta$ *следует неравенство*

$$\|F(x+h)-F(x)-L_x h\|_{E^n} \leqslant \varepsilon \|h\|_{E^m}.$$

То же самое сокращенно можно записать так:

$$F(x+h)-F(x)-L_x h=\mathbf{o}(h),$$

где величина $\|\mathbf{o}(h)\|/\|h\| \to 0$ при $\|h\| \to 0$. Аналогично общему случаю вводится и *производная* $F'(x)$ *отображения* $F(x)$.

Из курса линейной алгебры известно, что всякое линейное отображение m-мерного евклидова пространства в n-мерное (линейный оператор) можно задать некоторой матрицей порядка $(m \times n)$. Поскольку производная $F'(x)$ отображения $F(x)$, действующего из E^m в E^n, есть оператор из пространства E^m в пространство E^n (элемент пространства $(E^m \to E^n)$), то $F'(x)$ есть зависящая от x матрица порядка $(m \times n)$. Найдем вид матрицы.

Если в E^m и E^n выбраны базисы, а именно базис e_1, e_2, ..., e_m в E^m и базис f_1, f_2, ..., f_n в E^n, то всякий вектор $x \in E^m$ запишется в виде $x=x_1 e_1+x_2 e_2+...+x_m e_m$, где x_1, x_2, ..., x_m — координаты вектора x в базисе e_1, e_2, ..., e_m. Всякий вектор $y \in E^n$ можно записать в виде $y=y_1 f_1+y_2 f_2+ ... +y_n f_n$, где y_1, y_2, ..., y_n — координаты вектора y в базисе f_1, f_2, ..., f_n.

---

* Мы учитываем, что при каждом x из E^m значение $F(x)$ представляет собой вектор $(F_1(x), F_2(x), ..., F_n(x))$ из пространства E^n.

Отображение $y=F(x)$ пространства E^m в пространство E^n $(x\in E^m,\ y\in E^n)$ можно записать в виде

$$\begin{aligned} y_1&=F_1(x_1,\ x_2,\ \dots,\ x_m),\\ y_2&=F_2(x_1,\ x_2,\ \dots,\ x_m),\\ &\dots\dots\dots\dots\\ y_n&=F_n(x_1,\ x_2,\ \dots,\ x_m). \end{aligned}$$

Здесь $F_1(x_1, x_2, \dots, x_m)$, $F_2(x_1,\ x_2,\ \dots,\ x_m)$, ..., $F_n(x_1,\ x_2,\ \dots,\ x_m)$ — координаты вектора $y=F(x)$ при фиксированном $x=(x_1,\ x_2,\ \dots,\ x_m)$.

Далее, отображение $F'(x)$, как мы уже говорили, является линейным оператором — элементом пространства $(E^m\to E^n)$. Покажем, что *если $F(x)$ дифференцируема, как вектор-функция, то*

$$F'(x)=\begin{Vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \dots & \frac{\partial y_1}{\partial x_m}\\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \dots & \frac{\partial y_2}{\partial x_m}\\ \dots & \dots & \dots & \dots\\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \dots & \frac{\partial y_n}{\partial x_m} \end{Vmatrix}. \qquad (***)$$

В самом деле, в этом случае $F(x)=(F_1(x),\ F_2(x),\ \dots,\ F_m(x))$, $F(x+h)=(F_1(x+h),\ F_2(x+h),\ \dots,\ F_m(x+h))$, $x=(x_1,\ x_2,\ \dots,\ x_m)$, $x+h=(x_1+h_1,\ x_2+h_2,\ \dots,\ x_m+h_m)$, и если записать вектор-функцию покомпонентно в виде столбца и воспользоваться тем, что каждая компонента является дифференцируемой функцией m переменных, то

$$F(x+h)-F(x)=$$

$$=\begin{pmatrix} F_1(x_1+h_1,\ x_2+h_2,\ \dots,\ x_m+h_m)-F_1(x_1,\ x_2,\ \dots,\ x_m)\\ F_2(x_1+h_1,\ x_2+h_2,\ \dots,\ x_m+h_m)-F_2(x_1,\ x_2,\ \dots,\ x_m)\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ F_n(x_1+h_1,\ x_2+h_2,\ \dots,\ x_m+h_m)-F_n(x_1,\ x_2,\ \dots,\ x_m) \end{pmatrix}=$$

$$=\begin{Vmatrix} \sum_{i=1}^{m}\frac{\partial F_1(x_1,\ x_2,\ \dots,\ x_m)}{\partial x_i}h_i+o(h)\\ \sum_{i=1}^{m}\frac{\partial F_2(x_1,\ x_2,\ \dots,\ x_m)}{\partial x_i}h_i+o(h)\\ \dots\dots\dots\dots\dots\dots\\ \sum_{i=1}^{m}\frac{\partial F_n(x_1,\ x_2,\ \dots,\ x_m)}{\partial x_i}h_i+o(h) \end{Vmatrix}=$$

$$=\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_1}{\partial x_m} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \dots & \frac{\partial F_2}{\partial x_m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \dots & \frac{\partial F_n}{\partial x_m} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix} + \begin{pmatrix} o(h) \\ o(h) \\ \vdots \\ o(h) \end{pmatrix}.$$

Здесь величины $o(h)/\|h\|$ *, принимающие вещественные значения, стремятся к нулю, если $\|h\|=(h^2_1+h^2_2+\ldots+h^2_m)^{1/2}\to 0$. Эта матричная запись допускает, как это легко видеть, такое, более простое представление

$$F(x+h)-F(x)=L_x h+\mathbf{o}(h),$$

где

$$F'(x)=L_x=\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_1}{\partial x_m} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \dots & \frac{\partial F_2}{\partial x_m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \dots & \frac{\partial F_n}{\partial x_m} \end{pmatrix},\quad h=\begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix},\quad \mathbf{o}(h)=\begin{pmatrix} o(h) \\ o(h) \\ \vdots \\ o(h) \end{pmatrix},$$

причем величины $o(h)/\|h\|\to 0$, если $\|h\|=(h^2_1+h^2_2+\ldots+h^2_m)^{1/2}\to 0$. Поскольку координаты вектора $\mathbf{o}(h)/\|h\|$ стремятся к нулю, если $\|h\|\to 0$, то и норма вектора $\mathbf{o}(h)/\|h\|$ в пространстве E^n (т. е. корень квадратный из суммы квадратов его компонент) также стремится к нулю, если $\|h\|\to 0$.

Поэтому можно заключить, что производная $F'(x)$ дифференцируемого отображения $F(x): E^m\to E^n$ в точке x находится по формуле (***). Матрица, задаваемая формулой (***) называется **якобиевой матрицей отображения** $F: E^m\to E^n$, а в случае $n=m$ ее определитель — **якобианом** этого отображения в данной точке x.

9. Производные и дифференциалы высших порядков. Аналогично тому, как это было сделано в п. 7 для второй производной отображения $F(x)$ одного нормированного пространства N_1 в другое нормированное пространство N_2, можно ввести понятие третьей, четвертой и вообще n-й производной отображения $F(x)$, *определив n-ю производную как производную от производной $(n-1)$-го порядка.* При этом, очевидно, n-я производная представляет собой элемент пространства $(N_1\to(N_1\to\ldots-(N_1\to N_2)))$, N_1 встречается n раз. Повторяя рассуждения, проведенные для второй производной, можно каждому элементу этого пространства поставить в соответствие эле-

* В случае дифференцируемой функции m переменных величину $o(h)$ мы обозначали $o(\rho)=o(\|h\|)$, где $\rho=\|h\|$.

мент пространства $(N_1^n \to N_2)$ полилинейных (а именно n-линейных) отображений $N_1 \to N_2$. При этом *под полилинейным отображением (n-линейным) отображением одного нормированного пространства в другое понимается такое соответствие $y = A(x_1, x_2, ..., x_n)$ между упорядоченными системами $(x_1, x_2, ..., x_n)$ элементов из N_1* * *и элементами пространства N_2, которое линейно по каждому из x_i, $i = 1, 2, ...$, при фиксированных остальных элементах и удовлетворяет при некотором $M > 0$ условию*

$$\|A(x_1, x_2, ..., x_n)\| \leqslant M\|x_1\|\|x_2\| ... \|x_n\|.$$

Таким образом, n-ю производную $F^{(n)}(x)$ отображения F можно считать элементом пространства $(N_1^n \to N_2)$.

Рассмотрим теперь дифференциалы высших порядков. Напомним, что мы определили сильный дифференциал dF отображения F как результат применения к элементу $h \in N_1$ линейного оператора $F'(x)$, т. е. в виде $dF = F'(x)h$.

Дифференциал второго порядка определяется как величина $d^2F = F''(x)(h, h)$, *где $F''(x)(h, h)$ является квадратичным выражением***, *отвечающим отображению $F''(x) \in \in (N_1^2 \to N_2)$.*

Аналогично *дифференциалом n-го порядка называется выражение $d^nF = F^{(n)}(x)(h, h, ..., h)$, т. е. дифференциалом n-го порядка называется элемент пространства N_2, который получается в результате применения оператора $F^{(n)}(x)$ к элементу $(\underbrace{h, h, ..., h}_{n \text{ раз}})$ пространства $N_1 \times N_1 \times ... \times N_1 = N_1^n$.*

10. Формула Тейлора для отображений одного нормированного пространства в другое. Согласно рассмотрениям п. 1 сильная дифференцируемость отображения $F(x)$ означает, что разность $F(x+h) - F(x)$ может быть представлена в виде линейного члена по h и слагаемого, имеющего порядок выше первого относительно $\|h\|$.

Этот факт обобщает, как мы знаем, соответствующее разложение для дифференцируемой функции $F(x) = F(x_1, x_2, ..., x_m)$ m переменных. Для функции $F(x_1, x_2, ..., x_m)$ справедлива, как было показано в этой главе, формула Тейлора с остаточным членом в форме Пеано. Поэтому, естественно, возникает вопрос о получении формулы Тейлора с остаточным членом в форме, аналогичной форме Пеано, и для отображений нормированных пространств. Справедлива следующая теорема.

Теорема. *Пусть F — отображение нормированного про-*

* Элемент $(x_1, x_2, ..., x_n)$ принадлежит пространству $N_1^n = N_1 \times ... \times N_1$ (ср. дополнение 2).

** Выражение $B(x, x)$ называется квадратичным, если получено из билинейного отображения $B(x, y)$ при совпадающих аргументах, т. е. при $x = y$.

странства N_1 *в нормированное пространство* N_2, *определенное на некотором открытом множестве* $\Sigma\subset N_1$ *и такое, что* $F^{(n)}(x)$ *существует и представляет собой равномерно непрерывную функцию** *от* x *в* Σ. *Тогда имеет место равенство*

$$F(x+h)-F(x)=F'(x)h+{}^1/_2F''(x)(h,\ h)+\ldots$$
$$\ldots+F^{(n)}(x)(h,\ h,\ \ldots,\ h)+r(x,\ h), \qquad (****)$$

где $\|r(x,\ h)\|=o(\|h\|^n)$ **.

Доказательство. Проведем доказательство этой теоремы по индукции. При $n=1$ равенство означает просто дифференцируемость отображения $F(x)$, и тем самым это равенство по условию теоремы выполнено. Рассмотрим теперь произвольное фиксированное целое число $n\geqslant 2$ и предположим, что равенство, получающееся из (4) заменой n на $n-1$, справедливо для всех отображений, удовлетворяющих условиям теоремы, в которых n заменено на $n-1$. Докажем равенство (****).

Рассмотрим отображение $F'(x)$ и применим к нему формулу Тейлора (****), в которой n заменено на $n-1$, а вместо приращения h рассмотрено приращение th, где t — число, принадлежащее сегменту $[0,\ 1]$. Очевидно, что

$$F'(x+th)=F'(x)+tF''(x)h+\frac{t^2}{2}F'''(x)(h,\ h)+\ldots$$
$$\ldots+\frac{t^{n-1}}{(n-1)!}F^{(n)}(x)(h,\ h,\ \ldots,\ h)+r_1(x,th)^{***},$$

где $\|r_1(x,\ th)\|=o(t^{n-1}\|h\|^{n-1})$, а форма $(h,\ h,\ldots,\ h)$ имеет $n-1$ аргументов. Проинтегрируем обе части формулы Тейлора для $F'(x+h)$ **** по сегменту $[x,\ x+h]$ и воспользуемся формулой

---

* Отображение $F:N_1\to N_2$ (или абстрактная функция $F(x)$) называется равномерно непрерывной функцией на множестве $M\subset N_1$, если для любого числа $\varepsilon>0$ найдется такое число $\delta>0$, что $\|F(x_1)-F(x_2)\|<\varepsilon$ для всех точек x_1, x_2 из множества M одновременно, если только $\|x_1-x_2\|<\delta$ (см. также примечание п. 6).

** Это равенство означает, что для любого числа $\varepsilon>0$ найдется такое число $\delta>0$, что $\|r(x,\ h)\|<\varepsilon\|h\|^n$, если $\|h\|<\delta$, т. е. $\dfrac{\|r(x,\ h)\|'}{\|h\|^n}\to 0$, если $n\to 0$.

*** Мы пользуемся тем, что формы $(h,\ h)$, ..., $(h,\ h,\ \ldots,\ h)$ линейны по каждому из своих аргументов. Поэтому $(th,\ th)=t^2(h,\ h)$,
$$\underbrace{(th,\ th,\ \ldots,\ th)}_{n-1}=t^{n-1}\underbrace{(h,\ h,\ \ldots,\ h)}_{n-1}.$$

**** Т. е. формулу

$$F'(x+th)=F'(x)+tF''(x)h+\frac{t^2}{2}F'''(x)(h,\ h)+$$
$$+\ldots+\frac{t^{n-1}}{(n-1)!}F^{(n)}(x)\underbrace{(h,\ h,\ \ldots,\ h)}_{n-1}+r_1(x,\ h).$$