Математический институт им. В. А. Стеклова
Российской академии наук

Лекционные курсы НОЦ

Выпуск 17

Издание выходит с 2006 года

С. А. Теляковский

Курс лекций по математическому анализу

Семестр II

Москва
2011

УДК 517
ББК (В)22.16
Л43

Редакционный совет:

С.И. Адян, Д.В. Аносов, О.В. Бесов, И.В. Волович,
А.М. Зубков, А.Д. Изаак (ответственный секретарь),
В.В. Козлов, С.П. Новиков,
В.П. Павлов (заместитель главного редактора),
А.Н. Паршин, Ю.В. Прохоров, А.Г. Сергеев, А.А. Славнов,
Д.В. Трещев (главный редактор), Е.М. Чирка

Л43 **Лекционные курсы НОЦ**/ Математический институт им. В. А. Стеклова РАН (МИАН). – М.: МИАН, 2011. Вып. 17: Курс лекций по математическому анализу. Семестр II / Теляковский С. А. – 192 с.

ISBN 5-98419-030-3

Серия “Лекционные курсы НОЦ” – рецензируемое продолжающееся издание Математического института им. В. А. Стеклова РАН. В серии “Лекционные курсы НОЦ” публикуются материалы специальных курсов, прочитанных в Математическом институте им. В. А. Стеклова Российской академии наук в рамках программы Научно-образовательный центр МИАН.

Настоящая брошюра содержит “Курс лекций по математическому анализу” С. А. Теляковского, читавшийся студентам первого курса механико-математического факультета МГУ в 1996–2006 годах.

ISBN 5-98419-030-3

© Математический институт им. В. А. Стеклова РАН, 2011
© Теляковский С. А., 2011

Оглавление

Введение

В настоящем выпуске материал излагается в основном в том виде, как он читался автором на механико-математическом факультете МГУ.

По традиции, сложившейся на мехмате, во втором семестре курс математического анализа составляют интегралы функций одной переменной и дифференциальное исчисление функций многих переменных.

При изложении математического анализа неизбежно встаёт вопрос, какое определение интеграла студентам следует изучать в первую очередь.

Разумеется, студенты, получающие серьёзное математическое образование, должны активно владеть классическими понятиями интеграла Римана и интеграла Лебега.

Автор настоящего курса придерживается распространённого традиционного мнения, что начинать нужно с интеграла Римана, оставив изучение интеграла Лебега на более позднее время в курсе действительного анализа или аналогичных ему курсов.

Наряду с другими понятиями интеграла последнее время большую популярность приобретает интеграл Хенстока–Курцвейля. Это несомненно удачное изобретение привлекает просто формулируемым определением и широтой круга задач, которые с помощью этого интеграла можно решить.

Несомненно, студенты, специализирующиеся по теории функций, должны быть знакомы с интегралом Хенстока–Курцвейля. Но по убеждению автора начинать с него изложение понятия интеграла даже для этой группы студентов было бы преждевременным.

Что касается изложения дифференциального исчисления функций многих переменных, то и здесь автор придерживается сложившейся традиционной точки зрения.

В конце некоторых глав помещены задачи и упражнения. Это – полезное по мнению автора добавление к первому изданию. К тому же включить в курс задачный материал автору советовали и некоторые коллеги. Конечно, количество задач не должно быть чрезмерным (чтобы не дублировать задачники), но сейчас оно явно недостаточно и здесь предстоит ещё дальнейшая работа.

Январь 2011 г. С.А.Теляковский

Глава 8. Неопределённый интеграл

§ 8.1. Первообразная. Табличные интегралы

Как и ранее, промежутки означают отрезки, интервалы или полуинтервалы (когда один конец принадлежит промежутку, а другой – нет). При этом интервалы и полуинтервалы могут быть конечными и бесконечными.

Определение. Пусть на промежутке I задана функция $f(x)$. Функцию $F(x)$ называют *первообразной функции* $f(x)$ на I, если в каждой точке $x \in I$ существует производная $F'(x)$ и $F'(x) = f(x)$.

Если концы промежутка принадлежат ему, то как обычно, в этих точках имеются в виду соответствующие односторонние производные.

Функцию $F(x)$ называют также примитивной для функции $f(x)$, но последнее время это название употребляется редко.

Если $F(x)$ является первообразной функции $f(x)$, то для произвольной постоянной C функция $F(x) + C$ также является первообразной $f(x)$, так как $(F(x) + C)' = F'(x) = f(x)$.

Верно и обратное: если $F_1(x)$ и $F_2(x)$ – две первообразные функции $f(x)$, то они отличаются на некоторую постоянную величину. В самом деле,

$$(F_1(x) - F_2(x))' = F_1'(x) - F_2'(x) = f(x) - f(x) = 0.$$

А функция, производная которой во всех точках равна нулю, согласно теореме 6.2.4 является константой.

Таким образом, справедливо следующее утверждение.

Теорема 8.1.1. *Если для функции $f(x)$ известна одна её первообразная $F(x)$, то любая другая первообразная функции $f(x)$ может быть получена из $F(x)$ добавлением к ней некоторой постоянной.*

Поэтому выражение $F(x) + C$ называют общим видом первообразной.

Не каждая функция имеет первообразную. Действительно, согласно теореме Дарбу о промежуточных значениях 6.1.3, если функция f является производной некоторой функции на отрезке $[a, b]$, то $f(x)$ принимает все значения между $f(a)$ и $f(b)$. Значит, чтобы функция f имела первообразную на отрезке I, необходимо, чтобы f принимала на I все промежуточные значения. А этим свойством обладают не все функции.

Вместе с тем, в дальнейшем будет установлено, что каждая непрерывная функция имеет первообразную.

Определение. *Неопределённым интегралом* функции на промежутке I называют произвольную фиксированную её первообразную.

В некоторых руководствах дают другое определение и называют неопределённым интегралом множество всех первообразных данной функции.

Неопределённый интеграл функции $f(x)$ обозначают

$$\int f(x)\,dx \tag{8.1.1}$$

или короче

$$\int f\,dx.$$

Знак $\int$ называют знаком интеграла, функцию f – подынтегральной функцией.

Термин интеграл ввёл Я. Бернулли (1690).

Выражение $f(x)\,dx$, стоящее под знаком интеграла (8.1.1), является дифференциалом этого неопределённого интеграла. Действительно,

$$\left(\int f(x)\,dx\right)' = f(x), \tag{8.1.2}$$

значит,

$$d\left(\int f(x)\,dx\right) = f(x)\,dx.$$

Итак, неопределённый интеграл функции $f(x)$ – это некоторая фиксированная её первообразная, а для любой другой первообразной $F(x)$ существует постоянная C такая, что

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

Операцию нахождения неопределённых интегралов называют *интегрированием*. Если найдена первообразная функции f, то говорят, что мы *взяли* интеграл (8.1.1).

Теорема 8.1.2. *Если функции f и g имеют первообразные, то при произвольных числах α и β функция $\alpha f(x)+\beta g(x)$ имеет первообразную и выполняется равенство*

$$\int (\alpha f(x) + \beta g(x))\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx + C, \qquad (8.1.3)$$

где C – некоторая постоянная.

Для доказательства проверяем, что производная суммы из правой части формулы (8.1.3) равна $\alpha f(x) + \beta g(x)$.

Свойство, выраженное теоремой 8.1.2, называют линейностью неопределённого интеграла.

Если $F'(x) = f(x)$ на некотором промежутке, то на этом промежутке $F(x)$ является первообразной функции $f(x)$ и, значит,

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

Поэтому из таблицы производных элементарных функций получаем таблицу неопределённых интегралов. Эти интегралы называют *табличными*.

Если $a > 0$ и $a \neq 1$, то

$$\int a^x\,dx = \frac{a^x}{\ln a} + C. \qquad (8.1.4)$$

В частности,

$$\int e^x\,dx = e^x + C. \qquad (8.1.5)$$

Формулы (8.1.4) и (8.1.5) имеют место на произвольном промежутке изменения независимой переменной x.

Следующие три формулы также справедливы на любых x:

$$\int \sin x\,dx = -\cos x + C, \qquad (8.1.6)$$

$$\int \cos x\,dx = \sin x + C, \qquad (8.1.7)$$

$$\int \frac{1}{1+x^2}\,dx = \operatorname{arctg} x + C. \qquad (8.1.8)$$

Неопределённый интеграл степенной функции x^a при $a \neq -1$ имеет вид:

$$\int x^a\,dx = \frac{x^{a+1}}{a+1} + C, \quad x > 0. \tag{8.1.9}$$

При $a = -1$ выражение в правой части формулы (8.1.9) не имеет смысла. Поэтому условие $a \neq -1$ существенно для её справедливости.

Сделаем несколько замечаний о‘формуле (8.1.9), а затем рассмотрим случай $a = -1$.

Если $a > -1$, то в (8.1.9) условие $x > 0$ можно заменить на $x \geqslant 0$.

Для целых a равенство (8.1.9) выполняется и при $x < 0$. Если a – натуральное число или нуль, то равенство (8.1.9) справедливо на произвольном промежутке изменения переменной x (напомним, что при рассмотрении степенных функций считают $x^0 \equiv 1$). А если a – целое отрицательное число, то равенство (8.1.9) имеет место на произвольном промежутке, не содержащем точку 0.

Для неопределённого интеграла функции x^{-1} на произвольном промежутке, не содержащем $x = 0$, справедливо равенство:

$$\int \frac{1}{x}\,dx = \ln|x| + C. \tag{8.1.10}$$

Доказательство обычное – проверяем, что производная суммы из правой части равна подынтегральной функции.

Для каждой из формул

$$\int \frac{dx}{\cos^2 x} = \operatorname{tg} x + C, \tag{8.1.11}$$

$$\int \frac{dx}{\sin^2 x} = -\operatorname{ctg} x + C \tag{8.1.12}$$

промежуток изменения x не должен содержать нули знаменателя подынтегральной функции.

Наконец,

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C, \qquad x \in (-1, 1). \tag{8.1.13}$$

Равенства (8.1.4)–(8.1.13) содержат основные табличные интегралы. Они представляют собой иначе прочитанную таблицу производных. Эти формулы необходимо помнить.

§ 8.2. Методы интегрирования

Основными приёмами, используемыми при интегрировании элементарных функций, являются интегрирование по частям и интегрирование с помощью замены переменной (интегрирование подстановкой).

Заметим, что взятие неопределённых интегралов элементарных функций является трудной часто неразрешимой задачей. Здесь требуются опыт и изобретательность.

Теорема 8.2.1 (Интегрирование по частям). *Пусть на некотором промежутке функции $u(x)$ и $v(x)$ имеют производные и существует интеграл $\int u'v\,dx$. Тогда на этом промежутке существует интеграл $\int uv'dx$ и справедливо равенство, которое называют формулой интегрирования по частям,*

$$\int uv'\,dx = uv - \int u'v\,dx + C, \tag{8.2.1}$$

где C – некоторая постоянная.

Доказательство. Найдём производную функции из правой части (8.2.1):

$$\left(uv - \int u'v\,dx + C\right)' = (uv)' - \left(\int u'v\,dx\right)' = u'v + uv' - u'v = uv'.$$

Таким образом, функция $uv - \int u'v\,dx + C$ является первообразной функции uv'. Значит, неопределённый интеграл $\int uv'dx$ существует и отличается от этой первообразной на некоторую постоянную, т.е. справедливо равенство (8.2.1). Теорема доказана.

Формулу интегрирования по частям записывают также следующим образом:

$$\int u\,dv = uv - \int v\,du + C, \tag{8.2.2}$$

понимая $\int u\,dv$ как интеграл $\int uv'\,dx$ и $\int v\,du$ как $\int vu'\,dx$.

Теорема 8.2.2 (Замена переменной). *Пусть функция $f(x)$ на некотором промежутке имеет первообразную*

$$F(x) = \int f(x)\,dx + C. \tag{8.2.3}$$

Если функция $x = \varphi(t)$ имеет производную на промежутке I и все значения $\varphi(t)$ принадлежат промежутку, на котором справедливо равенство (8.2.3), то на I функция $f(\varphi(t))\varphi'(t)$ имеет первообразную и выполняется равенство

$$\int f(\varphi(t))\varphi'(t)\,dt = F(\varphi(t)) + C. \tag{8.2.4}$$

Доказательство. В силу условий теоремы сложная функция $f(\varphi(t))$ имеет смысл на промежутке I, а значит, имеет смысл и сложная функция $F(\varphi(t))$.

Согласно теореме 5.4.1 о производной сложной функции имеем для $t \in I$

$$\frac{d}{dt}F(\varphi(t)) = \frac{dF(x)}{dx}\bigg|_{x=\varphi(t)} \frac{d\varphi}{dt}(t) = f(x)\Big|_{x=\varphi(t)} \frac{d\varphi}{dt}(t) = f(\varphi(t))\,\varphi'(t)$$

(в концевых точках имеются в виду односторонние производные).

Отсюда вытекает равенство (8.2.4). Теорема доказана.

Формулу (8.2.4) можно записать иначе. Сделаем в (8.2.3) замену $x = \varphi(t)$:

$$F(\varphi(t)) = \int f(x)\,dx\bigg|_{x=\varphi(t)} + C.$$

Таким образом,

$$\int f(\varphi(t))\varphi'(t)\,dt = \int f(x)\,dx\bigg|_{x=\varphi(t)} + C, \tag{8.2.5}$$

т.е. для вычисления интеграла $\int f(\varphi(t))\varphi'(t)\,dt$ находим интеграл $\int f(x)\,dx$ и делаем затем замену $x = \varphi(t)$. Равенство (8.2.5) называют формулой *интегрирования подстановкой.*

Если функция $x = \varphi(t)$ на промежутке I строго монотонна, то существует обратная функция $t = \varphi^{-1}(x)$ и, если при отображении $x = \varphi(t)$ образом I является промежуток, равенство (8.2.5) можно записать на нём так:

$$\int f(x)\,dx = \int f(\varphi(t))\varphi'(t)\,dt\bigg|_{t=\varphi^{-1}(x)} + C. \tag{8.2.6}$$

§ 8.3. Интегрирование рациональных дробей

Рациональной дробью называется отношение двух многочленов $P(x)/Q(x)$. Далее будем считать, что P и Q – многочлены с действительными коэффициентами без общих корней, и рассматривать промежуток, не содержащий нулей знаменателя.

Покажем, что для каждой такой рациональной дроби существует элементарная функция, являющаяся её первообразной.

При этом достаточно рассматривать только правильные рациональные дроби (когда степень многочлена в числителе меньше степени многочлена в знаменателе). В самом деле, произвольную рациональную дробь можно представить в виде суммы многочлена и правильной рациональной дроби. Затем воспользоваться линейностью неопределённых интегралов и тем, что интеграл от многочлена является многочленом.

Из курса алгебры известно, что правильную рациональную дробь можно представить в виде суммы так называемых простейших дробей.

Именно, если a – действительный корень кратности k знаменателя $Q(x)$, то среди таких простейших дробей будут дроби

$$\frac{B_k}{(x-a)^k}, \quad \frac{B_{k-1}}{(x-a)^{k-1}}, \quad \ldots, \quad \frac{B_1}{x-a}, \tag{8.3.1}$$

где $B_k, B_{k-1}, \ldots, B_1$ – некоторые действительные числа.

А если квадратный трёхчлен x^2+px+q является произведеним двух комплексных сопряжённых корней кратности j многочлена $Q(x)$, то в представлении рациональной дроби $P(x)/Q(x)$ участвуют простейшие дроби вида

$$\frac{M_jx+N_j}{(x^2+px+q)^j}, \quad \frac{M_{j-1}x+N_{j-1}}{(x^2+px+q)^{j-1}}, \quad \ldots, \quad \frac{M_1x+N_1}{x^2+px+q}, \tag{8.3.2}$$

где $M_j, N_j, \ldots, M_1, N_1$ – действительные числа.

Таким образом, в силу линейности неопределённых интегралов задача сводится к интегрированию дробей (8.3.1) и (8.3.2).

Для дробей (8.3.1), пользуясь табличными интегралами, получаем

$$\int \frac{B_1}{x-a}\,dx = B_1 \ln|x-a| + C,$$

и если $l \neq 1$,

$$\int \frac{B_l}{(x-a)^l}\,dx = \frac{B_l}{(-l+1)(x-a)^{l-1}} + C. \tag{8.3.3}$$

Строго говоря, здесь нужно было бы сделать замену $x-a=t$, после которой воспользоваться табличными интегралами и вернуться к переменной x. Но эти промежуточные шаги ввиду их простоты можно явно не выписывать.

Рассмотрим теперь интегралы от дробей (8.3.2). Преобразуем трёхчлен из знаменателей

$$x^2+px+q=\left(x+\frac{p}{2}\right)^2+q-\frac{p^2}{4}\,.$$

Здесь

$$q-\frac{p^2}{4}>0.$$

Введём обозначение

$$a^2:=q-\frac{p^2}{4}\,.$$

После замены

$$x+\frac{p}{2}=t$$

дробь

$$\frac{M_rx+N_r}{(x^2+px+q)^r}$$

переходит в дробь вида

$$\frac{M_rt+L_r}{(t^2+a^2)^r}\,,$$

значение числа L_r в которой сейчас не нужно.

Представив интеграл от суммы в виде суммы интегралов, получим

$$\int\frac{M_rt+L_r}{(t^2+a^2)^r}\,dt=M_r\int\frac{t}{(t^2+a^2)^r}\,dt+L_r\int\frac{dt}{(t^2+a^2)^r}+C.$$

Интеграл

$$\int\frac{t}{(t^2+a^2)^r}\,dt$$

приводится к табличному заменой $y=t^2+a^2$. Тогда $dy=2t\,dt$ и

$$\int\frac{t}{(t^2+a^2)^r}\,dt=\frac{1}{2}\int\frac{1}{y^r}\,dy\bigg|_{y=t^2+a^2}+C.$$

Значит, этот интеграл равен $\ln y$, если $r = 1$, и $y^{1-r}/(1-r)$, если $r \geqslant 2$.

Рассмотрим теперь интеграл

$$A_r := \int \frac{dt}{(t^2+a^2)^r}\,.$$

Если $r = 1$, то с помощью замены $t = au$ находим

$$A_1 = \int \frac{dt}{t^2+a^2} = \int \frac{a\,du}{a^2(u^2+1)}\bigg|_{u=t/a} + C = \frac{1}{a}\operatorname{arctg}\frac{t}{a} + C.$$

Здесь при каждом переходе к новому выражению нужно было бы писать новую постоянную C. Но обычно в промежуточных преобразованиях произвольную постоянную вообще не пишут, а добавляют её только на последнем шаге.

Так и сделаем, рассматривая интеграл A_r при $r \geqslant 2$. Имеем

$$A_r = \frac{1}{a^2}\int \frac{t^2+a^2-t^2}{(t^2+a^2)^r}\,dt = \frac{1}{a^2}A_{r-1} - \frac{1}{2a^2}\int t\cdot\frac{2t}{(t^2+a^2)^r}\,dt.$$

К последнему интегралу применим формулу интегрирования по частям, положив

$$u = t, \quad dv = \frac{2t}{(t^2+a^2)^r}\,dt.$$

Тогда

$$du = dt, \quad v = \frac{1}{(-r+1)(t^2+a^2)^{r-1}}$$

и согласно (8.2.2)

$$\begin{aligned}\int t\cdot\frac{2t}{(t^2+a^2)^r}\,dt &= \\ &= \frac{t}{(-r+1)(t^2+a^2)^{r-1}} - \int\frac{1}{(-r+1)(t^2+a^2)^{r-1}}\,dt = \\ &= -\frac{t}{(r-1)(t^2+a^2)^{r-1}} + \frac{1}{r-1}A_{r-1}\,.\end{aligned}$$

Таким образом,

$$A_r = \frac{t}{2a^2(r-1)(t^2+a^2)^{r-1}} + \frac{2r-3}{2a^2(r-1)}A_{r-1} + C. \tag{8.3.4}$$

Пользуясь этой формулой, от A_r переходим к A_{r-1}, затем от A_{r-1} к A_{r-2} и т.д. Так понижая шаг за шагом индекс у A_r, приходим к интегралу A_1.

Теперь остаётся только с помощью обратной замены вернуться к исходной переменной x.

Проведенные преобразования не только доказывают, что неопределённый интерал от рациональной дроби может быть представлен в виде суммы многочлена, рациональных дробей, логарифмов и арктангенсов, но и показывают, как это можно сделать.

На практике при интегрировании рациональных дробей основной трудностью является нахождение корней знаменателя. Если эти корни известны, то для представления рациональной дроби в виде суммы простейших дробей обычно пользуются методом неопределённых коэффициентов.

Но кроме метода неопределённых коэффициентов могут быть полезны и другие соображения.

Будем считать, что дробь $P(x)/Q(x)$ несократима, т.е. многочлены $P(x)$ и $Q(x)$ не имеют общих корней.

Если a – простой корень знаменателя $Q(x)$, то $Q'(a) \neq 0$ и коэффициент B соответствующей простейшей дроби

$$\frac{B}{x-a}$$

можно найти следующим образом.

Пусть

$$\frac{P(x)}{Q(x)} = \frac{B}{x-a} + R(x),$$

где $R(x)$ – сумма остальных простейших дробей, знаменатели которых при $x = a$ не обращаются в нуль.

Пользуясь тем, что $Q(a) = 0$, находим

$$B + R(x)(x-a) = \frac{P(x)}{Q(x)}(x-a) = P(x) : \frac{Q(x) - Q(a)}{x-a}\,.$$

Переходя в этом равенстве к пределу при $x \to a$, получим

$$B = \frac{P(a)}{Q'(a)}\,.$$

Подобные соображения можно применить и в случае кратных корней знаменателя. Естественно, формулы при этом будут сложнее.

§ 8.4. Метод Остроградского интегрирования рациональных дробей

Интегрировать рациональную дробь $P(x)/Q(x)$, знаменатель которой имеет кратные корни, можно более простым способом, чем описанный в § 8.3. Его называют интегрированием методом Остроградского. Упрощения будут более значительными, когда кратность корней знаменателя $Q(x)$ велика.

Изложим метод Остроградского на примере правильной рациональной дроби

$$\frac{P(x)}{(x-a)^k(x^2+px+q)^j}\,, \tag{8.4.1}$$

где $k > 1$, $j > 1$ и $x^2+px+q > 0$ при всех x.

Согласно (8.3.3) интегрирование дробей (8.3.1) при $k > 1$ даёт дроби, у которых показатель степени разности $x-a$ в знаменателях на единицу меньше, чем у исходной дроби. В силу (8.3.4) интегрирование дробей (8.3.2) при $j > 1$ даёт дроби, у которых показатель степени трехчлена x^2+px+q в знаменателе на единицу меньше.

Таким образом, интеграл от дроби (8.4.1) представляет собой сумму некоторой постоянной C, логарифмов, арктангенсов и дроби вида

$$\frac{P_1(x)}{(x-a)^{k-1}(x^2+px+q)^{j-1}}\,, \tag{8.4.2}$$

где $P_1(x)$ – многочлен, степень которого меньше $k+2j-3$ (степени знаменателя), т.е. дробь (8.4.2) – правильная.

Так как логарифмы и арктангенсы появляются при интегрировании дробей (8.3.1) и (8.3.2), соответствующих $k=1$ и $j=1$, то

$$\int \frac{P(x)}{(x-a)^k(x^2+px+q)^j}\,dx = \\ = \frac{P_1(x)}{(x-a)^{k-1}(x^2+px+q)^{j-1}} + \\ + \int\left(\frac{B}{x-a} + \frac{Mx+N}{x^2+px+q}\right)dx + C. \tag{8.4.3}$$

Коэффициенты многочлена $P_1(x)$ и числа B, M и N можно найти методом неопределённых коэффициентов. В самом деле, приравняем производные функций из обеих частей равенства

(8.4.3):

$$\begin{aligned}
&\frac{P(x)}{(x-a)^k(x^2+px+q)^j} = \\
&\quad = \left(\frac{P_1(x)}{(x-a)^{k-1}(x^2+px+q)^{j-1}}\right)' + \frac{B}{x-a} + \frac{Mx+N}{x^2+px+q} = \\
&\quad = \frac{P_1'(x)}{(x-a)^{k-1}(x^2+px+q)^{j-1}} - P_1(x)\frac{k-1}{(x-a)^k(x^2+px+q)^{j-1}} - \\
&\qquad - P_1(x)\frac{(j-1)(2x+p)}{(x-a)^{k-1}(x^2+px+q)^j} + \frac{B}{x-a} + \frac{Mx+N}{x^2+px+q}.
\end{aligned} \tag{8.4.4}$$

Приведём дроби из (8.4.4) к общему знаменателю, который равен $(x-a)^k(x^2+px+q)^j$, и приравняем числители полученных дробей:

$$\begin{aligned}
P(x) = P_1'(x)(x-a)(x^2+px+q) - P_1(x)(k-1)(x^2+px+q) - \\
- P_1(x)(j-1)(2x+p)(x-a) + \\
+ B(x-a)^{k-1}(x^2+px+q)^j + \\
+ (Mx+N)(x-a)^k(x^2+px+q)^{j-1}.
\end{aligned} \tag{8.4.5}$$

Так как дробь (8.4.2) правильная, степень многочлена $P_1(x)$ меньше или равна $k+2j-4$. Запишем $P_1(x)$ как многочлен степени $k+2j-4$ с неизвестными коэффициентами, число которых равно $k+2j-3$. Тогда в правой части равенства (8.4.5) получим многочлен степени не выше $k+2j-1$, коэффициенты которого являются линейными комбинациями $k+2j-3$ коэффициентов многочлена $P_1(x)$ и чисел B, M, N, т.е. всего $k+2j$ неизвестных величин.

По предположению дробь (8.4.1) – правильная. Значит, степень многочлена $P(x)$ меньше или равна $k+2j-1$. Запишем $P(x)$ как многочлен степени $k+2j-1$ (коэффициенты которого при старших степенях x могут быть нулями).

Приравняв коэффициенты многочленов из левой и правой частей равенства (8.4.5) при степенях x^r, $r=0,1,\ldots,k+2j-1$, получим систему $k+2j$ линейных уравнений с указанными выше $k+2j$ неизвестными.

Из сказанного в § 8.3 следует, что для каждого многочлена $P(x)$ справедливо равенство вида (8.4.3). Поэтому построенная система линейных уравнений всегда разрешима и её решение единственно.

Это показывает возможность применения метода неопределённых коэффициентов при интегрировании рациональных дробей методом Остроградского.

Выигрыш в количестве проводимых при этом вычислений обусловлен тем, что сразу находят дробь вида (8.4.2), а не простейшие дроби в разложении дроби $P(x)/Q(x)$, которые затем ещё нужно интегрировать.

§ 8.5. Интегрирование некоторых других выражений

Был подробно рассмотрен вопрос об интегрировании рациональных дробей. Приёмы, позволяющие находить неопределённые интегралы, известны и для некоторых других элементарных функций. Приведём несколько примеров, когда с помощью замен переменных задача сводится к интегрированию рациональных дробей.

Если $R(x)$ – функция, полученная из $\sin x$, $\cos x$ и чисел с помощью сложения, вычитания, умножения и деления, то интеграл

$$\int R(x)\,dx$$

можно привести к интегралу от рациональной дроби с помощью подстановки $t = \operatorname{tg}(x/2)$. В самом деле, в этом случае

$$\sin x = \frac{2\sin(x/2)\cdot\cos(x/2)}{\cos^2(x/2)+\sin^2(x/2)} = \frac{2\operatorname{tg}(x/2)}{1+\operatorname{tg}^2(x/2)} = \frac{2t}{1+t^2}\,,$$
$$\cos x = \frac{\cos^2(x/2)-\sin^2(x/2)}{\cos^2(x/2)+\sin^2(x/2)} = \frac{1-\operatorname{tg}^2(x/2)}{1+\operatorname{tg}^2(x/2)} = \frac{1-t^2}{1+t^2}$$

и, так как $x = 2\operatorname{arctg} t$, то

$$dx = \frac{2dt}{1+t^2}\,.$$

В результате указанной замены будут получены функции, преобразующиеся в рациональные дроби относительно переменной t.

Иногда функцию $R(x)$ рассмотренного вида можно привести к рациональной дроби с помощью более простых подстановок

$t = \sin x$, $t = \cos x$ или $t = \operatorname{tg} x$, но это возможно не для всех функций $R(x)$. А замена $t = \operatorname{tg} x/2$ является универсальной, всегда приводящей к цели.

Следующий пример – интегрирование дифференциального бинома

$$\int x^m (a + bx^n)^p \, dx, \tag{8.5.1}$$

где m, n, p, a и b – действительные не равные нулю числа.

С помощью замены $x = t^{1/n}$ получаем

$$\int x^m (a + bx^n)^p \, dx = \int t^{m/n} (a + bt)^p \, \frac{1}{n} \, t^{1/n-1} \, dt \bigg|_{t=x^n} + C.$$

Положим

$$q := \frac{m+1}{n} - 1$$

и покажем, как берётся интеграл

$$\int t^q (a + bt)^p \, dt,$$

если p и q – рациональные числа и какое-либо из чисел p, q или $p + q$ целое.

Пусть p – целое и $q = r/s$, где r – целое и s – натуральное. Тогда после подстановки $t = y^s$ получим интеграл

$$\int y^r (a + by^s)^p s y^{s-1} \, dy,$$

в котором все показатели степени – целые числа и, значит, функция под знаком интеграла или многочлен или рациональная дробь.

Если q – целое и $p = r/s$, то в результате замены $a + bt = y^s$ приходим к интегралу

$$\int \left(\frac{y^s - a}{b} \right)^q y^r \, \frac{s \, y^{s-1}}{b} \, dy$$

с целыми показателями степени.

Наконец, если целое $p + q$, то

$$\int t^q (a + bt)^p \, dt = \int t^{p+q} \left(\frac{a + bt}{t} \right)^p dt$$

и если $p = r/s$, то делаем замену $a/t + b = y^s$, после которой, как и в других случаях, получаем интеграл от многочлена или от рациональной дроби.

Ещё один пример – интегралы

$$\int \sqrt{ax^2 + bx + c}\, dx, \tag{8.5.2}$$

где трёхчлен $ax^2 + bx + c$ на рассматриваемом промежутке положителен.

Замены переменной интегрирования, с помощью которых интеграл (8.5.2) приводится к интегралу от рациональной дроби, называют *подстановками Эйлера.*

Если $a > 0$, то положим

$$\sqrt{ax^2 + bx + c} = \sqrt{a}\, x \pm t,$$

где можно взять любой знак $+$ или $-$. Тогда

$$ax^2 + bx + c = ax^2 \pm 2\sqrt{a}\, xt + t^2.$$

Отсюда

$$x = \frac{t^2 - c}{b \mp 2\sqrt{a}\, t}\,,$$

т.е. переменная x и её производная представляют собой рациональные дроби переменной t.

Если $c > 0$, то можно положить

$$\sqrt{ax^2 + bx + c} = xt \pm \sqrt{c},$$

откуда $ax + b = xt^2 \pm 2\sqrt{c}\, t$, т.е.

$$x = \frac{b \mp 2\sqrt{c}t}{t^2 - a}$$

и вновь получен интеграл от рациональной дроби.

Если многочлен $ax^2 + bx + c$ имеет действительные различные корни x_1 и x_2, то в результате замены

$$\sqrt{ax^2 + bx + c} = (x - x_1)t$$

получим

$$ax^2 + bx + c = a(x - x_1)(x - x_2) = (x - x_1)^2 t^2,$$

откуда

$$x = \frac{2tx_1 - ax_2}{t^2 - a}\,.$$

Эта подстановка приводит интеграл (8.5.2) к интегралу от рациональной дроби относительно переменной t.

Интегралы от элементарных функций часто не являются элементарными функциями. Такие интегралы называют *неберущимися*.

Так, интеграл (8.5.1) представляет собой элементарную функцию только в тех случаях, когда целым является какое-либо из чисел p, q и $p+q$. Этот факт для рациональных m, n и p установил П. Л. Чебышев, а затем он был доказан для любых действительных m, n и p.

Некоторые неберущиеся интегралы играют важную роль и специально изучаются. Таковы, например, интеграл вероятностей

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

интегральные синус и косинус

$$\int \frac{\sin x}{x}\,dx, \qquad \int \frac{\cos x}{x}\,dx,$$

интегральный логарифм

$$\int \frac{dx}{\ln x}\,,$$

эллиптические интегралы

$$\int \sqrt{1 - k^2 \sin^2\varphi}\,d\varphi, \quad \int \frac{d\varphi}{\sqrt{1 - k^2 \sin^2\varphi}}\,, \qquad |k| < 1.$$

Эти интегралы не являются элементарными функциями. Доказательство основывается на глубоких алгебраических результатах и выходит за рамки настоящего курса.

Глава 9. Определённый интеграл

§ 9.1. Определение интеграла Римана

Говорят, что точки $x_0, x_1, \dots, x_n$ образуют разбиение отрезка $[a, b]$, если

$$a = x_0 < x_1 < \dots < x_n = b,$$

т.е. $x_0, x_1, \dots, x_n$ – строго возрастающая последовательность точек и концы отрезка принадлежат этой последовательности. Разбиения отрезка будем обозначать T.

Таким образом, отрезок $[a, b]$ разделён точками $x_0, x_1, \dots, x_n$ на n отрезков $[x_{k-1}, x_k]$, $k = 1, 2, \dots, n$. Длину отрезка $[x_{k-1}, x_k]$ обозначим $\Delta x_k := x_k - x_{k-1}$. Наибольшую из длин отрезков, полученных при разбиении T, т.е. число

$$\lambda_T := \max_{k=1,\dots,n} \Delta x_k,$$

называют *диаметром* разбиения T.

Напомним, что разбиения промежутков использовались в § 7.2.

Пусть на отрезке $[a, b]$ задана функция $f(x)$ и T – некоторое разбиение $[a, b]$. В каждом отрезке $[x_{k-1}, x_k]$ разбиения T возьмём произвольно точку ξ_k и составим сумму

$$\sum_{k=1}^{n} f(\xi_k)\Delta x_k. \tag{9.1.1}$$

Сумму (9.1.1) называют *интегральной суммой Римана* функции f, соответствующей разбиению T и выбору точек $\{\xi_k\}$, и обозначают $S_T(f, \xi)$ или $S_T(\xi)$.

Определение. Функцию f, заданную на отрезке $[a, b]$, называют *интегрируемой по Риману* на этом отрезке, если существует число I такое, что для каждого $\varepsilon > 0$ имеется число $\delta(\varepsilon) > 0$, при котором для любого разбиения T отрезка $[a, b]$ диаметра, меньшего δ, и произвольном выборе точек $\xi_k \in [x_{k-1}, x_k]$, $k = 1, 2, \dots, n$, справедлива оценка

$$|S_T(f, \xi) - I| < \varepsilon. \tag{9.1.2}$$

Число I называют *определённым интегралом Римана* от функции f на отрезке $[a, b]$ и обозначают

$$\int_a^b f(x)\,dx. \tag{9.1.3}$$

Числа a и b называют соответственно *нижним* и *верхним пределами интегрирования*, а $f(x)$ – подынтегральной функцией.

Это определение интеграла было дано Б. Риманом (1854). Ранее подобное определение для непрерывных функций дал О. Коши (1823). В § 9.2 будет доказано, что непрерывные на отрезке функции интегрируемы по Риману.

Если функция f интегрируема на отрезке $[a, b]$, будем писать $f \in \mathcal{R}[a, b]$.

Значениями функции f в определении интеграла могут быть как действительные, так и комплексные числа. В дальнейшем всюду, если не оговорено противное, будем иметь в виду интегралы действительнозначных функций.

Понятно, что если функция интегрируема по Риману, то значение её интеграла определяется однозначно.

Для краткости будем называть интеграл (9.1.3) определённым интегралом, не добавляя, что это интеграл Римана.

В качестве простейшего примера вычислим определённый интеграл от функции, принимающей постоянное значение L на всем отрезке $[a, b]$. В этом случае $f(\xi_k) = L$ для любых точек ξ_k и, значит, для каждого разбиения T

$$S_T(\xi) = \sum_{k=1}^{n} L\Delta x_k = L(b-a).$$

Таким образом,

$$\int_a^b L\,dx = L(b-a).$$

Покажем, что определённый интеграл может существовать только для ограниченных функций.

Теорема 9.1.1. *Если функция интегрируема на отрезке, то она ограничена на этом отрезке.*

Доказательство. Предположим, что функция $f(x)$ неограничена на отрезке $[a, b]$, и рассмотрим для произвольного разбиения T отрезка $[a, b]$ интегральную сумму Римана при каком-либо выборе точек $\xi_k \in [x_{k-1}, x_k]$, $k = 1, 2, \dots, n$:

$$S_T(\xi) = \sum_{k=1}^{n} f(\xi_k)\Delta x_k.$$

Функция $f(x)$ неограничена по крайней мере на одном из отрезков разбиения. Пусть это отрезок $[x_{j-1}, x_j]$. Тогда

$$S_T(\xi) = f(\xi_j)\Delta x_j + \sum_{k=1,\, k\neq j}^{n} f(\xi_k)\Delta x_k$$

и, значит,

$$|S_T(\xi)| \geqslant |f(\xi_j)|\Delta x_j - \left|\sum_{k=1,\, k\neq j}^{n} f(\xi_k)\Delta x_k\right|.$$

Считая числа ξ_k при $k \neq j$ фиксированными, для каждого числа M точку ξ_j можно выбрать так, чтобы выполнялось неравенство $|S_T(\xi)| > M$. Таким образом, интегральные суммы Римана неограничены и, значит, функция f не интегрируема.

Теорема доказана.

Таким образом, ограниченность функции необходима для её интегрируемости. Но это условие не является достаточным, что видно на примере функции Дирихле

$$D(x) := \begin{cases} 1, & \text{если } x \text{ рационально,} \\ 0, & \text{если } x \text{ иррационально.} \end{cases}$$

Для каждого разбиения T отрезка $[a, b]$ можно выбрать точки ξ_k, для которых $D(\xi_k) = 1$ при всех k. Тогда $S_T(\xi) = b - a$. А взяв точки ξ_k так, чтобы при всех k выполнялись равенства $D(\xi_k) = 0$, получим $S_T(\xi) = 0$. Таким образом, функция Дирихле не интегрируема ни на каком отрезке $[a, b]$.

§ 9.2. Условия интегрируемости. Суммы Дарбу

В силу теоремы 9.1.1 вопрос о существовании интеграла нужно рассматривать только для ограниченных функций.

Если функция f ограничена на $[a, b]$ и её значениями являются действительные числа, то для каждого разбиения T конечны величины

$$M_k(f) = M_k := \sup_{x\in[x_{k-1},x_k]} f(x), \qquad m_k(f) = m_k := \inf_{x\in[x_{k-1},x_k]} f(x). \tag{9.2.1}$$

Определение. Сумму

$$\overline{S_T}(f) = \overline{S_T} := \sum_{k=1}^{n} M_k(f)\Delta x_k$$

называют *верхней*, а сумму

$$\underline{S_T}(f) = \underline{S_T} := \sum_{k=1}^{n} m_k(f)\Delta x_k$$

– *нижней интегральными суммами Дарбу* функции $f(x)$, соответствующими разбиению T.

Говоря о суммах Дарбу, слово "интегральная" будем для краткости опускать.

В отличие от интегральных сумм Римана, которые определяются разбиением T и выбором точек ξ, суммы Дарбу зависят только от разбиения T.

Понятно, что если функции $f(x)$ и $g(x)$ ограничены на отрезке и $f(x) \leqslant g(x)$ при всех x, то $\overline{S_T}(f) \leqslant \overline{S_T}(g)$ и $\underline{S_T}(f) \leqslant \underline{S_T}(g)$ для любого разбиения T.

Сравнивая интегральные суммы Дарбу и Римана, видим, что для каждого разбиения T при произвольном выборе точек $\xi_k \in [x_{k-1}, x_k]$ справедливы неравенства

$$\underline{S_T} \leqslant S_T(\xi) \leqslant \overline{S_T}.$$

При этом согласно определению величин M_k и m_k

$$\sup_{\xi} S_T(\xi) = \overline{S_T} \tag{9.2.2}$$

и

$$\inf_{\xi} S_T(\xi) = \underline{S_T}. \tag{9.2.3}$$

Лемма 9.2.1. *Пусть для функции f на отрезке выполнено условие $|f(x)| \leqslant B$. Если к произвольному разбиению T этого отрезка добавить ещё одну точку, то верхняя сумма Дарбу не увеличится, а уменьшиться может не более, чем на $2B\lambda_T$. Нижняя сумма Дарбу при этом может только увеличиться, но не более, чем на $2B\lambda_T$.*

Доказательство. Пусть точка x^*, которую добавили к разбиению T, попала в отрезок $[x_{m-1}, x_m]$ разбиения T. Тогда по крайней мере на одном из отрезков $[x_{m-1}, x^*]$ и $[x^*, x_m]$ верхняя грань значений $f(x)$ не изменится, а на другом верхняя грань значений $f(x)$ может уменьшиться, но не более, чем на $2B$. Поэтому верхняя сумма Дарбу не может увеличиться, а уменьшиться может не более, чем на $2B\Delta x_m$, т.е. не более, чем на $2B\lambda_T$.

Для нижней суммы Дарбу доказательство аналогично.

Теорема 9.2.2. *Если функция ограничена на отрезке, то для любых двух разбиений T_1 и T_2 этого отрезка справедливо неравенство*

$$\underline{S_{T_1}} \leqslant \overline{S_{T_2}}. \tag{9.2.4}$$

Доказательство. Рассмотрим разбиение, образованное всеми точками разбиений T_1 и T_2. Обозначим его $T := T_1 \cup T_2$.

Так как разбиение T получено за счёт добавления новых точек и к T_1, и к T_2, то согласно лемме 9.2.1 $\overline{S_T} \leqslant \overline{S_{T_2}}$ и $\underline{S_{T_1}} \leqslant \underline{S_T}$. Поскольку $\underline{S_T} \leqslant \overline{S_T}$, отсюда вытекает оценка (9.2.4). Теорема доказана.

Из теоремы 9.2.2 следует, что для каждой ограниченной функции конечна нижняя грань $\inf_T \overline{S_T}$ и $\underline{S_{T_1}} \leqslant \inf_T \overline{S_T}$ для любого разбиения T_1. Обозначим

$$I^*(f) := \inf_T \overline{S_T}.$$

Число $I^*(f)$ называют *верхним интегралом Дарбу* функции f. Аналогично, величину

$$I_*(f) := \sup_T \underline{S_T}$$

называют *нижним интегралом Дарбу* функции f.

Таким образом, для каждой ограниченной функции f верхний и нижний интегралы Дарбу существуют и для произвольного разбиения T выполняются неравенства

$$\underline{S_T}(f) \leqslant I_*(f) \leqslant I^*(f) \leqslant \overline{S_T}(f). \tag{9.2.5}$$

Теорема 9.2.3. *Для интегрируемости ограниченной на отрезке функции f необходимо и достаточно, чтобы для каждого $\varepsilon > 0$ существовало $\delta > 0$ такое, что для каждого разбиения T, диаметр которого $\lambda_T < \delta$, справедлива оценка*

$$\overline{S_T}(f) - \underline{S_T}(f) < \varepsilon. \tag{9.2.6}$$

Доказательство. Пусть функция f интегрируема и I – её интеграл.

Согласно определению интеграла для каждого $\varepsilon > 0$ существует $\delta > 0$ такое, что для произвольного разбиения T, диаметр которого $\lambda_T < \delta$, и любого набора точек ξ_k, принадлежащих отрезкам разбиения $[x_{k-1}, x_k]$, для интегральных сумм Римана $S_T(f, \xi)$ выполняются оценки

$$I - \frac{\varepsilon}{3} < S_T(f, \xi) < I + \frac{\varepsilon}{3}\,.$$

Переходя в этом двойном неравенстве к верхним и нижним граням по ξ и пользуясь равенствами (9.2.2) и (9.2.3), находим

$$I - \frac{\varepsilon}{3} \leqslant \underline{S_T}(f) \leqslant \overline{S_T}(f) \leqslant I + \frac{\varepsilon}{3}\,.$$

Отсюда следует, что

$$\overline{S_T}(f) - \underline{S_T}(f) \leqslant \left(I + \frac{\varepsilon}{3}\right) - \left(I - \frac{\varepsilon}{3}\right) = \frac{2\varepsilon}{3} < \varepsilon$$

и необходимость условия теоремы установлена.

Докажем достаточность. По $\varepsilon > 0$ находим $\delta > 0$ такое, что $\overline{S_T} - \underline{S_T} < \varepsilon$ для любого разбиения T, для которого $\lambda_T < \delta$.

Согласно (9.2.5)

$$\underline{S_T}(f) \leqslant I^* \leqslant \overline{S_T}(f),$$

а для интегральных сумм Римана при любом выборе точек ξ имеем

$$\underline{S_T}(f) \leqslant S_T(f, \xi) \leqslant \overline{S_T}(f).$$

Значит, для произвольных точек ξ

$$|S_T(f,\xi) - I^*| \leqslant \overline{S_T}(f) - \underline{S_T}(f) < \varepsilon.$$

Так как эта оценка имеет место для каждого разбиения T, для которого $\lambda_T < \delta$, то функция f интегрируема и её интеграл Римана равен верхнему интегралу Дарбу.

Теорема доказана.

Если при доказательстве достаточности брать не верхний, а нижний интеграл Дарбу, то получим равенство значений интеграла Римана и нижнего интеграла Дарбу. Следовательно, для интегрируемых функций значения интеграла Римана, верхнего и нижнего интегралов Дарбу совпадают.

В дальнейшем будет показано, что из равенства верхнего и нижнего интегралов Дарбу вытекает интегрируемость функции, т.е. справедливо следующее утверждение.

Теорема 9.2.4. *Для интегрируемости ограниченной на отрезке $[a,b]$ функции f необходимо и достаточно равенство её верхнего и нижнего интегралов Дарбу. Если это условие выполнено, то*

$$\int_a^b f(x)\,dx = I^*(f) = I_*(f).$$

Согласно теореме 9.2.3 для интегрируемости функции f необходимо и достаточно, чтобы условие (9.2.6) выполнялось для всех разбиений, диаметр которых мал. Покажем, что выполнение этого условия можно требовать только для одного разбиения.

Теорема 9.2.5. *Для интегрируемости ограниченной на отрезке $[a,b]$ функции f необходимо и достаточно, чтобы для каждого положительного числа ε существовало разбиение T отрезка $[a,b]$ такое, что*

$$\overline{S_T}(f) - \underline{S_T}(f) < \varepsilon.$$

Доказательство. Необходимость вытекает из теоремы 9.2.3. Докажем достаточность.

Задав $\varepsilon > 0$, найдём разбиение T^* такое, что

$$\overline{S_{T^*}} - \underline{S_{T^*}} < \frac{\varepsilon}{2}\,. \tag{9.2.7}$$

Пусть N – число отрезков, на которые T^* разбивает отрезок $[a,b]$, и $|f(x)| < B$ при всех $x \in [a,b]$.

Выберем $\delta > 0$ так, чтобы выполнялось неравенство

$$2BN\delta < \frac{\varepsilon}{4}, \tag{9.2.8}$$

и рассмотрим произвольное разбиение T, диаметр которого $\lambda_T < \delta$.

Введём разбиение $T' := T \cup T^*$. Так как T' получено из T^* добавлением точек разбиения T, то

$$\underline{S_{T^*}} \leqslant \underline{S_{T'}} \leqslant \overline{S_{T'}} \leqslant \overline{S_{T^*}}.$$

Значит, в силу (9.2.7)

$$\overline{S_{T'}} - \underline{S_{T'}} < \frac{\varepsilon}{2}. \tag{9.2.9}$$

Теперь воспользуемся тем, что разбиение T' получено из T добавлением к нему точек разбиения T^*, лежащих внутри отрезка $[a, b]$, т.е. добавлением не более $N-1$ точки.

Согласно лемме 9.2.1 от добавления к разбиению T одной точки верхняя сумма Дарбу $\overline{S_T}$ может уменьшиться не более, чем на $2B\lambda_T$. Поэтому

$$\overline{S_T} \leqslant \overline{S_{T'}} + 2B\lambda_T(N-1) < \overline{S_{T'}} + 2B\delta N$$

и в силу (9.2.8)

$$\overline{S_T} < \overline{S_{T'}} + \frac{\varepsilon}{4}.$$

Поскольку при добавлении к T одной точки нижняя сумма Дарбу $\underline{S_T}$ может увеличиться не более, чем на $2B\lambda_T$, то точно так же находим

$$\underline{S_T} > \underline{S_{T'}} - \frac{\varepsilon}{4}.$$

Из этих оценок и (9.2.9) получаем

$$\overline{S_T} - \underline{S_T} < \overline{S_{T'}} + \frac{\varepsilon}{4} - \left(\underline{S_{T'}} - \frac{\varepsilon}{4}\right) = \overline{S_{T'}} - \underline{S_{T'}} + \frac{\varepsilon}{2} < \varepsilon.$$

Так как T – произвольное разбиение, для которого $\lambda_T < \delta$, то согласно теореме 9.2.3 функция f интегрируема.

Теорема доказана.

Теперь завершим доказательство теоремы 9.2.4. Согласно сказанному выше осталось показать, что функция интегрируема, если равны её верхний и нижний интегралы Дарбу.

Пусть I обозначает общее значение верхнего и нижнего интегралов Дарбу.

Согласно определению интегралов Дарбу для каждого $\varepsilon > 0$ существуют такие разбиения T_1 и T_2, что

$$I - \frac{\varepsilon}{2} < \underline{S_{T_1}} \leqslant I, \qquad I \leqslant \overline{S_{T_2}} < I + \frac{\varepsilon}{2}\,.$$

Для разбиения $T^* := T_1 \cup T_2$ имеем

$$I - \frac{\varepsilon}{2} < \underline{S_{T_1}} \leqslant \underline{S_{T^*}} \leqslant \overline{S_{T^*}} \leqslant \overline{S_{T_2}} < I + \frac{\varepsilon}{2}$$

и, значит,

$$\overline{S_{T^*}} - \underline{S_{T^*}} < I + \frac{\varepsilon}{2} - \left(I - \frac{\varepsilon}{2}\right) = \varepsilon.$$

Отсюда в силу теоремы 9.2.5 вытекает интегрируемость функции f.

Теперь теорема 9.2.4 доказана полностью.

С помощью доказанных теорем легко установить интегрируемость непрерывных функций и интегрируемость монотонных функций.

Теорема 9.2.6. *Функция, непрерывная на отрезке, интегрируема на этом отрезке.*

Доказательство. Непрерывная на отрезке $[a, b]$ функция f по теореме Кантора 4.4.1 равномерно непрерывна на этом отрезке. Значит, согласно теореме 4.4.4 для модуля непрерывности функции f на $[a, b]$ имеем $\omega(f, +0) = 0$.

Для произвольного разбиения T отрезка $[a, b]$ справедливо равенство

$$\overline{S_T}(f) - \underline{S_T}(f) = \sum_{k=1}^{n} (M_k(f) - m_k(f))\Delta x_k,$$

где числа $M_k(f)$ и $m_k(f)$ заданы формулами (9.2.1). Так как функция f непрерывна, то в силу теоремы 4.3.2 для каждого k существуют точки ξ_k и η_k из отрезка $[x_{k-1}, x_k]$, для которых $M_k(f) =$

$f(\xi_k)$ и $m_k(f) = f(\eta_k)$. Обе точки ξ_k и η_k принадлежат отрезку $[x_{k-1}, x_k]$, значит, $|\xi_k - \eta_k| \leqslant \Delta x_k \leqslant \lambda_T$ и

$$M_k(f) - m_k(f) = f(\xi_k) - f(\eta_k) \leqslant \omega(f, \lambda_T).$$

Поэтому

$$\overline{S_T}(f) - \underline{S_T}(f) \leqslant \sum_{k=1}^{n} \omega(f, \lambda_T)\Delta x_k = \omega(f, \lambda_T)(b-a). \tag{9.2.10}$$

Поскольку $\omega(f, +0) = 0$, для каждого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$ такое, что $\omega(f, \lambda_T)(b-a) < \varepsilon$, если только $\lambda_T < \delta$. Значит, для таких разбиений T

$$\overline{S_T}(f) - \underline{S_T}(f) < \varepsilon,$$

откуда согласно теореме 9.2.3 вытекает интегрируемость функции f.

Теорема доказана.

Теорема 9.2.7. *Функция, монотонная на отрезке, интегрируема на этом отрезке.*

Доказательство. Будем для определённости считать, что функция f возрастает (не обязательно строго) на отрезке $[a, b]$ и $f(a) < f(b)$, иначе утверждение теоремы очевидно.

Для любого разбиения T имеем

$$\begin{aligned}\overline{S_T}(f) - \underline{S_T}(f) &= \sum_{k=1}^{n} (M_k(f) - m_k(f))\Delta x_k = \\ &= \sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))\Delta x_k \leqslant \\ &\leqslant \sum_{k=1}^{n} (f(x_k) - f(x_{k-1}))\lambda_T = (f(b) - f(a))\lambda_T.\end{aligned}$$

По $\varepsilon > 0$ выберем разбиение T, для которого

$$\lambda_T < \frac{\varepsilon}{f(b) - f(a)}\,.$$

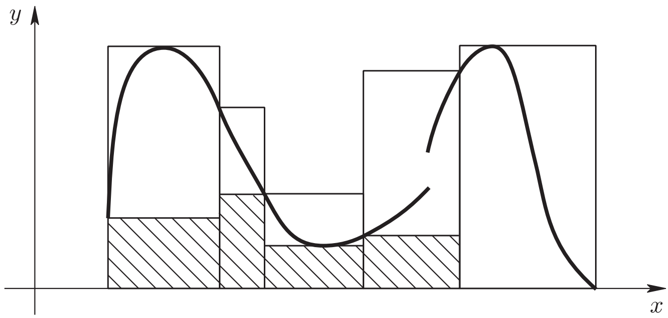
Тогда $\overline{S_T}(f) - \underline{S_T}(f) < \varepsilon$ и теорема доказана.

Выясним геометрический смысл интегральных сумм Дарбу неотрицательных функций.

Пусть функция $f(x)$ неотрицательна на отрезке $[a, b]$. Рассмотрим множество A точек (x, y) плоскости, удовлетворяющих условиям $a \leqslant x \leqslant b$, $0 \leqslant y \leqslant f(x)$. Таким образом, множество A ограничено сверху графиком функции $f(x)$, снизу отрезком $[a, b]$ оси OX, слева – отрезком, соединяющим точки $(a, 0)$ и $(a, f(a))$, справа – отрезком, соединяющим $(b, 0)$ и $(b, f(b))$, Такое множество называют *криволинейной трапецией*.

Пусть T – разбиение отрезка $[a, b]$ точками $a = x_0 < x_1 < \cdots < x_n = b$, а числа $M_k(f)$ и $m_k(f)$ определены формулами (9.2.1). Произведение $M_k(f)\Delta x_k$ можно рассматривать как площадь прямоугольника, основанием которого является отрезок $[x_{k-1}, x_k]$, а высота равна $M_k(f)$. Точно также, произведение $m_k(f)\Delta x_k$ равно площади прямоугольника с тем же основанием, высота которого равна $m_k(f)$.

Таким образом, верхняя сумма Дарбу $\overline{S_T}(f)$ равна сумме площадей прямоугольников, порождённых разбиением T, содержащих множество A. Нижняя сумма Дарбу $\underline{S_T}(f)$ равна сумме площадей прямоугольников, содержащихся во множестве A.



Если функция f интегрируема на отрезке $[a, b]$, то для каждого $\varepsilon > 0$ можно найти разбиение T такое, что $\overline{S_T}(f) - \underline{S_T}(f) < \varepsilon$. Значит, разность суммы площадей прямоугольников, содержащих множество A, и прямоугольников, содержащихся в A, меньше ε. Для интегрируемых функций существуют разбиения, для которых верхняя и нижняя суммы Дарбу как угодно мало отличаются от интеграла $\int_a^b f(x)\,dx$. Поэтому значение этого интеграла называют площадью множества A.

В гл. 20 вопрос о площадях будет рассматриваться специально и будет видно, что это определение площади множества A согласуется с общим определением площади. А сейчас сформулируем вывод: *определённый интеграл от неотрицательной интегрируемой функции равен площади криволинейной трапеции, ограниченной сверху графиком этой функции.*

§ 9.3. Линейные свойства определённого интеграла

Теорема 9.3.1 (Аддитивность интеграла относительно промежутка интегрирования). *Пусть функция $f(x)$ задана на отрезке $[a,b]$ и $a<c<b$. Если функция f интегрируема на отрезках $[a,c]$ и $[c,b]$, то она интегрируема на $[a,b]$. Если функция f интегрируема на $[a,b]$, то она интегрируема и на отрезках $[a,c]$ и $[c,b]$ и справедливо равенство*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx. \tag{9.3.1}$$

Доказательство. Если T – разбиение отрезка $[a,b]$, содержащее точку c, то точки T, принадлежащие отрезку $[a,c]$, образуют разбиение этого отрезка, которое обозначим T'. Аналогичное разбиение отрезка $[c,b]$ обозначим T''.

Имеем

$$\overline{S_T} - \underline{S_T} = (\overline{S_{T'}} - \underline{S_{T'}}) + (\overline{S_{T''}} - \underline{S_{T''}}). \tag{9.3.2}$$

Если f интегрируема на $[a,b]$, то согласно теореме 9.2.5 существует разбиение T, при котором как угодно мала разность в левой части (9.3.2). Значит, малы и обе разности верхних и нижних сумм Дарбу из правой части (9.3.2), так как такие разности всегда неотрицательны. Поэтому f интегрируема на отрезках $[a,c]$ и $[c,b]$. Наоборот, если малы разности в правой части (9.3.2), то мала и разность в левой части. Значит, из интегрируемости функции на $[a,c]$ и $[c,b]$ следует её интегрируемость на $[a,b]$.

Докажем равенство (9.3.1).

Задав положительное ε, найдём $\delta>0$ такое, что для каждого разбиения T отрезка $[a,b]$ с $\lambda_T<\delta$ при любом выборе точек ξ_k

$$\left|\int_a^b f(x)\,dx - S_T(f,\xi)\right| < \frac{\varepsilon}{3}\,. \tag{9.3.3}$$

Возьмём интегральные суммы Римана $S_{T'}(f,\xi')$ и $S_{T''}(f,\xi'')$ функции f на отрезках $[a,c]$ и $[c,b]$ соответственно такие, что $\lambda_{T'} < \delta$ и $\lambda_{T''} < \delta$ и

$$\left|\int_a^c f(x)\,dx - S_{T'}(f,\xi')\right| < \frac{\varepsilon}{3}, \qquad \left|\int_c^b f(x)\,dx - S_{T''}(f,\xi'')\right| < \frac{\varepsilon}{3}.$$

Тогда $S_{T'}(f,\xi')+S_{T''}(f,\xi'')$ является интегральной суммой Римана функции f на отрезке $[a,b]$, соответствующей разбиению, диаметр которого меньше δ. Значит, для этой интегральной суммы справедлива оценка (9.3.3). Таким образом,

$$\begin{aligned}&\left|\int_a^b f(x)\,dx - \int_a^c f(x)\,dx - \int_c^b f(x)\,dx\right| \leqslant \\ &\qquad\leqslant \left|\int_a^b f(x)\,dx - (S_{T'}(f,\xi') + S_{T''}(f,\xi''))\right| + \\ &\qquad\qquad + \left|\int_a^c f(x)\,dx - S_{T'}(f,\xi')\right| + \left|\int_c^b f(x)\,dx - S_{T''}(f,\xi'')\right| < \varepsilon.\end{aligned}$$

Отсюда вытекает (9.3.1), так как левая часть в полученном неравенстве не зависит от ε.

Теорема доказана.

Из теоремы 9.3.1 следует, что функция, интегрируемая на отрезке, интегрируема и на любом содержащемся в нём отрезке.

В определении интеграла $\int_a^b f dx$ нижний предел интегрирования a был меньше верхнего предела b. Но нередко нужны и интегралы, у которых нижний предел интегрирования больше верхнего. По определению полагают при $a > b$

$$\int_a^b f(x)dx := -\int_b^a f(x)dx$$

и

$$\int_a^a f(x)dx := 0.$$

Эти определения позволяют распространить равенство (9.3.1) на любое расположение точек a, b, c при условии, что функция f интегрируема на наибольшем отрезке с концами в этих точках.

Теорема 9.3.2 (Линейность интеграла). *Если функции $f(x)$ и $g(x)$ интегрируемы на отрезке $[a,b]$, то для любых чисел α и*

β функция $\alpha f(x)+\beta g(x)$ интегрируема на $[a,b]$ и имеет место равенство

$$\int_a^b (\alpha f(x)+\beta g(x))\,dx = \alpha\int_a^b f(x)\,dx + \beta\int_a^b g(x)\,dx. \tag{9.3.4}$$

Доказательство. Для каждого разбиения T отрезка $[a,b]$ при любом выборе точек ξ_k справедливо равенство интегральных сумм Римана

$$S_T(\alpha f+\beta g,\xi)=\sum_{k=1}^{n}[\alpha f(\xi_k)+\beta g(\xi_k)]\Delta x_k = \alpha S_T(f,\xi)+\beta S_T(g,\xi).$$

Пользуясь интегрируемостью функций f и g, по заданному $\varepsilon>0$ выбираем $\delta>0$ такое, что для каждого разбиения T, диаметр которого $\lambda_T<\delta$, и любых точек ξ_k сумма $S_T(f,\xi)$ отличается от интеграла $\int_a^b f\,dx$ меньше, чем на ε, и $S_T(g,\xi)$ отличается от $\int_a^b g\,dx$ меньше, чем на ε. Тогда сумма $S_T(\alpha f+\beta g,\xi)$ отличается от числа

$$\alpha\int_a^b f(x)\,dx+\beta\int_a^b g(x)\,dx$$

меньше, чем на $(|\alpha|+|\beta|)\varepsilon$. Отсюда вытекают интегрируемость функции $\alpha f+\beta g$ и равенство (9.3.4).

Теорема доказана.

Эта теорема показывает, что функции, интегрируемые на отрезке, образуют линейное пространство.

Возможность вынесения числового множителя из-под знака определённого интеграла называют *однородностью* интеграла.

Понятно, что равенство интегралов (9.3.4) выполняется и при $a\geqslant b$.

Отметим, что в доказательстве теоремы 9.3.2 суммы Дарбу не использовались, оно опиралось только на определение интеграла. Следовательно, эта теорема справедлива и для комплекснозначных функций. Поэтому, если $f(x)=u(x)+iv(x)$ – разложение функции f на действительную и мнимую составляющие, то

$$\int_a^b f(x)\,dx=\int_a^b u(x)\,dx+i\int_a^b v(x)\,dx.$$

Эту формулу можно было бы использовать как определение интеграла от комплекснозначных функций.

Докажем теорему об интегрировании неравенств.

Теорема 9.3.3. *Если функции $f(x)$ и $g(x)$ интегрируемы на отрезке $[a,b]$ и $f(x) \leqslant g(x)$ при всех $x \in [a,b]$, то справедливо неравенство*

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx. \tag{9.3.5}$$

Доказательство. Так как разность $g(x) - f(x)$ неотрицательна, то неотрицательны и все её интегральные суммы Римана, поэтому

$$\int_a^b (g(x) - f(x))\,dx \geqslant 0.$$

Отсюда, пользуясь линейностью интеграла, получаем (9.3.5).

Понятно, что для справедливости оценки (9.3.5) условие $a < b$ существенно.

Следующее утверждение иногда называют “основной леммой вариационного исчисления”.

Теорема 9.3.4. *Если для непрерывной неотрицательной на $[a,b]$ функции $f(x)$*

$$\int_a^b f(x)dx = 0,$$

то $f(x)$ тождественно равна нулю на $[a,b]$.

Доказательство. Если $f(x_0) > 0$ в некоторой точке $x_0 \in [a,b]$, то существует окрестность точки x_0 такая, что для всех x из этой окрестности $f(x) > f(x_0)/2$.

Пусть отрезок $[c,d]$ принадлежит указанной окрестности. В силу неотрицательности функции f имеем

$$\int_a^b f(x)dx \geqslant \int_c^d f(x)dx \geqslant \int_c^d \frac{f(x_0)}{2}dx = \frac{f(x_0)}{2}(d-c) > 0,$$

что противоречит условию теоремы. Значит, $f(x) \equiv 0$.

Теорема доказана.

Теорема 9.3.5. *Если изменить значения интегрируемой на отрезке функции в конечном числе точек, то получим интегрируемую функцию и значение интеграла не изменится.*

Доказательство. Рассмотрим функцию $g(x)$, которая равна 1 в некоторой точке x^0 отрезка и равна нулю во всех остальных его точках. Тогда для каждого разбиения T имеем $\underline{S_T}(g) = 0$ и $\overline{S_T}(g) \leqslant 2\lambda_T$. Поэтому функция g интегрируема и интеграл от g равен нулю.

Изменить значение интегрируемой функции $f(x)$ в точке x^0 можно, прибавив к ней функцию $g(x)$, умноженную на соответствующее число. Так как интеграл от функции g равен нулю, то, пользуясь линейностью интеграла, видим, что полученная функция интегрируема и значение интеграла от функции f не изменится. Повторив это рассуждение для каждой точки, в которой изменено значение функции, получим утверждение теоремы.

Таким образом, можно говорить об интегрировании на отрезке функций, которые не определены на конечном множестве точек отрезка.

Теоремы 9.3.1 и 9.3.5 вместе с теоремами 9.2.6 и 9.2.7 позволяют сделать вывод об интегрируемости на отрезке кусочно непрерывных и кусочно монотонных функций.

§ 9.4. Интегрируемость сложной функции

Теорема 9.4.1. *Пусть функция $y = f(x)$ интегрируема на отрезке $[a, b]$ и её значения принадлежат отрезку $[c, d]$. Если функция $\varphi(y)$ непрерывна на $[c, d]$, то сложная функция $\psi(x) := \varphi(f(x))$ интегрируема на $[a, b]$.*

Доказательство. Для произвольного $\varepsilon > 0$ в силу равномерной непрерывности функции φ на $[c, d]$ существует $\delta > 0$ такое, что

$$|\varphi(s) - \varphi(t)| < \varepsilon, \tag{9.4.1}$$

если $s, t \in [c, d]$ и $|s - t| < \delta$. Будем также считать $\delta < \varepsilon$.

Пусть T – разбиение отрезка $[a, b]$ точками $a = x_0 < x_1 < \dots < x_n = b$, для которого $\lambda_T < \delta$ и

$$\overline{S_T}(f) - \underline{S_T}(f) < \delta^2. \tag{9.4.2}$$

В представлении разности

$$\overline{S_T}(f) - \underline{S_T}(f) = \sum_{k=1}^{n} (M_k(f) - m_k(f))\Delta x_k$$

разобьём индексы k на две группы I и II. Положим $k \in \mathrm{I}$, если

$$M_k(f) - m_k(f) < \delta,$$

и $k \in \mathrm{II}$, если

$$M_k(f) - m_k(f) \geqslant \delta. \tag{9.4.3}$$

Если $k \in \mathrm{I}$, то для произвольных точек ξ и η из $[x_{k-1}, x_k]$ имеем

$$|f(\xi) - f(\eta)| \leqslant M_k(f) - m_k(f) < \delta$$

и, значит, согласно (9.4.1)

$$|\psi(\xi) - \psi(\eta)| = |\varphi(f(\xi)) - \varphi(f(\eta))| < \varepsilon.$$

Отсюда следует оценка $M_k(\psi) - m_k(\psi) \leqslant \varepsilon$, поскольку как нетрудно убедиться с помощью равенств (9.2.2) и (9.2.3),

$$\sup_{\xi,\eta\in[x_{k-1},x_k]} |\psi(\xi) - \psi(\eta)| = M_k(\psi) - m_k(\psi).$$

Таким образом,

$$\sum_{k\in I}(M_k(\psi) - m_k(\psi))\Delta x_k \leqslant \varepsilon \sum_{k\in I} \Delta x_k \leqslant \varepsilon(b-a). \tag{9.4.4}$$

Рассмотрим теперь $k \in \mathrm{II}$. В силу (9.4.2)

$$\sum_{k\in \mathrm{II}}(M_k(f) - m_k(f))\Delta x_k \leqslant \overline{S_T}(f) - \underline{S_T}(f) < \delta^2.$$

Но согласно (9.4.3)

$$\sum_{k\in \mathrm{II}}(M_k(f) - m_k(f))\Delta x_k \geqslant \delta \sum_{k\in \mathrm{II}} \Delta x_k.$$

Следовательно,

$$\sum_{k\in \mathrm{II}} \Delta x_k < \delta < \varepsilon$$

и, таким образом, если $|\varphi(y)| \leqslant L$ на $[c, d]$, то

$$\sum_{k\in \mathrm{II}}(M_k(\psi) - m_k(\psi))\Delta x_k \leqslant 2L \sum_{k\in \mathrm{II}} \Delta x_k < 2L\varepsilon.$$

Отсюда и из (9.4.4) находим

$$\overline{S_T}(\psi) - \underline{S_T}(\psi) < (b - a + 2L)\varepsilon.$$

Ввиду произвольности ε из этой оценки согласно теореме 9.2.5 вытекает интегрируемость функции $\psi(x)$.

Теорема доказана.

Следствие 9.4.2. *Если функция $f(x)$ интегрируема на отрезке, то для каждого $r > 0$ функция $|f(x)|^r$ интегрируема на этом отрезке.*

Это утверждение следует из теоремы 9.4.1 при $\varphi(y) = |y|^r$.

В частности, при $r = 1$ получаем: если $f(x) \in \mathcal{R}[a, b]$, то $|f(x)| \in \mathcal{R}[a, b]$. При этом справедлива оценка

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx, \tag{9.4.5}$$

вытекающая из неравенства $|f(x)| - f(x) \geqslant 0$.

Заметим, что из интегрируемости модуля $|f(x)|$ интегрируемость самой функции $f(x)$ не следует. Действительно, функция

$$g(x) := \begin{cases} 1, & \text{если } x \text{ рационально}, \\ -1, & \text{если } x \text{ иррационально}, \end{cases}$$

не интегрируема ни на каком отрезке, в то же время как её модуль $|g(x)| \equiv 1$ интегрируем.

Для справедливости оценки (9.4.5) условие $a < b$ существенно. Если на взаимное расположение точек a и b не накладывать ограничений, то вместо (9.4.5) нужно писать

$$\left|\int_a^b f(x)\,dx\right| \leqslant \left|\int_a^b |f(x)|\,dx\right|.$$

Теорема 9.4.3. *Если на отрезке $[a, b]$ функция $f(x)$ интегрируема и $|f(x)| \geqslant d$ для некоторого положительного числа d, то функция $1/f(x)$ интегрируема на $[a, b]$.*

Доказательство. Пусть T – разбиение отрезка $[a, b]$ на отрезки $[x_{k-1}, x_k]$, $k = 1, \ldots, n$. Для любых точек ξ и η из $[x_{k-1}, x_k]$ справедлива оценка

$$\frac{1}{f(\xi)} - \frac{1}{f(\eta)} \leqslant \left|\frac{f(\eta) - f(\xi)}{f(\xi)f(\eta)}\right| \leqslant \frac{1}{d^2}(M_k(f) - m_k(f)).$$

Отсюда, опираясь, как и в доказательстве теоремы 9.4.1, на равенства (9.2.2) и (9.2.3), находим

$$M_k\left(\frac{1}{f}\right) - m_k\left(\frac{1}{f}\right) \leqslant \frac{1}{d^2}(M_k(f) - m_k(f)).$$

Значит,

$$\overline{S_T}\left(\frac{1}{f}\right) - \underline{S_T}\left(\frac{1}{f}\right) \leqslant \frac{1}{d^2}(\overline{S_T}(f) - \underline{S_T}(f)).$$

Так как за счёт выбора разбиения T выражение в правой части этого неравенства можно сделать как угодно малым, то мала будет и разность в левой части, что согласно теореме 9.2.5 обеспечивает интегрируемость функции $1/f(x)$.

Теорема доказана.

Теорема 9.4.4. *Если на отрезке функции $f(x)$ и $g(x)$ интегрируемы, то интегрируемо и их произведение $f(x)g(x)$.*

Доказательство. Имеем

$$4f(x)g(x) \equiv (f(x) + g(x))^2 - (f(x) - g(x))^2.$$

Пользуясь линейностью интеграла и следствием 9.4.2, получаем интегрируемость функции в правой части этого тождества. Значит, интегрируема функция из левой части и теорема доказана.

В дальнейшем понадобится следующее простое утверждение.

Лемма 9.4.5. *Пусть функция $f(x)$ интегрируема на отрезке $[a, b]$ и для некоторого положительного числа r*

$$\int_a^b |f(x)|^r\, dx = 0. \tag{9.4.6}$$

Тогда для произвольного положительного q

$$\int_a^b |f(x)|^q\, dx = 0.$$

Доказательство. Пусть для упрощения записей $f(x) \geqslant 0$. Из (9.4.6) следует, что нижний интеграл Дарбу функции $f^r(x)$ равен нулю. Значит, $\underline{S_T}(f^r) = 0$ для любого разбиения T. Отсюда следует, что $m_k(f^r) = 0$ на каждом отрезке $[x_{k-1}, x_k]$ разбиения T. Поэтому $m_k(f^q) = 0$ на каждом таком отрезке и равен нулю нижний интеграл Дарбу функции $f^q(x)$, а так как эта функция интегрируема, равен нулю и её интеграл Римана. Лемма доказана.

§ 9.5. Приближение интегрируемых функций ступенчатыми и непрерывными функциями

Определение. Функция $f(x)$ называется *ступенчатой* на отрезке $[a,b]$, если она кусочно постоянна, т.е. существует такое разбиение $a = x_0 < x_1 < \dots < x_n = b$, что $f(x)$ постоянна на каждом интервале (x_{k-1}, x_k), $k = 1, \dots, n$.

Ступенчатые функции кусочно непрерывны и, следовательно, интегрируемы.

Теорема 9.5.1. *Для интегрируемости на отрезке $[a,b]$ функции $f(x)$ необходимо и достаточно, чтобы для каждого $\varepsilon > 0$ существовали такие ступенчатые функции $p(x)$ и $q(x)$, что $p(x) \leqslant f(x) \leqslant q(x)$ на $[a,b]$ и*

$$\int_a^b (q(x) - p(x))\,dx < \varepsilon. \tag{9.5.1}$$

Доказательство. Если функция f интегрируема, то согласно теореме 9.2.5 для каждого $\varepsilon > 0$ существует разбиение T отрезка $[a,b]$ такое, что

$$\overline{S_T}(f) - \underline{S_T}(f) < \varepsilon. \tag{9.5.2}$$

На интервалах (x_{k-1}, x_k) положим $q(x) := M_k(f)$ и $p(x) := m_k(f)$. В точках x_k разбиения T пусть функция $q(x)$ равна большему из значений, принимаемых ею на примыкающих к x_k интервалах разбиения T, а $p(x)$ равна меньшему из её значений на таких интервалах.

Тогда $q(x)$ и $p(x)$ – ступенчатые на $[a,b]$ функции, $p(x) \leqslant f(x) \leqslant q(x)$ и

$$\int_a^b q(x)\,dx = \overline{S_T}(f), \qquad \int_a^b p(x)\,dx = \underline{S_T}(f).$$

Таким образом, из (9.5.2) следует оценка (9.5.1) и необходимость условия теоремы доказана.

Докажем достаточность. Пусть $p(x)$ и $q(x)$ – произвольные ступенчатые на отрезке $[a,b]$ функции, удовлетворяющие условию $p(x) \leqslant f(x) \leqslant q(x)$.

Если T_q– разбиение отрезка $[a, b]$, содержащее концы отрезка и точки разрыва функции $q(x)$, то

$$\int_a^b q(x)\,dx = \overline{S_{T_q}}(q).$$

Точно также строим разбиение T_p, для которого

$$\int_a^b p(x)\,dx = \underline{S_{T_p}}(p).$$

Для разбиения $T := T_p \cup T_q$ имеем

$$\overline{S_T}(f) - \underline{S_T}(f) \leqslant \overline{S_{T_q}}(q) - \underline{S_{T_p}}(p) = \int_a^b (q(x) - p(x))\,dx.$$

Поэтому, если для функций $p(x)$ и $q(x)$ справедлива оценка (9.5.1), то

$$\overline{S_T}(f) - \underline{S_T}(f) < \varepsilon$$

и согласно теореме 9.2.5 функция f интегрируема.

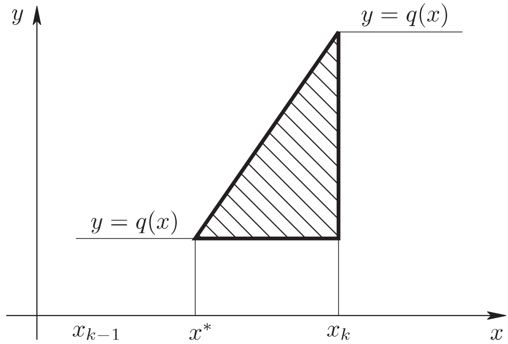
Теорема доказана.

Теорема 9.5.2. *Для интегрируемости на отрезке $[a, b]$ функции $f(x)$ необходимо и достаточно, чтобы для каждого $\varepsilon > 0$ существовали такие непрерывные на $[a, b]$ функции $u(x)$ и $v(x)$, что $u(x) \leqslant f(x) \leqslant v(x)$ и*

$$\int_a^b (v(x) - u(x))\,dx < \varepsilon.$$

Доказательство. Пусть функция f интегрируема на $[a, b]$, $\varepsilon > 0$, $p(x)$ и $q(x)$ – ступенчатые функции, построенные при доказательстве теоремы 9.5.1. Покажем, как с помощью этих функций можно определить функции $u(x)$ и $v(x)$.

Пусть T – разбиение отрезка $[a, b]$, содержащие точки a, b и все точки разрыва функции $q(x)$. Не теряя общности, будем считать, что значение функции $q(x)$ в точках разрыва равно большему из её односторонних пределов. Рассмотрим одну из таких точек разрыва x_k, считая для определённости, что в правой окрестности точки x_k значение $q(x)$ больше, чем в левой. Возьмём точку $x^* \in (x_{k-1}, x_k)$ и соединим отрезком точки $(x^*, q(x^*))$ и $(x_k, q(x_k))$ графика функции $q(x)$.



Разность $x_k - x^*$ будем считать настолько малой, что площадь треугольника, заштрихованного на рисунке, меньше ε/N, где N – число точек разбиения T.

Пусть $v(x)$ – функция, график которой включает построенные наклонные отрезки, а в остальных точках совпадает с графиком функции $q(x)$. Тогда функция $v(x)$ непрерывна, справедливо неравенство $q(x) \leqslant v(x)$ и интеграл $\int_a^b (v(x) - q(x))dx$ равен сумме площадей треугольников, построенных вблизи точек $(x_k, q(x_k))$ графика функции $q(x)$. Число точек разрыва функции $q(x)$ не превосходит N, а площадь каждого треугольника меньше ε/N. Поэтому

$$0 \leqslant \int_a^b (v(x) - q(x))\,dx < \varepsilon.$$

Аналогичным образом, отправляясь от функции $p(x)$, находим непрерывную функцию $u(x)$ такую, что $u(x) \leqslant p(x)$ и

$$0 \leqslant \int_a^b (p(x) - u(x))\,dx < \varepsilon.$$

По построению функции $u(x)$ и $v(x)$ удовлетворяют условиям $u(x) \leqslant f(x) \leqslant v(x)$ и справедлива оценка

$$\begin{aligned} 0 \leqslant \int_a^b (v(x) - u(x))\,dx &= \\ = \int_a^b (v(x) - q(x))\,dx &+ \\ + \int_a^b (q(x) - p(x))\,dx &+ \int_a^b (p(x) - u(x))\,dx < 3\varepsilon. \end{aligned}$$

Множитель 3 в этом неравенстве не играет существенной роли, поэтому необходимость условий теоремы доказана.

Достаточность легко получить, построив для функций $u(x)$ и $v(x)$ ступенчатые функции $p(x)$ и $q(x)$ такие, что

$$p(x) \leqslant u(x) \leqslant v(x) \leqslant q(x)$$

и

$$\int_a^b (q(x) - p(x))\, dx < \varepsilon.$$

Детали этого рассуждения приводить не будем.

Заметим, что функции $u(x)$ и $v(x)$ в теореме 9.5.2 можно считать имеющими любое число производных или даже бесконечно дифференцируемыми. Для этого в качестве фрагментов графиков функций $u(x)$ и $v(x)$ в окрестностях точек $(x_k, q(x_k))$ нужно брать достаточно гладкие кривые, а не наклонные отрезки, как было в приведенном доказательстве.

§ 9.6. Связь определённого и неопределённого интегралов

Если функция f интегрируема на отрезке $[a, b]$, то согласно теореме 9.3.1 об аддитивности интеграла относительно промежутка интегрирования f интегрируема на любом отрезке $[a, x]$ при $x \leqslant b$. Значит, на $[a, b]$ можно определить интеграл с переменным верхним пределом интегрирования

$$\varPhi(x) := \int_a^x f(t)\, dt.$$

Будем изучать свойства функции $\varPhi(x)$.

Теорема 9.6.1. *Если на отрезке $[a, b]$ функция $f(t)$ интегрируема и $|f(t)| \leqslant B$, то для функции $\varPhi(x)$ при всех $x', x'' \in [a, b]$ справедлива оценка*

$$|\varPhi(x') - \varPhi(x'')| \leqslant B|x' - x''|, \tag{9.6.1}$$

т.е. $\varPhi(x)$ удовлетворяет условию Липшица первого порядка с константой B. В частности, функция $\varPhi(x)$ непрерывна на $[a, b]$.

Доказательство. Пусть x и $x+h$ – произвольные точки отрезка $[a,b]$. Тогда

$$\Phi(x+h)-\Phi(x)=\int_a^{x+h} f(t)\,dt-\int_a^{x} f(t)\,dt=\int_x^{x+h} f(t)\,dt. \quad (9.6.2)$$

Никаких условий на знак h здесь не накладывается и верхний предел интегрирования может быть меньше нижнего.

Из (9.6.2) находим

$$|\Phi(x+h)-\Phi(x)|= \\ =\left|\int_x^{x+h} f(t)\,dt\right| \leqslant \left|\int_x^{x+h} |f(t)|\,dt\right| \leqslant \left|\int_x^{x+h} B\,dt\right| = B|h|.$$

Эта оценка только формой записи отличается от (9.6.1).

Теорема доказана.

Отметим, что функция $\Phi(x)$ не обязательно имеет производную в каждой точке. Например, если на отрезке $[-1,1]$

$$f(t):=\begin{cases}0 & \text{при } t\in[-1,0],\\ 1 & \text{при } t\in(0,1],\end{cases}$$

то

$$\Phi(x)=\begin{cases}0 & \text{для } x\in[-1,0],\\ x & \text{для } x\in[0,1]\end{cases}$$

и функция Φ не имеет производной в нуле.

Теорема 9.6.2. *Если интегрируемая на отрезке $[a,b]$ функция $f(x)$ непрерывна в некоторой точке $x_0\in[a,b]$, то функция $\Phi(x)$ верхнего предела интегрирования имеет в точке x_0 производную и*

$$\Phi'(x_0)=f(x_0).$$

Доказательство. Рассмотрим отношение приращения функции Φ к приращению аргумента в точке x_0. Если $x_0+h\in[a,b]$, то

$$\frac{\Phi(x_0+h)-\Phi(x_0)}{h}=\frac{1}{h}\int_{x_0}^{x_0+h} f(t)\,dt= \\ =f(x_0)+\frac{1}{h}\int_{x_0}^{x_0+h}[f(t)-f(x_0)]\,dt. \quad (9.6.3)$$

В силу непрерывности f в точке x_0 для каждого $\varepsilon > 0$ существует $\delta > 0$ такое, что $|f(t)-f(x_0)| < \varepsilon$, если $|t-x_0| < \delta$. Поэтому, для $|h| < \delta$

$$\left|\int_{x_0}^{x_0+h} [f(t) - f(x_0)]\, dt\right| \leqslant \\ \leqslant \left|\int_{x_0}^{x_0+h} |f(t) - f(x_0)|\, dt\right| \leqslant \left|\int_{x_0}^{x_0+h} \varepsilon\, dt\right| = \varepsilon|h|. \qquad (9.6.4)$$

Из (9.6.3) и (9.6.4) следует, что

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

Теорема доказана.

Если x_0 – один из концов отрезка, то в этой теореме, как обычно, имеются в виду односторонняя непрерывность и односторонняя производная.

Из теоремы 9.6.2 вытекает следующее утверждение.

Следствие 9.6.3. *Непрерывная на отрезке функция $f(x)$ имеет на этом отрезке первообразную. Одной из первообразных является функция $\Phi(x)$ верхнего предела интегрирования.*

О том, что каждая непрерывная функция имеет первообразную, было сказано в § 8.1.

Следствие 9.6.3 показывает, в частности, что теорема 4.3.3 Коши о промежуточных значениях (для непрерывных функций) вытекает из теоремы 6.1.3 Дарбу о промежуточных значениях. Об этом говорилось в § 6.1.

Теорема 9.6.4 (Формула Ньютона–Лейбница). *Пусть на отрезке $[a,b]$ функция $f(x)$ интегрируема и имеет первообразную $F(x)$. Тогда справедливо равенство*

$$\int_a^b f(x)\, dx = F(b) - F(a), \qquad (9.6.5)$$

которое называют формулой Ньютона–Лейбница.

Доказательство. Пусть T_n – разбиение отрезка $[a,b]$ точками $a = x_0 < x_1 < \dots < x_n = b$ на n отрезков $[x_{k-1}, x_k]$ равной длины $(b-a)/n$.

Имеем
$$F(b)-F(a)=\sum_{k=1}^{n}(F(x_k)-F(x_{k-1}).$$
Согласно формуле конечных приращений Лагранжа для каждого k существует точка $\xi_k \in (x_{k-1}, x_k)$ такая, что
$$F(x_k)-F(x_{k-1})=F'(\xi_k)\frac{b-a}{n}=f(\xi_k)\frac{b-a}{n}\,.$$

Таким образом,
$$F(b)-F(a)=\sum_{k=1}^{n}f(\xi_k)\frac{b-a}{n}\,.$$
Эта сумма является интегральной суммой Римана $S_{T_n}(f,\xi)$, где ξ – набор точек $\xi_1, \xi_2, \ldots, \xi_n$.

Так как функция f интегрируема, то для каждого $\varepsilon > 0$ существует n такое, что при любом выборе точек ξ^*
$$\left|\int_a^b f(x)\,dx - S_{T_n}(f,\xi^*)\right| < \varepsilon.$$

Отсюда вытекает справедливость равенства (9.6.5).

Теорема доказана.

Заметим, что первообразную могут иметь и неинтегрируемые функции. Например, функция $F(x)$, заданная на $[0,1]$ формулами
$$F(x) := x^2\sin\frac{1}{x^2}$$
при $x\in(0,1]$ и $F(0):=0$, является первообразной неограниченной и, значит, неинтегрируемой функции
$$F'(x)=2x\sin\frac{1}{x^2}-\frac{2}{x}\cos\frac{1}{x^2}\,,\qquad 0<x\leqslant 1,$$
$F'(0)=0$.

Разность $F(b)-F(a)$ часто обозначают
$$F(x)\Big|_a^b \qquad \text{или} \qquad [F(x)]_a^b.$$

В этих обозначениях формула (9.6.5) имеет вид:

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b = [F(x)]_a^b.$$

Формула Ньютона–Лейбница является одним из основных результатов интегрального исчисления и математического анализа вообще. Её значение в том, что она связывает определённый и неопределённый интегралы и даёт возможность вычислять определённые интегралы функций, для которых известна первообразная.

Следующее утверждение вытекает из формулы Ньютона– Лейбница.

Теорема 9.6.5. *Если функция $F(x)$ имеет на отрезке $[a,b]$ непрерывную производную, то для $x \in [a,b]$ справедливо равенство*

$$F(x) = F(a) + \int_a^x F'(t)dt. \tag{9.6.6}$$

Равенство (9.6.6) решает задачу о восстановлении функции по её производной, если производная непрерывна. Оно показывает, что дифференцирование и интегрирование являются взаимно обратными операциями.

Отметим вместе с тем, что интеграл Римана не даёт полного решения задачи о восстановлении функции по производной, так как производная может быть, например, неограниченной. Восстановление функции по производной обеспечивает более общее понятие интеграла, изучение которого выходит за рамки настоящего курса.

Применим теорему 9.6.5 к вопросу о длине кривой в трёхмерном пространстве. Согласно теореме 7.3.2, если кривая

$$\Gamma = \{\overline{r}(\tau), \tau \in [a,b]\}$$

непрерывно дифференцируема, то функция $s(t)$ – длина дуги кривой Γ, когда параметр τ пробегает отрезок $[a,t]$, $t \leqslant b$, – имеет непрерывную производную и

$$\frac{ds}{dt} = \left|\frac{d\overline{r}(t)}{dt}\right| = \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2},$$

где $\overline{r}(t) = (x(t), y(t), z(t))$.

Значит,

$$s(t) = \int_a^t \left| \frac{d\overline{r}(\tau)}{d\tau} \right| d\tau$$

и

$$S_\Gamma = \int_a^b \left| \frac{d\overline{r}(\tau)}{d\tau} \right| d\tau = \int_a^b \sqrt{x'(\tau)^2 + y'(\tau)^2 + z'(\tau)^2}\, d\tau.$$

Покажем, что в теореме 9.6.5 условия на функцию $F(x)$ можно несколько ослабить.

Теорема 9.6.6. *Равенство* (9.6.6) *справедливо, если функция* $F(x)$ *на отрезке* $[a,b]$ *непрерывна и кусочно непрерывно дифференцируема, т.е. имеет кусочно непрерывную производную.*

Доказательство. В силу наложенных на F условий существует разбиение $a = x_0 < x_1 < \dots < x_n = b$ отрезка $[a,b]$ такое, что производная $F'(x)$ на каждом интервале (x_{k-1}, x_k) непрерывна и имеет конечные односторонние пределы в концах этих интервалов. Значит, согласно следствию 6.2.12 след функции $F(x)$ на отрезке $[x_{k-1}, x_k]$ является непрерывно дифференцируемой функцией.

Представим интеграл по отрезку $[a,b]$ в виде суммы интегралов по отрезкам $[x_{k-1}, x_k]$, в концах каждого из этих отрезков доопределим $F'(x)$ по непрерывности и воспользуемся формулой (9.6.6). Тогда получим

$$\int_a^b F'(x)\,dx = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} F'(x)\,dx = \\ = \sum_{k=1}^n (F(x_k) - F(x_{k-1})) = F(b) - F(a).$$

Теорема доказана.

Возможно дальнейшее ослабление условий на функцию $F(x)$, при которых справедливо равенство (9.6.6), но не будем здесь на этом останавливаться.

Применим основные приёмы вычисления неопределённых интегралов (интегрирование с помощью замены переменной и интегрирование по частям) к определённым интегралам.

Теорема 9.6.7 (Формула замены переменной). *Пусть функция* $\varphi(t)$ *на отрезке* $[c,d]$ *непрерывно дифференцируема и все её* *значения принадлежат отрезку, на котором непрерывна функция $f(x)$. Тогда справедливо равенство*

$$\int_a^b f(x)\,dx = \int_c^d f(\varphi(t))\varphi'(t)\,dt, \tag{9.6.7}$$

где $a := \varphi(c)$ и $b := \varphi(d)$.

Доказательство. Сначала заметим, что сложная функция $f(\varphi(t))$ существует на $[c,d]$. Значит, если $F(x)$ – первообразная функции $f(x)$ в области её задания, то на $[c,d]$ сложная функция $F(\varphi(t))$ также существует и является первообразной функции $f(\varphi(t))\varphi'(t)$.

Пользуясь формулой Ньютона–Лейбница, находим

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

и

$$\int_c^d f(\varphi(t))\varphi'(t)\,dt = F(\varphi(d)) - F(\varphi(c)) = F(b) - F(a).$$

Из этих равенств вытекает (9.6.7).

Теорема доказана.

Для неопределённых интегралов формула замены переменной (8.2.5) имела вид

$$\int f(x)\,dx|_{x=\varphi(t)} = \int f(\varphi(t))\varphi'(t)\,dt + C,$$

где интеграл в правой части является функцией t, а интеграл в левой части – функцией x. Поэтому в интеграле слева, чтобы получить функцию t, делается замена $x = \varphi(t)$.

А формула (9.6.7) для определённых интегралов представляет собой равенство чисел, переход к другой переменной учтён при замене пределов интегрирования. Отметим, что в (9.6.7) a не обязательно меньше b.

Теорема 9.6.8 (Интегрирование по частям). *Если функции $u(x)$ и $v(x)$ на отрезке $[a,b]$ непрерывны и кусочно непрерывно дифференцируемы, то справедливо равенство*

$$\int_a^b u(x)v'(x)\,dx = u(x)v(x)|_a^b - \int_a^b u'(x)v(x)\,dx, \tag{9.6.8}$$

которое называют формулой интегрирования по частям.

Доказательство. Оба интеграла в (9.6.8) существуют как интегралы от кусочно непрерывных функций. Производные функций $u(x)$ и $v(x)$ не определены в конечном числе точек, но согласно теореме 9.3.5 их значения в этих точках можно задать произвольно.

Тождество

$$u(x)v'(x) \equiv (u(x)v(x))' - u'(x)v(x)$$

справедливо на всём отрезке $[a, b]$, за исключением конечного числа точек, в которых производные функций $u(x)$ и $v(x)$ не существуют.

Проинтегрируем это тождество по отрезку $[a, b]$:

$$\int_a^b u(x)v'(x)\,dx = \int_a^b (u(x)v(x))'\,dx - \int_a^b u'(x)v(x)\,dx.$$

Так как в силу условий теоремы функция $u(x)v(x)$ непрерывна и кусочно непрерывно дифференцируема, то согласно теореме 9.6.6

$$\int_a^b (u(x)v(x))'\,dx = u(x)v(x)\Big|_a^b$$

и мы получили равенство (9.6.8).

Теорема доказана.

Формула интегрирования по частям (9.6.8) справедлива и при менее ограничительных условиях на функции $u(x)$ и $v(x)$, но сейчас нас ограничивает та степень общности, с какой была доказана теорема 9.6.6.

Используем формулу интегрирования по частям для получения ещё одного представления остаточного члена формулы Тейлора в добавление к тем, какие были доказаны в § 6.4.

Предположим, что функция $f(t)$ на отрезке с концами в точках x_0 и x имеет непрерывные производные до порядка $n \geqslant 1$.

Согласно формуле Ньютона–Лейбница

$$f(x) = f(x_0) + \int_{x_0}^x f'(t)\,dt.$$

Если $n \geqslant 2$, то в силу формулы интегрирования по частям

$$\int_{x_0}^x f'(t)\,dt = -\int_{x_0}^x f'(t)\,d(x-t) =$$

$$= -\left[f'(t)(x-t)\Big|_{x_0}^{x} - \int_{x_0}^{x} f''(t)(x-t)\,dt\right] =$$
$$= f'(x_0)(x-x_0) + \int_{x_0}^{x} f''(t)(x-t)\,dt.$$

Если $n \geqslant 3$, то этот интеграл проинтегрируем по частям ещё раз:

$$\int_{x_0}^{x} f''(t)(x-t)\,dt =$$
$$= -\int_{x_0}^{x} f''(t)\,d\left(\frac{(x-t)^2}{2}\right) =$$
$$= -\left[f''(t)\frac{(x-t)^2}{2}\Big|_{x_0}^{x} - \frac{1}{2}\int_{x_0}^{x} f'''(t),(x-t)^2\,dt\right] =$$
$$= \frac{f''(x_0)}{2!}(x-x_0)^2 + \frac{1}{2!}\int_{x_0}^{x} f'''(t)(x-t)^2\,dt.$$

Следовательно,

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) +$$
$$+ \frac{f''(x_0)}{2!}(x-x_0)^2 + \frac{1}{2!}\int_{x_0}^{x} f'''(t)(x-t)^2\,dt.$$

Продолжая интегрирование по частям пока под знаком интеграла не будет получена производная $f^{(n)}$, приходим к равенству

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}\,dt. \tag{9.6.9}$$

Поскольку равенство (9.6.9) получено с помощью интегрирования по частям, условие непрерывности производной $f^{(n)}$ можно ослабить и предполагать, что на отрезке с концами в точках x и x_0 непрерывна производная функции f порядка $n-1$, а производная порядка n кусочно непрерывна.

Таким образом, имеет место следующее утверждение.

Теорема 9.6.9. *Пусть на отрезке с концами в точках x и x_0 функция f непрерывна вместе со всеми своими производными* *до порядка $n-1$, $n \geqslant 1$, включительно, а производная f порядка n кусочно непрерывна. Тогда справедливо равенство* (9.6.9), *которое называют формулой Тейлора с остаточным членом в интегральной форме.*

Связь остаточного члена формулы Тейлора в интегральной форме с представлением остаточного члена в форме Лагранжа будет рассмотрена в следующем параграфе.

§ 9.7. Теоремы о среднем

Теорема 9.7.1 (Первая теорема о среднем). *Пусть на отрезке $[a, b]$ функции $f(x)$ и $\varphi(x)$ интегрируемы,*

$$m \leqslant f(x) \leqslant M \tag{9.7.1}$$

и $\varphi(x) \geqslant 0$. Тогда существует такое число $\mu \in [m, M]$, что

$$\int_a^b f(x)\varphi(x)\,dx = \mu \int_a^b \varphi(x)\,dx. \tag{9.7.2}$$

Доказательство. Пользуясь неотрицательностью функции $\varphi(x)$, умножим неравенства (9.7.1) на $\varphi(x)$:

$$m\varphi(x) \leqslant f(x)\varphi(x) \leqslant M\varphi(x).$$

Проинтегрируем полученное двойное неравенство:

$$m \int_a^b \varphi(x)\,dx \leqslant \int_a^b f(x)\varphi(x)\,dx \leqslant M \int_a^b \varphi(x)\,dx. \tag{9.7.3}$$

Если

$$\int_a^b \varphi(x)\,dx = 0,$$

то из (9.7.3) следует, что интеграл в левой части (9.7.2) также равен нулю и, значит, равенство (9.7.2) справедливо при любом μ.

Если же

$$\int_a^b \varphi(x)\,dx \neq 0,$$

то равенство (9.7.2) имеет место при

$$\mu := \int_a^b f(x)\varphi(x)\,dx \Bigg/ \int_a^b \varphi(x)\,dx,$$

причём в силу (9.7.3) $m \leqslant \mu \leqslant M$.

Теорема доказана.

Понятно, что в этой теореме условие $\varphi(x) \geqslant 0$ можно заменить на $\varphi(x) \leqslant 0$.

Если в теореме 9.7.1 функция $f(x)$ непрерывна, то в качестве чисел m и M естественно взять минимальное и максимальное значения f на отрезке $[a, b]$. Тогда, используя теорему Коши о промежуточных значениях, получим следующее утверждение.

Следствие 9.7.2. *Если на отрезке* $[a, b]$ *функция* $f(x)$ *непрерывна, а функция* $\varphi(x)$ *интегрируема и неотрицательна, то существует точка* $\xi \in [a, b]$ *такая, что*

$$\int_a^b f(x)\varphi(x)\,dx = f(\xi)\int_a^b \varphi(x)\,dx. \tag{9.7.4}$$

В частности, для $\varphi(x) \equiv 1$ *имеем*

$$\int_a^b f(x)\,dx = f(\xi)(b-a). \tag{9.7.5}$$

Выясним с помощью теоремы 9.7.1 связь остаточного члена формулы Тейлора (9.6.9) в интегральной форме с другими представлениями остаточного члена.

Будем считать, что производная $f^{(n)}(t)$ непрерывна на отрезке с концами в точках x и x_0.

Так как функция $(x-t)^{n-1}$ переменной t на этом отрезке не меняет знак, то согласно (9.7.4) между x и x_0 имеется точка ξ, для которой

$$\frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}\,dt =$$
$$= \frac{1}{(n-1)!}f^{(n)}(\xi)\int_{x_0}^{x}(x-t)^{n-1}\,dt = \frac{1}{n!}f^{(n)}(\xi)\,(x-x_0)^n.$$

Таким образом, из остаточного члена формулы Тейлора в интегральной форме получен остаточный член в форме Лагранжа.

Напомним (см. § 6.4), что в точке, где производная $f^{(n)}$ непрерывна, остаточный член в форме Пеано можно получить из остаточного члена в форме Лагранжа.

Выведем из остаточного члена формулы Тейлора в интегральной форме для случая, когда производная $f^{(n)}$ непрерывна, ещё два варианта записи остаточного члена, о которых раньше не говорилось.

Согласно (9.7.5) между x_0 и x существует точка ξ такая, что

$$\frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}\,dt = \\ = \frac{1}{(n-1)!} f^{(n)}(\xi)(x-\xi)^{n-1}(x-x_0). \tag{9.7.6}$$

Записав ξ в виде $\xi = x_0 + \theta(x - x_0)$, $0 < \theta < 1$, получим

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

Это выражение, как и правую часть формулы (9.7.6), называют остаточным членом формулы Тейлора в *форме Коши*.

Применив для произвольного натурального $p < n-1$ первую теорему о среднем к произведению $f^{(n)}(t)(x-t)^p \cdot (x-t)^{n-1-p}$, находим

$$\frac{1}{(n-1)!}\int_{x_0}^{x} f^{(n)}(t)(x-t)^{n-1}\,dt = \\ = \frac{1}{(n-1)!} f^{(n)}(\xi)(x-\xi)^p \frac{(x-x_0)^{n-p}}{n-p}\,.$$

Такое представление называют остаточным членом в форме Шлёмильха–Роша.

Выше остаточный член в форме Лагранжа был получен при более сильных требованиях на функцию f, чем в § 6.4. Точно также представление остаточного члена в форме Коши и в форме Шлёмильха–Роша справедливо при меньших условиях на функцию f, чем это сделано сейчас. Остаточный член в интегральной форме такхе можно получить при меньших ограничениях на f, чем в теореме 9.6.9.

Однако, разница в этих условиях не так велика, а наша цель была – показать связь остаточного члена формулы Тейлора в интегральной форме с другими представлениями остаточного члена.

Теорема 9.7.3 (Вторая теорема о среднем). *Пусть на отрезке $[a, b]$ функции $f(x)$ монотонно убывает, а функция $\varphi(x)$ интегрируема. Тогда существует такая точка $\xi \in [a, b]$, что*

$$\int_a^b f(x)\,\varphi(x)\,dx = f(a)\int_a^{\xi} \varphi(x)\,dx + f(b)\int_{\xi}^{b} \varphi(x)\,dx. \tag{9.7.7}$$

Если при этом функция $f(x)$ неотрицательна, то существует точка $\xi \in [a,b]$, для которой

$$\int_a^b f(x)\,\varphi(x)\,dx = f(a)\int_a^\xi \varphi(x)\,dx. \tag{9.7.8}$$

Доказательство. Рассмотрим случай, когда $f(x) \geqslant 0$ на $[a,b]$.

Пусть T_n – разбиение отрезка $[a,b]$ на n равных частей точками x_k, $k = 0, 1, \ldots, n$. Положим для сокращения записей

$$I := \int_a^b f(x)\,\varphi(x)\,dx$$

и представим интеграл I в виде суммы интегралов по отрезкам $[x_{k-1}, x_k]$

$$\begin{aligned} I &= \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(x)\,\varphi(x)\,dx = \\ &= \sum_{k=1}^n f(x_{k-1}) \int_{x_{k-1}}^{x_k} \varphi(x)\,dx + \sum_{k=1}^n \int_{x_{k-1}}^{x_k} [f(x) - f(x_{k-1})]\varphi(x)\,dx. \end{aligned}$$

Обозначим полученные суммы α_n и β_n соответственно.

Если $|\varphi(x)| \leqslant B$ при всех x, то

$$\begin{aligned} |\beta_n| &\leqslant \sum_{k=1}^n \int_{x_{k-1}}^{x_k} |f(x) - f(x_{k-1})||\varphi(x)|\,dx \leqslant \\ &\leqslant B \sum_{k=1}^n \int_{x_{k-1}}^{x_k} |f(x) - f(x_{k-1})|\,dx \leqslant \\ &\leqslant B \sum_{k=1}^n \int_{x_{k-1}}^{x_k} (f(x_{k-1}) - f(x_k))\,dx = \\ &= B \sum_{k=1}^n (f(x_{k-1}) - f(x_k))\Delta x_k = B(\overline{S_{T_n}}(f) - \underline{S_{T_n}}(f)). \end{aligned}$$

Значит, $\beta_n \to 0$ при $n \to \infty$ в силу интегрируемости функции $f(x)$.

Поэтому

$$I = \lim_{n\to\infty} \alpha_n. \tag{9.7.9}$$

Оценим величины α_n. Введём функцию

$$\psi(x) := \int_a^x \varphi(t)\,dt.$$

Тогда

$$\alpha_n = \sum_{k=1}^{n} f(x_{k-1})(\psi(x_k) - \psi(x_{k-1})) =$$
$$= \sum_{k=1}^{n} f(x_{k-1})\psi(x_k) - \sum_{k=0}^{n-1} f(x_k)\psi(x_k)$$

и так как $\psi(x_0) = \psi(a) = 0$, то

$$\alpha_n = \sum_{k=1}^{n-1} (f(x_{k-1}) - f(x_k))\psi(x_k) + f(x_{n-1})\psi(x_n). \tag{9.7.10}$$

Функция $\psi(x)$ на $[a, b]$ непрерывна, пусть m и M – её минимальное и максимальное значения на $[a, b]$. В силу монотонного убывания и неотрицательности функции $f(x)$ из (9.7.10) вытекают оценки α_n сверху

$$\alpha_n \leqslant \sum_{k=1}^{n-1} (f(x_{k-1}) - f(x_k))M + f(x_{n-1})M = Mf(a)$$

и снизу

$$\alpha_n \geqslant \sum_{k=1}^{n-1} (f(x_{k-1}) - f(x_k))m + f(x_{n-1})m = mf(a).$$

Таким образом,

$$mf(a) \leqslant \alpha_n \leqslant Mf(a).$$

Согласно (9.7.9) отсюда следует, что

$$mf(a) \leqslant I \leqslant Mf(a)$$

и, значит,

$$I = cf(a), \tag{9.7.11}$$

где число $c \in [m, M]$.

В силу непрерывности функции $\psi(x)$ найдётся точка ξ такая, что $c = \psi(\xi)$. Поэтому равенство (9.7.11) доказывает (9.7.8).

Пусть теперь функция $f(x)$ может принимать значения разных знаков.

Функция $f(x) - f(b)$ убывает и неотрицательна. Значит, по уже доказанному существует такая точка $\xi \in [a, b]$, что

$$\int_a^b (f(x) - f(b))\varphi(x)\,dx = (f(a) - f(b)) \int_a^\xi \varphi(x)\,dx.$$

Следовательно,

$$\int_a^b f(x)\varphi(x)\,dx =$$
$$= f(b)\int_a^b \varphi(x)\,dx + f(a)\int_a^\xi \varphi(x)\,dx - f(b)\int_a^\xi \varphi(x)\,dx =$$
$$= f(a)\int_a^\xi \varphi(x)\,dx + f(b)\int_\xi^b \varphi(x)\,dx,$$

а это и есть равенство (9.7.7).

Теорема доказана.

Понятно, что утверждение, аналогичное теореме 9.7.3, имеет место и для монотонно возрастающих функций $f(x)$.

Отметим, что несмотря на условность названий “первая теорема о среднем”, “вторая теорема о среднем” эти названия устоялись и их можно считать общепринятыми. Вторую теорему о среднем называют также теоремой Бонне.

§ 9.8. Некоторые классические неравенства для интегралов

Установим интегральные аналоги неравенств из § 6.8.

Теорема 9.8.1 (Неравенство Гёльдера). *Если функции $f(x)$ и $g(x)$ интегрируемы на отрезке $[a, b]$, $p > 1$ и q – сопряженное с p число, т.е.*

$$\frac{1}{p} + \frac{1}{q} = 1,$$

то справедливо неравенство

$$\left|\int_a^b f(x)g(x)\,dx\right| \leqslant \int_a^b |f(x)g(x)|\,dx \leqslant$$

$$\leqslant \left(\int_a^b |f(x)|^p\,dx\right)^{1/p} \left(\int_a^b |g(x)|^q\,dx\right)^{1/q}, \quad (9.8.1)$$

которое называют неравенством Гёльдера для интегралов.

Доказательство. Интегралы из неравенств (9.8.1) существуют согласно теореме 9.4.4 и следствию 9.4.2.

Введём обозначения

$$\|f\|_p := \left(\int_a^b |f(x)|^p\,dx\right)^{1/p}, \qquad \|g\|_q := \left(\int_a^b |g(x)|^q\,dx\right)^{1/q}. \quad (9.8.2)$$

Эти величины называют нормами функций, но не будем говорить сейчас об этом подробнее.

Будем считать, что величины (9.8.2) отличны от нуля, так как если какая-либо из них равна нулю, то равен нулю и интеграл

$$\int_a^b |f(x)g(x)|\,dx.$$

В самом деле, если $\|f\|_p = 0$ и $|g| \leqslant B$, то $|f(x)g(x)| \leqslant B|f(x)|$ и применяем лемму 9.4.5.

Положим

$$c(x) := \frac{|f(x)|}{\|f\|_p}, \qquad d(x) := \frac{|g(x)|}{\|g\|_q}.$$

Согласно неравенству Янга (6.8.2) для любых неотрицательных чисел c и d имеет место оценка

$$cd \leqslant \frac{1}{p}\,c^p + \frac{1}{q}\,d^q.$$

Поэтому

$$c(x)d(x) \leqslant \frac{|f(x)|^p}{p\|f\|_p^p} + \frac{|g(x)|^q}{q\|g\|_q^q}.$$

Проинтегрируем это неравенство по отрезку $[a, b]$:

$$\int_a^b c(x)d(x)\,dx \leqslant \frac{1}{p\|f\|_p^p}\int_a^b |f(x)|^p\,dx + \frac{1}{q\|g\|_q^q}\int_a^b |g(x)|^q\,dx = \\ = \frac{1}{p} + \frac{1}{q} = 1.$$

Таким образом,

$$\int_a^b \frac{|f(x)|}{\|f\|_p} \frac{|g(x)|}{\|g\|_q}\, dx \leqslant 1.$$

Отсюда следует (9.8.1).

Теорема доказана.

Неравенство (9.8.1) при $p = 2$ было доказано В. Я. Буняковским (1859), оно является интегральным аналогом установленного О. Коши неравенства (6.8.4).

В § 6.8 отмечалось, что и неравенство (6.8.4) и неравенство (9.8.1) при $p = 2$ принято называть неравенствами Коши–Буняковского. В иностранной литературе оценку (9.8.1) при $p = 2$ часто называют неравенством Шварца, хотя Г. Шварц доказал её на 25 лет позднее Буняковского.

Теорема 9.8.2 (Неравенство Минковского). *Для функций $f(x)$ и $g(x)$, интегрируемых на отрезке $[a, b]$, при $p > 1$ справедливо неравенство*

$$\left(\int_a^b |f(x)+g(x)|^p\, dx\right)^{1/p} \leqslant \\ \leqslant \left(\int_a^b |f(x)|^p\, dx\right)^{1/p} + \left(\int_a^b |g(x)|^p\, dx\right)^{1/p}, \qquad (9.8.3)$$

которое называют неравенством Минковского или неравенством треугольника.

Доказательство. Существование интегралов в (9.8.3), вытекает из свойств интегрируемых функций. Будем считать, что интеграл в левой части неравенства (9.8.3) не равен нулю, иначе это неравенство очевидно.

Оценим интеграл из левой части (9.8.3) следующим образом:

$$\int_a^b |f(x)+g(x)|^p\, dx = \int_a^b |f(x)+g(x)| \cdot |f(x)+g(x)|^{p-1}\, dx \leqslant \\ \leqslant \int_a^b |f(x)|\, |f(x)+g(x)|^{p-1}\, dx + \\ + \int_a^b |g(x)|\, |f(x)+g(x)|^{p-1}\, dx. \qquad (9.8.4)$$

Применим к интегралам из правой части оценки (9.8.4) неравенство Гёльдера с показателями p и $q = p/(p-1)$. Тогда получим

$$\begin{aligned}\int_a^b |f(x)+g(x)|^p\,dx \leqslant \\ \leqslant \left(\int_a^b |f(x)|^p\,dx\right)^{1/p}\left(\int_a^b |f(x)+g(x)|^p\,dx\right)^{(p-1)/p} + \\ + \left(\int_a^b |g(x)|^p\,dx\right)^{1/p}\left(\int_a^b |f(x)+g(x)|^p\,dx\right)^{(p-1)/p} = \\ = \left[\left(\int_a^b |f(x)|^p\,dx\right)^{1/p} + \left(\int_a^b |g(x)|^p\,dx\right)^{1/p}\right] \times \\ \times \left(\int_a^b |f(x)+g(x)|^p\,dx\right)^{(p-1)/p}.\end{aligned} \tag{9.8.5}$$

Разделим обе части этого неравенства на

$$\left(\int_a^b |f(x)+g(x)|^p\,dx\right)^{(p-1)/p}.$$

Тогда в левой части (9.8.5) получим интеграл в степени $1/p$ и, таким образом, неравенство (9.8.3) доказано.

Заметим, что в обозначениях (9.8.2) неравенство (9.8.3) имеет вид

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p.$$

Теорема 9.8.3 (Неравенство Иенсена). *Пусть на отрезке* $[a,b]$ *функции* $f(x)$ *и* $\alpha(x)$ *интегрируемы,* $m \leqslant f(x) \leqslant M$, $\alpha(x) \geqslant 0$,

$$\int_a^b \alpha(x)\,dx = 1,$$

а функция φ *выпукла и непрерывна на отрезке* $[m, M]$. *Тогда имеет место неравенство*

$$\varphi\left(\int_a^b f(x)\,\alpha(x)\,dx\right) \leqslant \int_a^b \varphi(f(x))\,\alpha(x)\,dx, \tag{9.8.6}$$

которое называют неравенством Иенсена для интегралов.

Доказательство. Положим
$$s := \int_a^b f(x)\,\alpha(x)\,dx.$$
Тогда $m \leqslant s \leqslant M$.

Пусть формула $y(u) = k(u-s) + \varphi(s)$ задаёт опорную прямую функции $\varphi(x)$ в точке s, т.е. $k = \varphi'(s)$, если эта производная существует, и k – произвольное число, заключённое между $\varphi'_-(s)$ и $\varphi'_+(s)$ в противном случае. Тогда согласно теореме 6.7.2 $\varphi'_-(s) < \varphi'_+(s)$.

При всех $u \in [m, M]$ имеем $\varphi(u) \geqslant k(u-s) + \varphi(s)$ и, значит, для $x \in [a, b]$
$$\varphi(f(x)) \geqslant k(f(x) - s) + \varphi(s).$$

Умножим это неравенство на $\alpha(x)$ и проинтегрируем его затем по x
$$\int_a^b \varphi(f(x))\alpha(x)\,dx \geqslant k\int_a^b (f(x)-s)\alpha(x)\,dx + \varphi(s)\int_a^b \alpha(x)\,dx = \varphi(s).$$

Таким образом, неравенство (9.8.6) доказано.

Если функция $\varphi(x)$ вогнута, то знак неравенства в (9.8.6) нужно заменить на противоположный.

Теорема 9.8.4 (Неравенство Чебышева). *Если функции $f(x)$ и $g(x)$ на отрезке $[a, b]$ ограничены и возрастают, то справедливо неравенство*
$$\int_a^b f(x)\,dx \cdot \int_a^b g(x)\,dx \leqslant (b-a)\int_a^b f(x)g(x)\,dx, \tag{9.8.7}$$
называемое неравенством Чебышева для интегралов.

Доказательство. Рассмотрим разбиение T_n отрезка $[a, b]$ на n равных частей точками $a = x_0 < x_1 < \dots < x_n = b$. Согласно неравенству Чебышева для сумм (6.8.13) имеем
$$\sum_{k=1}^n f(x_k) \cdot \sum_{k=1}^n g(x_k) \leqslant n \sum_{k=1}^n f(x_k)g(x_k),$$
откуда
$$\sum_{k=1}^n f(x_k)\frac{b-a}{n} \cdot \sum_{k=1}^n g(x_k)\frac{b-a}{n} \leqslant (b-a)\sum_{k=1}^n f(x_k)g(x_k)\frac{b-a}{n}\,. \tag{9.8.8}$$

Суммы в этой оценке представляют собой интегральные суммы Римана соответствующих функций для разбиения T_n и выбора в отрезках $[x_{k-1}, x_k]$ точек $\xi_k = x_k$. В этих обозначениях оценка (9.8.8) имеет вид

$$S_{T_n}(f,\xi) \cdot S_{T_n}(g,\xi) \leqslant (b-a) S_{T_n}(fg,\xi). \tag{9.8.9}$$

Переходя в (9.8.9) к пределу при $n \to \infty$, получаем (9.8.7).

Теорема доказана.

Неравенство (9.8.7) справедливо и в случае, когда обе функции $f(x)$ и $g(x)$ убывают на $[a, b]$. А если одна из этих функций возрастает, а другая убывает, знак неравенства в (9.8.7) нужно заменить на противоположный.

§ 9.9. Приближённое вычисление интегралов

Значения многих определённых интегралов нельзя выразить, используя элементарные функции. Но даже в тех случаях, когда это возможно, найти численное значение интеграла часто бывает затруднительно.

Поэтому разрабатываются методы приближённого вычисления определённых интегралов. Эта задача является одной из основных в интегральном исчислении.

Для приближённого вычисления интеграла

$$\int_a^b f(x)\,dx$$

обычно его заменяют суммой вида

$$\sum_{k=1}^n p_k f(\xi_k),$$

где ξ_k – некоторые точки отрезка $[a, b]$, напоминающей интегральные суммы Римана. При этом пишут

$$\int_a^b f(x)\,dx \approx \sum_{k=1}^n p_k f(\xi_k). \tag{9.9.1}$$

Знак $\approx$ означает здесь приближённое равенство.

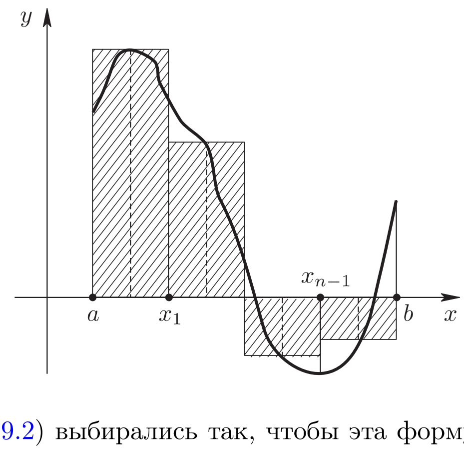
Приближённое равенство (9.9.1) называют *квадратурной формулой* с *узлами* ξ_k и *весами* p_k.

В простейшем случае отрезок $[a, b]$ разбивают на n равных частей, в качестве узлов ξ_k берут середины полученных отрезков, а веса p_k выбирают так, чтобы для функции $f(x)$, постоянной на $[a, b]$, в (9.9.1) получалось бы точное равенство, т.е. полагают $p_k = (b - a)/n$ при всех k.

Полученную таким образом квадратурную формулу

$$\int_a^b f(x)\, dx \approx \sum_{k=1}^{n} \frac{b-a}{n}\, f\left(\frac{b-a}{n}\left(k - \frac{1}{2}\right)\right) \tag{9.9.2}$$

называют *формулой прямоугольников*, так как значение интеграла заменяется в этом случае суммой площадей прямоугольников, основаниями которых служат отрезки разбиения, а высоты равны значениям функции в серединах оснований. При этом площади прямоугольников, расположенных ниже оси OX, берутся со знаком минус.



Веса в (9.9.2) выбирались так, чтобы эта формула была точна для функций, принимающих на $[a, b]$ постоянное значение. Но нетрудно убедиться, что формула прямоугольников даёт точное значение интеграла и для линейных функций, т.е. для многочленов первой степени.

Правая часть формулы (9.9.2) представляет собой интегральную сумму Римана функции $f(x)$. Поэтому для каждой интегрируемой функции погрешность формулы прямоугольников стремится к нулю при $n \to \infty$.

Оценим погрешность формулы (9.9.2) при вычислении интеграла от произвольной непрерывной функции $f(x)$.

Теорема 9.9.1. *Если функция $f(x)$ на отрезке $[a,b]$ непрерывна и $\omega(f,\delta)$ – её модуль непрерывности, то справедлива оценка*

$$\left|\int_a^b f(x)\,dx - \sum_{k=1}^{n} \frac{b-a}{n}\, f\left(\frac{b-a}{n}\left(k-\frac{1}{2}\right)\right)\right| \leqslant (b-a)\,\omega\left(f, \frac{b-a}{2n}\right). \tag{9.9.3}$$

Доказательство. Так как расстояние каждой точки отрезка $[x_{k-1}, x_k]$ до точки $\xi_k = (b-a)(k-1/2)/n$ не превосхолит $(b-a)/2n$, то

$$\left|\int_{x_{k-1}}^{x_k} f(x)\,dx - \frac{b-a}{n} f(\xi_k)\right| = \left|\int_{x_{k-1}}^{x_k} (f(x)-f(\xi_k))\,dx\right| \leqslant$$
$$\leqslant \omega\left(f, \frac{b-a}{2n}\right)\frac{b-a}{n}\,.$$

Просуммировав такие оценки по всем k, приходим к (9.9.3).
Теорема доказана.

Если функция $f(x)$ не только непрерывна, а например, непрерывно дифференцируема, оценку погрешности формулы прямоугольников можно улучшить.

На практике помимо формулы прямоугольников применяются многие другие квадратурные формулы, в том числе такие, в которых участвуют не только значения самих функций, но и значения их производных.

§ 9.10. Несобственные интегралы

В определении интеграла Римана было существенно, что промежуток интегрирования конечен, так как в интегральных суммах используются длины отрезков, полученных при разбиении исходного промежутка на конечное число частей. Кроме того, интегрируемая функция должна быть ограниченной.

Этот параграф посвящён одному из обобщений интеграла Римана, которое позволит рассматривать интегралы по бесконечным промежуткам, а также интегралы от неограниченных функций.

Начнём с интеграла по бесконечному промежутку. Если на $[a, +\infty)$ задана функция $f(x)$, интегрируемая на каждом отрезке $[a, b]$, и существует конечный предел

$$\lim_{b \to +\infty} \int_a^b f(x)\,dx, \tag{9.10.1}$$

то функцию f называют интегрируемой на $[a, +\infty)$ по Риману *в несобственном смысле* и по определению полагают

$$\int_a^{+\infty} f(x)\,dx := \lim_{b \to +\infty} \int_a^b f(x)\,dx.$$

Интегралы, которые рассматривались до сих пор, называют собственными.

Если предел (9.10.1) существует, то говорят, что интеграл

$$\int_a^{+\infty} f(x)\,dx$$

сходится. В противном случае такой интеграл называют *расходящимся*.

Подобным образом вводится несобственный интеграл от функций, заданных на конечном промежутке и неограниченных в окрестности одного из его концов.

Именно, если на конечном промежутке $[a, d)$ задана функция $f(x)$ такая, что при любом $b < d$ существует собственный интеграл по отрезку $[a, b]$ и, кроме того, существует конечный предел

$$\lim_{b \to d-0} \int_a^b f(x)\,dx,$$

то говорят, что существует или *сходится* несобственный интеграл

$$\int_a^d f(x)\,dx, \tag{9.10.2}$$

и полагают

$$\int_a^d f(x)\,dx := \lim_{b \to d-0} \int_a^b f(x)\,dx.$$

В противном случае говорят, что интеграл (9.10.2) не существует или *расходится*.

Заметим, что если интеграл (9.10.2) существует как собственный, то он существует и как несобственный, так как тогда в силу теоремы 9.6.1 интеграл

$$\int_a^b f(x)\,dx$$

является непрерывной функцией от b.

Оба приведенные варианта определения несобственных интегралов по существу аналогичны. Чтобы рассматривать их одновременно, будем говорить о несобственном интеграле

$$\int_a^d f(x)\,dx, \tag{9.10.3}$$

где d может быть некоторым числом, большим a, или $+\infty$. Считаем, что интеграл по любому отрезку $[a,b] \subset [a,d)$ существует как собственный и существует конечный предел

$$\lim_{b\to d} \int_a^b f(x)\,dx, \tag{9.10.4}$$

который называется несобственным интегралом (9.10.3).

Эти определения и обозначения легко переносятся на случай, когда особенность у интеграла не в верхнем, а в нижнем пределе интегрирования.

Поскольку сходимость несобственного интеграла (9.10.3) определена как существование предела (9.10.4), сразу получаем критерий Коши сходимости интеграла (9.10.3).

Теорема 9.10.1 (Критерий Коши). *Пусть функция $f(x)$ задана на промежутке $[a,d)$ и интегрируема по Риману на любом отрезке $[a,b]$, содержащемся в этом промежутке. Тогда для сходимости несобственного интеграла (9.10.3) необходимо и достаточно условие Коши, состоящее в том, что для каждого $\varepsilon > 0$ существует такое число $b_\varepsilon < d$, что для произвольных чисел b' и b'', удовлетворяющих условию $b_\varepsilon < b' < b'' < d$, имеет место оценка*

$$\left| \int_{b'}^{b''} f(x)\,dx \right| < \varepsilon.$$

Многие свойства собственных интегралов переносятся на несобственные.

Например, аддитивность интеграла относительно промежутка интегрирования: если $a < c < d$, то интегралы по промежуткам $[a, d)$ и $[c, d)$ сходятся или расходятся одновременно и, если эти интегралы сходятся, то справедливо равенство

$$\int_a^d f(x)\,dx = \int_a^c f(x)\,dx + \int_c^d f(x)\,dx.$$

Это вытекает из того, что для любого $b \in [a, d)$

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

Точно также имеет место линейность интеграла:

$$\int_a^d (\alpha f(x) + \beta g(x))\,dx = \alpha \int_a^d f(x)\,dx + \beta \int_a^d g(x)\,dx.$$

Здесь α и β – произвольные числа и предполагается, что несобственные интегралы из правой части равенства сходятся, а утверждается сходимость несобственного интеграла из левой части и справедливость равенства.

Таким образом, сходящиеся несобственные интегралы с особенностью в верхнем пределе интегрирования образуют линейное пространство.

Но не все свойства собственных интегралов переносятся на несобственные.

Так, если функции интегрируемы в несобственном смысле, то их произведение может быть неинтегрируемым. Например, несобственный интеграл с особенностью в нуле

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx$$

сходится, так как для любого $\alpha \in (0, 1]$

$$\int_\alpha^1 \frac{1}{\sqrt{x}}\,dx = 2\sqrt{x}\Big|_\alpha^1 = 2 - 2\sqrt{\alpha}$$

и предел полученного выражения при $\alpha \to 0$ существует. Но интеграл от произведения двух таких функций расходится, так как при $\alpha > 0$

$$\int_\alpha^1 \frac{1}{\sqrt{x}} \cdot \frac{1}{\sqrt{x}}\,dx = \int_\alpha^1 \frac{1}{x}\,dx = \ln x\Big|_\alpha^1 = -\ln \alpha,$$

а предел $\lim_{\alpha \to 0} \ln \alpha$ не существует.

Далее, для собственных интегралов из интегрируемости функции следует интегрируемость её модуля. Вскоре будет приведен пример, показывающий, что для несобственных интегралов это не так.

Для несобственных интегралов вводится понятие абсолютной сходимости.

Определение. Несобственный интеграл

$$\int_a^d f(x)\,dx$$

называют *абсолютно сходящимся*, если сходится интеграл

$$\int_a^d |f(x)|\,dx.$$

Теорема 9.10.2. *Если несобственный интеграл сходится абсолютно, то он сходится.*

Для любых чисел b' и b'' справедлива оценка

$$\left|\int_{b'}^{b''} f(x)\,dx\right| \leqslant \left|\int_{b'}^{b''} |f(x)|\,dx\right|,$$

из которой утверждение теоремы 9.10.2 выводится стандартными рассуждениями с помощью критерия Коши.

Покажем, что сходящийся несобственный интеграл может не сходиться абсолютно. Сходящиеся, но не сходящиеся абсолютно несобственные интегралы называют сходящимися *условно*.

Пусть функция $h(x)$ задана на $[1,+\infty)$ следующим образом. Каждый промежуток $[k,k+1)$, где k – натуральное число, разбит на $2k$ равных частей вида $[\alpha,\beta)$ и функция $h(x)$ на этих частях равна попеременно $+1$ и -1.

Тогда $|h(x)| \equiv 1$ и интеграл

$$\int_1^{+\infty} h(x)\,dx \qquad (9.10.5)$$

не сходится абсолютно. Вместе с тем для любого натурального k имеем

$$\int_1^k h(x)\,dx = 0.$$

А если $b \in (k, k+1)$, то

$$\int_1^b h(x)\,dx = \int_k^b h(x)\,dx,$$

значит,

$$\left|\int_1^b h(x)\,dx\right| \leqslant \frac{1}{2k}\,.$$

Следовательно, интеграл (9.10.5) сходится и его значение равно нулю.

Отметим некоторые свойства несобственных интегралов от неотрицательных функций.

Если $f(x) \geqslant 0$, то интеграл

$$\int_a^b f(x)\,dx \tag{9.10.6}$$

как функция переменной b возрастает и сходимость интеграла

$$\int_a^d f(x)\,dx$$

равносильна ограниченности интеграла (9.10.6) как функции переменной b.

Поэтому для обозначения сходимости несобственных интегралов от неотрицательных функций (и только для таких функций) пишут

$$\int_a^d f(x)\,dx < +\infty,$$

а для обозначения их расходимости –

$$\int_a^d f(x)\,dx = +\infty.$$

Следующие две теоремы называют признаками сравнения сходимости несобственных интегралов.

Теорема 9.10.3. *Пусть функции $f(x)$ и $g(x)$ неотрицательны на $[a, d)$, интегрируемы на каждом отрезке $[a, b]$, $b < d$, и $f(x) \leqslant g(x)$ при всех $x \in [a, d)$. Тогда*

1°) *если*

$$\int_a^d g(x)\,dx < +\infty,$$

то

$$\int_a^d f(x)\,dx < +\infty$$

и справедливы неравенства

$$0 \leqslant \int_a^d f(x)\,dx \leqslant \int_a^d g(x)\,dx;$$

2°) *если*

$$\int_a^d f(x)\,dx = +\infty,$$

то

$$\int_a^d g(x)\,dx = +\infty.$$

Эти утверждения вытекают из оценок для собственных интегралов

$$0 \leqslant \int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx, \qquad b < d.$$

Разумеется, если говорить только о сходимости интегралов и не сравнивать их значения, то справедливость неравенств $0 \leqslant f(x) \leqslant g(x)$ можно требовать не при всех $x \in [a, d)$, а только при x, близких к d.

Теорема 9.10.4. *Пусть функции $f(x)$ и $g(x)$ положительны на $[a, d)$ и интегрируемы на каждом отрезке $[a, b]$, $b < d$. Если существуют такие положительные числа c_1 и c_2, что при всех $x \in [a, d)$ справедливы оценки*

$$c_1 \leqslant \frac{f(x)}{g(x)} \leqslant c_2, \tag{9.10.7}$$

то для произвольной неотрицательной функции $\varphi(x)$, интегрируемой на каждом отрезке $[a, b] \subset [a, d)$, интегралы

$$\int_a^d f(x)\varphi(x)\,dx, \qquad \int_a^d g(x)\varphi(x)\,dx$$

сходятся или расходятся одновременно.

Это утверждение следует из теоремы 9.10.3 в силу неравенств

$$c_1 g(x)\varphi(x) \leqslant f(x)\varphi(x) \leqslant c_2 g(x)\varphi(x),$$

вытекающих из (9.10.7).

Далее будем рассматривать интегралы от функций, которые могут принимать значения разных знаков.

Формулу Ньютона–Лейбница для несобственных интегралов запишем как утверждение об интегрировании производной.

Теорема 9.10.5 (Формула Ньютона–Лейбница). *Пусть функция $F(x)$ непрерывна на $[a, d)$ и кусочно непрерывно дифференцируема на каждом отрезке $[a, b] \subset [a, d)$. Если существует конечный предел*

$$\lim_{b \to d-0} F(b),$$

который обозначим $F(d)$, то сходится интеграл

$$\int_a^d F'(x)\, dx$$

и справедливо равенство

$$\int_a^d F'(x)\, dx = F(d) - F(a).$$

Согласно теореме 9.6.6 для каждого $b < d$ имеет место равенство

$$\int_a^b F'(x)\, dx = F(b) - F(a)$$

и нужно перейти в этом равенстве к пределу при $b \to d - 0$.

В случае, когда промежуток $[a, d)$ конечен, теорема 9.10.5 несколько ослабляет условия теоремы 9.6.6 на функцию $F(x)$, при которых справедливо равенство (9.6.6). В теореме 9.6.6 функция $F(x)$ могла иметь конечное число точек разрыва на всем промежутке $[a, d]$, а теперь – конечное число точек разрыва на каждом отрезке $[a, b] \subset [a, d)$, а на всём промежутке $[a, d)$ точек разрыва может быть бесконечно много.

Теорема 9.10.6 (Интегрирование по частям). *Пусть на $[a, d)$ функции $\varphi(x)$ и $\psi(x)$ непрерывны, $\Phi(x)$ обозначает первообразную функции $\varphi(x)$, а $\psi(x)$ имеет непрерывную производную. Предположим, что сходится интеграл*

$$\int_a^d \Phi(x)\, \psi'(x)\, dx \tag{9.10.8}$$

и существует конечный предел

$$\lim_{x\to d-0} \Phi(x)\,\psi(x), \tag{9.10.9}$$

который обозначим $\Phi(d)\,\psi(d)$. *Тогда сходится интеграл*

$$\int_a^d \varphi(x)\,\psi(x)\,dx \tag{9.10.10}$$

и справедливо равенство

$$\int_a^d \varphi(x)\,\psi(x)\,dx = \Phi(x)\,\psi(x)\Big|_a^d - \int_a^d \Phi(x)\,\psi'(x)\,dx. \tag{9.10.11}$$

Доказательство. Согласно формуле интегрирования по частям для собственных интегралов при каждом $b < d$ справедливо равенство

$$\int_a^b \varphi(x)\,\psi(x)\,dx = \Phi(x)\,\psi(x)\Big|_a^b - \int_a^b \Phi(x)\,\psi'(x)\,dx,$$

переходя в котором к пределу при $b \to d-0$, приходим к (9.10.11). Теорема доказана.

В теореме 9.10.6 можно было предполагать, что существуют какие-либо два из выражений (9.10.8)–(9.10.10), и утверждать существование третьего выражения и справедливость равенства (9.10.11).

На теореме 9.10.6 основываются признаки сходимости несобственных интегралов от произведения двух функций.

Теорема 9.10.7. *Пусть на промежутке* $[a, d)$ *функция* $\varphi(x)$ *непрерывна, её первообразная* $\Phi(x)$ *ограничена, а функция* $\psi(x)$ *имеет непрерывную производную. Если* $\psi(x) \to 0$ *при* $x \to d-0$ *и сходится интеграл*

$$\int_a^d |\psi'(x)|\,dx,$$

то сходится интеграл

$$\int_a^d \varphi(x)\,\psi(x)\,dx. \tag{9.10.12}$$

Доказательство. Если $|\varPhi(x)| \leqslant M$ на $[a, d)$, то при каждом $b \in [a, d)$

$$\int_a^b |\varPhi(x)\,\psi'(x)|\,dx \leqslant M \int_a^b |\psi'(x)|\,dx.$$

Отсюда следует абсолютная сходимость, а значит, и сходимость интеграла

$$\int_a^d \varPhi(x)\,\psi'(x)\,dx.$$

Далее, из ограниченности функции $\varPhi(x)$ и стремления $\psi(x)$ при $x \to d-0$ к нулю следует, что и $\varPhi(x)\psi(x) \to 0$. Значит, выполнены условия теоремы 9.10.6 и интеграл (9.10.12) сходится.

Заметим, что в отличие от признаков сравнения из теорем 9.10.3 и 9.10.4, которые фактически были признаками абсолютной сходимости, в теореме 9.10.7 нельзя говорить об абсолютной сходимости интеграла (9.10.12).

Приведём два следствия теоремы 9.10.7, которыми часто пользуются при исследовании сходимости несобственных интегралов.

Теорема 9.10.8 (Признак Дирихле). *Если на $[a, d)$ функция $\varphi(x)$ непрерывна, её первообразная ограничена, а функция $\psi(x)$ имеет непрерывную производную и при $x \to d-0$ монотонно стремится к нулю, то интеграл (9.10.12) сходится.*

Доказательство. В силу монотонности функции $\psi(x)$ её производная не меняет знак. Значит,

$$\int_a^d |\psi'(x)|\,dx = \left|\int_a^d \psi'(x)\,dx\right| = \left|\psi(x)\Big|_a^d\right| = |\psi(a)|. \tag{9.10.13}$$

Здесь первые два равенства нужно было бы сначала записать для интегралов по отрезкам $[a, b] \subset [a, d)$, а затем перейти в них к пределу при $b \to d-0$. Но обычно сразу пишут несобственные интегралы, понимая подобные равенства условно, т.е. что они справедливы, если интегралы сходятся.

Из (9.10.13) следует, что выполнены условия теоремы 9.10.7, и, таким образом, интеграл (9.10.12) сходится.

Теорема 9.10.9 (Признак Абеля). *Пусть на $[a, d)$ функция $\varphi(x)$ непрерывна и её первообразная $\varPhi(x)$ имеет предел при $x \to d-0$, а функция $\psi(x)$ непрерывно дифференцируема и при $x \to d-0$ имеет предел, к которому она стремится монотонно. Тогда сходится интеграл (9.10.12).*

Доказательство. Для функции $\psi(x) - \alpha$, где

$$\alpha := \lim_{b \to d-0} \psi(x),$$

выполнены условия признака Дирихле, согласно которому сходится интеграл

$$\int_a^d \varphi(x)[\psi(x) - \alpha]\, dx.$$

Интеграл от функции $\alpha\varphi(x)$ сходится в силу существования предела функции $\Phi(x)$, поэтому утверждение теоремы справедливо.

Отметим, что из существования предела первообразной $\Phi(x)$ следует её ограниченность. Таким образом, в признаке Абеля требования на функцию $\varphi(x)$ немного усилены по сравнению с условиями из признака Дирихле. В то же время условия на функцию $\psi(x)$ ослаблены: там предполагалось её монотонное стремление к нулю, а теперь – монотонное стремление к пределу, который не обязательно равен нулю.

На несобственные интегралы переносятся неравенства Гёльдера и Минковского, причём в этом случае они дают и признаки сходимости несобственных интегралов.

Теорема 9.10.10 (Неравенство Гёльдера). *Пусть* $p > 1$ *и* $1/p + 1/q = 1$. *Если функции* $f(x)$ *и* $g(x)$ *интегрируемы на каждом отрезке* $[a, b] \subset [a, d)$ *и сходятся несобственные интегралы*

$$\int_a^d |f(x)|^p\, dx, \qquad \int_a^d |g(x)|^q\, dx,$$

то интеграл

$$\int_a^d f(x)g(x)\, dx$$

абсолютно сходится и справедливы оценки

$$\left|\int_a^d f(x)g(x)\, dx\right| \leqslant \int_a^d |f(x)g(x)|\, dx \leqslant \\ \leqslant \left(\int_a^d |f(x)|^p\, dx\right)^{1/p} \left(\int_a^d |g(x)|^q\, dx\right)^{1/q}. \tag{9.10.14}$$

Теорема 9.10.11 (Неравенство Минковского). *Пусть $p > 1$ и функции $f(x)$ и $g(x)$ интегрируемы на каждом отрезке $[a, b] \subset [a, d)$. Если сходятся несобственные интегралы*

$$\int_a^d |f(x)|^p \, dx, \qquad \int_a^d |g(x)|^p \, dx,$$

то сходится интеграл

$$\int_a^d |f(x) + g(x)|^p \, dx$$

и справедливо неравенство

$$\left(\int_a^d |f(x) + g(x)|^p \, dx\right)^{1/p} \leqslant \\ \leqslant \left(\int_a^d |f(x)|^p \, dx\right)^{1/p} + \left(\int_a^d |g(x)|^p \, dx\right)^{1/p}. \qquad (9.10.15)$$

Для доказательства этих теорем записываем неравенства вида (9.10.14) и (9.10.15) для произвольного отрезка $[a, b] \subset [a, d)$ и, пользуясь неотрицательностью подынтегральных функций, переходим к пределу при $b \to d - 0$.

До сих пор рассматривались несобственные интегралы, когда особенность была в одном из пределов интегрирования, обычно это был верхний предел.

Рассмотрим теперь интегралы, имеющие особенность и в верхнем и в нижнем пределах интегрирования. Пусть для определённости это интегралы по промежутку $(-\infty, +\infty)$.

Говорят, что интеграл

$$\int_{-\infty}^{+\infty} f(x) \, dx \qquad (9.10.16)$$

сходится, если для некоторого числа c сходятся оба интеграла

$$\int_{-\infty}^{c} f(x) \, dx, \qquad \int_{c}^{+\infty} f(x) \, dx. \qquad (9.10.17)$$

В этом случае по определению полагают

$$\int_{-\infty}^{+\infty} f(x) \, dx := \int_{-\infty}^{c} f(x) \, dx + \int_{c}^{+\infty} f(x) \, dx.$$

Из аддитивности собственных интегралов относительно промежутка интегрирования следует, что сходимость интеграла (9.10.16) и его значение не зависят от того, какое число c было взято в этих определениях.

Если хотя бы один из интегралов (9.10.17) расходится, то интеграл (9.10.16) называют расходящимся.

Рассмотрим интеграл по конечному отрезку $[a, b]$, когда подынтегральная функция имеет особенность в окрестности некоторой внутренней точки $c \in [a, b]$. Тогда несобственный интеграл

$$\int_a^b f(x)\,dx \tag{9.10.18}$$

называют сходящимся, если сходятся оба интеграла

$$\int_a^c f(x)\,dx, \qquad \int_c^b f(x)\,dx. \tag{9.10.19}$$

При этом полагают по определению

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx. \tag{9.10.20}$$

А если хотя бы один из интегралов (9.10.19) расходится, то интеграл (9.10.18) называют расходящимся.

Обсудим это определение подробнее. Если интеграл (9.10.18) сходится, то равенство (9.10.20) означает, что

$$\int_a^b f(x)\,dx = \lim_{\varepsilon_1 \to +0} \int_a^{c-\varepsilon_1} f(x)\,dx + \lim_{\varepsilon_2 \to +0} \int_{c+\varepsilon_2}^b f(x)\,dx.$$

Здесь ε_1 и ε_2 стремятся к $+0$ независимо друг от друга.

Но в некоторых естественно возникающих задачах (одна из них встретится в главе 19) нужно более слабое свойство, когда такие пределы рассматриваются при $\varepsilon_1 = \varepsilon_2$, т.е. когда существует предел

$$\lim_{\varepsilon \to +0} \left[\int_a^{c-\varepsilon} f(x)\,dx + \int_{c+\varepsilon}^b f(x)\,dx \right]. \tag{9.10.21}$$

Если предел (9.10.21) существует, то говорят, что интеграл (9.10.18) сходится *в смысле главного значения* или *по Коши*.

В этом случае пишут

$$\lim_{\varepsilon\to+0}\left[\int_a^{c-\varepsilon} f(x)\,dx + \int_{c-\varepsilon}^b f(x)\,dx\right] := \text{v.p.}\int_a^b f(x)\,dx.$$

Здесь v.p. – сокращение от французского valeur principal.

Например, интеграл

$$\int_{-1}^1 \frac{1}{x}\,dx$$

как несобственный расходится, а как интеграл в смысле главного значения он сходится и равен нулю.

Интеграл в смысле главного значения вводят и для случая, когда интегрирование ведётся по всей оси $(-\infty, +\infty)$: по определению полагают

$$\text{v.p.}\int_{-\infty}^{+\infty} f(x)\,dx := \lim_{M\to+\infty}\int_{-M}^{M} f(x)\,dx,$$

если этот предел существует.

Наконец, функция $f(x)$ может иметь на промежутке $[a, b]$ конечное число особенностей, которые могут быть в концевых точках a и b и в некоторых внутренних точках из (a, b).

Тогда (a, b) разбивают на конечное число таких промежутков, что функция $f(x)$ имеет на каждом из них единственную особенность в каком-либо из концов промежутка. Если сходятся несобственные интегралы по каждому из этих промежутков, то говорят, что сходится несобственный интеграл по промежутку $[a, b]$. В противном случае интеграл по промежутку $[a, b]$ называют расходящимся. Если несобственный интеграл

$$\int_a^b f(x)\,dx$$

сходится, то в качестве его значения берут сумму всех интегралов по построенным промежуткам.

§ 9.11. Задачи и упражнения

9.11.1. Докажите, что если функция $f(x)$ непрерывна на $[-1, 1]$ и для каждой непрерывной на $[-1, 1]$ чётной функции $g(x)$

$$\int_{-1}^1 f(x)g(x)\,dx = 0,$$

то функция $f(x)$ нечётна.

9.11.2. Докажите, что функция Римана

$$R(x) := \begin{cases} 0, & \text{если } x \text{ иррационально}, \\ 1/q, & \text{если } x = p/q \text{ рационально} \\ & \qquad \text{и дробь } p/q \text{ несократима}, \end{cases}$$

интегрируема на $[0, 1]$.

9.11.3. Докажите, что если на отрезке функции $f(x)$ и $g(x)$ интегрируемы, то интегрируемы также функции $\max(f(x), g(x))$ и $\min(f(x), g(x))$.

9.11.4. Докажите, что если функция $f(x)$ интегрируема, то

$$\int_0^x \left(\int_0^y f(t)\,dt\right) dy = \int_0^x f(t)(x-t)\,dt.$$

9.11.5. Докажите, что если функция $f(x)$ непрерывна на отрезке $[0, 1]$, то

$$\lim_{x \to +0} x \int_x^1 \frac{f(t)}{t^2}\,dt = f(0).$$

9.11.6. Докажите, что если функция $f(x)$ непрерывна на отрезке $[-1, 1]$, то

$$\lim_{h \to +0} \int_{-1}^1 \frac{h}{h^2 + x^2} f(x)\,dx = \pi f(0).$$

9.11.7. Докажите, что если на $[a, b]$ функция $f(x)$ непрерывна и неотрицательна, то

$$\max_{x \in [a,b]} f(x) = \lim_{p \to \infty} \left(\int_a^b f^p(x)\,dx\right)^{1/p}.$$

9.11.8. Докажите, что если функция $f(x)$ интегрируема и

$$f(x) = \int_0^x f(t)\,dt,$$

то $f(x) \equiv 0$.

9.11.9. Докажите, что для функции

$$f(x) := \int_x^{x+1} \sin(t^2)\, dt, \quad x > 0,$$

справедлива оценка

$$|f(x)| \leqslant \frac{1}{x}\,.$$

9.11.10. Покажите, что для функций $\varphi(x)$, принимающих на $[a,b]$ значения разных знаков, теорема 9.7.1, вообще говоря, не верна.

9.11.11. При каких значениях параметра s сходится интеграл

$$\text{a)} \quad \int_0^{+\infty} \frac{x^s}{1+x}\, dx; \qquad \text{b)} \quad \int_0^{+\infty} \frac{\sin x}{x^s}\, dx?$$

9.11.12. Докажите, что если $f(x)$ – убывающая на $[0,+\infty)$ функция, то из сходимости интеграла

$$\int_0^{+\infty} f(x)\, dx$$

следует, что $\lim_{x\to+\infty} xf(x) = 0$.

9.11.13. Докажите, что если функция $f(x)$ интегрируема на отрезке $[a,b]$, а вне этого отрезка равна нулю, то

$$\lim_{\alpha\to 0} \int_{-\infty}^{+\infty} |f(x+\alpha) - f(x)|\, dx = 0.$$

9.11.14. Постройте на $[0,+\infty)$ функцию $f(x)$, для которой интеграл

$$\int_0^{+\infty} f(x)\, dx$$

сходится и $\lim_{x\to+\infty} f(x) = \infty$.

9.11.15. Докажите формулу Фруллани:

$$\int_0^{+\infty} \frac{f(\alpha x) - f(\beta x)}{x}\, dx = (f(+\infty) - f(0)) \ln \frac{\alpha}{\beta}\,,$$

где $\alpha > 0$, $\beta > 0$, а $f(x)$ – непрерывная на $[0,+\infty)$ функция, для которой существует предел $f(+\infty) := \lim_{x\to+\infty} f(x)$.

Глава 10. Интеграл Римана–Стилтьеса

§ 10.1. Функции ограниченной вариации

Пусть функция $\varphi(x)$ задана на отрезке $[a,b]$. Для произвольного разбиения T этого отрезка

$$a = x_0 < x_1 < \dots < x_n = b$$

составим сумму

$$\nu_T := \sum_{k=1}^{n} |\varphi(x_{k-1}) - \varphi(x_k)|. \qquad (10.1.1)$$

Определение. Функция $\varphi(x)$ называется *функцией ограниченной вариации* (функцией с ограниченным изменением) на отрезке $[a,b]$, если суммы ν_T ограничены некоторым числом, не зависящим от разбиения T.

Для функции φ ограниченной вариации точную верхнюю грань значений сумм ν_T, взятую по всем разбиениям T, называют *вариацией* (полным изменением) φ на отрезке $[a,b]$ и обозначают $V(\varphi, q[a,b])$.

Множество функций, имеющих ограниченную вариацию на отрезке $[a,b]$, обозначают $V[a,b]$.

Когда ясно, какой отрезок имеется в виду, вместо $V(\varphi,[a,b])$ и $V[a,b]$ пишут $V(\varphi)$ и V.

Монотонные на отрезке $[a,b]$ функции φ имеют ограниченную вариацию, поскольку у таких функций разности $\varphi(x_{k-1}) - \varphi(x_k)$, $k = 1,\dots,n$, не могут принимать значения разных знаков и

$$\nu_T = \left|\sum_{k=1}^{n} \varphi(x_{k-1}) - \varphi(x_k)\right| = |\varphi(a) - \varphi(b)|.$$

Поэтому $V(\varphi) = |\varphi(a) - \varphi(b)|$.

Если функция φ удовлетворяет на отрезке $[a,b]$ условию Липшица первого порядка

$$|\varphi(x') - \varphi(x'')| \leqslant L|x' - x''| \qquad \text{для всех} \quad x', x'' \in [a,b],$$

то φ имеет ограниченную вариацию, так как

$$\sum_{k=1}^{n} |\varphi(x_{k-1}) - \varphi(x_k)| \leqslant \sum_{k=1}^{n} L|x_{k-1} - x_k| = L(b-a).$$

Таким образом, справедлива оценка

$$V(\varphi) \leqslant L(b-a).$$

Отметим некоторые простые свойства функций ограниченной вариации.

1°. Если $\varphi \in V$, то для любого числа α имеем $\alpha\varphi \in V$ и $V(\alpha\varphi) = |\alpha| V(\varphi)$.

2°. Если функции $\varphi, \psi \in V$, то $\varphi \pm \psi \in V$ и $V(\varphi \pm \psi) \leqslant V(\varphi) + V(\psi)$.

Эти свойства вытекают непосредственно из определения.

3°. Если функция φ имеет ограниченную вариацию, то функция $|\varphi|$ также имеет ограниченную вариацию и $V(|\varphi|) \leqslant V(\varphi)$.

Это утверждение справедливо в силу оценок

$$\big||\varphi(x_{k-1})| - |\varphi(x_k)|\big| \leqslant |\varphi(x_{k-1}) - \varphi(x_k)|, \qquad k = 1, \ldots, n.$$

4°. Если функция $\varphi \in V[a,b]$, то она ограничена на $[a,b]$.

В самом деле, для каждой точки $x \in [a,b]$

$$|\varphi(x)| \leqslant |\varphi(a)| + |\varphi(x) - \varphi(a)| \leqslant |\varphi(a)| + V(\varphi, [a,b]).$$

Теорема 10.1.1 (Аддитивность вариации). *Если функция φ имеет ограниченную вариацию на отрезке $[a,b]$ и точка $c \in (a,b)$, то φ имеет ограниченную вариацию на отрезках $[a,c]$ и $[c,b]$. Наоборот, если функция φ имеет ограниченную вариацию на отрезках $[a,c]$ и $[c,b]$, $a < c < b$, то она имеет ограниченную вариацию на $[a,b]$. При этом имеет место равенство*

$$V(\varphi, [a,b]) = V(\varphi, [a,c]) + V(\varphi, [c,b]). \tag{10.1.2}$$

Действительно, от добавления к произвольному разбиению T точки c значение суммы (10.1.1) может только увеличиться, а верхние грани сумм (10.1.1) по отрезкам $[a,c]$ и $[c,b]$ берутся независимо друг от друга.

Из (10.1.2) вытекает возрастание функции $V(\varphi, [a, x])$ на отрезке $[a, b]$.

Из теоремы 10.1.1 следует, что кусочно монотонные функции имеют ограниченную вариацию. Понятие функций ограниченной вариации введено К. Жорданом (1880) как обобщение кусочно монотонных функций.

Теорема 10.1.2. *Функция имеет ограниченную вариацию в том и только том случае, когда она равна разности двух возрастающих функций.*

Доказательство. Достаточность вытекает из того, что монотонные функции имеют ограниченную вариацию и разность функций ограниченной вариации также имеет ограниченную вариацию.

Докажем необходимость. Пусть $\varphi \in V[a, b]$. Запишем тождество

$$\varphi(x) \equiv V(\varphi, [a, x]) - (V(\varphi, [a, x]) - \varphi(x)).$$

Функция $V(\varphi, [a, x])$ с ростом x возрастает. Докажем возрастание функции $V(\varphi, [a, x]) - \varphi(x)$.

Для точек $x_1 < x_2$ с помощью (10.1.2) получаем

$$\begin{aligned} V(\varphi, [a, x_2]) - \varphi(x_2) - (V(\varphi, [a, x_1]) - \varphi(x_1)) \\ = V(\varphi, [x_1, x_2]) - (\varphi(x_2) - \varphi(x_1)), \end{aligned}$$

а эта разность неотрицательна, так как

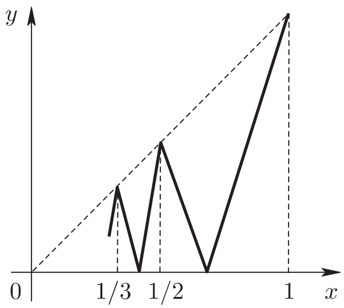
$$|\varphi(x_2) - \varphi(x_1)| \leqslant V(\varphi, [x_1, x_2]).$$

Теорема доказана.

Монотонные функции могут иметь разрывы только первого рода, причём множество их точек разрыва не более чем счётно (см. § 4.2). Согласно теореме 10.1.2 функции ограниченной вариации также могут иметь разрывы только первого рода и не более, чем в счётном множестве точек.

Покажем, что непрерывные функции не обязательно имеют ограниченную вариацию.

Зададим на отрезке $[0, 1]$ функцию $f(x)$ следующим образом. Пусть $f(0) := 0$, $f(1/k) := 1/k$, $k = 1, 2, \ldots$ и в каждом промежутке $[1/(k + 1), 1/k]$ функция f кусочно линейна и один раз обращается в нуль.



Тогда

$$V\left(f,\left[\frac{1}{k+1},\frac{1}{k}\right]\right)=\frac{1}{k+1}+\frac{1}{k}$$

и согласно (10.1.2)

$$V\left(f,\left[\frac{1}{n+1},1\right]\right)=\sum_{k=1}^{n}\left(\frac{1}{k+1}+\frac{1}{k}\right)=2\sum_{k=1}^{n}\frac{1}{k}-1+\frac{1}{n+1}\,. \tag{10.1.3}$$

Так как

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{1}{k}=+\infty$$

(см. § 2.9), то с ростом n выражение в правой части равенства (10.1.3) неограниченно возрастает и, таким образом, функция f не является функцией ограниченной вариации на $[0,1]$. Вместе с тем, она непрерывна на $[0,1]$.

Теорема 10.1.3. *Если функция φ имеет ограниченную вариацию на отрезке $[a,b]$ и непрерывна в некоторой точке $x_0\in[a,b]$, то функция $v(x):=V(\varphi,[a,x])$ также непрерывна в точке x_0.*

Доказательство. Если x_0 – один из концов отрезка $[a,b]$, то как обычно, имеется в виду соответствующая односторонняя непрерывность в точке x_0.

Докажем непрерывность $v(x)$ в точке $x_0\in[a,b)$ справа.

В силу непрерывности φ в точке x_0 для каждого $\varepsilon>0$ существует $\delta>0$ такое, что при всех $x\in(x_0,x_0+\delta)$ имеем

$$|\varphi(x)-\varphi(x_0)|<\frac{\varepsilon}{2}\,. \tag{10.1.4}$$

Выберем разбиение отрезка $[x_0, b]$ точками $x_0 < x_1 < \dots < x_n = b$ такое, что

$$V(\varphi, [x_0, b]) < \sum_{k=1}^{n} |\varphi(x_{k-1}) - \varphi(x_k)| + \frac{\varepsilon}{2}$$

и $x_1 - x_0 < \delta$. С помощью (10.1.4) получаем

$$V(\varphi, [x_0, b]) < \sum_{k=2}^{n} |\varphi(x_{k-1}) - \varphi(x_k)| + \varepsilon \leqslant V(\varphi, [x_1, b]) + \varepsilon.$$

Отсюда

$$V(\varphi, [x_0, b]) - V(\varphi, [x_1, b]) < \varepsilon. \qquad (10.1.5)$$

Так как x_1 – произвольная точка интервала $(x_0, x_0 + \delta)$, то оценка (10.1.5) доказывает непрерывность функции $v(x)$ в точке x_0 справа.

Точно так же устанавливается непрерывность $v(x)$ в точках из $(a, b]$ слева.

Теорема доказана.

Поскольку в теореме 10.1.2 представление функции ограниченной вариации φ в виде разности возрастающих функций строилось с помощью функции $V(\varphi, [a, x])$, то справедливо следующее добавление к теореме 10.1.2.

Следствие 10.1.4. *Функция является непрерывной функцией ограниченной вариации в том и только том случае, когда её можно представить в виде разности двух непрерывных возрастающих функций.*

Рассмотрим функции, имеющие интегрируемую производную.

Теорема 10.1.5. *Если на отрезке $[a, b]$ функция φ непрерывна, а её производная интегрируема, то φ имеет ограниченную вариацию и*

$$V(\varphi, [a, b]) = \int_a^b |\varphi'(x)|\, dx. \qquad (10.1.6)$$

Доказательство. Для каждого разбиения T отрезка $[a, b]$ точками $a = x_0 < \dots < x_n = b$ согласно формуле конечных приращений Лагранжа имеем

$$\sum_{k=1}^{n} |\varphi(x_{k-1}) - \varphi(x_k)| = \sum_{k=1}^{n} |\varphi'(\xi_k)|(x_k - x_{k-1}),$$

где ξ_k – некоторые точки из (x_{k-1}, x_k).

Сумма в правой части полученного равенства является интегральной суммой Римана функции $|\varphi'(x)|$. Так как эти интегральные суммы неотрицательны, то верхняя грань таких сумм по всем разбиениям T равна значению интеграла из (10.1.6).

Теорема доказана.

§ 10.2. Определение интеграла Римана–Стилтьеса

Наряду с несобственными интегралами, рассматривавшимися в § 9.10, интеграл Римана–Стилтьеса является ещё одним обобщением интеграла Римана.

Пусть на отрезке $[a, b]$ заданы функции $f(x)$ и $\varphi(x)$. Для произвольного разбиения T отрезка $[a, b]$

$$a = x_0 < x_1 < \cdots < x_n = b$$

в каждом отрезке $[x_{k-1}, x_k]$ возьмём некоторую точку ξ_k и составим сумму

$$\sigma_T(f\,d\varphi, \xi) := \sum_{k=1}^{n} f(\xi_k)[\varphi(x_k) - \varphi(x_{k-1})].$$

Эту сумму называют *интегральной суммой Римана–Стилтьеса* (соответствующей разбиению T и выбору точек ξ) *функции* f *по функции* φ.

Определение. Пусть на отрезке $[a, b]$ заданы функции $f(x)$ и $\varphi(x)$. Функцию f называют *интегрируемой по функции* φ на $[a, b]$ в смысле интеграла Римана–Стилтьеса, если существует такое число I, что для любого $\varepsilon > 0$ имеется число $\delta(\varepsilon) > 0$, при котором для каждого разбиения T отрезка $[a, b]$ диаметра $\lambda_T < \delta$ при произвольном выборе точек ξ справедлива оценка

$$|I - \sigma_T(f\,d\varphi, \xi)| < \varepsilon.$$

Число I называют *интегралом Римана–Стилтьеса функции* f *по функции* φ на отрезке $[a, b]$ и обозначают

$$\int_a^b f(x)\,d\varphi(x). \tag{10.2.1}$$

Функцию f называют интегрируемой, а функцию φ – интегрирующей.

В случае $\varphi(x) \equiv x$ интегральная сумма Римана–Стилтьеса $\sigma_T(f\,d\varphi,\xi)$ является интегральной суммой Римана $S_T(f,\xi)$, а интеграл Римана–Стилтьеса – интегралом Римана.

Понятно, что если интеграл Римана–Стилтьеса существует, то его значение определяется однозначно.

Теорема 10.2.1 (Критерий Коши существования интеграла Римана–Стилтьеса). *Пусть на отрезке $[a,b]$ заданы функции $f(x)$ и $\varphi(x)$. Интеграл Римана–Стилтьеса функции f по φ существует в том и только том случае, когда выполнено условие Коши, состоящее в том, что для каждого $\varepsilon > 0$ существует $\delta = \delta(\varepsilon) > 0$ такое, что для любых разбиений T' и T'' отрезка $[a,b]$, диаметры которых меньше δ, и произвольных наборах точек ξ', ξ'' справедлива оценка*

$$|\sigma_{T'}(fd\varphi,\xi') - \sigma_{T''}(fd\varphi,\xi'')| < \varepsilon. \qquad (10.2.2)$$

Доказательство. Необходимость устанавливается с помощью стандартных рассуждений. Именно, если интеграл (10.2.1) существует, то для заданного $\varepsilon > 0$ находим $\delta > 0$ такое, что для каждого разбиения T отрезка $[a,b]$, диаметр которого $\lambda_T < \delta$, при любом выборе точек ξ выполняется оценка

$$|I - \sigma_T(f\,d\varphi,\xi)| < \frac{\varepsilon}{2}\,.$$

Тогда для произвольных разбиений T' и T'', диаметры которых меньше δ, справедлива оценка (10.2.2), т.е. выполнено условие Коши.

Докажем теперь достаточность.

Для $\varepsilon_p = 1/p$, $p \in \mathbb{N}$, найдём $\delta_p > 0$, при котором выполнено условие Коши. Не теряя общности, будем считать, что числа δ_p с ростом p монотонно убывают.

Для каждого p возьмём некоторое разбиение T_p отрезка $[a,b]$ с диаметром $\lambda_{T_p} < \delta_p$ и точки $\xi^{(p)}$ из отрезков этого разбиения. Значения полученных интегральных сумм $\sigma_{T_p}(f\,d\varphi,\xi^{(p)})$ обозначим σ_p.

Покажем, что для последовательности $\{\sigma_p\}$ выполнено условие Коши сходимости числовых последовательностей.

Для произвольного $\varepsilon > 0$ находим натуральное N такое, что $\varepsilon_N < \varepsilon$. Тогда для любых $p, q > N$ имеем $\lambda_{T_p} < \delta_p \leqslant \delta_N$ и $\lambda_{T_q} < \delta_q \leqslant \delta_N$ и согласно (10.2.2)

$$|\sigma_p - \sigma_q| = |\sigma_{T_p}(f\,d\varphi,\xi^{(p)}) - \sigma_{T_q}(f\,d\varphi,\xi^{(q)})| < \varepsilon.$$

Таким образом, последовательность $\{\sigma_p\}$ сходится. Покажем, что её предел

$$I := \lim_{p\to\infty} \sigma_{T_p}(f\,d\varphi, \xi^{(p)}).$$

является интегралом Римана–Стилтьеса (10.2.1).

Для заданного $\varepsilon > 0$ найдём n такое, что $\varepsilon_n < \varepsilon$. Тогда для каждого $p > n$ имеем $\lambda_{T_p} < \delta_p \leqslant \delta_n$. Поэтому в силу условия Коши для любого разбиения T с $\lambda_T < \delta_n$ и любого набора точек ξ при всех $p > n$ справедлива оценка

$$|\sigma_{T_p}(f\,d\varphi, \xi^{(p)}) - \sigma_T(f\,d\varphi, \xi)| < \varepsilon_n.$$

Переходя здесь к пределу при $p \to \infty$, получаем

$$|I - \sigma_T(f\,d\varphi, \xi)| \leqslant \varepsilon_n < \varepsilon.$$

Таким образом, I является интегралом Римана–Стилтьеса (10.2.1).

Теорема доказана.

При $\varphi(x) \equiv x$ теорема 10.2.1 даёт критерий Коши существования собственного интеграла Римана. В главе 9 об этом не говорилось.

На интеграл Римана–Стилтьеса легко переносятся некоторые свойства интеграла Римана. Например, линейность относительно интегрируемой функции:

$$\int_a^b (\alpha f(x) + \beta g(x))\,d\varphi(x) = \alpha \int_a^b f(x)\,d\varphi(x) + \beta \int_a^b g(x)\,d\varphi(x).$$

Здесь предполагается, что функции $f(x)$ и $g(x)$ интегрируемы по $\varphi(x)$, α и β – произвольные числа, а утверждается интегрируемость функции $\alpha f(x) + \beta g(x)$ и справедливость равенства.

Свойство линейности интеграла Римана–Стилтьеса имеет место и по отношению к интегрирующей функции:

$$\int_a^b f(x)\,d(\alpha\varphi(x) + \beta\psi(x)) = \alpha \int_a^b f(x)\,d\varphi(x) + \beta \int_a^b f(x)\,d\psi(x). \tag{10.2.3}$$

Здесь также предполагается, что существуют интегралы из правой части и утверждается существование интеграла в левой части и справедливость равенства.

Эти свойства основываются на линейности интегральных сумм Римана–Стилтьеса и критерии Коши. Доказательства не приводим ввиду их простоты.

Теорема 10.2.2. *Интеграл*

$$\int_a^b f(x)\, d\varphi(x) \tag{10.2.4}$$

существует, если на отрезке $[a, b]$ *функция* $f(x)$ *непрерывна, а функция* $\varphi(x)$ *имеет ограниченную вариацию.*

Доказательство. Так как функции ограниченной вариации представимы в виде разности двух возрастающих функций, то согласно (10.2.3) достаточно доказать теорему для возрастающей на $[a, b]$ функции φ.

Рассуждения следуют схеме доказательства с помощью сумм и интегралов Дарбу интегрируемости по Риману непрерывных функций.

Пусть T – разбиение отрезка $[a, b]$

$$a = x_0 < x_1 < \cdots < x_n = b$$

и $M_k(f)$ и $m_k(f)$ – числа, определённые для функции f формулами (9.2.1).

Рассмотрим верхние и нижние суммы Дарбу–Стилтьеса

$$\overline{S_T}(f\, d\varphi) := \sum_{k=1}^{n} M_k(f)(\varphi(x_k) - \varphi(x_{k-1})),$$

$$\underline{S_T}(f\, d\varphi) := \sum_{k=1}^{n} m_k(f)(\varphi(x_k) - \varphi(x_{k-1})).$$

В силу возрастания φ от добавления к разбиению T новой точки верхняя сумма $\overline{S_T}(f\, d\varphi)$ может только уменьшиться, а нижняя сумма $\underline{S_T}(f\, d\varphi)$ может только увеличиться. Поэтому справедлив аналог теоремы 9.2.2, согласно которому $\underline{S_{T_1}}(f\, d\varphi) \leqslant \overline{S_{T_2}}(f\, d\varphi)$ для любых разбиений T_1 и T_2.

Следовательно, существует верхний интеграл Дарбу–Стилтьеса

$$I^*(f\, d\varphi) := \inf_T \overline{S_T}(f\, d\varphi)$$

и для каждого разбиения T имеем

$$\underline{S_T}(f\,d\varphi) \leqslant I^*(f\,d\varphi) \leqslant \overline{S_T}(f\,d\varphi). \tag{10.2.5}$$

Для разности верхней и нижней сумм Дарбу–Стилтьеса справедливо неравенство, аналогичное оценке (9.2.10), установленной при доказательстве теоремы 9.2.6:

$$\overline{S_T}(f\,d\varphi) - \underline{S_T}(f\,d\varphi) \leqslant \omega(f, \lambda_T)(\varphi(b) - \varphi(a)). \tag{10.2.6}$$

где λ_T – диаметр разбиения T.

Для возрастающих функций φ при произвольных точках ξ интегральные суммы Римана–Стилтьеса связаны с суммами Дарбу–Стилтьеса неравенствами

$$\underline{S_T}(f\,d\varphi) \leqslant \sigma_T(f d\varphi, \xi) \leqslant \overline{S_T}(f\,d\varphi).$$

Поэтому согласно (10.2.5) и (10.2.6) при любом выборе точек ξ

$$\begin{aligned} |I^*(f\,d\varphi) - \sigma_T(f d\varphi, \xi)| \leqslant \overline{S_T}(f\,d\varphi) - \underline{S_T}(f\,d\varphi) \leqslant \\ \leqslant \omega(f, \lambda_T)(\varphi(b) - \varphi(a)). \end{aligned} \tag{10.2.7}$$

Функция f равномерно непрерывна на отрезке $[a, b]$, поэтому $\omega(f, +0) = 0$ и выражение в правой части (10.2.7) можно сделать как угодно малым за счёт малости диаметра разбиения T. Значит, интеграл (10.2.4) существует и теорема доказана.

Если φ не произвольная функция ограниченной вариации, а удовлетворяет условию Липшица первого порядка, то в теореме 10.2.2 требование непрерывности функции f можно ослабить.

Теорема 10.2.3. *Интеграл (10.2.4) существует, если на отрезке $[a, b]$ функция $\varphi(x)$ удовлетворяет условию Липшица первого порядка, а функция $f(x)$ интегрируема по Риману.*

Доказательство. Рассуждение аналогично доказательству теоремы 10.2.2.

Не теряя общности, функцию φ можно считать возрастающей. В самом деле, если $|\varphi(x') - \varphi(x'')| \leqslant L|x' - x''|$ для всех x' и x'', то функция $Lx - \varphi(x)$ на $[a, b]$ возрастает, так как при $a \leqslant x_1 < x_2 \leqslant b$

$$Lx_2 - \varphi(x_2) - (Lx_1 - \varphi(x_1)) = L(x_2 - x_1) - (\varphi(x_2) - \varphi(x_1)) \geqslant 0.$$

Поэтому тождество

$$\varphi(x) \equiv Lx - (Lx - \varphi(x))$$

даёт представление $\varphi(x)$ в виде разности двух возрастающих функций, удовлетворяющих, как легко видеть, условию Липшица первого порядка.

Оценим разность верхней и нижней сумм Дарбу–Стилтьеса функции f по φ. Если $|\varphi(x') - \varphi(x'')| \leqslant L|x' - x''|$, то для каждого разбиения T имеем

$$\begin{aligned} \overline{S_T}(f\,d\varphi) - \underline{S_T}(f\,d\varphi) &= \sum_{k=1}^{n}(M_k(f) - m_k(f))(\varphi(x_k) - \varphi(x_{k-1})) \leqslant \\ &\leqslant L\sum_{k=1}^{n}(M_k(f) - m_k(f))\Delta x_k = \\ &= L(\overline{S_T}(f) - \underline{S_T}(f)). \end{aligned}$$

Таким образом, разность сумм Дарбу–Стилтьеса функции f по φ оценена через разность сумм Дарбу–Римана функции f.

Следовательно, для любых точек ξ вместо (10.2.7) получаем оценку

$$|I^*(f\,d\varphi) - \sigma_T(f\,d\varphi, \xi)| \leqslant L(\overline{S_T}(f) - \underline{S_T}(f)),$$

и используя интегрируемость функции f по Риману, приходим к существованию интеграла (10.2.4).

§ 10.3. Свойства интеграла Римана–Стилтьеса

Аддитивность интеграла Римана относительно промежутка интегрирования (теорема 9.3.1) в отличие от свойств линейности на интеграл Римана–Стилтьеса переносится не в полном объёме.

Теорема 10.3.1. *Если функция $f(x)$ интегрируема по функции $\varphi(x)$ на отрезке $[a, b]$ и $c \in (a, b)$, то $f(x)$ интегрируема по $\varphi(x)$ на отрезках $[a, c]$ и $[c, b]$ и справедливо равенство*

$$\int_a^b f(x)\,d\varphi(x) = \int_a^c f(x)\,d\varphi(x) + \int_c^b f(x)\,d\varphi(x). \qquad (10.3.1)$$

Доказательство. Так как функция f интегрируема по φ на $[a,b]$, то согласно критерию Коши для каждого $\varepsilon > 0$ существует $\delta > 0$ такое, что для произвольных разбиений T_* и T_{**} отрезка $[a,b]$, диаметры которых меньше δ, при любых наборах точек ξ_* и ξ_{**} справедливо неравенство

$$|\sigma_{T_*}(f\,d\varphi,\xi_*) - \sigma_{T_{**}}(f\,d\varphi,\xi_{**})| < \varepsilon. \tag{10.3.2}$$

Чтобы с помощью критерия Коши доказать существование интеграла по отрезку $[a,c]$, рассмотрим произвольные разбиения T'_* и T'_{**} отрезка $[a,c]$ с диаметрами, меньшими указаного $\delta > 0$.

Оценим разность сумм Дарбу–Стилтьеса по отрезку $[a,c]$ для этих разбиений при произвольных наборах точек ξ^* и ξ^{**}

$$\sigma_{T'_*}(f\,d\varphi,\xi'_*) - \sigma_{T'_{**}}(f\,d\varphi,\xi'_{**}).$$

Дополним разбиения T'_* и T'_{**} одними и теми же точками из отрезка $[c,b]$ до соответственно разбиений T^* и T^{**} отрезка $[a,b]$ так, чтобы выполнялись оценки $\lambda_{T^*} < \delta$ и $\lambda_{T^{**}} < \delta$. На каждом из добавленных при этом отрезков разбиений возьмём одни и те же точки ξ_k. Полученные наборы точек обозначим ξ^* и ξ^{**}. Тогда

$$\sigma_{T'_*}(f\,d\varphi,\xi'_*) - \sigma_{T'_{**}}(f\,d\varphi,\xi'_{**}) = \sigma_{T^*}(f\,d\varphi,\xi^*) - \sigma_{T^{**}}(f\,d\varphi,\xi^{**}).$$

Согласно (10.3.2) имеем

$$|\sigma_{T^*}(f\,d\varphi,\xi^*) - \sigma_{T^{**}}(f\,d\varphi,\xi^{**})| < \varepsilon.$$

Таким образом, для интеграла

$$\int_a^c f(x)\,d\varphi(x)$$

выполнено условие Коши и, значит, этот интеграл сходится. Точно так же устанавливается сходимость интеграла по отрезку $[c,b]$.

Равенство (10.3.1) доказывается так же, как и соответствующее равенство (9.3.1) для интеграла Римана. Не будем повторять эти рассуждения.

Таким образом, если функция $f(x)$ интегрируема по функции $\varphi(x)$ на отрезке $[a,b]$, то интеграл от f по φ существует и на любом отрезке $[c,d] \subset [a,b]$.

Но в отличие от интеграла Римана из существования интегралов в правой части равенства (10.3.1) не следует существование интеграла в левой части.

Пусть, например,

$$g(x) := \begin{cases} 0 & \text{при } -1 \leqslant x \leqslant 0, \\ 1 & \text{при } 0 < x \leqslant 1 \end{cases}$$

и

$$\psi(x) := \begin{cases} 0 & \text{при } -1 \leqslant x < 0, \\ 1 & \text{при } 0 \leqslant x \leqslant 1. \end{cases}$$

Тогда интегралы

$$\int_{-1}^{0} g(x)\, d\psi(x), \qquad \int_{0}^{1} g(x)\, d\psi(x)$$

существуют и оба они равны нулю – первый в силу равенства нулю функции $g(x)$, а второй в силу постоянства функции $\psi(x)$.

Вместе с тем, интеграл

$$\int_{-1}^{1} g(x)\, d\psi(x)$$

не существует. В самом деле, если x_{k-1} и x_k – такие точки разбиения T, что $0 \in (x_{k-1}, x_k)$, то справедливо равенство

$$\sigma_T(g\, d\psi, \xi) = g(\xi_k)(\psi(x_k) - \psi(x_{k-1})) = g(\xi_k).$$

Значит, для $\xi_k > 0$ эта интегральная сумма равна 1, а для $\xi_k < 0$ она равна 0.

Аналогично устанавливается, что интеграл $\int f(x)\, d\varphi(x)$ не существует, если функции $f(x)$ и $\varphi(x)$ имеют общую точку разрыва первого рода, даже когда разрыв является устранимым.

Приведём теоремы об интегрировании неравенств для интеграла Римана–Стилтьеса.

Теорема 10.3.2. *Пусть на отрезке $[a, b]$ для функций $f(x)$ и $g(x)$ справедливо неравенство $f(x) \leqslant g(x)$ и эти функции интегрируемы по функции $\varphi(x)$, которая на $[a, b]$ возрастает. Тогда*

$$\int_{a}^{b} f(x)\, d\varphi(x) \leqslant \int_{a}^{b} g(x)\, d\varphi(x).$$

Теорема 10.3.3. *Пусть на отрезке* $[a,b]$ *функция* $\varphi(x)$ *имеет ограниченную вариацию и* $v(x) := V(\varphi,[a,x])$. *Тогда*

$$\left|\int_a^b f(x)\,d\varphi(x)\right| \leqslant \int_a^b |f(x)|\,dv(x)$$

для каждой функции $f(x)$, *для которой эти интегралы существуют.*

Утверждения теорем 10.3.2 и 10.3.3 вытекают из соответствующих оценок интегральных сумм Римана–Стилтьеса.

Теорема 10.3.4 (Формула интегрирования по частям). *Если на отрезке* $[a,b]$ *функция* $f(x)$ *интегрируема по функции* $\varphi(x)$, *то* $\varphi(x)$ *интегрируема по* $f(x)$ *и справедливо равенство*

$$\int_a^b \varphi(x)\,df(x) = f(x)\varphi(x)\Big|_a^b - \int_a^b f(x)\,d\varphi(x), \tag{10.3.3}$$

которое называют формулой интегрирования по частям.

Доказательство. Пользуясь существованием интеграла из правой части равенства (10.3.3), по $\varepsilon > 0$ находим $\delta > 0$ такое, что для каждого разбиения T' с $\lambda_{T'} < \delta$ и любого набора точек ξ' выполняется неравенство

$$\left|\int_a^b f(x)\,d\varphi(x) - \sigma_{T'}(f\,d\varphi,\xi')\right| < \varepsilon. \tag{10.3.4}$$

Для произвольного разбиения T

$$a = x_0 < x_1 < \dots < x_n = b,$$

диаметр которого $\lambda_T < \delta/2$, представим сумму Римана–Стилтьеса $\sigma_T(\varphi\,df,\xi)$ следующим образом:

$$\begin{aligned}\sigma_T(\varphi\,df,\xi) &= \sum_{k=1}^{n} \varphi(\xi_k)(f(x_k) - f(x_{k-1})) = \\ &= \sum_{k=1}^{n} \varphi(\xi_k)f(x_k) - \sum_{k=0}^{n-1} \varphi(\xi_{k+1})f(x_k).\end{aligned}$$

Положив $\xi_0 := a$ и $\xi_{n+1} := b$, получим

$$\sigma_T(\varphi\,df,\xi) = \sum_{k=0}^{n} f(x_k)[\varphi(\xi_k) - \varphi(\xi_{k+1})] - \varphi(a)f(a) + \varphi(b)f(b) =$$

$$= -\,\sigma_{T^*}(f\,d\varphi, x^*) + f(x)\varphi(x)\big|_a^b, \tag{10.3.5}$$

где T^* – разбиение отрезка $[a,b]$

$$a = \xi_0 \leqslant \xi_1 \leqslant \dots \leqslant \xi_n \leqslant \xi_{n+1} = b,$$

а x^* – набор точек x_k. Если какие-либо две точки ξ_k совпадают, считаем, что они задают одну точку разбиения T^*. При этом $x_k \in [\xi_k, \xi_{k+1}]$.

Для диаметра разбиения T^* справдлива оценка $\lambda_{T^*} \leqslant 2\lambda_T$. Поэтому $\lambda_{T^*} < \delta$ и согласно (10.3.4) для интегральной суммы $\sigma_{T^*}(f\,d\varphi, x^*)$ имеем

$$\left|\int_a^b f(x)\,d\varphi(x) - \sigma_{T^*}(f\,d\varphi, x^*)\right| < \varepsilon.$$

Таким образом, из (10.3.5) следует, что для произвольного разбиения T, для которого $\lambda_T < \delta/2$, и любого набора точек ξ имеет место оценка

$$\left|\int_a^b f(x)\,d\varphi(x) - f(x)\varphi(x)\big|_a^b + \sigma_T(\varphi\,df, \xi)\right| < \varepsilon.$$

Это обеспечивает существование интеграла

$$\int_a^b \varphi(x)\,df(x)$$

и выполнение равенства (10.3.3).

Теорема доказана.

Из теорем 10.3.4 и 10.2.2 следует, что функции ограниченной вариации интегрируемы по непрерывным функциям.

В некоторых случаях интеграл Римана–Стилтьеса можно представить как интеграл Римана.

Теорема 10.3.5. *Пусть функции $f(x)$ и $\varphi(x)$ интегрируемы по Риману на отрезке $[a,b]$ и*

$$\varPhi(x) := \int_a^x \varphi(t)\,dt.$$

Тогда функция $f(x)$ интегрируема по $\varPhi(x)$ и справедливо равенство

$$\int_a^b f(x)\,d\varPhi(x) = \int_a^b f(x)\varphi(x)\,dx. \tag{10.3.6}$$

Доказательство. Согласно теореме 9.6.1 функция $\varPhi(x)$ удовлетворяет условию Липшица первого порядка, а в силу теоремы 10.2.3 интеграл в левой части (10.3.6) существует.

Рассмотрим интегральную сумму Римана–Стилтьеса функции f по $\varPhi$:

$$\sigma_T(f\,d\varPhi,\xi) = \sum_{k=1}^{n} f(\xi_k)[\varPhi(x_k) - \varPhi(x_{k-1})],$$

где $x_0, x_1, \ldots, x_n$ – точки разбиения T и $\xi_k \in [x_{k-1}, x_k]$.

Пусть

$$M_k = M_k(\varphi) := \sup_x \varphi(x); \qquad m_k = m_k(\varphi) := \inf_x \varphi(x),$$

где верхняя и нижняя грани берутся по $x \in [x_{k-1}, x_k]$.

Поскольку

$$\varPhi(x_k) - \varPhi(x_{k-1}) = \int_{x_{k-1}}^{x_k} \varphi(t)\,dt,$$

имеем

$$m_k \Delta x_k \leqslant \varPhi(x_k) - \varPhi(x_{k-1}) \leqslant M_k \Delta x_k.$$

Так как

$$m_k \Delta x_k \leqslant \varphi(\xi_k)\Delta x_k \leqslant M_k \Delta x_k,$$

то

$$|[\varPhi(x_k) - \varPhi(x_{k-1})] - \varphi(\xi_k)\Delta x_k| \leqslant (M_k - m_k)\Delta x_k.$$

Таким образом, если $|f(x)| < B$ при $x \in [a, b]$, то

$$\begin{aligned}
&\left|\sigma_T(f\,d\varPhi,\xi) - \sum_{k=1}^{n} f(\xi_k)\varphi(\xi_k)\Delta x_k\right| \leqslant \\
&\quad \leqslant \sum_{k=1}^{n} |f(\xi_k)|\,|[\varPhi(x_k) - \varPhi(x_{k-1})] - \varphi(\xi_k)\Delta x_k| \leqslant \\
&\quad \leqslant \sum_{k=1}^{n} |f(\xi_k)|(M_k - m_k)\Delta x_k \leqslant B\sum_{k=1}^{n}(M_k - m_k)\Delta x_k = \\
&\quad = B(\overline{S_T}(\varphi) - \underline{S_T}(\varphi)). \qquad (10.3.7)
\end{aligned}$$

Здесь $\overline{S_T}(\varphi)$ и $\underline{S_T}(\varphi)$ – суммы Дарбу–Римана функции φ.

Зададим произвольное $\varepsilon > 0$. Согласно теореме 9.2.3 существует число $\delta_1 > 0$ такое, что для любого разбиения T с $\lambda_T < \delta_1$ справедлива оценка

$$\overline{S_T}(\varphi) - \underline{S_T}(\varphi) < \frac{\varepsilon}{2B}\,.$$

Так как функция $f(x)\varphi(x)$ интегрируема по Риману, существует $\delta_2 > 0$ такое, что для каждого разбиения T с $\lambda_T < \delta_2$ при любом наборе точек ξ

$$\left|\sum_{k=1}^{n} f(\xi_k)\varphi(\xi_k)\Delta x_k - \int_a^b f(x)\varphi(x)\,dx\right| < \frac{\varepsilon}{2}\,.$$

Положив $\delta := \min(\delta_1, \delta_2)$, получаем, пользуясь (10.3.7), что для каждого разбиения T с $\lambda_T < \delta$ и любом наборе точек ξ выполняется неравенство

$$\left|\sigma_T(f\,d\varPhi, \xi) - \int_a^b f(x)\varphi(x)\,dx\right| < \varepsilon,$$

откуда следует (10.3.6).

Теорема доказана.

Часто используется следующее утверждение, вытекающее из теоремы 10.3.5.

Следствие 10.3.6. *Если на отрезке $[a, b]$ функция $\varPhi(x)$ непрерывно дифференцируема, то каждая интегрируемая по Риману функция $f(x)$ интегрируема по функции $\varPhi(x)$ и справедливо равенство*

$$\int_a^b f(x)\,d\varPhi(x) = \int_a^b f(x)\,\varPhi'(x)\,dx.$$

Интеграл Римана–Стилтьеса выражается через интеграл Римана не всегда.

Например, пусть $c \in (a, b)$ и

$$g(x) := \begin{cases} 0 & \text{при } a \leqslant x < c, \\ 1 & \text{при } c \leqslant x \leqslant b. \end{cases}$$

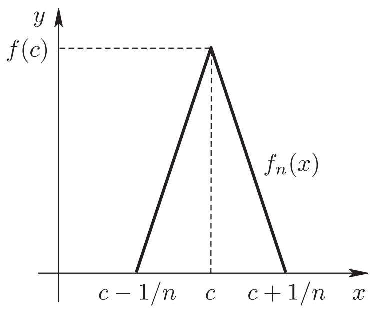
Легко видеть, что для каждой непрерывной на $[a, b]$ функции $f(x)$

$$\int_a^b f(x)\,dg(x) = f(c).$$

Но ни при какой интегрируемой по Риману функции $h(x)$ равенство
$$\int_a^b f(x)h(x)\,dx = f(c) \tag{10.3.8}$$
не может выполняться для всех непрерывных функций $f(x)$.

Покажем, что если бы такая интегрируемая по Риману функция $h(x)$ существовала, то из (10.3.8) следовало бы, что $f(c) = 0$.

Рассмотрим функцию $f_n(x)$, график которой изображен на рисунке.



Согласно (10.3.8) для каждого n
$$\int_a^b f_n(x)h(x)\,dx = f(c).$$

Но если $|h(x)| \leqslant H$ для $x \in [a,b]$, то
$$|f(c)| = \left|\int_a^b f_n(x)h(x)\,dx\right| \leqslant H|f(c)|\frac{1}{n}\,.$$
Так как эта оценка имеет место при всех n, то $f(c) = 0$.

Таким образом, равенство вида (10.3.8) не может выполняться для всех непрерывных функций $f(x)$.

Продолжим изучение свойств интеграла Римана–Стилтьеса.

Пусть функция $\psi(x)$ равна 1 в точке c, являющейся внутренней точкой отрезка $[a,b]$, и равна нулю во всех остальных его точках. Поскольку $\psi(x)$ – функция ограниченной вариации, согласно теореме 10.2.2 для каждой непрерывной функции $f(x)$ существует интеграл
$$\int_a^b f(x)\,d\psi(x). \tag{10.3.9}$$

Так как для каждого разбиения T отрезка $[a,b]$, не содержащего точку c, интегральная сумма Римана–Стилтьеса $\sigma_T(f\,d\psi,\xi)$ равна нулю, интеграл (10.3.9) равен нулю.

Поэтому справедливо следующее утверждение.

Теорема 10.3.7. *Пусть функция $f(x)$ непрерывна на $[a,b]$, а функция $\varphi(x)$ такова, что существует интеграл*

$$\int_a^b f(x)\,d\varphi(x). \tag{10.3.10}$$

Если значения функции $\varphi(x)$ произвольным образом изменить в конечном числе внутренних точек отрезка $[a,b]$, то полученный интеграл будет существовать и значение интеграла (10.3.10) не изменится.

Выше отмечалось, что если функция $f(x)$ кусочно непрерывна и функция ограниченной вариации $\varphi(x)$ разрывна в точках разрыва f, то интеграл (10.3.10) не существует. Покажем, что этот интеграл существует, если f и φ не имеют общих точек разрыва.

Теорема 10.3.8. *Интеграл Римана–Стилтьеса (10.3.10) существует, если на отрезке $[a,b]$ функция $f(x)$ кусочно непрерывна, а функция $\varphi(x)$ имеет ограниченную вариацию и непрерывна в точках разрыва функции $f(x)$.*

Доказательство. Рассмотрим сначала случай, когда $f(x)$ имеет на $[a,b]$ одну точку разрыва (первого рода). Пусть это точка c.

Если разрыв f является неустранимым, то прибавив к $f(x)$ функцию

$$g(x) := \begin{cases} 0 & \text{при } a \leqslant x < c, \\ 1 & \text{при } c \leqslant x \leqslant b, \end{cases}$$

умноженную на соответствующее число, можно получить непрерывную функцию, которая интегрируема по φ.

Функция g интегрируема по φ, так как если c принадлежит отрезку $[x_{m-1},x_m]$ разбиения T, то

$$\sigma_T(g\,d\varphi,\xi) = g(\xi_m)(\varphi(x_m)-\varphi(x_{m-1})) + \sum_{k=m+1}^{n}(\varphi(x_k)-\varphi(x_{k-1})) =$$

$$= g(\xi_m)(\varphi(x_m)-\varphi(x_{m-1})) + \varphi(b)-\varphi(x_m),$$

а при $\Delta x_m \to 0$ в силу непрерывности $\varphi(x)$ в точке c предел выражения в правой части этого равенства равен $\varphi(b) - \varphi(c)$.

Если c – точка устранимого разрыва функции f, то вводим функцию

$$h(x) := \begin{cases} 1 & \text{при } x = c, \\ 0 & \text{при } x \neq c \end{cases}$$

и проводим аналогичные рассуждения с функцией $h(x)$ вместо $g(x)$. Тогда в тех же обозначениях имеем

$$\sigma_T(h\,d\varphi, \xi) = h(\xi_m)(\varphi(x_m) - \varphi(x_{m-1})),$$

а это выражение стремится к нулю при $\Delta x_m \to 0$.

Итак, теорема доказана, если $f(x)$ имеет одну точку разрыва.

В общем случае теорема вытекает из линейности интеграла Римана–Стилтьеса и того, что функцию $f(x)$ можно сделать непрерывной на $[a, b]$, прибавив к ней линейную комбинацию функций $g(x + \alpha)$ и $h(x + \beta)$ при соответствующих сдвигах аргумента α и β.

§ 10.4. Задачи и упражнения

10.4.1. Докажите, что если на отрезке функция $f(x)$ имеет ограниченную вариацию и $|f(x)| \geqslant d > 0$, то функция $1/f(x)$ также имеет ограниченную вариацию.

10.4.2. Является ли ограниченность функции $f(x)$ необходимым условием существования интеграла Римана–Стилтьеса

$$\int_a^b f(x)\,d\varphi(x)?$$

10.4.3. Докажите, что интеграл Римана–Стилтьеса от непрерывной функции по непрерывной функции может не существовать.

10.4.4. Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и существует интеграл Римана–Стилтьеса

$$\int_a^b f(x)\,d\varphi(x). \tag{10.4.1}$$

Согласно теореме 10.3.7 значение интеграла (10.4.1) сохраняется при изменении функции φ во внутренней точке отрезка $[a, b]$. Изменится ли и, если изменится, то как, значение интеграла (10.4.1), если изменить функцию φ в концевых точках a или b?

Глава 11. Функции многих переменных

§ 11.1. Многомерные евклидовы пространства

До сих пор изучались функции одной переменной, аргументами которых были числа. При этом широко использовались наглядные геометрические представления, хотя доказательства на них не опирались. Аргументы считались точками числовой прямой, рассматривались расстояния между точками, окрестности точек и т.п.

Для функций двух переменных аргументами являются упорядоченные пары чисел x, y. Значение функции f, которое она принимает при заданных x и y, обозначают $f(x, y)$. Как и в случае функций одной переменной, функции двух переменных будем обозначать и f и $f(x, y)$.

При изучении функций двух переменных $f(x, y)$ аргументы естественно рассматривать как точки плоскости, на которой выбрана декартова прямоугольная система координат, и считать, что областью определения функции является некоторое множество точек этой плоскости.

Аналогично, функции трёх переменных $f(x, y, z)$ считают заданными на множествах точек трёхмерного пространства. Если в этом пространстве выбрана декартова прямоугольная система координат, то каждую точку можно рассматривать как упорядоченную тройку чисел (x, y, z), являющихся координатами этой точки. В курсе аналитической геометрии было доказано, что расстояние между точками (x', y', z') и (x'', y'', z'') равно

$$\sqrt{(x' - x'')^2 + (y' - y'')^2 + (z' - z'')^2}.$$

Чтобы пользоваться геометрической интерпретацией и для функций большего, чем три, числа переменных, используют многомерные евклидовы пространства и считают, что функции заданы на множествах точек таких пространств.

По определению точками m-мерного евклидова пространства $\mathbb{E}^m$, $m \in \mathbb{N}$, называют упорядоченные наборы из m чисел. Если точка обозначена $\mathbf{x}$, то соответствующий ей набор чисел обозначают $(x_1, x_2, \dots, x_m)$, а сами эти числа называют координатами точки $\mathbf{x}$. Как и в трёхмерном случае, точки совпадают тогда и только тогда, когда равны все их соответствующие координаты. Расстоянием между точками $\mathbf{x}$ и $\mathbf{y}$ называется число

$$\rho(\mathbf{x}, \mathbf{y}) := \left(\sum_{k=1}^{m} (x_k - y_k)^2 \right)^{1/2}. \qquad (11.1.1)$$

В трёхмерном случае эта формула была доказана, а в $\mathbb{E}^m$ при $m > 3$ она принимается в качестве определения расстояния.

Итак, $\mathbb{E}^m$ – это множество точек $\mathbf{x} = (x_1, x_2, \dots, x_m)$, расстояние между которыми задаётся формулой (11.1.1).

Свойства расстояния:

1°) $\rho(\mathbf{x}, \mathbf{y}) \geqslant 0$ для любых точек $\mathbf{x}$ и $\mathbf{y}$, причём $\rho(\mathbf{x}, \mathbf{y}) = 0$ в том и только том случае, когда $\mathbf{x} = \mathbf{y}$;
2°) $\rho(\mathbf{x}, \mathbf{y}) = \rho(\mathbf{y}, \mathbf{x})$ для каждой пары точек $\mathbf{x}, \mathbf{y}$;
3°) для любых трёх точек $\mathbf{x}, \mathbf{y}, \mathbf{z}$ справедливо *неравенство треугольника*:

$$\rho(\mathbf{x}, \mathbf{y}) \leqslant \rho(\mathbf{x}, \mathbf{z}) + \rho(\mathbf{z}, \mathbf{y}). \qquad (11.1.2)$$

Свойства 1° и 2° следуют сразу из определения расстояния.

Неравенство (11.1.2) вытекает из неравенства Минковского (6.8.5), согласно которому для любых $2m$ чисел a_k, b_k, $k = 1, \dots, m$,

$$\left(\sum_{k=1}^{m} (a_k + b_k)^2 \right)^{1/2} \leqslant \left(\sum_{k=1}^{m} a_k^2 \right)^{1/2} + \left(\sum_{k=1}^{m} b_k^2 \right)^{1/2}. \qquad (11.1.3)$$

Полагаем $a_k := x_k - z_k$, $b_k := z_k - y_k$, $k = 1, \dots, m$. Тогда $a_k + b_k = x_k - y_k$ и, подставив эти значения в (11.1.3), приходим к (11.1.2).

Многомерные евклидовы пространства являются частным случаем метрических пространств.

Определение. Множество X элементов произвольной природы называется *метрическим пространством*, если на элементах X определена функция двух переменных $\rho(\mathbf{x}, \mathbf{y})$, обладающая свойствами расстояния 1°–3°.

Говорят, что пара (X, ρ) задаёт метрическое пространство или является метрическим пространством. Функцию ρ называют *метрикой*.

Свойства 1°–3° играют роль аксиом метрики.

Некоторые определения и утверждения, которые будут рассматриваться далее в многомерных евклидовых пространств, могут быть отнесены к произвольным метрическим пространствам, но не будем на этом останавливаться.

В m-мерном евклидовом пространстве $\mathbb{E}^m$ замкнутым шаром радиуса $r > 0$ с центром в точке $\mathbf{a}$ называют множество точек $\mathbf{x}$, для которых $\rho(\mathbf{x}, \mathbf{a}) \leqslant r$. Граница этого шара, т.е. ограничивающая его сфера $\rho(\mathbf{x}, \mathbf{a}) = r$, принадлежит шару. Множество точек $\mathbf{x}$, для которых $\rho(\mathbf{x}, \mathbf{a}) < r$, называют открытым шаром радиуса r с центром в точке $\mathbf{a}$. Граница такого шара ему не принадлежит.

Определяются также m-мерные прямоугольные параллелепипеды, которые для краткости будем называть m-мерными прямоугольниками. Пусть имеются два набора по m чисел α_k, β_k, $k = 1, \ldots, m$, причём $\alpha_k < \beta_k$ при каждом k. Множество точек $\mathbf{x} \in \mathbb{E}^m$, для координат которых выполняются неравенства $\alpha_k \leqslant x_k \leqslant \beta_k$, $k = 1, \ldots, m$, называют замкнутым m-мерным прямоугольником. Множество точек $\mathbf{x} \in \mathbb{E}^m$, для координат которых выполняются неравенства $\alpha_k < x_k < \beta_k$, $k = 1, \ldots, m$, называют открытым m-мерным прямоугольником. Как и в трёхмерном случае, точки прямоугольников, у которых хотя бы при одном k имеет место равенство $x_k = \alpha_k$ или равенство $x_k = \beta_k$ образуют грани прямоугольника. Они параллельны координатным плоскостям.

Если разности $\beta_k - \alpha_k$ при всех k равны между собой, то говорят об m-мерных квадратах или m-мерных кубах.

При этом появляется некоторая нечёткость терминологии, когда в m-мерном пространстве используются названия, соответствующие иногда двумерному, а иногда трёхмерному случаю.

Как и в трёхмерном случае, в $\mathbb{E}^m$ каждый шар можно поместить в куб, а каждый куб окружить шаром. Действительно, если $\mathbf{a}$ – центр шара радиуса r, то сам шар (для определённости открытый) состоит из точек $\mathbf{x}$ таких, что

$$\sum_{k=1}^{m} (x_k - a_k)^2 < r^2.$$

Отсюда $|x_k - a_k| < r$, $k = 1, \ldots, m$, т.е. $a_k - r < x_k < a_k + r$ при каждом k. Поэтому все точки шара принадлежат кубу с центром в точке $\mathbf{a}$, длина ребра которого равна $2r$ (рёбра определяются по аналогии с трёмерным случаем).

Наоборот, рассмотрим куб с центром в точке $\mathbf{a}$ и длиной ребра $2r$, т.е. точки $\mathbf{x}$, для которых $|x_k - a_k| < r$, $k = 1, \ldots, m$. Тогда

$$\rho(\mathbf{x}, \mathbf{a}) = \left(\sum_{k=1}^{m}(x_k - a_k)^2\right)^{1/2} < \left(\sum_{k=1}^{m} r^2\right)^{1/2} = r\sqrt{m}.$$

Значит, этот куб содержится в шаре радиуса $r\sqrt{m}$ с центром в точке $\mathbf{a}$.

Открытый шар радиуса $\varepsilon > 0$ с центром в точке $\mathbf{x} \in \mathbb{E}^m$ называют *шаровой* ε-окрестностью точки $\mathbf{x}$. Аналогично, открытые кубы с центром в точке $\mathbf{x}$ называют кубическими окрестностями точки $\mathbf{x}$.

Согласно сказанному каждая шаровая окрестность содержит кубическую окрестность и в свою очередь содержится в некоторой кубической окрестности. Поэтому, когда говорится о достаточно малых окрестностях точки, можно иметь в виду как шаровые, так и кубические окрестности.

Определение. Пусть в пространстве $\mathbb{E}^m$ задана последовательность точек $\{\mathbf{x}^{(n)}\}$, $n = 1, 2, \ldots$. Точку $\mathbf{a} \in \mathbb{E}^m$ называют *пределом последовательности* $\{\mathbf{x}^{(n)}\}$, если для каждого положительного ε существует число $N(\varepsilon)$ такое, что при всех $n > N$ справедливо неравенство

$$\rho(\mathbf{x}^{(n)}, \mathbf{a}) < \varepsilon.$$

В этом случае пишут

$$\lim_{n\to\infty} \mathbf{x}^{(n)} = \mathbf{a}.$$

Таким образом, сходимость последовательности $\{\mathbf{x}^{(n)}\}$ к $\mathbf{a}$ означает, что числовая последовательность $\rho(\mathbf{x}^{(n)}, \mathbf{a})$ сходится к нулю.

В определении предела всё выглядит, как и для числовых последовательностей, когда $m = 1$. Тогда расстояние между соответствующими точками (числами) равнялось $|x^{(n)} - a|$.

Теорема 11.1.1. *Последовательность точек* $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots$ *пространства* $\mathbb{E}^m$ *сходится к точке* $\mathbf{a} \in \mathbb{E}^m$ *в том и только том случае, когда для каждого* $k, k = 1, \ldots, m$, *числовая последовательность* $x_k^{(1)}, x_k^{(2)}, \ldots$ *сходится к* a_k.

Доказательство. Так как

$$\rho(\mathbf{x}^{(n)}, \mathbf{a}) = \left(\sum_{k=1}^{m} (x_k^{(n)} - a_k)^2 \right)^{1/2}, \tag{11.1.4}$$

то $|x_k^{(n)} - a_k| \leqslant \rho(\mathbf{x}^{(n)}, \mathbf{a})$ для каждого k. Поэтому из $\rho(\mathbf{x}^{(n)}, \mathbf{a}) \to 0$ при $n \to \infty$, следует, что $x_k^{(n)} \to a_k$, $n \to \infty$.

Наоборот, если при каждом k имеем $x_k^{(n)} \to a_k$, $n \to \infty$, то согласно (11.1.4) $\rho(\mathbf{x}^{(n)}, \mathbf{a}) \to 0$, что заканчивает доказательство теоремы.

Множество точек пространства $\mathbb{E}^m$ называется *ограниченным*, если существует шар, содержащий все точки этого множества. Конечно, здесь можно говорить о кубе или прямоугольнике.

Для непустых ограниченных множеств $A \subset \mathbb{E}^m$ вводится *диаметр множества*, который обозначают $\operatorname{diam} A$. По определению полагают

$$\operatorname{diam} A := \sup_{\mathbf{x}, \mathbf{y} \in A} \rho(\mathbf{x}, \mathbf{y}).$$

Ясно, что для прямоугольника $I := \{\mathbf{x} \in \mathbb{E}^m : a_k \leqslant x_k \leqslant b_k,\ k = 1, \ldots, m\}$ имеем

$$\operatorname{diam} I = \left(\sum_{k=1}^{m} (b_k - a_k)^2 \right)^{1/2}.$$

Такое равенство справедливо и для диаметра соответствующего открытого прямоугольника.

Малость размера множества характеризуют малостью его диаметра.

Установим в пространстве $\mathbb{E}^m$ свойство, аналогичное теореме 1.7.1 о вложенных отрезках числовой прямой.

Теорема 11.1.2. *Для каждой последовательности вложенных замкнутых m-мерных прямоугольников $\{I^{(n)}\}$, диаметры которых стремятся к нулю, существует единственная точка, принадлежащая всем этим прямоугольникам.*

Доказательство. Понятно, что достаточно рассмотреть случай, когда грани всех прямоугольников $I^{(n)}$ параллельны координатным плоскостям некоторой декартовой прямоугольной системы координат.

Пусть прямоугольники $I^{(n)}$ заданы неравенствами $a_k^{(n)} \leqslant x_k \leqslant b_k^{(n)}$, $k = 1, \ldots, m$. Согласно условиям теоремы при каждом k имеем последовательность вложенных отрезков $\{[a_k^{(n)}, b_k^{(n)}]\}$, длины которых стремятся к нулю. Значит, существует единственная точка x_k^*, принадлежащая всем отрезкам $[a_k^{(n)}, b_k^{(n)}]$, $n = 1, 2, \ldots$. Точка $\mathbf{x}^* = (x_1^*, \ldots, x_m^*)$ принадлежит всем прямоугольникам $I^{(n)}$, а из стремления к нулю диаметров прямоугольников следует, что такая точка только одна. Теорема доказана.

Теорема 11.1.3 (Теорема Больцано–Вейерштрасса). *Каждая ограниченная последовательность точек пространства $\mathbb{E}^m$ имеет сходящуюся подпоследовательность.*

Доказательство. Пусть $\{\mathbf{x}^{(n)}\}$ – ограниченная последовательность точек из $\mathbb{E}^m$. Рассмотрим куб $I = \{\mathbf{x} \in \mathbb{E}^m : a_k \leqslant x_k \leqslant a_k + c,\ k = 1, \ldots, m\}$, содержащий все точки $\mathbf{x}^{(n)}$. Ясно, что $\operatorname{diam} I = \sqrt{m}\, c$.

Разделим каждое ребро куба I пополам и рассмотрим замкнутые кубики, построенные на полученных при этом половинках рёбер. Всего таких кубиков 2^m штук. Обозначим $I^{(1)}$ любой из этих кубиков, содержащий бесконечно много точек последовательности $\{\mathbf{x}^{(n)}\}$. По крайней мере один такой кубик обязательно существует. Заметим, что $\operatorname{diam} I^{(1)} = \sqrt{m}\, c/2 = 2^{-1} \operatorname{diam} I$. Выберем некоторый элемент $\mathbf{x}^{(n_1)}$ последовательности $\{\mathbf{x}^{(n)}\}$, принадлежащий $I^{(1)}$.

Делим теперь каждое ребро куба $I^{(1)}$ пополам, получаем 2^m более мелких замкнутых кубиков следующего поколения и в качестве $I^{(2)}$ берём любой из таких кубиков, содержащий бесконечно много точек исходной последовательности. Выбираем в $I^{(2)}$ элемент $\mathbf{x}^{(n_2)}$ последовательности $\{\mathbf{x}^{(n)}\}$, для которого $n_2 > n_1$. При этом $\operatorname{diam} I^{(2)} = 2^{-2} \operatorname{diam} I$.

Продолжив неограниченно такое построение, получим последовательность вложенных кубов $\{I^{(p)}\}$, причём $\operatorname{diam} I^{(p)} \to 0$, $p \to \infty$, и подпоследовательность $\{\mathbf{x}^{(n_p)}\}$ последовательности $\{\mathbf{x}^{(n)}\}$. Согласно теореме 11.1.2 существует единственная точка $\mathbf{x}^*$, принадлежащая всем кубам $I^{(p)}$.

Эта точка $\mathbf{x}^*$ является пределом последовательности $\{\mathbf{x}^{(n_p)}\}$, так как любая ε-окрестность точки $\mathbf{x}^*$ содержит все кубики $I^{(n_p)}$, начиная с некоторого.

Теорема доказана.

В пространствах $\mathbb{E}^m$ при $m > 3$ по аналогии с трёхмерным случаем определяются непрерывные кривые.

Пусть на некотором (конечном или бесконечном) промежутке $[a, b]$ заданы непрерывные функции $\varphi_1(t), \ldots, \varphi_m(t)$. Множество точек с координатами $(\varphi_1(t), \ldots, \varphi_m(t))$, $t \in [a, b]$, называют *непрерывной кривой* в $\mathbb{E}^m$.

Прямой в пространстве $\mathbb{E}^m$ называют множество точек $\mathbf{x}$, для координат которых справедливы равенства

$$\begin{gathered} x_1 = x_1^0 + \alpha_1 t, \\ \ldots\ldots\ldots\ldots\ldots \\ x_m = x_m^0 + \alpha_m t, \end{gathered}$$

где $\mathbf{x}^0 = (x_1^0, \ldots, x_m^0)$ – некоторая точка (этой прямой), $\alpha_1, \ldots, \alpha_m$ – числа, не равные нулю одновременно (их считают компонентами направляющего вектора прямой), а параметр t принимает всевозможные значения.

В трёхмерном случае такие равенства рассматривались как уравнение прямой в параметрическом виде. А в пространствах большего числа измерений эти равенства принимают за определение прямой.

Если здесь параметр t пробегает некоторый отрезок, интервал или полуотрезок, то говорят, что указанные формулы задают в $\mathbb{E}^m$ соответственно отрезок, интервал или полуотрезок.

Понятно, что в $\mathbb{E}^m$ через каждые две точки можно провести прямую.

Множество $A \subset \mathbb{E}^m$ называют *выпуклым*, если для любой пары его точек отрезок, соединяющий эти точки, принадлежит A.

В пространствах $\mathbb{E}^m$ при $m > 3$ наряду с прямыми рассматривают двумерные, трёхмерные, ..., $(m-1)$-мерные плоскости, при этом плоскости размерности $m - 1$ называют *гиперплоскостями*. Дадим точные определения.

Гиперплоскость задаётся уравнением относительно $x_1, \ldots, x_m$ вида

$$\sum_{k=1}^{m} a_k (x_k - x_k^0) = 0, \tag{11.1.5}$$

в котором среди чисел $a_1, \ldots, a_m$ есть отличные от нуля.

Пересечение l гиперплоскостей, заданных уравнениями

$$\sum_{k=1}^{m} a_{j,k}(x_k - x_{j,k}^0) = 0, \qquad j = 1, \dots, l, \tag{11.1.6}$$

называют $(m-l)$-мерной плоскостью в $\mathbb{E}^m$, если ранг матрицы

$$\begin{pmatrix} a_{1,1} & \dots & a_{1,m} \\ \dots & \dots & \dots \\ a_{l,1} & \dots & a_{l,m} \end{pmatrix}$$

из коэффициентов $a_{j,k}$ равен l. В этом случае решения системы линейных уравнений (11.1.6) можно представить в виде

$$\begin{gathered} x_1 = x_1^0 + \alpha_{1,1}t_1 + \dots + \alpha_{1,m-l}t_{m-l}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_m = x_m^0 + \alpha_{m,1}t_1 + \dots + \alpha_{m,m-l}t_{m-l}, \end{gathered}$$

где параметры $t_1, \dots, t_{m-l}$ независимо друг от друга принимают всевозможные значения от $-\infty$ до $+\infty$, а $\alpha_{1,1}, \dots, \alpha_{1,m-l}, \dots, \alpha_{m,1}, \dots, \alpha_{m,m-l}$ – такие числа, что ранг матрицы

$$\begin{pmatrix} \alpha_{1,1} & \dots & \alpha_{1,m-l} \\ \dots & \dots & \dots \\ \alpha_{m,1} & \dots & \alpha_{m,m-l} \end{pmatrix}$$

равен $m-l$.

В пространствах $\mathbb{E}^m$ вводят m-мерные векторы.

Вектор – это направленный отрезок. Компонентами вектора с началом в точке $(a_1, \dots, a_m)$ и концом в точке $(b_1, \dots, b_m)$ по определению являются числа $(b_1 - a_1, \dots, b_m - a_m)$.

Скалярное произведение векторов $\overline{\mathbf{x}} = (x_1, \dots, x_m)$ и $\overline{\mathbf{y}} = (y_1, \dots, y_m)$ задаётся формулой

$$(\overline{\mathbf{x}}, \overline{\mathbf{y}}) := \sum_{k=1}^{m} x_k y_k.$$

При $m = 3$ так выражается скалярное произведение векторов через их компоненты в прямоугольной декартовой системе координат.

Длину вектора $\overline{\mathbf{x}}$ определяют равенством

$$|\overline{\mathbf{x}}| := \sqrt{(\overline{\mathbf{x}}, \overline{\mathbf{x}})}.$$

Как и при $m = 3$, точки в $\mathbb{E}^m$ не отличают от их радиусов-векторов.

Выведем формулу для расстояния точки до гиперплоскости, которое понимается как точная нижняя грань расстояний этой точки до точек гиперплоскости.

Будем считать, что числа $a_1, \ldots, a_m$ в уравнении гиперплоскости (11.1.5) удовлетворяют условию

$$\sum_{k=1}^{m} a_k^2 = 1 \tag{11.1.7}$$

и введём вектор $\overline{n} := (a_1, \ldots, a_m)$. Тогда $|\overline{n}| = 1$.

Если $\overline{r}$ – произвольная точка плоскости (11.1.5), а $\overline{r}_0$ – некоторая фиксированная точка этой плоскости, то уравнение (11.1.5) можно записать с помощью скалярного произведения

$$(\overline{r} - \overline{r}_0, \overline{n}) = 0. \tag{11.1.8}$$

Найдём расстояние фиксированной точки $\overline{R} \in \mathbb{E}^m$ до плоскости (11.1.8).

Проведём через точку $\overline{R}$ прямую

$$\overline{r} = \overline{R} + t\overline{n}, \qquad -\infty < t < +\infty,$$

параллельную вектору $\overline{n}$, и найдём точку её пересечения с плоскостью (11.1.8). Для t, соответствующего точке пересечения, имеем

$$(\overline{R} + t\overline{n} - \overline{r}_0, \overline{n}) = 0,$$

откуда $t = -(\overline{R} - \overline{r}_0, \overline{n})$. Таким образом, прямая пересекает плоскость в точке

$$\overline{r}^* = \overline{R} - (\overline{R} - \overline{r}_0, \overline{n})\overline{n}. \tag{11.1.9}$$

Покажем, что для произвольной точки $\overline{r}$ плоскости (11.1.8) справедливо равенство

$$\rho^2(\overline{R}, \overline{r}) = \rho^2(\overline{R}, \overline{r}^*) + \rho^2(\overline{r}^*, \overline{r}), \tag{11.1.10}$$

являющееся многомерным аналогом теоремы Пифагора. Обе точки $\overline{r}$ и $\overline{r}^*$ принадлежат плоскости, поэтому $(\overline{r}-\overline{r}^*,\overline{n}) = 0$. Значит, $(\overline{R}-\overline{r}^*,\overline{r}-\overline{r}^*) = 0$ и

$$\begin{aligned}\rho^2(\overline{R},\overline{r}) &= (\overline{R}-\overline{r},\overline{R}-\overline{r}) = (\overline{R}-\overline{r}^*+\overline{r}^*-\overline{r},\overline{R}-\overline{r}^*+\overline{r}^*-\overline{r}) = \\ &= (\overline{R}-\overline{r}^*,\overline{R}-\overline{r}^*) + (\overline{r}^*-\overline{r},\overline{r}^*-\overline{r}) = \\ &= \rho^2(\overline{R},\overline{r}^*) + \rho^2(\overline{r}^*,\overline{r}).\end{aligned}$$

Из (11.1.10) следует, что минимальное расстояние точки $\overline{R}$ от точек плоскости (11.1.8) равно $\rho(\overline{R},\overline{r}^*)$.

С помощью (11.1.9) находим

$$\rho^2(\overline{R},\overline{r}^*) = (\overline{R}-\overline{r}^*,\overline{R}-\overline{r}^*) = (\overline{R}-\overline{r}_0,\overline{n})^2.$$

Следовательно, расстояние точки $\overline{R}$ до плоскости (11.1.8) равно

$$\rho(\overline{R},\overline{r}^*) = |(\overline{R}-\overline{r}_0,\overline{n})|.$$

Если $(R_1,\ldots,R_m)$ – координаты точки $\overline{R}$, то через коэффициенты уравнения (11.1.5) это равенство записывается так:

$$\rho(\overline{R},\overline{r}^*) = \left|\sum_{k=1}^{m} a_k(R_k - x_k^0)\right|. \tag{11.1.11}$$

Напомним, что числа $a_1,\ldots,a_m$ здесь удовлетворяли условию (11.1.7).

В трёхмерном случае проведенные рассуждения означают, что из точки $\overline{R}$ проведен перпендикуляр к плоскости (11.1.5) и найдена длина этого перпендикуляра.

§ 11.2. Открытые и замкнутые множества

Определение. Точка $\mathbf{x}$ называется *внутренней точкой* множества $A \subset \mathbb{E}^m$, если существует ε-окрестность точки $\mathbf{x}$, принадлежащая A.

Множество называется *открытым*, если каждая его точка является внутренней.

Покажем, что открытый шар является открытым множеством. Этот факт нуждается в доказательстве, так как определения открытого шара и открытого множества давались по-разному.

Обычно в подобных случаях целесообразно сначала обосновать соответствующее утверждение для множеств на плоскости из элементарных геометрических соображений. А затем провести такие рассуждения в m-мерном пространстве.

Пусть S_0 – открытый шар радиуса r с центром в точке $\mathbf{x}_0$ и точка $\mathbf{x} \in S_0$, т.е. $\rho(\mathbf{x}_0, \mathbf{x}) < r$. Чтобы показать, что $\mathbf{x}$ – внутренняя точка шара S_0, построим открытый шар S_1 с центром в $\mathbf{x}$ радиуса $r - \rho(\mathbf{x}_0, \mathbf{x})$. Для каждой точки $\mathbf{y}$ шара S_1 согласно неравенству треугольника

$$\rho(\mathbf{x}_0, \mathbf{y}) \leqslant \rho(\mathbf{x}_0, \mathbf{x}) + \rho(\mathbf{x}, \mathbf{y}) < \rho(\mathbf{x}_0, \mathbf{x}) + r - \rho(\mathbf{x}_0, \mathbf{x}) = r.$$

Значит, $\mathbf{y}$ лежит в S_0, поэтому $\mathbf{x}$ – внутренняя точка шара S_0 и S_0 – открытое множество.

Аналогично можно доказать, что открытые m-мерные прямоугольники также являются открытыми множествами.

Теорема 11.2.1. *Объединение произвольного набора открытых множеств является открытым множеством.*

Пересечение конечного набора открытых множеств является открытым множеством.

Доказательство. Пусть множество G является объединением открытых множеств G_α

$$G = \bigcup_{\alpha \in \mathcal{A}} G_\alpha,$$

где $\mathcal{A}$ – некоторый набор индексов α.

Если точка $\mathbf{x} \in G$, то $\mathbf{x} \in G_{\alpha_0}$ при некотором α_0. Так как множество G_{α_0} – открытое, то вместе с $\mathbf{x}$ ему принадлежит и некоторая ε-окрестность точки $\mathbf{x}$. Эта окрестность принадлежит G и, значит, множество G является открытым.

Докажем теперь утверждение о пересечениях. Пусть G является пересечением конечного набора открытых множеств G_k, $k = 1, \ldots, n$:

$$G = \bigcap_{k=1}^{n} G_k.$$

Если точка $\mathbf{x} \in G$, то $\mathbf{x} \in G_k$ при всех $k = 1, \ldots, n$. Каждому G_k принадлежит некоторая ε_k-окрестность точки $\mathbf{x}$. Положим

$\varepsilon := \min(\varepsilon_1, \ldots, \varepsilon_n)$. Тогда $\varepsilon > 0$ и ε-окрестность точки $\mathbf{x}$ принадлежит каждому множеству G_k, а значит, и множеству G. Таким образом, $\mathbf{x}$ является внутренней точкой G.

Теорема доказана.

В теореме 11.2.1 нельзя брать пересечение бесконечного набора открытых множеств. В самом деле, если в качестве множеств G_k, $k \in \mathbb{N}$, взять открытые шары радиуса $1/k$ с центром в точке $\mathbf{0}$, то пересечением G_k является точка $\mathbf{0}$, т.е. пересечение не является открытым множеством.

Наряду шаровыми и кубическими окрестностями вводится общее определение окрестности точки, как произвольного открытого множества, содержащего эту точку. Ясно, что каждая такая окрестность содержит шаровую окрестность. При этом окрестностями являются объединения произвольного набора окрестностей и пересечения конечного набора окрестностей.

Множество $A \subset \mathbb{E}^m$ называют *связным*, если любые две его точки можно соединить непрерывной кривой, лежащей в A. Точнее здесь было бы говорить о линейной (от слова линия) связности, чтобы отличать её от других определений связности, которых в этом курсе касаться не будем.

Открытое связное множество в $\mathbb{E}^m$ называют *областью*. Впрочем, слово область часто используется более широко. Например, область задания и область значений функции, когда слово “область” заменяет слово “множество”.

Определение. Точка $\mathbf{x}$ называется *предельной точкой* множества A, если каждая окрестность точки $\mathbf{x}$ содержит бесконечно много точек множества A.

Предельная точка может как принадлежать, так и не принадлежать множеству A.

Легко убедиться, что точка $\mathbf{x}$ является предельной точкой множества A в том и только том случае, когда существует сходящаяся к $\mathbf{x}$ последовательность точек из A, не содержащая точку $\mathbf{x}$.

Определение. Множество называется *замкнутым*, если оно содержит все свои предельные точки.

Эквивалентная формулировка – множество A называется замкнутым, если предел любой сходящейся последовательности точек из A принадлежит A.

Множества из конечного числа элементов, а также пустое множество не имеют предельных точек, они являются замкнутыми множествами.

Покажем, что замкнутый шар является замкнутым множеством. Необходимость такого доказательства и подход к доказательству – те же, что и для открытых шаров.

Рассмотрим замкнутый шар радиуса $r > 0$ с центром в точке $\mathbf{x}^0$, т.е. множество таких точек $\mathbf{x}$, что $\rho(\mathbf{x}, \mathbf{x}^0) \leqslant r$, и покажем, что никакая точка вне шара не может быть его предельной точкой.

Пусть точка $\mathbf{x}^*$ лежит вне шара, т.е. $\rho(\mathbf{x}^*, \mathbf{x}^0) > r$. Положим $\varepsilon := \rho(\mathbf{x}^*, \mathbf{x}^0) - r$ и рассмотрим точки $\mathbf{y}$ из ε-окрестности точки $\mathbf{x}^*$. По неравенству треугольника

$$\rho(\mathbf{x}^0, \mathbf{x}^*) \leqslant \rho(\mathbf{x}^0, \mathbf{y}) + \rho(\mathbf{y}, \mathbf{x}^*).$$

Отсюда

$$\rho(\mathbf{x}^0, \mathbf{y}) \geqslant \rho(\mathbf{x}^0, \mathbf{x}^*) - \rho(\mathbf{y}, \mathbf{x}^*) > \rho(\mathbf{x}^0, \mathbf{x}^*) - (\rho(\mathbf{x}^0, \mathbf{x}^*) - r) = r.$$

Значит, все такие точки $\mathbf{y}$ лежат вне исходного шара и $\mathbf{x}^*$ не может быть его предельной точкой. Поэтому рассматриваемый шар является замкнутым множеством.

Аналогично можно доказать замкнутость замкнутых прямоугольников.

Теорема 11.2.2. *Пересечение произвольного набора замкнутых множеств является замкнутым множеством.*

Объединение конечного набора замкнутых множеств является замкнутым множеством.

Доказательство. Пусть

$$F = \bigcap_{\alpha \in \mathcal{A}} F_\alpha,$$

где все множества F_α замкнуты. Рассмотрим произвольную последовательность точек $\{\mathbf{x}^{(n)}\}$ из F, сходящуюся к некоторой точке $\mathbf{x}^*$. Каждая точка $\mathbf{x}^{(n)}$ принадлежит всем множествам F_α и в силу замкнутости этих множеств точка $\mathbf{x}^*$ принадлежит всем F_α, а значит, принадлежит и F.

Докажем второе утверждение теоремы. Пусть

$$F = \bigcup_{k=1}^{n} F_k$$

и множества F_k замкнуты. Для каждой последовательности точек, принадлежащих F и сходящихся к некоторой точке $\mathbf{x}^*$, имеется множество F_k, содержащее бесконечно много точек этой последовательности. В силу замкнутости F_k точка $\mathbf{x}^*$ принадлежит этому множеству, а значит, принадлежит F.

Теорема доказана.

Здесь существенно, что рассматривалось объединение именно конечного набора замкнутых множеств. В самом деле, на оси $(-\infty, +\infty)$ множества

$$F_k = \left[0, 1 - \frac{1}{k}\right], \qquad k \in \mathbb{N},$$

замкнуты, а их объединение

$$\bigcup_{k=1}^{\infty} F_k = [0, 1)$$

не замкнуто.

Теорема 11.2.3. *Если множество F замкнуто, а множество G открыто, то разность $F \setminus G$ замкнута.*

Доказательство. Покажем, что предельные точки множества $F \setminus G$ принадлежат этому множеству.

Если $\mathbf{x}^0$ – предельная точка $F \setminus G$, то $\mathbf{x}^0$ – тем более предельная точка F и, значит, $\mathbf{x}^0 \in F$. При этом точка $\mathbf{x}^0$ не может принадлежать G, иначе она принадлежала бы G вместе с некоторой своей окрестностью, т.е. существовала бы окрестность точки $\mathbf{x}^0$, свободная от точек $F \setminus G$.

Теорема доказана.

Если к произвольному множеству A присоединить его предельные точки, то полученное множество называют *замыканием* A и обозначают $\overline{A}$.

Определение. Точка называется *граничной точкой* множества A, если каждая её окрестность содержит как точки, принадлежащие A, так и точки, не принадлежащие A. Множество всех граничных точек множества A называют его *границей* и обозначают ∂A.

Теорема 11.2.4. *Множество A открыто в том и только том случае, когда $\partial A \cap A = \varnothing$.*

Множество A замкнуто в том и только том случае, когда $\partial A \subset A$.

Доказательство. Пусть A – открытое множество. Если $\mathbf{x} \in \partial A$, то $\mathbf{x}$ не может принадлежать A, так как каждая точка множества A принадлежит ему вместе с некоторой своей окрестностью, а любая окрестность точки из ∂A должна содержать точки, не принадлежащие A.

Наоборот, пусть $\partial A \cap A = \varnothing$. Если точка $\mathbf{x} \in A$, то $\mathbf{x} \notin \partial A$ и существует окрестность $\mathbf{x}$, все точки которой принадлежат A, – иначе точка $\mathbf{x}$ принадлежала бы границе множества A. Значит, множество A открыто.

Докажем утверждение о замкнутых множествах.

Каждая точка границы произвольного множества A либо принадлежит A, либо является его предельной точкой. Поэтому, если A замкнуто, то $\partial A \subset A$.

С другой стороны, каждая предельная точка множества A принадлежит или A или ∂A. Поэтому, если $\partial A \subset A$, то множество A замкнуто.

Теорема доказана.

Следствие 11.2.5. *Множество $A \subset \mathbb{E}^m$ замкнуто в том и только том случае, когда его дополнение, т.е. множество $\mathbb{E}^m \setminus A$, открыто.*

Это утверждение вытекает из теоремы 11.2.4, так как граница множества является также границей его дополнения.

Любое множество из $\mathbb{E}^m$, кроме всего пространства $\mathbb{E}^m$ и пустого множества $\varnothing$, имеет граничные точки. Только эти два множества являются одновременно и замкнутыми и открытыми. Так как некоторые точки границы множества могут принадлежать множеству, а другие точки границы ему не принадлежать, то множество может быть и не замкнутым и не открытым.

Определение. Непустое множество точек пространства $\mathbb{E}^m$ называется *компактным* или *компактом*, если оно замкнуто и ограничено.

Нам понадобится ещё следующее понятие. Пусть даны множество A и семейство множеств B_α. Если каждая точка A принадлежит некоторому множеству этого семейства, т.е. если

$$A \subset \bigcup_\alpha B_\alpha,$$

то говорят, что семейство множеств B_α покрывает множество A.

Теорема 11.2.6 (Лемма Гейне–Бореля). *Если семейство открытых множеств $\{G_\alpha\}$ покрывает компакт $K \subset \mathbb{E}^m$, то существует конечный набор множеств семейства $\{G_\alpha\}$, которые покрывают K.*

Доказательство. Проведём рассуждения от противного. Допустим, что компакт K нельзя покрыть конечным набором множеств семейства $\{G_\alpha\}$.

Рассмотрим замкнутый куб I, содержащий компакт K. Разделив каждое ребро куба I пополам, построим на полученных рёбрах 2^m одинаковых замкнутых кубиков. Порции компакта K (т.е. части K), попавшие в каждый из этих кубиков, являются замкнутыми множествами. По крайней мере одну из этих порций множества K нельзя покрыть конечным набором множеств семейства $\{G_\alpha\}$.

Берём теперь кубик, содержащий такую порцию компакта K, этот кубик также делим на 2^m одинаковых замкнутых кубиков следующего поколения и находим среди них такой, что содержащуюся в нём порцию компакта K нельзя покрыть конечным набором множеств семейства $\{G_\alpha\}$.

Продолжив неограниченно такое построение, получим последовательность замкнутых вложенных кубов, диаметры которых стремятся к нулю и в каждом кубе есть точки множества K . Согласно теореме 11.1.2 существует точка $\mathbf{x}^*$, принадлежащая всем этим кубам, которая в силу замкнутости множества K принадлежит K.

В семействе $\{G_\alpha\}$ имеется множество G_{α^*}, которому принадлежит точка $\mathbf{x}^*$, причём является его внутренней точкой. Значит, достаточно малая окрестность точки $\mathbf{x}^*$ содержится в G_{α^*}.

Так как диаметры кубиков построенной последовательности стремятся к нулю, то все эти кубики, начиная с некоторого, попадут в указанную окрестность точки $\mathbf{x}^*$ и, таким образом, будут покрыты множеством G_{α^*}.

Получено противоречие, поскольку кубики выбирались так, что содержащиеся в них порции компакта K нельзя покрыть конечным набором множеств семейства $\{G_\alpha\}$.

Теорема доказана.

Теорему 11.2.6 кратко формулируют так: из каждого покрытия компакта открытыми множествами можно выбрать конечное подпокрытие и называют её *теоремой о конечном покрытии*.

§ 11.3. Пределы функций многих переменных

Будем рассматривать функции, заданные на множествах m-мерного евклидова пространства $\mathbb{E}^m$ и принимающие числовые значения, т.е. функции вида

$$f(\mathbf{x})\colon D \to \mathbb{R},$$

где $D \subset \mathbb{E}^m$. Для обозначения функций, а также их значений будем использовать как равноправные записи $f(\mathbf{x})$ и $f(x_1, \ldots, x_m)$ и называть такие функции функциями m переменных $x_1, \ldots, x_m$.

Определим предел функции по множеству. Это определение является новым и для функций одной переменной.

Определение. Пусть $\mathbf{x}^0$ – предельная точка множества $D \subset \mathbb{E}^m$. Число a называют *пределом функции* $f(\mathbf{x})$ *в точке* $\mathbf{x}^0$ *по множеству* D, если:

1°) функция f определена во всех точках множества D, принадлежащих некоторой окрестности точки $\mathbf{x}^0$, за исключением, быть может, самой точки $\mathbf{x}^0$;

2°) для каждого положительного числа ε существует число $\delta(\varepsilon) > 0$ такое, что для точек $\mathbf{x} \in D$, удовлетворяющих условию $0 < \rho(\mathbf{x}, \mathbf{x}^0) < \delta$, справедлива оценка

$$|f(\mathbf{x}) - a| < \varepsilon.$$

В этом случае пишут

$$a = \lim_{\mathbf{x} \to \mathbf{x}^0,\, \mathbf{x} \in D} f(\mathbf{x}).$$

Условие 1° можно переформулировать так: функция f определена во всех точках множества D, принадлежащих некоторой проколотой окрестности точки $\mathbf{x}^0$. Как и ранее, проколотая окрестность – это окрестность точки без самой этой точки.

В 2° участвуют точки $\mathbf{x} \in D$ из проколотой δ-окрестности точки $\mathbf{x}^0$.

Если существует проколотая окрестность точки $\mathbf{x}^0$, целиком принадлежащая множеству D, слова “по множеству D” в определении предела опускают.

Заметим, что односторонние пределы функций одной переменной фактически являются пределами по множеству.

Приведенное определение было определением предела по Коши. В определении предела по Гейне условие 2^0 формулируется так:

$2^{\circ'}$) для любой последовательности точек $\{\mathbf{x}^{(n)}\}$, принадлежащих D, сходящейся к точке $\mathbf{x}^0$, имеет место равенство

$$\lim_{n\to\infty} f(\mathbf{x}^{(n)}) = a.$$

Эквивалентность этих определений устанавливается точно так же, как для функций одной переменной. Нужно только говорить о точках множества D, поскольку рассматривается предел по множеству.

Теорема 11.3.1 (Критерий Коши существования предела функции). *Пусть $\mathbf{x}^0$ – предельная точка множества $D \subset \mathbb{E}^m$ и функция $f(\mathbf{x})$ определена в точках множества D из некоторой проколотой окрестности точки $\mathbf{x}^0$. Для того чтобы функция f имела в точке $\mathbf{x}^0$ предел по множеству D, необходимо и достаточно условие Коши, состоящее в том, что для каждого положительного ε существует число $\delta(\varepsilon) > 0$ такое, что для любых точек $\mathbf{x}'$ и $\mathbf{x}''$ из проколотой δ-окрестности точки $\mathbf{x}^0$, принадлежащих множеству D, справедливо неравенство*

$$|f(\mathbf{x}') - f(\mathbf{x}'')| < \varepsilon.$$

Доказательство теоремы 11.3.1 не отличается от доказательства критерия Коши для функций одной переменной, приведенного в § 3.4. Нужно только говорить о точках, принадлежащих множеству D.

Наряду с конечными пределами функций рассматриваются бесконечные пределы. Для предела, равного $+\infty$, условие 2° в определении предела по Коши имеет вид:

2°) для каждого числа M существует число $\delta(M) > 0$ такое, что для всех точек $\mathbf{x} \in D$, для которых $0 < \rho(\mathbf{x}, \mathbf{x}^0) < \delta$, выполняется условие $f(\mathbf{x}) > M$.

Если здесь оценку $f(\mathbf{x}) > M$ заменить на $f(\mathbf{x}) < M$, получим определение предела, равного $-\infty$, а если заменить эту оценку на $|f(\mathbf{x})| > M$, получим определение предела, равного ∞.

Вводится также предел функции при $\mathbf{x} \to \infty$ по множеству D.

Определение. Пусть для каждого числа B существует бесконечно много точек множества $D \subset \mathbb{E}^m$, для которых $\rho(\mathbf{x}, \mathbf{0}) > B$. Число a называют *пределом функции* $f(\mathbf{x})$ *при* $\mathbf{x} \to \infty$ *по множеству* D, если

1°) существует такое число B, что функция f определена во всех точках множества D, для которых $\rho(\mathbf{x}, \mathbf{0}) > B$;
2°) для каждого $\varepsilon > 0$ существует число $M(\varepsilon)$ такое, что для всех $\mathbf{x} \in D$, для которых $\rho(\mathbf{x}, \mathbf{0}) > M$, справедлива оценка $|f(\mathbf{x}) - a| < \varepsilon$.

В этом случае пишут

$$\lim_{\mathbf{x} \to \infty,\, \mathbf{x} \in D} f(\mathbf{x}) = a.$$

Соответствующим образом даются определения бесконечных пределов функции по неограниченному множеству D при $\mathbf{x} \to \infty$.

По сути все приведенные определения предела одинаковы. По-разному понимаются окрестности точек. Когда говорят о пределе при $\mathbf{x} \to \infty$, считают, что окрестностями символа ∞ являются точки $\mathbf{x}$, удовлетворяющие условию вида $\rho(\mathbf{x}, \mathbf{0}) > B$.

Чтобы не говорить о каждом из этих случаев по отдельности, а охватить их все сразу, вводят предел по базе.

Определение. Бесконечный набор $\mathcal{A}$ непустых подмножеств A_α множества X

$$\mathcal{A} := \{A_\alpha\}$$

называют *базой в* X, если для каждой пары множеств A_α и A_β этого набора в $\mathcal{A}$ имеется множество A_γ такое, что

$$A_\gamma \subset \left(A_\alpha \bigcap A_\beta\right). \tag{11.3.1}$$

Например, если $\mathbf{x}^0$ – предельная точка множества D, то базу в D образует совокупность подмножеств множества D, содержащихся в проколотых ε-окрестностях точки $\mathbf{x}^0$, где ε – положительные числа. В этом случае для любых A_α и A_β существует элемент A_γ базы, для которого в (11.3.1) имеет место равенство.

Другую базу в D образует совокупность подмножеств множества D, содержащихся в произвольных открытых шарах, которым принадлежит точка $\mathbf{x}^0$, из которых сама точка $\mathbf{x}^0$ исключена, т.е. это проколотые шары. Но теперь пересечение не каждой пары элементов базы ей принадлежит.

Определение. Пусть функция $f(\mathbf{x})$ задана на множестве $D \subset \mathbb{E}^m$ и $\mathcal{A}$ – база в D. Число a называют *пределом функции* $f(\mathbf{x})$ *по базе* $\mathcal{A}$, если для каждого $\varepsilon > 0$ существует элемент базы A_ε такой, что для всех $\mathbf{x} \in A_\varepsilon$ справедлива оценка $|f(\mathbf{x}) - a| < \varepsilon$.

Таким образом, предел функции по множеству D в точке $\mathbf{x}^0$, являющейся предельной точкой D, соответствует случаю, когда в качестве базы взяты подмножества множества D, содержащиеся в проколотых δ-окрестностях точки $\mathbf{x}^0$.

Для предела функции при $\mathbf{x} \to \infty$ по неограниченному множеству D в качестве базы берутся подмножества D, содержащиеся вне замкнутых шаров произвольного положительного радиуса.

Перечислим свойства пределов функций многих переменных по множеству, аналогичные свойствам обычных пределов функций одной переменной.

Если функция f имеет конечный предел при $\mathbf{x} \to \mathbf{x}^0$ по множеству D, то f ограничена в точках некоторой проколотой окрестности точки $\mathbf{x}^0$, принадлежащих множеству D.

Если

$$\lim_{\mathbf{x} \to \mathbf{x}^0,\, \mathbf{x} \in D} f(\mathbf{x}) = a \neq 0,$$

то во всех точках некоторой проколотой окрестности точки $\mathbf{x}^0$, принадлежащих D, выполняется неравенство $|f(\mathbf{x})| > |a|/2$.

Арифметические свойства пределов, т.е. пределы суммы, разности, произведения и частного (а также предел модуля) функций, имеющих пределы, справедливы и для пределов функций многих переменных по множеству.

И в формулировках и в доказательствах этих свойств нет ничего нового по сравнению с одномерным случаем.

Перейдём теперь к вопросам, характерным именно для функций многих переменных.

Пусть (для простоты) функция $f(\mathbf{x})$ определена во всех точках некоторой проколотой окрестности точки $\mathbf{x}^0$.

Рассмотрим луч с вершиной в точке $\mathbf{x}^0$, параллельный ненулевому m-мерному вектору $(\alpha_1, \ldots, \alpha_m)$, т.е. множество точек $\mathbf{x}$, для координат которых выполняются равенства

$$x_1 = x_1^0 + \alpha_1 t,$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$x_m = x_m^0 + \alpha_m t,$$

при $t \geqslant 0$.

Предел функции f в точке $\mathbf{x}^0$ по этому лучу называют *пределом по направлению*, определяемому вектором $(\alpha_1, \ldots, \alpha_m)$. Предел по направлению является пределом функции $f(x_1^0 + \alpha_1 t, \ldots, x_m^0 + \alpha_m t)$ одной переменной t при $t \to +0$.

Понятно, что если функция имеет предел в точке, то она имеет в этой точке равные ему пределы по всем направлениям.

Если не предполагать, что функция имеет предел в точке, и рассматривать только пределы в этой точке по направлениям, то могут быть разные случаи.

Предел по одним направлениям может существовать, а по другим – нет. Могут существовать пределы по всем направлениям, но значения пределов по разным направлениям различны. Соответствующие примеры строятся легко, так как на каждом луче значения функции можно выбирать произвольно.

Менее очевиден пример, когда пределы по всем направлениям существуют и их значения равны, а предела нет.

Пусть функция двух переменных $f(x, y)$ определена всюду и равна 1, если $y = x^2$, а в остальных точках равна нулю. Тогда предел f в точке $(0, 0)$ по любому направлению равен 0. А так как в каждой окрестности начала координат есть точки, в которых функция равна 1, и точки, где она равна нулю, то предел f в точке $(0, 0)$ не существует.

Для функций многих переменных вводится ещё одно важное понятие предела – повторные пределы. Чтобы не затемнять дело несущественными деталями, будем рассматривать этот вопрос для функций двух переменных, заданных в проколотой окрестности некоторой точки.

Определение. Пусть функция $f(x, y)$ определена в проколотой окрестности точки (x^0, y^0). *Повторными пределами* функции $f(x, y)$ в точке (x^0, y^0) называют пределы

$$\lim_{x \to x^0} \Big(\lim_{y \to y^0} f(x, y) \Big) \qquad (11.3.2)$$

и
$$\lim_{y\to y^0}\Bigl(\lim_{x\to x^0} f(x,y)\Bigr).$$

Здесь участвуют только пределы функций одной переменной. Заметим, что в (11.3.2) предел при $y \to y_0$ рассматривается только для $x \neq x_0$.

Понятно, что повторные пределы не всегда существуют.

Покажем на примерах, что существование повторных пределов не связано, вообще говоря, с существованием обычного предела.

Пусть функция $g(x,y)$ задана во всех точках плоскости, кроме начала координат, формулой
$$g(x,y) := \frac{2xy}{x^2+y^2}\,.$$
Тогда предел
$$\lim_{(x,y)\to(0,0)} g(x,y)$$
не существует, так как g равна нулю, если $x = 0$, $y \neq 0$, или $y = 0$, $x \neq 0$, а при $x = y$ значения g равны 1. Вместе с тем, легко убедиться, что оба повторных предела g в точке $(0,0)$ существуют и их значения равны (они равны нулю, так как равны нулю внутренние пределы).

Рассмотрим теперь функцию
$$h(x,y) := \begin{cases} x\sin(1/y), & \text{если } y \neq 0,\\ 0, & \text{если } y = 0.\end{cases}$$

Так как $|h(x,y)| \leqslant |x|$ во всех точках (x,y), то
$$\lim_{(x,y)\to(0,0)} h(x,y) = 0.$$

Повторный предел
$$\lim_{y\to 0}\Bigl(\lim_{x\to 0} h(x,y)\Bigr)$$
существует и равен нулю, а предел
$$\lim_{x\to 0}\Bigl(\lim_{y\to 0} h(x,y)\Bigr)$$
не существует, так как не существует внутренний предел.

Пример, когда внутренний предел существует, а внешний – нет, даёт функция

$$u(x,y) := \begin{cases} \sin(1/y), & \text{если } y \neq 0, \\ 0, & \text{если } y = 0. \end{cases}$$

В самом деле, для $y \neq 0$

$$\lim_{x\to 0} u(x,y) = \sin\frac{1}{y}\,,$$

поэтому предел

$$\lim_{y\to 0}\Big(\lim_{x\to 0} u(x,y)\Big)$$

не существует.

Самое поучительное свойство повторных пределов иллюстрирует функция

$$v(x,y) := \frac{y^2}{x^2+y^2}\,, \qquad (x,y) \neq (0,0).$$

Нетрудно убедиться, что

$$\lim_{x\to 0}\Big(\lim_{y\to 0} v(x,y)\Big) = 0$$

и

$$\lim_{y\to 0}\Big(\lim_{x\to 0} v(x,y)\Big) = 1.$$

Приведенные примеры показывают следующее важное свойство повторных пределов.

Если в каком-либо выражении поочередно перейти к пределу по двум параметрам (двум переменным), то и существование пределов и их значения могут зависеть от того, в каком порядке берутся пределы.

Покажем, что при некоторых условиях из существования предела функции следует существование повторных пределов.

Теорема 11.3.2. *Пусть существует предел*

$$\lim_{(x,y)\to(x^0,y^0)} f(x,y). \tag{11.3.3}$$

Если для каждого y *из некоторой проколотой окрестности точки* y^0 *существует предел*

$$\lim_{x\to x^0} f(x,y), \tag{11.3.4}$$

то существует повторный предел

$$\lim_{y\to y^0}\Bigl(\lim_{x\to x^0} f(x,y)\Bigr) \tag{11.3.5}$$

и справедливо равенство

$$\lim_{y\to y^0}\Bigl(\lim_{x\to x^0} f(x,y)\Bigr) = \lim_{(x,y)\to(x^0,y^0)} f(x,y). \tag{11.3.6}$$

Доказательство. Обозначим предел (11.3.3) через a. Так как этот предел существует, то для $\varepsilon > 0$ имеется $\delta > 0$ такое, что для каждой точки (x,y) из проколотой δ-окрестности точки (x^0,y^0) выполняется неравенство

$$|a - f(x,y)| < \frac{\varepsilon}{2}\,. \tag{11.3.7}$$

Будем считать δ настолько малым, что предел (11.3.4) существует при всех y, для которых $0 < |y-y^0| < \delta$. При каждом таком y в неравенстве (11.3.7) можно перейти к пределу при $x \to x^0$:

$$|a - \lim_{x\to x^0} f(x,y)| \leqslant \frac{\varepsilon}{2} < \varepsilon.$$

Так как эта оценка имеет место для всех y из проколотой δ-окрестности точки y^0, то из неё следует и существование предела (11.3.5) и равенство (11.3.6).

Теорема доказана.

Для функций m переменных можно определить $m!$ повторных пределов, когда в различном порядке переходят к пределу по каждой из переменных поочередно.

При $m > 2$ можно также разделить все переменные, скажем, на две группы и переходить к пределу сначала по переменным одной из этих групп, а затем – по переменным другой группы.

§ 11.4. Непрерывные функции многих переменных

Определение. Пусть точка $\mathbf{x}^0 \in D \subset \mathbb{E}^m$. Говорят, что функция $f(\mathbf{x})$ *непрерывна в точке* $\mathbf{x}^0$ *по множеству* D, если f определена во всех точках множества D из некоторой окрестности точки $\mathbf{x}^0$ и для каждого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$ такое, что для всех $\mathbf{x} \in D$ из δ-окрестности точки $\mathbf{x}^0$, справедливо неравенство

$$|f(\mathbf{x}) - f(\mathbf{x}^0)| < \varepsilon.$$

Это – определение "на языке $\varepsilon - \delta$". На "языке последовательностей" при тех же предположениях об области определения функции f соответствующее условие формулируется так: для любой сходящейся к точке $\mathbf{x}^0$ последовательности точек $\{\mathbf{x}^{(n)}\}$, принадлежащих множеству D, имеет место равенство

$$\lim_{n\to\infty} f(\mathbf{x}^{(n)}) = f(\mathbf{x}^0).$$

Если $\mathbf{x}^0$ – изолированная точка множества D, т.е. в некоторой окрестности $\mathbf{x}^0$ нет других точек из D, то каждая функция, которая определена в точке $\mathbf{x}^0$, непрерывна в этой точке по множеству D.

Если точка $\mathbf{x}^0 \in D$ не является изолированной точкой множества D, то она – предельная точка D. В этом случае непрерывность функции в точке $\mathbf{x}^0$ можно определить с помощью предела: функция f непрерывна в точке $\mathbf{x}^0 \in D$ по множеству D, если

$$\lim_{\mathbf{x}\to\mathbf{x}^0,\, \mathbf{x}\in D} f(\mathbf{x}) = f(\mathbf{x}^0).$$

В предельных точках множества D многие свойства функций, непрерывных по D, вытекают из свойств пределов.

Например, если функция непрерывна в точке $\mathbf{x}^0$ по множеству D, то в точках множества D, принадлежащих достаточно малой окрестности точки $\mathbf{x}^0$ она ограничена. Если функция f непрерывна в точке $\mathbf{x}^0$ по множеству D и $f(\mathbf{x}^0) \neq 0$, то f сохраняет знак в точках D из некоторой окрестности точки $\mathbf{x}^0$. На этот случай переносятся также теоремы о непрерывности функций, полученных в результате арифметических действий над непрерывными функциями, и теорема о непрерывности модуля непрерывной функции.

Остановимся подробнее на вопросе о непрерывности сложной функции многих переменных, хотя и здесь нет ничего по существу нового по сравнению с функциями одной переменной.

Пусть функция $\varphi(\mathbf{x})$ определена на множестве $D \subset \mathbb{E}^m$ и имеется ещё некоторое множество $A \subset \mathbb{E}^p$, в точках $\mathbf{t}$ которого заданы m функций

$$f_1(\mathbf{t}), \dots, f_m(\mathbf{t}),$$

таких, что точки $(f_1(\mathbf{t}), \dots, f_m(\mathbf{t}))$ принадлежат множеству D.

Тогда на множестве A можно определить функцию

$$\psi(\mathbf{t}) := \varphi(f_1(\mathbf{t}), \dots, f_m(\mathbf{t})) \tag{11.4.1}$$

или в координатных обозначениях

$$\psi(t_1, \dots, t_p) = \varphi(f_1(t_1, \dots, t_p), \dots, f_m(t_1, \dots, t_p)),$$

которую называют *сложной функцией*.

Теорема 11.4.1. *Пусть функции $f_1(\mathbf{t}), \dots, f_m(\mathbf{t})$ непрерывны в точке $\mathbf{t}^0 \in A \subset \mathbb{E}^p$ по множеству A и для всех $\mathbf{t} \in A$ из некоторой окрестности точки $\mathbf{t}^0$ точки $(f_1(\mathbf{t}), \dots, f_m(\mathbf{t}))$ принадлежат множеству $D \subset \mathbb{E}^m$. Если функция $\varphi(\mathbf{x}) = \varphi(x_1, \dots, x_m)$ непрерывна в точке $\mathbf{x}^0 := (f_1(\mathbf{t}^0), \dots, f_m(\mathbf{t}^0))$ по множеству D, то сложная функция (11.4.1) непрерывна в точке $\mathbf{t}^0$ по множеству A.*

Доказательство. Задав $\varepsilon > 0$, выбираем $\sigma > 0$ так, что для всех точек $\mathbf{x} \in D$ из m-мерной σ-окрестности точки $\mathbf{x}^0$, выполняется неравенство

$$|\varphi(\mathbf{x}) - \varphi(\mathbf{x}^0)| < \varepsilon.$$

Так как функции $f_1(\mathbf{t}), \dots, f_m(\mathbf{t})$ непрерывны в точке $\mathbf{t}^0$ по множеству A, то по этому σ можно выбрать $\delta > 0$ такое, что для всех $\mathbf{t} \in A$, принадлежащих p-мерной δ-окрестности точки $\mathbf{t}^0$, справедливы неравенства

$$|f_k(\mathbf{t}) - f_k(\mathbf{t}^0)| < \frac{\sigma}{\sqrt{m}}, \quad k = 1, \dots, m.$$

Такое δ существует, так как функций f_k конечное число.

В результате по ε выбрано $\delta > 0$ такое, что для всех точек $\mathbf{t} \in A$ из p-мерной δ-окрестности точки $\mathbf{t}^0$

$$\rho_m((f_1(\mathbf{t}), \ldots, f_m(\mathbf{t})), (f_1(\mathbf{t}^0), \ldots, f_m(\mathbf{t}^0))) =$$
$$= \sqrt{\sum_{k=1}^{m} (f_k(\mathbf{t}) - f_k(\mathbf{t}^0))^2} < \sqrt{\sum_{k=1}^{m} \left(\frac{\sigma}{\sqrt{m}}\right)^2} = \sigma.$$

Можно считать δ настолько малым, что точки $(f_1(\mathbf{t}), \ldots, f_m(\mathbf{t}))$ принадлежат множеству D. А так как их расстояния от точки $(f_1(\mathbf{t}^0), \ldots, f_m(\mathbf{t}^0))$ меньше σ, то

$$|\psi(\mathbf{t}) - \psi(\mathbf{t}^0)| = |\varphi(f_1(\mathbf{t}), \ldots, f_m(\mathbf{t})) - \varphi(f_1(\mathbf{t}^0), \ldots, f_m(\mathbf{t}^0))| < \varepsilon.$$

Теорема доказана.

Определение. Функция $f(\mathbf{x})$ называется *непрерывной на* множестве $D \subset \mathbb{E}^m$, если она непрерывна по множеству D в каждой точке $\mathbf{x} \in D$.

Для обозначения непрерывности функции f на D будем писать $f \in C(D)$. Ранее такое обозначение использовалось для функций одной переменной, непрерывных на промежутке.

Отметим, что каждая непрерывная функция l переменных $x_1, \ldots, x_l$ при $m > l$ является непрерывной и как функция m переменных $x_1, \ldots, x_m$. Поэтому элементарные функции одной переменной дают примеры непрерывных функций многих переменных. А если рассматривать ещё сложные функции m переменных, то получим довольно значительный запас непрерывных функций многих переменных.

Приведём для непрерывных функций многих переменных теоремы, аналогичные теоремам из главы 4 для функций одной переменной.

Теорема 11.4.2. *Если множество $D \subset \mathbb{E}^m$ компактно и функция $f(\mathbf{x})$ непрерывна на D, то f ограничена на D.*

Доказательство. Предположим противное – что функция $f \in C(D)$ не является ограниченной на компакте D, т.е. не существует такого числа M, что $|f(\mathbf{x})| \leqslant M$ для всех $\mathbf{x} \in D$.

Тогда для каждого натурального числа n существует точка $\mathbf{x}^{(n)} \in D$ такая, что $|f(\mathbf{x}^{(n)})| > n$. Значит,

$$\lim_{n \to \infty} f(\mathbf{x}^{(n)}) = \infty. \tag{11.4.2}$$

Так как точки последовательности $\{\mathbf{x}^{(n)}\}$ принадлежат ограниченному множеству D, то согласно теореме Больцано–Вейерштрасса 11.1.3 существует подпоследовательность $\{\mathbf{x}^{(n_k)}\}$ последовательности $\{\mathbf{x}^{(n)}\}$, сходящаяся к некоторой точке $\mathbf{x}^*$. В силу замкнутости множества D точка $\mathbf{x}^*$ принадлежит D. Из непрерывности функции f на D заключаем, что

$$\lim_{k\to\infty} f(\mathbf{x}^{(n_k)}) = f(\mathbf{x}^*).$$

Значит, последовательность $\{f(\mathbf{x}^{(n_k)})\}$ ограничена, что противоречит (11.4.2).

Теорема доказана.

Теорема 11.4.3 (Теорема Вейерштрасса). *Если функция $f(\mathbf{x})$ непрерывна на компактном множестве $D \subset \mathbb{E}^m$, то f в некоторых точках множества D достигает точной верхней и точной нижней граней своих значений на D.*

Доказательство. Докажем достижимость точной верхней грани значений f на D.

Из теоремы 11.4.2 следует, что эта точная верхняя грань существует. Пусть $M = \sup_{\mathbf{x}\in D} f(\mathbf{x})$. Для каждого натурального n существует точка $\mathbf{x}^{(n)} \in D$ такая, что

$$M - \frac{1}{n} < f(\mathbf{x}^{(n)}) \leqslant M.$$

Значит,

$$\lim_{n\to\infty} f(\mathbf{x}^{(n)}) = M. \tag{11.4.3}$$

В силу компактности множества D последовательность точек $\{\mathbf{x}^{(n)}\}$ содержит подпоследовательность $\{\mathbf{x}^{(n_k)}\}$, сходящуюся к некоторой точке $\mathbf{x}^* \in D$. Поскольку функция f непрерывна в точке $\mathbf{x}^*$, то

$$\lim_{k\to\infty} f(\mathbf{x}^{(n_k)}) = f(\mathbf{x}^*).$$

Отсюда и из (11.4.3) получим $f(\mathbf{x}^*) = M$ и для точной верхней грани теорема доказана.

Для точной нижней грани доказательство аналогично.

Определение. Функция $f(\mathbf{x})$ называется *равномерно непрерывной на множестве* $D \subset \mathbb{E}^m$, если для каждого $\varepsilon > 0$ существует $\delta(\varepsilon) > 0$ такое, что для любой пары точек $\mathbf{x}', \mathbf{x}'' \in D$, для которых $\rho(\mathbf{x}', \mathbf{x}'') < \delta$, выполняется неравенство

$$|f(\mathbf{x}') - f(\mathbf{x}'')| < \varepsilon.$$

Теорема 11.4.4 (Теорема Кантора). *Если функция $f(\mathbf{x})$ непрерывна на компакте $D \subset \mathbb{E}^m$, то она равномерно непрерывна на D.*

Доказательство. Предположим, что теорема не верна, т.е. существует непрерывная, но не равномерно непрерывная на D функция f.

Тогда найдётся положительное число ε_0 такое, что для каждого $\delta > 0$ существует пара точек $\mathbf{x}, \mathbf{y}$ из D, для которых $\rho(\mathbf{x}, \mathbf{y}) < \delta$, но

$$|f(\mathbf{x}) - f(\mathbf{y})| \geqslant \varepsilon_0.$$

Выбирая в качестве δ числа $1/n$, $n \in \mathbb{N}$, получим две последовательности принадлежащих D точек $\{\mathbf{x}^{(n)}\}$ и $\{\mathbf{y}^{(n)}\}$, таких, что

$$\rho(\mathbf{x}^{(n)}, \mathbf{y}^{(n)}) < \frac{1}{n}\,,$$

но в то же время

$$|f(\mathbf{x}^{(n)}) - f(\mathbf{y}^{(n)})| \geqslant \varepsilon_0. \tag{11.4.4}$$

Так как D – компакт, существует подпоследовательность $\{\mathbf{x}^{(n_k)}\}$ последовательности $\{\mathbf{x}^{(n)}\}$, сходящаяся к некоторой точке $\mathbf{x}^* \in D$.

Из неравенства

$$\rho(\mathbf{y}^{(n_k)}, \mathbf{x}^*) \leqslant \rho(\mathbf{y}^{(n_k)}, \mathbf{x}^{(n_k)}) + \rho(\mathbf{x}^{(n_k)}, \mathbf{x}^*) < \frac{1}{n_k} + \rho(\mathbf{x}^{(n_k)}, \mathbf{x}^*)$$

следует, что последовательность $\{\mathbf{y}^{(n_k)}\}$ также сходится к $\mathbf{x}^*$.

В силу непрерывности функции f в точке $\mathbf{x}^*$ по множеству D имеем

$$f(\mathbf{x}^{(n_k)}) \to f(\mathbf{x}^*), \quad f(\mathbf{y}^{(n_k)}) \to f(\mathbf{x}^*), \qquad k \to \infty.$$

Эти соотношения противоречат неравенству (11.4.4).

Теорема доказана.

Для функций m переменных, заданных на множестве $D \subset \mathbb{E}^m$, вводится *модуль непрерывности*

$$\omega(f, \delta) := \sup |f(\mathbf{x}') - f(\mathbf{x}'')|,$$

где верхняя грань берется по всем точкам $\mathbf{x}'$ и $\mathbf{x}''$, принадлежащим D, расстояние между которыми не превышает δ. Так же, как для функций, заданных на отрезке, доказывается, что равномерная непрерывность функции f на множестве D равносильна условию $\omega(f, +0) = 0$.

Доказательства теорем 11.4.2–11.4.4 фактически повторяли доказательства соответствующих утверждений из главы 4 для функций одной переменной, непрерывных на отрезке. Тем не менее, теоремы 11.4.2–11.4.4 дают новые результаты и для функций одной переменной, так как они относятся к функциям, непрерывным на произвольном компакте, а не только на отрезке.

Следующая теорема для функций одной переменной не даёт ничего нового.

Теорема 11.4.5 (Теорема Коши о промежуточных значениях). *Пусть функция $f(\mathbf{x})$ непрерывна на связном множестве $D \in \mathbb{E}^m$. Тогда для произвольных точек $\mathbf{x}'$ и $\mathbf{x}''$, принадлежащих D, и каждого значения, промежуточного между числами $f(\mathbf{x}')$ и $f(\mathbf{x}'')$, существует точка множества D, в которой функция f принимает это значение.*

Доказательство. Так как D – связное множество, то точки $\mathbf{x}'$ и $\mathbf{x}''$ можно соединить непрерывной кривой, целиком лежащей в D. Пусть такая кривая задана уравнениями

$$x_1 = \varphi_1(t), \ldots, x_m = \varphi_m(t), \tag{11.4.5}$$

где переменная t пробегает некоторый отрезок $[a, b]$, причём при $t = a$ получаем точку $\mathbf{x}'$, а при $t = b$ – точку $\mathbf{x}''$.

Функция переменной t

$$g(t) := f(\varphi_1(t), \ldots, \varphi_m(t))$$

в силу теоремы 11.4.1 о непрерывности сложной функции непрерывна на отрезке $[a, b]$ и в концевых его точках принимает значения $f(\mathbf{x}')$ и $f(\mathbf{x}'')$. Значит, согласно теореме Коши 4.3.3 о промежуточных значениях для функций одной переменной функция $g(t)$ каждое значение между $f(\mathbf{x}')$ и $f(\mathbf{x}'')$ принимает при некотором t из отрезка $[a, b]$. Этому t соответствует точка на кривой (11.4.5), которая принадлежит множеству D.

Теорема доказана.

§ 11.5. Задачи и упражнения

11.1. Докажите, что в пространстве $\mathbb{E}^m$ сфера, т.е. множество точек $\mathbf{x}$ таких, что $\rho(\mathbf{x}, \mathbf{a}) = r > 0$, замкнута.

11.2. Докажите, что замыкание любого множества замкнуто.

11.3. В теореме 11.2.6 предполагалось, что 1) множество K ограничено, 2) K замкнуто, 3) множества G_α – открыты.

Приведите примеры, показывающие, что ни от одного из этих условий отказаться нельзя, т.е. выполнены только какие-либо два из этих условий и утверждение теоремы не верно.

11.4. Пусть функция $f(x, y)$, заданная на некотором квадрате, при каждом фиксированном y непрерывна как функция от x и при каждом фиксированном x непрерывна как функция от y. Следует ли отсюда непрерывность $f(x, y)$ как функции двух переменных?

Глава 12. Дифференциальное исчисление функций многих переменных

§ 12.1. Частные производные и дифференцируемость функций многих переменных

Пусть функция $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ задана в некоторой окрестности точки $\mathbf{x}^0 = (x_1^0, \ldots, x_m^0) \in \mathbb{E}^m$.

Зафиксировав значения переменных $x_2, \ldots, x_m$, равными $x_2^0, \ldots, x_m^0$, получим функцию $f(x_1, x_2^0, \ldots, x_m^0)$ одной переменной x_1. Если эта функция имеет производную в точке x_1^0, то такую производную называют *частной производной первого порядка по переменной* x_1 функции $f(x_1, \ldots, x_m)$ в точке $\mathbf{x}^0$ и обозначают

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^0)$$

или $f'_{x_1}(\mathbf{x}^0)$. Таким образом, по определению

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^0) := \frac{df}{dx_1}(x_1, x_2^0, \ldots, x_m^0)\Big|_{x_1 = x_1^0}.$$

Выражение

$$\frac{\partial f}{\partial x_1}$$

является единым символом, его нельзя рассматривать как дробь (в отличие от соответствующей записи производных функций одной переменной, равных отношению дифференциалов).

Частные производные функции $f(x_1, \ldots, x_m)$ по другим переменным вводятся аналогично.

Из свойств производных функций одной переменной следует, что если функция $f(\mathbf{x})$ имеет в точке $\mathbf{x}^0$ частную производную $\partial f / \partial x_1$, то f непрерывна в точке $\mathbf{x}^0$ по направлению первого базисного вектора. Но как функция m переменных f может при этом не быть непрерывной в точке $\mathbf{x}^0$. В самом деле, в определении производной $\partial f/\partial x_1$ участвуют только значения функции на малом отрезке, проходящем через точку $\mathbf{x}^0$ параллельно оси OX_1. Таким образом, из существования частных производных по всем переменным непрерывность функции многих переменных не вытекает.

При определении частной производной $\partial f/\partial x_1(\mathbf{x}^0)$ фактически рассматривается разность

$$f(x_1^0 + \Delta x_1, x_2^0, \dots, x_m^0) - f(x_1^0, x_2^0, \dots, x_m^0),$$

когда функции f в точке $\mathbf{x}^0$ даётся "частное приращение" Δx_1 вдоль прямой, параллельной оси OX_1, и ставится вопрос о существовании предела

$$\lim_{\Delta x_1 \to 0} \frac{1}{\Delta x_1}[f(x_1^0 + \Delta x_1, x_2^0, \dots, x_m^0) - f(x_1^0, x_2^0, \dots, x_m^0)].$$

Наряду с частными приращениями функции $f(x_1, \dots, x_m)$ по одной переменной рассматриваются "полные приращения", когда точка $\mathbf{x}^0 = (x_1^0, \dots, x_m^0)$ получает приращение $\Delta\mathbf{x} := (\Delta x_1, \dots, \Delta x_m)$ по, вообще говоря, всем переменным.

Приращением функции $f(\mathbf{x})$ в точке $\mathbf{x}^0$, соответствующим приращению аргумента $\Delta\mathbf{x}$, называется величина

$$\begin{aligned}\Delta f &:= f(\mathbf{x}^0 + \Delta\mathbf{x}) - f(\mathbf{x}^0) = \\ &= f(x_1^0 + \Delta x_1, \dots, x_m^0 + \Delta x_m) - f(x_1^0, \dots, x_m^0).\end{aligned}$$

Расстояние между точками $\mathbf{x}^0 + \Delta\mathbf{x}$ и $\mathbf{x}^0$ равно

$$|\Delta\mathbf{x}| := \left(\sum_{k=1}^{m} (\Delta x_k)^2\right)^{1/2}.$$

Определение. Функция $f(\mathbf{x})$ называется *дифференцируемой в точке* $\mathbf{x}$, если существуют такие числа $A_1, \dots, A_m$, что приращение f в этой точке можно представить в виде

$$\Delta f = f(\mathbf{x} + \Delta\mathbf{x}) - f(\mathbf{x}) = \sum_{k=1}^{m} A_k \Delta x_k + o(|\Delta\mathbf{x}|), \qquad |\Delta\mathbf{x}| \to 0. \tag{12.1.1}$$

Иногда приращение функции f нужно рассматривать и при $\Delta\mathbf{x} = \overline{0}$. Считают, что в этом случае $\Delta f = 0$.

Рассмотрим связь дифференцируемости функции и свойств её частных производных.

Теорема 12.1.1. *Если функция f дифференцируема в точке $\mathbf{x}$, то*

1°) *f непрерывна в точке $\mathbf{x}$;*
2°) *f имеет в точке $\mathbf{x}$ частные производные первого порядка по всем переменным и для чисел A_k из (12.1.1) справедливы равенства*

$$A_k = \frac{\partial f}{\partial x_k}(\mathbf{x}), \qquad k = 1, \ldots, m.$$

Доказательство. Непрерывность функции в точке, где она дифференцируема, следует из того, что если $|\Delta\mathbf{x}| \to 0$, то $\Delta x_k \to 0$ при всех $k = 1, \ldots, m$, значит, согласно (12.1.1) $\Delta f \to 0$.

Докажем утверждение 2°. Придадим функции $f(\mathbf{x})$ приращение только по переменной x_1. Тогда согласно (12.1.1)

$$f(x_1 + \Delta x_1, x_2, \ldots, x_m) - f(x_1, x_2, \ldots, x_m) =$$
$$= A_1 \Delta x_1 + o(\Delta x_1), \qquad \Delta x_1 \to 0.$$

Разделив обе части этой оценки на Δx_1, находим

$$\frac{1}{\Delta x_1}[f(x_1 + \Delta x_1, x_2, \ldots, x_m) - f(x_1, x_2, \ldots, x_m)] =$$
$$= A_1 + o(1), \qquad \Delta x_1 \to 0.$$

Так как при $\Delta x_1 \to 0$ предел выражения в правой части существует и равен A_1, то частная производная $\partial f/\partial x_1$ существует и равна A_1.

Для частных производных по переменным $x_2, \ldots, x_m$ доказательство такое же.

Таким образом, формулу (12.1.1) можно записать в виде:

$$\Delta f = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}) \Delta x_k + o(|\Delta\mathbf{x}|), \qquad |\Delta\mathbf{x}| \to 0. \tag{12.1.2}$$

Теорема 12.1.1 даёт необходимые условия дифференцируемости функций многих переменных. Получим теперь достаточное условие.

Теорема 12.1.2. *Если функция $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ имеет в точке $\mathbf{x}$ частные производные по всем m переменным и $m-1$ из этих производных непрерывны в точке $\mathbf{x}$, то f дифференцируема в точке $\mathbf{x}$.*

Доказательство. Для простоты записей рассуждения проведём на примере функции трёх переменных x, y, z. Пусть z – та переменная, непрерывность частной производной по которой не предполагается.

Непрерывность частных производных в точке (x, y, z) предполагает их существование в некоторой окрестности этой точки. Будем рассматривать приращения аргументов, при которых точки $(x+\Delta x, y+\Delta y, z+\Delta z)$ не выходят за пределы этой окрестности.

Имеем

$$\begin{aligned}\Delta f &= f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x, y, z) = \\ &= [f(x+\Delta x, y+\Delta y, z+\Delta z) - f(x, y+\Delta y, z+\Delta z)] + \\ &\quad + [f(x, y+\Delta y, z+\Delta z) - f(x, y, z+\Delta z)] + \\ &\quad + [f(x, y, z+\Delta z) - f(x, y, z)].\end{aligned}$$

К первым двум разностям в полученном выражении применим формулу конечных приращений Лагранжа для функций одной переменной, пользуясь тем, что значения двух других переменных в этих разностях одинаковы. Тогда получим

$$\begin{aligned}\Delta f = \frac{\partial f}{\partial x}&(x+\theta_1\Delta x, y+\Delta y, z+\Delta z)\Delta x + \\ &+ \frac{\partial f}{\partial y}(x, y+\theta_2\Delta y, z+\Delta z)\Delta y + \\ &+ [f(x, y, z+\Delta z) - f(x, y, z)].\end{aligned} \tag{12.1.3}$$

где числа θ_i, $i = 1, 2$, удовлетворяют условиям $0 < \theta_i < 1$.

В силу непрерывности в точке (x, y, z) частных производных функции f по переменным x и y

$$\frac{\partial f}{\partial x}(x+\theta_1\Delta x, y+\Delta y, z+\Delta z) = \frac{\partial f}{\partial x}(x, y, z) + \varepsilon_1(\Delta x, \Delta y, \Delta z),$$
$$\frac{\partial f}{\partial y}(x, y+\theta_2\Delta y, z+\Delta z) = \frac{\partial f}{\partial y}(x, y, z) + \varepsilon_2(\Delta y, \Delta z),$$

где величины ε_i, $i = 1, 2$, стремятся к нулю, когда приращения $\Delta x, \Delta y$, Δz стремятся к нулю.

Так как в точке (x, y, z) существует частная производная $\partial f/\partial z$, то

$$f(x, y, z + \Delta z) - f(x, y, z) = \frac{\partial f}{\partial z}(x, y, z)\Delta z + \varepsilon_3(\Delta z)\Delta z,$$

где $\varepsilon_3(\Delta z) \to 0$ при $\Delta z \to 0$.

Таким образом, из (12.1.3) получаем

$$\Delta f = \frac{\partial f}{\partial x}(x, y, z)\Delta x + \frac{\partial f}{\partial y}(x, y, z)\Delta y + \frac{\partial f}{\partial z}(x, y, z)\Delta z + \\ + \varepsilon_1(\Delta x, \Delta y, \Delta z)\Delta x + \varepsilon_2(\Delta y, \Delta z)\Delta y + \varepsilon_3(\Delta z)\Delta z.$$

Оценим сумму последних трёх слагаемых в полученном выражении. Согласно неравенству Коши–Буняковского (6.8.4)

$$|\varepsilon_1\Delta x + \varepsilon_2\Delta y + \varepsilon_3\Delta z| \leqslant \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2}\,\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} = \\ = o\left(\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}\,\right),$$

когда приращения переменных Δx, Δy, Δz стремятся к нулю.

Таким образом, для функций трёх переменных теорема доказана. Доказательство в общем случае не отличается от приведенного.

Отметим, что утверждение теоремы 5.2.1 о дифференцируемости функции одной переменной в точках, где она имеет производную, можно рассматривать как частный случай теоремы 12.1.2, соответствующий $m = 1$.

Необходимые условия дифференцируемости функций многих переменных из теоремы 12.1.1 и достаточное условие из теоремы 12.1.2 отличаются друг от друга. Покажем, что условия ни одной из этих теорем не являются необходимыми и достаточными.

Сначала убедимся, что функция может быть недифференцируемой в точке, в которой она непрерывна и имеет частные производные первого порядка по всем переменным.

Пусть функция двух переменных при $(x, y) \neq (0, 0)$ задана равенством

$$f(x, y) := \frac{2xy}{\sqrt{x^2 + y^2}}\,, \tag{12.1.4}$$

и $f(0, 0) := 0$.

Так как $|2xy| \leqslant x^2 + y^2$, то

$$|f(x,y) - f(0,0)| = |f(x,y)| \leqslant \sqrt{x^2+y^2}.$$

Отсюда следует непрерывность f в точке $(0,0)$.

Найдём частные производные функции f. Так как $f(x,0) = 0$ и $f(0,y) = 0$, то

$$\frac{\partial f}{\partial x}(0,0) = 0, \qquad \frac{\partial f}{\partial y}(0,0) = 0. \tag{12.1.5}$$

А для точек $(x,y) \neq (0,0)$ имеем

$$\frac{\partial f}{\partial x}(x,y) = \frac{2y}{\sqrt{x^2+y^2}} - 2xy\,\frac{2x}{2(\sqrt{x^2+y^2})^3} = \frac{2y^3}{(x^2+y^2)^{3/2}}$$

и аналогично

$$\frac{\partial f}{\partial y}(x,y) = \frac{2x^3}{(x^2+y^2)^{3/2}}\,.$$

Таким образом, f имеет частные производные всюду, но в точке $(0,0)$ они не являются непрерывными.

Функция f в точке $(0,0)$ недифференцируема. В самом деле, согласно (12.1.5) дифференцируемость f в точке $(0,0)$ означала бы, что при $\sqrt{x^2+y^2} \to 0$

$$f(x,y) = o(\sqrt{x^2+y^2}\,). \tag{12.1.6}$$

А, например, при $y = x$ имеем

$$f(x,y) = \frac{2x^2}{\sqrt{2x^2}} = \sqrt{2}\,|x|$$

и оценка (12.1.6) не выполняется.

В теореме 12.1.2 одна из частных производных функции m переменных не предполагалась непрерывной. Функция (12.1.4) показывает, что нельзя отказаться от требования непрерывности двух частных производных.

Покажем теперь, что функция может быть дифференцируемой, когда все её частные производные существуют всюду, но не непрерывны.

Пусть

$$g(x,y) := (x^2+y^2)\sin\frac{1}{x^2+y^2}\,,$$

если $(x,y) \neq (0,0)$, и $g(0,0) := 0$.

Так как
$$|g(x,y)| \leqslant x^2 + y^2,$$
то при $x^2 + y^2 \to 0$
$$g(x,y) = o(\sqrt{x^2+y^2}\,).$$
Поэтому функция g дифференцируема в точке $(0,0)$.

При $x \neq 0$
$$g(x,0) = x^2 \sin\frac{1}{x^2}$$
и
$$\frac{\partial g}{\partial x}(x,0) = 2x\sin\frac{1}{x^2} + x^2\cos\frac{1}{x^2}\cdot\frac{-2}{x^3} = 2x\sin\frac{1}{x^2} - \frac{2}{x}\cos\frac{1}{x^2}\,.$$
Таким образом, частная производная g по x разрывна в точке $(0,0)$. Разрывна также и частная производная по y.

Приведём необходимое и достаточное условие дифференцируемости функций многих переменных.

Теорема 12.1.3. *Функция $f(\mathbf{x})$, имеющая в точке $\mathbf{x}^0 \in \mathbb{E}^m$ частные производные по всем переменным, дифференцируема в точке $\mathbf{x}^0$ в том и только том случае, когда при $|\Delta\mathbf{x}| \to 0$*
$$\begin{aligned}f(x_1^0 + \Delta x_1, \ldots, x_m^0 + \Delta x_m) = \\ = \sum_{k=1}^{m} f(x_1^0, \ldots, x_{k-1}^0, x_k^0 + \Delta x_k, x_{k+1}^0, \ldots, x_m^0) + \\ + (m-1) f(x_1^0, \ldots, x_m^0) = \\ = o(|\Delta\mathbf{x}|).\end{aligned} \tag{12.1.7}$$

Доказательство. Из существования частных производных следует, что при $\Delta x_1 \to 0$
$$\begin{aligned}f(x_1^0 + \Delta x_1, x_2^0, \ldots, x_m^0) - f(x_1^0, x_2^0, \ldots, x_m^0) = \\ = \frac{\partial f}{\partial x_1}(x_1^0, x_2^0, \ldots, x_m^0)\Delta x_1 + o(\Delta x_1)\end{aligned}$$
и аналогичные оценки для приращений по остальным аргументам. Поэтому оценка (12.1.7) эквивалентна тому, что при $|\Delta\mathbf{x}| \to 0$
$$\Delta f = f(x_1^0 + \Delta x_1, \ldots, x_m^0 + \Delta x_m) - f(x_1^0, \ldots, x_m^0) =$$

$$= \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x})\Delta x_k + o(|\Delta \mathbf{x}|).$$

Отсюда вытекает утверждение теоремы.

Для функций двух переменных формула (12.1.7) имеет вид

$$f(x+\Delta x, y+\Delta y) - f(x+\Delta x, y) - f(x, y+\Delta y) + f(x,y) = \\ = o(\sqrt{(\Delta x)^2 + (\Delta y)^2}\,), \qquad |\Delta x| + |\Delta y| \to 0. \tag{12.1.8}$$

Определение. Пусть функция $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ дифференцируема в точке $\mathbf{x}$, т.е. справедлива оценка (12.1.2). Часть приращения функции f в этой точке, линейно зависящую от приращений аргументов $\Delta x_1, \ldots, \Delta x_m$, называют *дифференциалом* (полным дифференциалом) f в точке $\mathbf{x}$.

Обозначают дифференциал $df(\mathbf{x})$ или df. Таким образом,

$$df(\mathbf{x}) := \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x})\Delta x_k.$$

Слово "полный" в определение дифференциала добавляют, чтобы подчеркнуть, что приращение получают, вообще говоря, все переменные $x_1, \ldots, x_m$. Впрочем, это слово обычно опускают и говорят просто дифференциал, имея в виду полный дифференциал.

Как и в случае функций одной переменной, приращения независимых переменных Δx_k называют их дифференциалами и обозначают dx_k. Таким образом,

$$df(\mathbf{x}) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x})\, dx_k.$$

Отметим важное отличие дифференциалов функций одной и многих переменных.

Если дифференциал функции одной переменной в некоторой точке не равен нулю, то в достаточно малой окрестности этой точки он даёт главную часть приращения функции. В частности, дифференциал имеет тот же знак, что и приращение функции.

Для функций многих переменных подобное свойство не имеет места и о знаке приращения функции по дифференциалу судить нельзя.

Действительно, пусть дифференциал функции $f(x_1, \ldots, x_m)$, $m \geqslant 2$, не равен нулю тождественно, т.е. среди её частных производных есть отличные от нуля. Рассмотрим равенство

$$\sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x})\Delta x_k = 0 \tag{12.1.9}$$

как линейное уравнение с m неизвестными $\Delta x_1, \ldots, \Delta x_m$. Уравнение (12.1.9) имеет ненулевые решения. Поэтому, для приращения $\Delta\mathbf{x} = (\Delta x_1, \ldots, \Delta x_m)$, где набор $\Delta x_1, \ldots, \Delta x_m$ является ненулевым решением уравнения (12.1.9), получим согласно (12.1.2)

$$\Delta f = o(|\Delta\mathbf{x}|), \qquad \Delta\mathbf{x} \to 0.$$

Таким образом, в этом случае дифференциал не даёт никакой информации о знаке приращения функции.

§ 12.2. Касательная плоскость

Выясним геометрический смысл частных производных и полного дифференциала функций многих переменных.

Пусть задана функция m переменных $f(x_1, \ldots, x_m)$. Рассмотрим в пространстве $\mathbb{E}^{m+1}$ множество точек с координатами $(x_1, \ldots, x_m, f(x_1, \ldots, x_m))$. Это множество называется графиком функции $z = f(x_1, \ldots, x_m)$.

При определении частной производной

$$\frac{\partial f}{\partial x_1}(x_1^0, \ldots, x_m^0)$$

в пространстве $\mathbb{E}^{m+1}$ точек $(x_1, \ldots, x_m, z)$ берётся двумерная плоскость, задаваемая $m-1$ уравнением $x_2 = x_2^0, \ \ldots, \ x_m = x_m^0$. В этой плоскости, точки которой характеризуются координатами (x_1, z), рассматривается кривая, являющаяся графиком функции $z = f(x_1, x_2^0, \ldots, x_m^0)$.

Согласно сказанному в § 5.2 существование производной функции $f(x_1, \ldots, x_m)$ по переменной x_1 равносильно существованию касательной к графику этой функции в точке x_1^0, а значением производной является тангенс угла наклона касательной к оси OX_1, причём угол берется со знаком $+$ или $-$.

Это показывает геометрический смысл частных производных.

Выясним теперь, какому геометрическому свойству соответствует дифференцируемость функции многих переменных.

График функции $z = f(x_1, \dots, x_m)$, т.е. множество точек $(m+1)$-мерного пространства $\mathbb{E}^{m+1} = \{(x_1, \dots, x_m, z)\}$, имеющих координаты

$$(x_1, \dots, x_m, f(x_1, \dots, x_m)),$$

будем назвать поверхностью в этом пространстве. Обозначим её S. Подробно о поверхностях будет говориться в главе 22, сейчас достаточно интуитивных представлений.

Определение. Пусть в некоторой окрестности точки $(x_1^0, \dots, x_m^0)$ задана функция $f(x_1, \dots, x_m)$ и L – гиперплоскость в пространстве $\mathbb{E}^{m+1} = \{(x_1, \dots, x_m, z)\}$, проходящая через точку $M^0(x_1^0, \dots, x_m^0, f(x_1^0, \dots, x_m^0))$. Расстояние произвольной точки $M(x_1, \dots, x_m, f(x_1, \dots, x_m))$ поверхности S до плоскости L обозначим через $\rho(M, L)$. Плоскость L называется *касательной* к поверхности S в точке M^0, если

$$\rho(M, L) = o(\rho(M, M^0)), \qquad \rho(M, M^0) \to 0, \tag{12.2.1}$$

где $\rho(M, M^0)$ – расстояние точки M до M^0.

Напомним, что подобным образом может быть определена касательная к кривой на плоскости, см. (5.2.9).

Теорема 12.2.1. *Функция $f(x_1, \dots, x_m)$ дифференцируема в точке $(x_1^0, \dots, x_m^0)$ в том и только том случае, когда в точке $M^0(x_1^0, \dots, x_m^0, f(x_1^0, \dots, x_m^0))$ существует не параллельная оси OZ касательная плоскость к поверхности $z = f(x_1, \dots, x_m)$.*

Если такая касательная плоскость существует, то она определяется однозначно и имеет уравнение

$$z - z^0 = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x_1^0, \dots, x_m^0)(x_k - x_k^0), \tag{12.2.2}$$

где $z^0 := f(x_1^0, \dots, x_m^0)$.

Доказательство. Предположим, что касательная плоскость, о которой говорится в теореме, существует. Обозначим эту плоскость L и пусть

$$\sum_{k=1}^{m} C_k(x_k - x_k^0) + D(z - z^0) = 0 \tag{12.2.3}$$

– её нормированное уравнение, т.е. выполнено условие

$$\sum_{k=1}^{m} C_k^2 + D^2 = 1. \tag{12.2.4}$$

Так как плоскость L не параллельна оси OZ, то $D \neq 0$. Будем считать, что $D > 0$.

Покажем, что из существования касательной плоскости следует, что если точки $M(x_1, \ldots, x_m, f(x_1, \ldots, x_m))$ и $M^0(x_1^0, \ldots, x_m^0, f(x_1^0, \ldots, x_m^0))$ поверхности S достаточно близки, то существует число α такое, что

$$|f(x_1, \ldots, x_m) - f(x_1^0, \ldots, x_m^0)| \leqslant \alpha r, \tag{12.2.5}$$

где

$$r := \sqrt{\sum_{k=1}^{m} (x_k - x_k^0)^2}. \tag{12.2.6}$$

Поскольку (12.2.3) – нормированное уравнение плоскости L, то согласно (11.1.11) расстояние точки M до L выражается формулой

$$\rho(M, L) = \left| \sum_{k=1}^{m} C_k(x_k - x_k^0) + D(f(x_1, \ldots, x_m) - z^0) \right|. \tag{12.2.7}$$

Согласно (12.2.1), если точка M достаточно близка к M^0, то

$$\left| \sum_{k=1}^{m} C_k(x_k - x_k^0) + D(f(x_1, \ldots, x_m) - z^0) \right| \leqslant \frac{D}{2} \rho(M, M^0) =$$
$$= \frac{D}{2} \sqrt{\sum_{k=1}^{m} (x_k - x_k^0)^2 + (f(x_1, \ldots, x_m) - z^0)^2} =$$
$$= \frac{D}{2} \sqrt{r^2 + (f(x_1, \ldots, x_m) - z^0)^2}.$$

Отсюда получаем

$$D|f(x_1, \ldots, x_m) - z^0| \leqslant$$
$$\leqslant \left| \sum_{k=1}^{m} C_k(x_k - x_k^0) \right| + \frac{D}{2} \sqrt{r^2 + (f(x_1, \ldots, x_m) - z^0)^2}. \tag{12.2.8}$$

С помощью неравенства Коши–Буняковского (6.8.4) и (12.2.4) находим

$$\left|\sum_{k=1}^{m} C_k(x_k - x_k^0)\right| \leqslant \sqrt{\sum_{k=1}^{m} C_k^2}\sqrt{\sum_{k=1}^{m}(x_k - x_k^0)^2} \leqslant r.$$

Поэтому из (12.2.8) следует, что

$$D|f(x_1, \ldots, x_m) - z^0| \leqslant r + \frac{D}{2}\sqrt{r^2 + (f(x_1, \ldots, x_m) - z^0)^2} \leqslant$$
$$\leqslant r + \frac{D}{2}(r + |f(x_1, \ldots, x_m) - z^0|).$$

Значит,

$$\frac{D}{2}|f(x_1, \ldots, x_m) - z^0| \leqslant r + \frac{D}{2}\,r.$$

Это доказывает оценку (12.2.5).

Из (12.2.1), (12.2.7) и (12.2.5) следует, что при $r \to 0$

$$\sum_{k=1}^{m} C_k(x_k - x_k^0) + D(f(x_1, \ldots, x_m) - z^0) =$$
$$= o\left(\sqrt{\sum_{k=1}^{m}(x_k - x_k^0)^2 + (f(x_1, \ldots, x_m) = -z^0)^2}\right) =$$
$$= o(\sqrt{r^2 + \alpha^2 r^2}\,) = o(r).$$

Значит,

$$f(x_1, \ldots, x_m) - z^0 = -\sum_{k=1}^{m} \frac{C_k}{D}(x_k - x_k^0) + o(r), \qquad r \to 0.$$

Таким образом, установлены дифференцируемость функции $f(x_1, \ldots, x_m)$ в точке $(x_1^0, \ldots, x_m^0)$ и равенства

$$-\frac{C_k}{D} = \frac{\partial f}{\partial x_k}(x_1^0, \ldots, x_m^0), \qquad k = 1, \ldots, m,$$

из которые следует, что уравнение касательной плоскости (12.2.3) имеет вид (12.2.2).

Итак, если в точке $(x_1^0, \dots, x_m^0)$ существует непараллельная оси OZ касательная плоскость, то функция $f(x_1, \dots, x_m)$ дифференцируема в этой точке и (12.2.2) является уравнением касательной плоскости.

Докажем теперь, что если функция $f(x_1, \dots, x_m)$ дифференцируема в точке $(x_1^0, \dots, x_m^0)$, то плоскость (12.2.2), которую обозначим L, является касательной к графику функции $z = f(x_1, \dots, x_m)$ в точке $M^0(x_1^0, \dots, x_m^0, f(x_1^0, \dots, x_m^0))$.

Согласно (11.1.11) для расстояния точки $M \in S$ до плоскости L справедливо равенство

$$\rho(M, L) = \frac{1}{a}\left| f(x_1, \dots, x_m) - z^0 - \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x_1^0, \dots, x_m^0)(x_k - x_k^0)\right|, \tag{12.2.9}$$

где

$$a := \left(1 + \sum_{k=1}^{m}\left(\frac{\partial f}{\partial x_k}(x_1^0, \dots, x_m^0)\right)^2\right)^{1/2}$$

– нормирующий множитель.

Так как функция $f(x_1, \dots, x_m)$ дифференцируема в точке $(x_1^0, \dots, x_m^0)$, то

$$f(x_1, \dots, x_m) - f(x_1^0, \dots, x_m^0) = \\ = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x_1^0, \dots, x_m^0)(x_k - x_k^0) + o(r), \qquad r \to 0,$$

где r определено равенством (12.2.6).

Значит, согласно (12.2.9)

$$\rho(M, L) = o(r), \qquad r \to 0. \tag{12.2.10}$$

Но

$$r \leqslant \sqrt{\sum_{k=1}^{m}(x_k - x_k^0)^2 + (f(x_1, \dots, x_m) - f(x_1^0, \dots, x_m^0))^2} = \\ = \rho(M, M^0).$$

Поэтому из (12.2.10) следует, что

$$\rho(M, L) = o(\rho(M, M^0)), \qquad r \to 0,$$

т.е. L является касательной плоскостью к поверхности $z = f(x_1, \ldots, x_m)$ в точке $(x_1^0, \ldots, x_m^0, f(x_1^0, \ldots, x_m^0))$.

Теорема доказана.

Из того, что касательная плоскость имеет уравнение (12.2.2), вытекает геометрический смысл дифференциала функций многих переменных: дифференциал равен приращению функции, графиком которой является касательная плоскость.

В теореме 12.2.1 существенно, что касательная плоскость не параллельна оси OZ.

В самом деле, рассмотрим в пространстве точек (x, y, z) поверхность, являющуюся графиком функции $z = \sqrt[3]{y}$. Эта функция не дифференцируема в точке $(0, 0)$.

Покажем, что плоскость L, заданная уравнением $y = 0$, является касательной к поверхности $z = \sqrt[3]{y}$ в точке $M^0 = (0, 0, 0)$.

Если $M = (x, y, \sqrt[3]{y})$ – произвольная точка поверхности, то $\rho(M, L) = |y|$ и $\rho(M, M^0) = \sqrt{x^2 + y^2 + y^{2/3}}$. Поэтому для $M \neq M^0$

$$\frac{\rho(M, L)}{\rho(M, M^0)} = \frac{|y|}{\sqrt{x^2 + y^2 + y^{2/3}}}$$

и, если $y \neq 0$, то

$$\frac{\rho(M, L)}{\rho(M, M^0)} < \frac{|y|}{|y|^{1/3}} = o(1), \qquad M \to M^0,$$

т.е. плоскость L является касательной.

§ 12.3. Дифференцируемость сложной функции

Установим дифференцируемость сложной функции многих переменных, построенной из дифференцируемых функций.

Теорема 12.3.1. *Пусть функции*

$$x_1(\mathbf{t}) = x_1(t_1, \ldots, t_p), \quad \ldots, \quad x_m(\mathbf{t}) = x_m(t_1, \ldots, t_p), \qquad p \geqslant 1,$$

дифференцируемы в точке $\mathbf{t}^0 = (t_1^0, \ldots, t_p^0)$ *и функция* $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ *дифференцируема в точке* $\mathbf{x}^0 = (x_1^0, \ldots, x_m^0)$, *где* $x_1^0 := x_1(\mathbf{t}^0)$, ..., $x_m^0 := x_m(\mathbf{t}^0)$. *Тогда сложная функция*

$$\varphi(\mathbf{t}) = \varphi(t_1, \ldots, t_p) := f(x_1(t_1, \ldots, t_p), \ldots, x_m(t_1, \ldots, t_p))$$

дифференцируема в точке $\mathbf{t}^0$ *и для её частных производных справедливы равенства*

$$\frac{\partial \varphi}{\partial t_j}(\mathbf{t}^0) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0) \cdot \frac{\partial x_k}{\partial t_j}(\mathbf{t}^0), \qquad j = 1, \ldots, p. \tag{12.3.1}$$

Доказательство. Из дифференцируемости функций $x_1(\mathbf{t}), \ldots, x_m(\mathbf{t})$ в точке $\mathbf{t}^0$ и функции $f(\mathbf{x})$ в точке $\mathbf{x}^0$ следует непрерывность этих функций в указанных точках. Поэтому согласно теореме 11.4.1 функция $\varphi(\mathbf{t})$ непрерывна в точке $\mathbf{t}^0$.

В силу дифференцируемости функции $f(\mathbf{x})$ в точке $\mathbf{x}^0$ существует такое число $\sigma > 0$, что для всех точек $\mathbf{x}^0+\Delta\mathbf{x}$, для которых $|\Delta\mathbf{x}| < \sigma$, справедлива оценка

$$\Delta f = f(\mathbf{x}^0 + \Delta\mathbf{x}) - f(\mathbf{x}^0) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)\Delta x_k + \varepsilon(\Delta\mathbf{x})|\Delta\mathbf{x}|, \tag{12.3.2}$$

где $\varepsilon(\Delta\mathbf{x}) \to 0$ при $|\Delta\mathbf{x}| \to 0$.

Выберем настолько малое $\delta > 0$, чтобы условие $|\Delta\mathbf{t}| < \delta$ обеспечивало попадание точек $(x_1(\mathbf{t}^0 + \Delta\mathbf{t}), \ldots, x_m(\mathbf{t}^0 + \Delta\mathbf{t}))$ в σ-окрестность точки $\mathbf{x}^0$. Для таких $\Delta\mathbf{t}$ при каждом $k = 1, \ldots, m$ имеем

$$\Delta x_k = x_k(\mathbf{t}^0 + \Delta\mathbf{t}) - x_k(\mathbf{t}^0) = \sum_{j=1}^{p} \frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\Delta t_j + \varepsilon_k(\Delta\mathbf{t})|\Delta\mathbf{t}|,$$

где $\varepsilon_k(\Delta\mathbf{t}) \to 0$ при $|\Delta\mathbf{t}| \to 0$.

Подставив эти выражения Δx_k в (12.3.2), получим

$$\begin{aligned}
\Delta\varphi &= \varphi(\mathbf{t}^0 + \Delta\mathbf{t}) - \varphi(\mathbf{t}^0) = f(\mathbf{x}^0 + \Delta\mathbf{x}) - f(\mathbf{x}^0) = \\
&= \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)\left[\sum_{j=1}^{p} \frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\Delta t_j + \varepsilon_k(\Delta\mathbf{t})|\Delta\mathbf{t}|\right] + \varepsilon(\Delta\mathbf{x})|\Delta\mathbf{x}| = \\
&= \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)\left[\sum_{j=1}^{p} \frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\Delta t_j\right] + \\
&\qquad + \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)\varepsilon_k(\Delta\mathbf{t})|\Delta\mathbf{t}| + \varepsilon(\Delta\mathbf{x})|\Delta\mathbf{x}|.
\end{aligned} \tag{12.3.3}$$

Так как все $\varepsilon_k(\Delta\mathbf{t}) \to 0$, $k = 1, \ldots, m$, при $|\Delta\mathbf{t}| \to 0$, то

$$\sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)\varepsilon_k(\Delta\mathbf{t})|\Delta\mathbf{t}| = o(|\Delta\mathbf{t}|), \qquad |\Delta\mathbf{t}| \to 0. \tag{12.3.4}$$

Покажем, что при $|\Delta\mathbf{t}| \to 0$

$$\varepsilon(\Delta\mathbf{x})|\Delta\mathbf{x}| = o(|\Delta\mathbf{t}|). \tag{12.3.5}$$

Поскольку $|\Delta\mathbf{x}| \to 0$ при $|\Delta\mathbf{t}| \to 0$ и

$$\varepsilon(\Delta\mathbf{x})|\Delta\mathbf{x}| = \varepsilon(\Delta\mathbf{x})\frac{|\Delta\mathbf{x}|}{|\Delta\mathbf{t}|}|\Delta\mathbf{t}|,$$

достаточно доказать ограниченность отношения $|\Delta\mathbf{x}|/|\Delta\mathbf{t}|$ при малых $|\Delta\mathbf{t}|$. Имеем

$$\begin{aligned}\frac{|\Delta\mathbf{x}|}{|\Delta\mathbf{t}|} &= \frac{1}{|\Delta\mathbf{t}|}\sqrt{\sum_{k=1}^{m}(\Delta x_k)^2} \leqslant \frac{1}{|\Delta\mathbf{t}|}\sum_{k=1}^{m}|\Delta x_k| = \\ &= \frac{1}{|\Delta\mathbf{t}|}\sum_{k=1}^{m}\left|\sum_{j=1}^{p}\frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\Delta t_j + \varepsilon_k(\Delta\mathbf{t})|\Delta\mathbf{t}|\right| \leqslant \\ &\leqslant \sum_{k=1}^{m}\left(\sum_{j=1}^{p}\left|\frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\right| + |\varepsilon_k(\Delta\mathbf{t})|\right).\end{aligned}$$

Из (12.3.3)–(12.3.5) вытекает, что

$$\begin{aligned}\Delta\varphi &= \sum_{k=1}^{m}\frac{\partial f}{\partial x_k}(\mathbf{x}^0)\sum_{j=1}^{p}\frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\Delta t_j + o(|\Delta\mathbf{t}|) = \\ &= \sum_{j=1}^{p}\left[\sum_{k=1}^{m}\frac{\partial f}{\partial x_k}(\mathbf{x}^0)\frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\right]\Delta t_j + o(|\Delta\mathbf{t}|), \qquad |\Delta\mathbf{t}| \to 0.\end{aligned} \tag{12.3.6}$$

Таким образом, функция $\varphi(\mathbf{t})$ дифференцируема. Так как согласно теореме 12.1.1 множители перед приращениями аргументов Δt_j в представлении дифференциала функции равны соответствующим частным производным этой функции, то из (12.3.6) следуют равенства (12.3.1).

Теорема доказана.

Приведём частный случай теоремы 12.3.1, когда $p = 1$.

Следствие 12.3.2. *Пусть функции $x_1 = x_1(t), \ldots, x_m = x_m(t)$ имеют производные в точке $t^0 \in [a, b]$, а функция $f(\mathbf{x}) =$*

$f(x_1, \ldots, x_m)$ *дифференцируема в точке* $\mathbf{x}^0 = (x_1^0, \ldots, x_m^0)$, *где* $\mathbf{x}^0 = (x_1(t^0), \ldots, x_m(t^0))$. *Тогда сложная функция* $\varphi(t) := f(x_1(t), \ldots, x_m(t))$ *имеет в точке* t^0 *производную и*

$$\frac{d\varphi}{dt}(t^0) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0) \cdot \frac{dx_k}{dt}(t^0). \tag{12.3.7}$$

Если выполнены условия теоремы 12.3.1, то функция $\varphi(\mathbf{t})$ дифференцируема в точке $\mathbf{t}^0$ и дифференциал $d\varphi$ выражается через дифференциалы независимых переменных по формуле

$$d\varphi = \sum_{j=1}^{p} \frac{\partial \varphi}{\partial t_j}(\mathbf{t}^0)\, dt_j. \tag{12.3.8}$$

Согласно (12.3.1)

$$d\varphi = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0) \sum_{j=1}^{p} \frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\, dt_j.$$

Так как

$$\sum_{j=1}^{p} \frac{\partial x_k}{\partial t_j}(\mathbf{t}^0)\, dt_j = dx_k,$$

то

$$d\varphi = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)\, dx_k. \tag{12.3.9}$$

Это выражение имеет тот же вид, что и формула (12.3.8), но в (12.3.9) dx_k являются дифференциалами зависимых переменных.

Тот факт, что формулы (12.3.8) и (12.3.9), имеют одинаковый вид и при выражении $d\varphi$ через дифференциалы независимых переменных и через дифференциалы зависимых переменных, как и для дифференциалов функций одной переменой, называют *инвариантностью формы первого дифференциала*.

В § 12.5 для функций многих переменных будут определены дифференциалы старших порядков, которые как и дифференциалы старших порядков функций одной переменной, свойством инвариантности не обладают.

С помощью равенств (12.3.1) легко выразить дифференциалы суммы, разности, произведения и частного функций через дифференциалы функций, над которыми производятся эти действия. Эти формулы аналогичны соответствующим формулам для функций одной переменной из § 5.2.

§ 12.4. Производная по направлению. Градиент

Пусть функция $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ задана на луче с вершиной в точке $\mathbf{x}^0 \in \mathbb{E}^m$, параллельном m-мерному единичному вектору $\overline{a} = (\cos\alpha_1, \ldots, \cos\alpha_m)$, компоненты которого $\cos\alpha_1, \ldots,$ $\cos\alpha_m$ называют направляющими косинусами этого вектора. Так как вектор $\overline{a}$ единичный, то

$$\sum_{k=1}^{m} \cos^2\alpha_k = 1.$$

Точки $\mathbf{x}$, принадлежащие лучу, имеют координаты

$$x_1 = x_1^0 + t\cos\alpha_1, \quad \ldots, \quad x_m = x_m^0 + t\cos\alpha_m, \qquad t \geqslant 0. \tag{12.4.1}$$

Придадим аргументу функции f в точке $\mathbf{x}^0$ приращение в направлении луча (12.4.1) и рассмотрим отношение приращения функции к приращению аргумента:

$$\frac{1}{t}\big(f(x_1^0 + t\cos\alpha_1, \ldots, x_m^0 + t\cos\alpha_m) - f(x_1^0, \ldots, x_m^0)\big). \tag{12.4.2}$$

Если дробь (12.4.2) имеет предел при $t \to +0$, то говорят, что функция $f(\mathbf{x})$ имеет в точке $\mathbf{x}^0$ *производную по направлению вектора* $\overline{a}$, и значение предела называют производной по этому направлению:

$$\frac{\partial f}{\partial \overline{a}}(\mathbf{x}^0) := \lim_{t\to+0} \frac{1}{t}\big(f(x_1^0+t\cos\alpha_1, \ldots, x_m^0+t\cos\alpha_m)-f(x_1^0, \ldots, x_m^0)\big).$$

Если в качестве вектора $\overline{a}$ взят один из базисных векторов, то производная по направлению является соответствующей односторонней частной производной.

Если функция $f(\mathbf{x})$ дифференцируема в точке $\mathbf{x}^0$, то f имеет в этой точке производную по любому направлению и

$$\frac{\partial f}{\partial \overline{a}}(\mathbf{x}^0) = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)\cos\alpha_k.$$

Введём вектор

$$\operatorname{grad} f(\mathbf{x}^0) := \left(\frac{\partial f}{\partial x_1}(\mathbf{x}^0), \ldots, \frac{\partial f}{\partial x_m}(\mathbf{x}^0)\right),$$

который называют *градиентом* функции f в точке $\mathbf{x}^0$. Производная по направлению равна скалярному произведению $\operatorname{grad} f$ и вектора $\overline{a}$

$$\frac{\partial f}{\partial \overline{a}} = (\operatorname{grad} f, \overline{a}). \tag{12.4.3}$$

Так как $\overline{a}$ – единичный вектор, то из (12.4.3) следует неравенство

$$\frac{\partial f}{\partial \overline{a}} \leqslant |\operatorname{grad} f| = \sqrt{\sum_{k=1}^{m} \left(\frac{\partial f}{\partial x_k}(\mathbf{x}^0) \right)^2}. \tag{12.4.4}$$

Оценка (12.4.4) справедлива для производной по любому направлению. Если вектор $\operatorname{grad} f$ не нулевой, то существует единственное направление, для производной по которому в (12.4.4) имеет место знак равенства. Это направление задаёт вектор $\overline{a}$, для которого

$$\cos \alpha_k := \frac{1}{|\operatorname{grad} f|} \frac{\partial f}{\partial x_k}(\mathbf{x}^0), \qquad k = 1, \ldots, m.$$

Отсюда вытекает геометрическая характеристика градиента: градиент – это вектор, длина которого равна максимальному значению производной по направлению, а направлением градиента является то, производная по которому имеет наибольшее значение.

В определении градиента участвуют частные производные. Поэтому формально он зависит от системы координат. Однако, приведенная геометрическая характеристика показывает, что на самом деле градиент выражает внутренние свойства функции и, таким образом, не зависит от выбора системы координат. Это свойство градиента называют его *инвариантностью*.

Для записи градиентов и действий над ними удобно пользоваться формальным символическим вектором ∇ (“набла"), который по определению имеет компоненты

$$\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_m} \right).$$

Вектор ∇ называют оператором Гамильтона.

С помощью вектора набла градиент записывают так:

$$\operatorname{grad} f = \nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_m} \right),$$

т.е. умножение ∇ на скалярную функцию производится по правилу умножения вектора на скаляр.

Необходимо при этом иметь в виду, что символ ∇ действует на стоящую за ним функцию как дифференциальный оператор и обладает некоторыми свойствами, аналогичными свойствам производной. Например, если функции f и g имеют в точке частные производные по всем переменным, то в этой точке справедливо равенство

$$\nabla fg = g\nabla f + f\nabla g,$$

т.е.

$$\operatorname{grad}(fg) = g\operatorname{grad} f + f\operatorname{grad} g.$$

§ 12.5. Частные производные и дифференциалы высших порядков

Если функция $f(x_1, \ldots, x_m)$ в некоторой окрестности точки $(x_1^0, \ldots, x_m^0)$ имеет производную $\partial f/\partial x_k$, то можно ставить вопрос о существовании у этой производной частных производных в точке $(x_1^0, \ldots, x_m^0)$.

Предположим, что существует частная производная

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right). \tag{12.5.1}$$

Если $j \neq k$, то эту производную называют *смешанной* частной производной второго порядка функции f по переменным x_k и x_j и обозначают

$$\frac{\partial^2 f}{\partial x_j \partial x_k}.$$

Обратите внимание на порядок, в котором записаны здесь ∂x_j и ∂x_k.

А если $j = k$, то производную (12.5.1) обозначают

$$\frac{\partial^2 f}{\partial x_k^2}$$

и называют *чистой* частной производной второго порядка функции f по переменной x_k. Здесь ∂x_k^2 понимается как $(\partial x_k)^2$.

Точно так же вводятся частные производные третьего, четвертого и вообще любого порядка. Их называют чистыми, если каждый раз производная берётся по одной и той же переменной, и называют смешанными, если среди переменных, по которым берутся производные, по крайней мере две различны.

Частные производные высших порядков, как и производные первого порядка, могут не существовать. Но даже если смешанные частные производные существуют, их значения могут зависеть от того, в каком порядке берутся переменные в определении этих производных. Поэтому необходимо различать производные

$$\frac{\partial^2 f}{\partial x_j \partial x_k}, \qquad \frac{\partial^2 f}{\partial x_k \partial x_j}.$$

Здесь имеется определённая аналогия с повторным пределами.

Вопрос о зависимости смешанных частных производных от порядка переменных, в котором вычисляются эти производные, рассмотрим подробно для частных производных второго порядка функций двух переменных.

Сначала убедимся, что значения смешанных частных производных могут зависеть от порядка переменных.

Положим для $(x, y) \neq (0, 0)$

$$f(x, y) := \frac{x^3 y}{x^2 + y^2} \tag{12.5.2}$$

и $f(0, 0) := 0$.

Функция f непрерывна в точке $(0, 0)$, так как $|f(x, y)| \leqslant |xy|$. Непрерывность f и всех её частных производных в остальных точках очевидны.

Найдём производные функции f. Если точка (x, y) отлична от $(0, 0)$, то

$$\frac{\partial f}{\partial x}(x, y) = \frac{3x^2 y(x^2 + y^2) - x^3 y \cdot 2x}{(x^2 + y^2)^2} = \frac{x^2 y(x^2 + 3y^2)}{(x^2 + y^2)^2}.$$

Так как $f(x, 0) = 0$ при всех x, то

$$\frac{\partial f}{\partial x}(0, 0) = 0.$$

Таким образом, во всех точках (x, y) справедлива оценка

$$\left| \frac{\partial f}{\partial x}(x, y) \right| \leqslant 3|y|$$

и эта частная производная непрерывна в начале координат.

Так как
$$\frac{\partial f}{\partial x}(0,y)=0,$$
то
$$\frac{\partial^2 f}{\partial y \partial x}(0,0)=0.$$

Найдём теперь значение в точке $(0,0)$ смешанной производной, взятой в другом порядке.

Если $(x,y)\neq(0,0)$, то
$$\frac{\partial f}{\partial y}(x,y)=\frac{x^3(x^2+y^2)-x^3y\cdot 2y}{(x^2+y^2)^2}=\frac{x^3(x^2-y^2)}{(x^2+y^2)^2}\,.$$
А так как
$$\frac{\partial f}{\partial y}(0,0)=0,$$
то всегда
$$\left|\frac{\partial f}{\partial y}(x,y)\right|\leqslant|x|.$$
Значит, частная производная $\partial f/\partial y$ непрерывна в начале координат.

В силу равенства
$$\frac{\partial f}{\partial y}(x,0)=x$$
имеем
$$\frac{\partial^2 f}{\partial x \partial y}(0,0)=1.$$

Итак, у функции (12.5.2) обе частные производные первого порядка непрерывны всюду, обе смешанные частные производные второго порядка в точке $(0,0)$ существуют, но их значения различны.

Рассмотрим достаточные условия равенства смешанных производных.

Теорема 12.5.1 (Теорема Янга). *Если частные производные первого порядка $\partial f/\partial x$ и $\partial f/\partial y$ функции $f(x,y)$ существуют в некоторой окрестности точки (x^0,y^0) и дифференцируемы в точке (x^0,y^0), то значения обеих смешанных частных производных второго порядка функции f в точке (x^0,y^0) равны:*
$$\frac{\partial^2 f}{\partial x \partial y}(x^0,y^0)=\frac{\partial^2 f}{\partial y \partial x}(x^0,y^0). \tag{12.5.3}$$

Доказательство. Из дифференцируемости производных $\partial f/\partial x$ и $\partial f/\partial y$ в точке (x^0, y^0) следует, что в этой точке существуют все частные производные второго порядка функции f.

Рассмотрим при достаточно малых h функцию

$$\varphi(y) := f(x^0 + h, y) - f(x^0, y)$$

и вторую смешанную разность функции f в точке (x^0, y^0):

$$\begin{aligned} \Delta^2 f := f(x^0 + h, y^0 + h) - f(x^0 + h, y^0) - \\ - f(x^0, y^0 + h) + f(x^0, y^0) = \\ = \varphi(y^0 + h) - \varphi(y^0). \end{aligned}$$

Применив к разности значений функции φ формулу конечных приращений Лагранжа, получим, что при некотором $\theta \in (0, 1)$

$$\begin{aligned} \Delta^2 f = \frac{d\varphi}{dy}(y^0 + \theta h)h = \\ = \left[\frac{\partial f}{\partial y}(x^0 + h, y^0 + \theta h) - \frac{\partial f}{\partial y}(x^0, y^0 + \theta h)\right]h. \end{aligned} \tag{12.5.4}$$

В силу дифференцируемости в точке (x^0, y^0) производной $\partial f/\partial y$ имеем при $h \to 0$

$$\begin{aligned} \frac{\partial f}{\partial y}(x^0 + h, y^0 + \theta h) - \frac{\partial f}{\partial y}(x^0, y^0) = \\ = \frac{\partial^2 f}{\partial x \partial y}(x^0, y^0)h + \frac{\partial^2 f}{\partial y^2}(x^0, y^0)\theta h + o(h) \end{aligned}$$

и

$$\frac{\partial f}{\partial y}(x^0, y^0 + \theta h) - \frac{\partial f}{\partial y}(x^0, y^0) = \frac{\partial^2 f}{\partial y^2}(x^0, y^0)\theta h + o(h).$$

Поэтому из (12.5.4) следует оценка

$$\Delta^2 f = \frac{\partial^2 f}{\partial x \partial y}(x^0, y^0)h^2 + o(h^2), \qquad h \to 0. \tag{12.5.5}$$

Точно также доказывается оценка

$$\Delta^2 f = \frac{\partial^2 f}{\partial y \partial x}(x^0, y^0)h^2 + o(h^2), \qquad h \to 0,$$

из которой и (12.5.5) вытекает равенство (12.5.3).

Теорема доказана.

Условия на функцию в теореме 12.5.1 обеспечивают не только равенство смешанных производных в точке, но и существование в этой точке чистых частных производных второго порядка. Но смешанные производные могут существовать и у функций, не имеющих чистых производных. Например, функция $f(x,y) := x|x|+y|y|$ не имеет чистых частных производных второго порядка в точке $(0,0)$, а смешанные частные производные f равны нулю всюду.

В следующем утверждении чистые частные производные функции могут не существовать.

Теорема 12.5.2 (Теорема Шварца). *Пусть функция $f(x,y)$ в некоторой окрестности точки (x^0,y^0) имеет частные производные $\partial f/\partial x$ и $\partial f/\partial y$ и смешанную производную*

$$\frac{\partial^2 f}{\partial y \partial x}\,. \tag{12.5.6}$$

Если производная (12.5.6) непрерывна в точке (x^0,y^0), то в этой точке существует смешанная производная

$$\frac{\partial^2 f}{\partial x \partial y}$$

и справедливо равенство

$$\frac{\partial^2 f}{\partial x \partial y}(x^0,y^0) = \frac{\partial^2 f}{\partial y \partial x}(x^0,y^0). \tag{12.5.7}$$

Доказательство. Рассмотрим для отличных от нуля чисел h и k вторую смешанную разность

$$\Delta^2_{h,k} f := f(x^0+h,y^0+k) - f(x^0+h,y^0) - f(x^0,y^0+k) + f(x^0,y^0).$$

Будем считать h и k настолько малыми, что все точки, участвующие в построении $\Delta^2_{h,k} f$, попадают в окрестность точки (x^0,y^0), в которой существуют производные первого порядка $\partial f/\partial x$, $\partial f/\partial y$ и смешанная производная (12.5.6).

Положим

$$\psi(x) := f(x,y^0+k) - f(x,y^0).$$

Тогда

$$\Delta^2_{h,k} f = \psi(x^0+h) - \psi(x^0). \tag{12.5.8}$$

Так как в достаточно малой окрестности точки x^0 функция $\psi(x)$ имеет производную, то согласно формуле конечных приращений Лагранжа при некотором $\theta \in (0,1)$

$$\Delta^2_{h,k} f = \frac{d\psi}{dx}(x^0 + \theta h)h = \\ = \left(\frac{\partial f}{\partial x}(x^0 + \theta h, y^0 + k) - \frac{\partial f}{\partial x}(x^0 + \theta h, y^0)\right) h.$$

Отсюда в силу существования смешанной производной (12.5.6), вновь пользуясь формулой конечных приращений Лагранжа, заключаем, что при некотором $\theta_1 \in (0,1)$ справедливо равенство

$$\Delta^2_{h,k} f = hk \frac{\partial^2 f}{\partial y \partial x}(x^0 + \theta h, y^0 + \theta_1 k). \tag{12.5.9}$$

Так как производная (12.5.6) непрерывна в точке (x^0, y^0), то для каждого положительного ε существует такое $\delta > 0$, что при всех $|h| < \delta$ и $|k| < \delta$ имеем

$$\left|\frac{\partial^2 f}{\partial y \partial x}(x^0 + \theta h, y^0 + \theta_1 k) - \frac{\partial^2 f}{\partial y \partial x}(x^0, y^0)\right| < \frac{\varepsilon}{2}\,.$$

Значит, согласно (12.5.9) для таких h и k

$$\left|\Delta^2_{h,k} f - hk \frac{\partial^2 f}{\partial y \partial x}(x^0, y^0)\right| < |hk| \frac{\varepsilon}{2}\,. \tag{12.5.10}$$

Отсюда в силу (12.5.8) получаем

$$\left|\frac{1}{h}\left(\frac{\psi(x^0 + h)}{k} - \frac{\psi(x^0)}{k}\right) - \frac{\partial^2 f}{\partial y \partial x}(x^0, y^0)\right| < \frac{\varepsilon}{2}\,. \tag{12.5.11}$$

Но

$$\frac{\psi(x^0 + h)}{k} = \frac{1}{k}(f(x^0 + h, y^0 + k) - f(x^0 + h, y^0)),$$
$$\frac{\psi(x^0)}{k} = \frac{1}{k}(f(x^0, y^0 + k) - f(x^0, y^0))$$

и частная производная $\partial f / \partial y$ существует в окрестности точки (x^0, y^0), поэтому переходя в неравенстве (12.5.11) к пределу при $k \to 0$, находим

$$\left|\frac{1}{h}\left(\frac{\partial f}{\partial y}(x^0 + h, y^0) - \frac{\partial f}{\partial y}(x^0, y^0)\right) - \frac{\partial^2 f}{\partial y \partial x}(x^0, y^0)\right| \leqslant \frac{\varepsilon}{2} < \varepsilon.$$

Эта оценка имеет место при всех достаточно малых h. Следовательно, из неё вытекают существование производной

$$\frac{\partial^2 f}{\partial x \partial y}(x^0, y^0)$$

и равенство (12.5.7).

Теорема доказана.

Для проверки равенства смешанных производных часто пользуются следующим простым достаточным условием, вытекающим из теоремы 12.5.2.

Следствие 12.5.3. *Если обе смешанные частные производные второго порядка функции $f(x, y)$*

$$\frac{\partial^2 f}{\partial x \partial y}(x, y), \qquad \frac{\partial^2 f}{\partial y \partial x}(x, y)$$

непрерывны в точке (x^0, y^0), то их значения в этой точке равны.

Это утверждение легко установить непосредственно, а не как следствие теоремы 12.5.2. В самом деле, в этом случае наряду с (12.5.10) справедлива оценка

$$\left|\Delta^2_{h,k} f - hk \frac{\partial^2 f}{\partial x \partial y}(x^0, y^0)\right| < |hk| \frac{\varepsilon}{2},$$

которая вместе с (12.5.10) приводит к равенству (12.5.7).

Рассмотрим дифференциалы высших порядков, будем при этом иметь в виду полные дифференциалы.

Если первые частные производные функции f дифференцируемы в некоторой точке, то в этой точке по определению полагают

$$d^2 f := d\,(df)$$

и функцию f называют дважды дифференцируемой. Отметим, что согласно теореме 12.5.1 смешанные производные второго порядка дважды дифференцируемой функции равны.

Выразим дифференциал второго порядка через частные производные функции и дифференциалы аргументов.

В § 12.3 доказано, что дифференциал первого порядка функции $f(x_1, \ldots, x_m)$ имеет вид

$$df = \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}\, dx_k \tag{12.5.12}$$

и в случае, когда x_k являются независимыми переменными, и в случае, когда x_k зависят от других переменных.

Как и для функций одной переменной, для дифференциалов второго порядка в качестве приращений аргумента будем брать те же дифференциалы $dx_1, \ldots, dx_m$, что и в (12.5.12).

Пользуясь правилами вычисления дифференциалов первого порядка, находим, что в точке, где функция f дважды дифференцируема

$$\begin{aligned} d^2 f &= d\left(\sum_{k=1}^{m} \frac{\partial f}{\partial x_k}\, dx_k\right) = \sum_{k=1}^{m} d\left(\frac{\partial f}{\partial x_k}\, dx_k\right) = \\ &= \sum_{k=1}^{m} \left[d\left(\frac{\partial f}{\partial x_k}\right) \cdot dx_k + \frac{\partial f}{\partial x_k}\, d\,(dx_k)\right] = \\ &= \sum_{k=1}^{m} \left[\sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_k}\, dx_j\right] dx_k + \sum_{k=1}^{m} \frac{\partial f}{\partial x_k} d^2 x_k = \\ &= \sum_{k=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_k}\, dx_j\, dx_k + \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}\, d^2 x_k. \end{aligned} \tag{12.5.13}$$

При выводе формулы (12.5.13) имелось в виду, что переменные x_k могут быть функциями других переменных.

Если же переменные x_k независимы, то их вторые дифференциалы $d^2 x_k$ равны нулю. Тогда в правой части равенства (12.5.13) последняя сумма пропадает и оно приобретает вид

$$d^2 f = \sum_{k=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_k}\, dx_j\, dx_k. \tag{12.5.14}$$

Так как для дважды непрерывно дифференцируемых функций их смешанные частные второго порядка не зависят от порядка, в котором берутся эти производные, то двойную сумму в правой части формулы (12.5.14) можно рассматривать как квадратичную форму дифференциалов $dx_1, \ldots, dx_m$.

Эту квадратичную форму можно символически записать в виде

$$\sum_{k=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_k}\, dx_j\, dx_k = \left(\sum_{k=1}^{m} dx_k \frac{\partial}{\partial x_k}\right)^2 f. \tag{12.5.15}$$

Здесь символы $\partial/\partial x_k$ действуют как дифференциальные операторы на функции, записанные после этих символов, выражения

$(\partial/\partial x_k)^2$, получающиеся при возведении в квадрат, понимают как $\partial^2/\partial x_k^2$, произведения $\partial/\partial x_k$ и $\partial/\partial x_j$ понимают как $\partial^2/\partial x_k \partial x_j$, а умножение на f означает, что берутся соответствующие производные функции f.

Дифференциалы третьего и более высоких порядков вводятся по индукции. Если все частные производные функции f до порядка $n-1$ включительно дифференцируемы в некоторой точке, то в этой точке по определению полагают

$$d^n f := d\,(d^{n-1} f), \qquad n = 2, 3, \ldots,$$

и функцию f называют n раз дифференцируемой. Если при этом все частные производные до порядка n включительно непрерывны в точке или в области, то функцию называют n раз непрерывно дифференцируемой соответственно в точке или области.

Вывод формул для дифференциалов третьего и более высоких порядков проводится аналогично тому, как это делалось для дифференциалов второго порядка. Естественно, что формулы при этом усложняются, если переменные x_k не являются независимыми.

Если же переменные x_k независимы и производные непрерывны, то

$$d^n f = \sum_{k_1=1}^{m} \cdots \sum_{k_n=1}^{m} \frac{\partial^n f}{\partial x_{k_1} \ldots \partial x_{k_n}} dx_{k_1} \ldots dx_{k_n} = \left(\sum_{k=1}^{m} dx_k \frac{\partial}{\partial x_k} \right)^n f.$$

§ 12.6. Формула Тейлора

Как и для функций одной переменной, формула Тейлора для функций многих переменных даёт с точностью до определенных остаточных членов представление функции в виде многочлена, который называют многочленом Тейлора.

Пусть функция $f(\mathbf{x})$ дифференцируема n раз в некоторой области пространства $\mathbb{E}^m$, содержащей отрезок, соединяющий точки $\mathbf{x}^0$ и $\mathbf{x}$.

Рассмотрим след функции f на этом отрезке, т.е. функцию

$$\varphi(t) := f(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0)) = \\ = f(x_1^0 + t(x_1 - x_1^0), \ldots, x_m^0 + t(x_m - x_m^0)), \qquad t \in [0, 1].$$

Тогда $\varphi(0) = f(\mathbf{x}^0)$, $\varphi(1) = f(\mathbf{x})$ и при $t \in (0,1)$ точки $\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0)$ являются внутренними точками рассматриваемого отрезка.

Из свойств сложных функций следует, что $\varphi(t)$ имеет на отрезке $[0,1]$ производные до порядка n включительно. Поэтому для функции $\varphi(t)$ согласно формуле Тейлора с остаточным членом в форме Лагранжа имеем

$$\varphi(t) = \sum_{k=0}^{n-1} \frac{\varphi^{(k)}(0)}{k!}\, t^k + \frac{\varphi^{(n)}(\theta t)}{n!}\, t^n, \qquad t \in [0,1], \tag{12.6.1}$$

где число θ удовлетворяет условию $0 < \theta < 1$.

Выразим производные функции φ через частные производные функции f. По правилу дифференцирования сложной функции имеем

$$\varphi'(t) = \sum_{j=1}^{m} \frac{\partial f}{\partial x_j}(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0))(x_j - x_j^0),$$

$$\varphi''(t) = \sum_{j_1=1}^{m} \sum_{j_2=1}^{m} \frac{\partial^2 f}{\partial x_{j_1} \partial x_{j_2}}(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0))(x_{j_1} - x_{j_1}^0)(x_{j_2} - x_{j_2}^0)$$

и вообще при $k \leqslant n$

$$\varphi^{(k)}(t) = \sum_{j_1=1}^{m} \cdots \sum_{j_k=1}^{m} \frac{\partial^k f}{\partial x_{j_1} \ldots \partial x_{j_k}}(\mathbf{x}^0 + t(\mathbf{x} - \mathbf{x}^0)) \times \\ \times (x_{j_1} - x_{j_1}^0) \ldots (x_{j_k} - x_{j_k}^0).$$

С помощью этих равенств из (12.6.1) при $t = 1$ получаем следующее утверждение.

Теорема 12.6.1. *Если функция $f(\mathbf{x})$ дифференцируема n раз в области пространства $\mathbb{E}^m$, содержащей отрезок с концами в точках $\mathbf{x}^0$ и $\mathbf{x}$, то при некотором $\theta \in (0,1)$ справедливо равенство*

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \sum_{k=1}^{n-1} \frac{1}{k!} \sum_{j_1=1}^{m} \cdots \sum_{j_k=1}^{m} \frac{\partial^k f(\mathbf{x}^0)}{\partial x_{j_1} \ldots \partial x_{j_k}} \times \\ \times (x_{j_1} - x_{j_1}^0) \cdots (x_{j_k} - x_{j_k}^0) +$$

$$+\frac{1}{n!}\sum_{j_1=1}^{m}\cdots\sum_{j_n=1}^{m}\frac{\partial^n f\left(\mathbf{x}^0+\theta(\mathbf{x}-\mathbf{x}^0)\right)}{\partial x_{j_1}\ldots\partial x_{j_n}}\times$$
$$\times(x_{j_1}-x_{j_1}^0)\ldots(x_{j_p}-x_{j_n}^0). \quad (12.6.2)$$

Равенство (12.6.2) называют *формулой Тейлора функции многих переменных с остаточным членом в форме Лагранжа.*

Отметим два частных случая формулы (12.6.2). При $n = 1$ получаем равенство

$$f(\mathbf{x})-f(\mathbf{x}^0)=\sum_{j=1}^{m}\frac{\partial f}{\partial x_j}(\mathbf{x}^0+\theta(\mathbf{x}-\mathbf{x}^0))(x_j-x_j^0), \qquad 0<\theta<1, \quad (12.6.3)$$

которое называют *формулой конечных приращений для функций многих переменных*. Равенство (12.6.3) имеет место для функций f, дифференцируемых в области, содержащей отрезок, соединяющий точки $\mathbf{x}$ и $\mathbf{x}^0$. Можно сказать, что приращение $f(\mathbf{x})-f(\mathbf{x}^0)$ равно значению первого дифференциала функции f в некоторой точке интервала с концами в точках $\mathbf{x}$ и $\mathbf{x}^0$.

При $n = 2$ формула (12.6.2) имеет вид

$$f(\mathbf{x})=f(\mathbf{x}^0)+\sum_{j=1}^{m}\frac{\partial f}{\partial x_j}(\mathbf{x}^0)(x_j-x_j^0)+$$
$$+\frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 f}{\partial x_k\partial x_j}(\mathbf{x}^0+\theta(\mathbf{x}-\mathbf{x}^0))\times$$
$$\times(x_k-x_k^0)(x_j-x_j^0), \qquad 0<\theta<1. \quad (12.6.4)$$

Как и для функций одной переменной, из формулы Тейлора с остаточным членом в форме Лагранжа можно получить формулу Тейлора с остаточным членом в форме Пеано.

Покажем это для простоты на примере формулы (12.6.4). В главе 14 будет использован именно этот случай формулы Тейлора с остаточным членом в форме Пеано.

Если функция $f(\mathbf{x})$ в точке $\mathbf{x}^0$ дважды непрерывно дифференцируема, то

$$\frac{\partial^2 f}{\partial x_k\partial x_j}(\mathbf{x}^0+\theta(\mathbf{x}-\mathbf{x}^0))=\frac{\partial^2 f}{\partial x_k\partial x_j}(\mathbf{x}^0)+\varepsilon_{k,j}(\Delta\mathbf{x}),$$

где $\varepsilon_{k,j}(\Delta\mathbf{x}) \to 0$ при $|\Delta\mathbf{x}| \to 0$.

Подставив эти выражения в (12.6.4), получим

$$\begin{aligned} f(\mathbf{x}) = f(\mathbf{x}^0) + \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)(x_k - x_k^0) + \\ + \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_k \partial x_j}(\mathbf{x}^0)(x_k - x_k^0)(x_j - x_j^0) + \\ + \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m} \varepsilon_{k,j}(\Delta\mathbf{x})(x_k - x_k^0)(x_j - x_j^0). \end{aligned} \tag{12.6.5}$$

Оценим последнюю сумму в этой формуле. Если $\varepsilon(\Delta\mathbf{x}) := \max_{k,j} |\varepsilon_{k,j}(\Delta\mathbf{x})|$, то при $|\Delta\mathbf{x}| \to 0$ имеем $\varepsilon(\Delta\mathbf{x}) \to 0$ и

$$\begin{aligned} \left|\sum_{k=1}^{m}\sum_{j=1}^{m} \varepsilon_{k,j}(\Delta\mathbf{x})(x_k - x_k^0)(x_j - x_j^0)\right| \leqslant \sum_{k=1}^{m}\sum_{j=1}^{m} |\varepsilon_{k,j}(\Delta\mathbf{x})||\Delta\mathbf{x}||\Delta\mathbf{x}| \leqslant \\ \leqslant \varepsilon(\Delta\mathbf{x}) m^2 |\Delta\mathbf{x}|^2 = o(|\Delta\mathbf{x}|^2). \end{aligned}$$

Таким образом, установлено следующее предложение.

Теорема 12.6.2. *Если функция $f(\mathbf{x})$ дважды непрерывно дифференцируема в области пространства $\mathbb{E}^m$, содержащей отрезок с концами в точках $\mathbf{x}^0$ и $\mathbf{x}$, то справедлива оценка*

$$\begin{aligned} f(\mathbf{x}) = f(\mathbf{x}^0) + \sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(\mathbf{x}^0)(x_k - x_k^0) + \\ + \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_k \partial x_j}(\mathbf{x}^0)(x_k - x_k^0)(x_j - x_j^0) + \\ + o(|\Delta\mathbf{x}|^2), \qquad |\Delta\mathbf{x}| \to 0. \end{aligned} \tag{12.6.6}$$

Это и есть *формула Тейлора с остаточным членом в форме Пеано*, когда многочлен Тейлора является многочленом второй степени.

По аналогии с (12.5.15) формулу (12.6.6) можно записать в виде

$$f(\mathbf{x}) - f(\mathbf{x}^0) = \left(\sum_{k=1}^{m} (x_k - x_k^0)\frac{\partial}{\partial x_k}\right) f(\mathbf{x}^0) +$$

$$+\frac{1}{2}\left(\sum_{i=1}^{m}(x_k - x_k^0)\frac{\partial}{\partial x_k}\right)^2 f(\mathbf{x}^0) +$$
$$+ o(|\Delta\mathbf{x}|^2), \qquad |\Delta\mathbf{x}| \to 0.$$

Остаточный член формулы Тейлора в форме Пеано для многочленов Тейлора степени n имеет вид $o(|\Delta\mathbf{x}|^n)$. В этом можно убедиться, рассуждая аналогично выводу формулы (12.6.6).

Глава 13. Неявные функции

§ 13.1. Свойства функций, заданных неявно

До сих пор изучались функции вида $y = \varphi(\mathbf{x})$, когда каждой точке из области задания функции каким-либо способом (например, с помощью формулы) ставилось в соответствие число y – значение функции. В таких случаях функции считают заданными явно.

О неявном задании функций, короче о *неявных функциях*, говорят, когда переменные $\mathbf{x}$ и y связаны равенством $f(\mathbf{x}, y) = 0$ и y считается функцией от $\mathbf{x}$, т.е. когда уравнение $f(\mathbf{x}, y) = 0$ нужно решить относительно y. Понятно, что это не всегда возможно.

Поясним на простом примере постановку вопроса и характер ожидаемых результатов.

Пусть числа x и y связаны уравнением единичной окружности

$$x^2 + y^2 - 1 = 0. \tag{13.1.1}$$

Точкам $x \in [-1, 1]$ в силу (13.1.1) соответствуют значения $y = \pm\sqrt{1 - x^2}$. Так как знак $+$ или $-$ в каждой точке можно выбрать произвольно, уравнение (13.1.1) задаёт бесконечно много функций.

Если же рассматривать только непрерывные функции, то уравнение (13.1.1) определяет на $[-1, 1]$ две функции

$$y = \sqrt{1 - x^2} \tag{13.1.2}$$

и

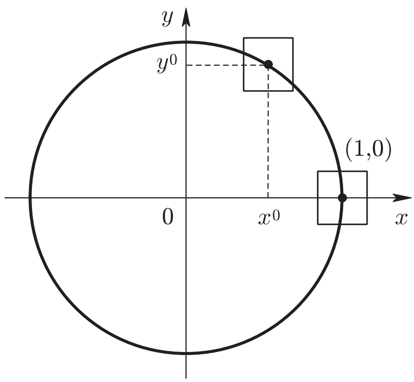
$$y = -\sqrt{1 - x^2}.$$

Графиком первой из них является верхняя полуокружность, а графиком второй – нижняя полуокружность. Таким образом, уравнение (13.1.1) не определяет на отрезке $[-1, 1]$ непрерывную функцию однозначно.

Перейдём от такой глобальной постановки вопроса к локальной и рассмотрим вопрос о разрешимости уравнения (13.1.1) не во всей полосе $|x| \leqslant 1$, а только в окрестности некоторой точки (x^0, y^0) окружности (13.1.1). Пусть для определённнсти эта точка лежит на верхней полуокружности.

Если $|x^0| < 1$, то в достаточно малой окрестности точки (x^0, y^0) уравнение (13.1.1) задаёт функцию формулой (13.1.2). Графиком этой функции является часть окружности, содержащаяся в рассматриваемой окрестности.

А если в качестве (x^0, y^0) взять точку $(1, 0)$, то ни в какой окрестности этой точки уравнение (13.1.1) не задаёт однозначную непрерывную функцию.



Заметим, что частная производная по y функции из левой части уравнения (13.1.1)

$$\frac{\partial}{\partial y}(x^2 + y^2 - 1) = 2y$$

в точках (x^0, y^0) при $|x^0| < 1$ не равна нулю, а в точке $(1, 0)$ эта производная обращается в нуль.

Отмеченные свойства функции (13.1.2), определяемой уравнением (13.1.1), характерны при неявном задании функций и в общем случае.

Будем рассматривать вопрос об y как функции m переменных $x_1, \ldots, x_m$, заданной неявно уравнением

$$f(x_1, \ldots, x_m, y) = 0. \tag{13.1.3}$$

Множество точек $(x_1, \ldots, x_m, y)$ пространства $\mathbb{E}^{m+1}$, координаты которых удовлетворяют уравнению (13.1.3), обозначим $\mathcal{M}$.

Выясним, при каких условиях в некоторой окрестности точки $(x_1^0, \ldots, x_m^0, y^0)$ множества $\mathcal{M}$ существует непрерывная функция $y = \varphi(x_1, \ldots, x_m)$, графиком которой являются все точки $\mathcal{M}$, содержащиеся в этой окрестности. Иначе говоря, в некоторой m-мерной окрестности точки $(x_1^0, \ldots, x_m^0)$ тождественно выполняется равенство

$$f(x_1, \ldots, x_m, \varphi(x_1, \ldots, x_m)) = 0.$$

Теорема 13.1.1. *Пусть в некоторой шаровой окрестности* $U \subset \mathbb{E}^{m+1}$ *точки* $(x_1^0, \ldots, x_m^0, y^0)$ *задана непрерывная функция* $f(x_1, \ldots, x_m, y)$, *которая в этой точке обращается в нуль и для каждой точки* $(x_1^*, \ldots, x_m^*)$ *проекции шара* U *на плоскость* $y = 0$ *функция* $f(x_1^*, \ldots, x_m^*, y)$ *переменной* y *строго возрастает.*

Тогда для каждого положительного числа ε *существуют такие положительные числа* ε_0 *и* δ, *что* $\varepsilon_0 < \varepsilon$ *и решением уравнения* (13.1.3) *относительно* y *в* $(m+1)$*-мерном прямоугольнике*

$$P_{m+1} := \{(x_1, \ldots, x_m, y) : |x_j - x_j^0| \leqslant \delta,\ j = 1, \ldots, m;\ |y - y^0| < \varepsilon_0\}$$

является функция $y = \varphi(x_1, \ldots, x_m)$, *непрерывная в* m*-мерном кубе*

$$K_m := \{(x_1, \ldots, x_m) : |x_j - x_j^0| \leqslant \delta,\ j = 1, \ldots, m\},$$

т.е. в точках $(x_1, \ldots, x_m)$ *этого куба справедливо равенство*

$$f(x_1, \ldots, x_m, \varphi(x_1, \ldots, x_m)) = 0. \tag{13.1.4}$$

Доказательство. По заданному ε выбираем положительное число $\varepsilon_0 < \varepsilon$ так, чтобы обе точки $(x_1^0, \ldots, x_m^0, y^0 + \varepsilon_0)$ и $(x_1^0, \ldots, x_m^0, y^0 - \varepsilon_0)$ лежали в шаре U.

Так как функция $f(x_1^0, \ldots, x_m^0, y)$ переменной y при $y = y^0$ обращается в нуль и на отрезке, соединяющем точки $(x_1^0, \ldots, x_m^0, y^0 - \varepsilon_0)$ и $(x_1^0, \ldots, x_m^0, y^0 + \varepsilon_0)$, строго возрастает, то

$$f(x_1^0, \ldots, x_m^0, y^0 + \varepsilon^0) > 0$$

и

$$f(x_1^0, \ldots, x_m^0, y^0 - \varepsilon^0) < 0.$$

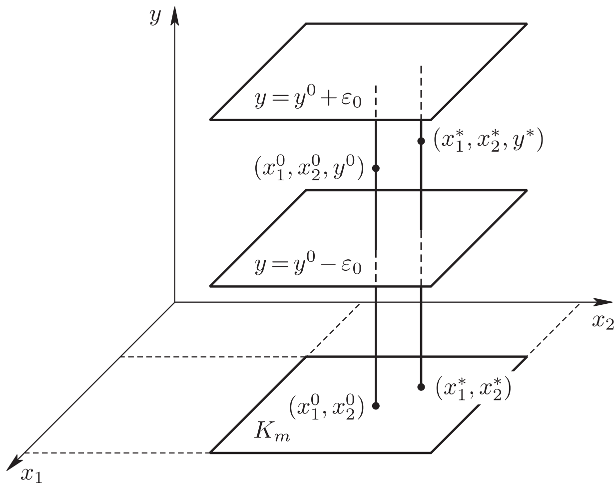
Поскольку функция $f(x_1, \ldots, x_m, y)$ в окрестности U непрерывна, существует такое положительное число δ, что для всех точек $(x_1, \ldots, x_m, y^0 + \varepsilon_0)$, принадлежащих пересечению гиперплоскости $y = y^0 + \varepsilon_0$ пространства $\mathbb{E}^{m+1}$ с U, первые m координат которых удовлетворяют условиям $|x_j - x_j^0| < \delta$, $j = 1, \ldots, m$, выполняется неравенство

$$f(x_1, \ldots, x_m, y^0 + \varepsilon_0) > 0$$

и для всех точек $(x_1, \ldots, x_m, y^0 - \varepsilon_0)$, принадлежащих пересечению гиперплоскости $y = y^0 - \varepsilon_0$ с U, при тех же условиях на x_j, $j = 1, \ldots, m$, выполняется неравенство

$$f(x_1, \ldots, x_m, y^0 - \varepsilon_0) < 0.$$

Таким образом получен $(m+1)$-мерный прямоугольник P_{m+1}, лежащий в окрестности U. Покажем, что в этом прямоугольнике выполняются утверждения теоремы.



Рассмотрим произвольную точку $(x_1^*, \ldots, x_m^*)$, принадлежащую кубу K_m. На отрезке, соединяющем точки $(x_1^*, \ldots, x_m^*, y^0 - \varepsilon_0)$ и $(x_1^*, \ldots, x_m^*, y^0 + \varepsilon_0)$, первая из которых принадлежит гиперплоскости $y = y^0 - \varepsilon_0$, а вторая – гиперплоскости $y = y^0 + \varepsilon_0$, функция $f(x_1^*, \ldots, x_m^*, y)$ строго возрастает, причём в нижнем конце этого отрезка она отрицательна, а в верхнем положительна. Значит, существует единственное значение переменной y, при котором функция $f(x_1^*, \ldots, x_m^*, y)$ равна нулю.

Обозначим это значение y^* и положим

$$\varphi(x_1^*, \ldots, x_m^*) := y^*.$$

Итак, в кубе K_m определена функция $\varphi(x_1, \ldots, x_m)$, для которой выполняется равенство (13.1.4). Докажем непрерывность функции φ.

Заметим сначала, что по построению функция $\varphi(x_1, \ldots, x_m)$ непрерывна в точке $(x_1^0, \ldots, x_m^0)$. В самом деле, для произвольного $\varepsilon_0 > 0$ было найдено $\delta > 0$ такое, что во всех точках куба K_m значения функции $\varphi(x_1, \ldots, x_m)$ заключены между $y^0 - \varepsilon_0$ и $y^0 + \varepsilon_0$, т.е. отличаются от $\varphi(x_1^0, \ldots, x_m^0)$ меньше, чем на ε_0.

Но в качестве исходной точки в проведенных рассуждениях можно было взять любую из построенных точек $(x_1^*, \ldots, x_m^*, y^*)$ прямоугольника P_{m+1}. Поэтому функция $\varphi(x_1, \ldots, x_m)$ непрерывна во всех точках куба K_m.

Теорема доказана.

Понятно, что в этой теореме строгое возрастание функции $f(x_1^*, \ldots, x_m^*, y)$ можно заменить на строгое убывание.

Покажем, что функция $\varphi(x_1, \ldots, x_m)$ из теоремы 13.1.1 при некоторых дополнительных условиях имеет частные производные.

Теорема 13.1.2. *Пусть функция $f(x_1, \ldots, x_m, y)$ удовлетворяет условиям теоремы 13.1.1. Если частные производные f по y и по некоторой переменной x_j непрерывны на множестве U и*

$$\frac{\partial f}{\partial y}(x_1^0, \ldots, x_m^0, y^0) \neq 0,$$

то функция $\varphi(x_1, \ldots, x_m)$, являющаяся решением уравнения (13.1.3), в достаточно малой окрестности каждой точки $\mathbf{x} \in K_m$ имеет непрерывную частную производную $\partial\varphi/\partial x_i$, для которой справедливо равенство

$$\frac{\partial \varphi}{\partial x_j}(\mathbf{x}) = -\frac{\partial f}{\partial x_j}(\mathbf{x}, \varphi(\mathbf{x})) \Big/ \frac{\partial f}{\partial y}(\mathbf{x}, \varphi(\mathbf{x})). \tag{13.1.5}$$

Доказательство. Для упрощения записи будем считать $f(x_1, \ldots, x_m, y)$ функцией только переменных x_j и y, так как все остальные аргументы f в проводимых рассуждениях будут оставаться фиксированными. Тогда функция φ зависит только от x_j.

В случае, когда f является функцией только переменных x_j и y, куб K_m представляет собой отрезок.

Будем придавать переменной x_j приращения Δx_j, при которых точки $x_j + \Delta x_j$ принадлежали K_m.

Если $y := \varphi(x_j)$ и $\Delta y := \varphi(x_j + \Delta x_j) - \varphi(x_j)$, то согласно определению функции φ имеем

$$f(x_j, y) = 0$$

и

$$f(x_j + \Delta x_j, y + \Delta y) = 0.$$

Значит,

$$f(x_j + \Delta x_j, y + \Delta y) - f(x_j, y) = 0.$$

Применив к этой разности формулу конечных приращений для функций многих переменных (12.6.3), находим, что при некотором $0 < \theta < 1$

$$\begin{aligned} f(x_j + \Delta x_j, y + \Delta y) - f(x_j, y) = \\ = \frac{\partial f}{\partial x_i}(x_j + \theta\Delta x_j, y + \theta\Delta y)\Delta x_i + \\ + \frac{\partial f}{\partial y}(x_j + \theta\Delta x_j, y + \theta\Delta y)\Delta y = 0. \end{aligned} \tag{13.1.6}$$

Но $\Delta y \to 0$ при $\Delta x_j \to 0$ в силу непрерывности функции $\varphi(x_j)$.

Так как частная производная функции f по y в рассматриваемой окрестности отлична от нуля и непрерывна, то из равенства (13.1.6) следует, что

$$\frac{\Delta y}{\Delta x_j} = -\frac{\partial f}{\partial x_j}(x_j + \theta\Delta x_j, y + \theta\Delta y) \Big/ \frac{\partial f}{\partial y}(x_j + \theta\Delta x_j, y + \theta\Delta y), \tag{13.1.7}$$

причём в правой части (13.1.7) можно перейти к пределу при $\Delta x_j \to 0$ отдельно в числителе и в знаменателе.

Таким образом,

$$\lim_{\Delta x_j \to 0} \frac{\Delta y}{\Delta x_j} = -\frac{\partial f}{\partial x_j}(x_j, y) \Big/ \frac{\partial f}{\partial y}(x_j, y).$$

Это показывает существование производной $\partial\varphi/\partial x_j$, её непрерывность и справедливость равенства (13.1.5).

Теорема доказана.

Формула (13.1.5) позволяет указать условия на функцию f, достаточные для существования частных производных функции φ второго и более высокого порядка.

Так, согласно теореме 12.3.1 о производных сложной функции из (13.1.5) следует, что функция $\varphi(\mathbf{x})$ имеет непрерывную чистую частную производную второго порядка по переменной x_j, если непрерывны частные производные

$$\frac{\partial^2 f}{\partial x_j^2}\,, \qquad \frac{\partial^2 f}{\partial x_j \partial y}\,, \qquad \frac{\partial^2 f}{\partial y^2}\,.$$

А если функция f имеет непрерывные частные производные второго порядка

$$\frac{\partial^2 f}{\partial x_k \partial x_j}\,, \qquad \frac{\partial^2 f}{\partial x_k \partial y}\,, \qquad \frac{\partial^2 f}{\partial x_j \partial y}\,, \qquad \frac{\partial^2 f}{\partial^2 y}$$

при $k \neq j$, то φ имеет непрерывную смешанную частную производную

$$\frac{\partial^2 \varphi}{\partial x_k \partial x_j}\,.$$

Проиллюстрируем применение полученных результатов на задаче об обратной функции одной переменной.

Рассмотрим функцию $y = F(x)$. Если значения независимой переменной обратной функции обозначать x, а значения зависимой переменной y, то нужно решить относительно y уравнение $x = F(y)$.

Запишем это уравнение в виде

$$x - F(y) = 0. \tag{13.1.8}$$

Согласно теоремам 13.1.1 и 13.1.2, если функция $F(y)$ и её производная dF/dy непрерывны в некоторой окрестности точки y^0, $x^0 := F(y^0)$ и $F'(y^0) \neq 0$, то существует окрестность точки (x^0, y^0), в которой уравнение (13.1.8) можно разрешить относительно y, т.е. это уравнение равносильно равенству $y = \varphi(x)$, где $\varphi(x)$ – непрерывная функция.

Так как функция из левой части уравнения (13.1.8) имеет непрерывную производную по x, то функция $\varphi(x)$ имеет непрерывную производную и согласно (13.1.5)

$$\varphi'(x^0) = -\frac{1}{-F'(\varphi(x^0))} = \frac{1}{F'(y^0)}\,.$$

Таким образом, теорема 5.3.1 получена как следствие теорем 13.1.1 и 13.1.2.

§ 13.2. Система неявных функций

В § 13.1 приведены достаточные условия разрешимости уравнения (13.1.3) относительно переменной y. Рассмотрим более общую задачу – о разрешимости системы n уравнений относительно n переменных.

Определение. Для функций n переменных $f_1(y_1,\dots,y_n),\dots,$ $f_n(y_1,\dots,y_n)$, имеющих в точке $\mathbf{y}$ частные производные первого порядка по всем переменным, определитель

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial t_1}(\mathbf{y}) & \dots & \dfrac{\partial f_1}{\partial t_n}(\mathbf{y}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial t_1}(\mathbf{y}) & \dots & \dfrac{\partial f_n}{\partial t_n}(\mathbf{y}) \end{vmatrix}$$

называют *якобианом* (в честь К. Якоби) и обозначают

$$\frac{\partial(f_1,\dots f_n)}{\partial(t_1,\dots,t_n)}(\mathbf{y}).$$

Это обозначение якобиана (как и обозначение частных производных) нужно рассматривать как единый символ, а не как дробь.

Теорема 13.2.1. *Пусть в шаровой окрестности* U, $U \subset \mathbb{E}^{m+n}$, *точки*

$$(\mathbf{x}^0;\mathbf{y}^0) = (x_1^0,\dots,x_m^0;\, y_1^0,\dots,y_n^0)$$

заданы n *функций*

$$f_k(x_k,\dots,x_m;\, y_1,\dots,y_n), \qquad k=1,\dots,n,$$

непрерывных вместе со всеми частными производными первого порядка по переменным y_j, $j=1,\dots,n$. *Если в точке* $(\mathbf{x}^0;\mathbf{y}^0)$ *выполняются условия*

$$f_k(\mathbf{x}^0;\mathbf{y}^0)=0, \quad k=1,\dots,n; \qquad \frac{\partial(f_1,\dots,f_n)}{\partial(y_1,\dots,y_n)}(\mathbf{x}^0;\mathbf{y}^0)\neq 0,$$

то для каждого числа $\varepsilon > 0$ существуют такие положительные числа ε_0 и δ, что $\varepsilon_0 < \varepsilon$ и в $(m+n)$-мерном прямоугольнике

$$\{(\mathbf{x};\mathbf{y}) : |x_k - x_k^o| < \delta,\ k = 1,\dots,m;\ |y_j - y_j^0| < \varepsilon_0,\ j = 1,\dots,n\},$$

принадлежащем окрестности U, система уравнений

$$f_p(x_1,\dots,x_m;y_1,\dots,y_n) = 0, \qquad p = 1,\dots,n, \tag{13.2.1}$$

относительно $y_1,\dots,y_n$ имеет решение, т.е. существуют такие непрерывные в кубе $|x_k - x_k^0| < \delta$, $k = 1,\dots,m$, функции $\varphi_1,\dots,\varphi_n$, что уравнения (13.2.1) равносильны равенствам

$$y_1 = \varphi_1(x_1,\dots,x_m), \quad \dots, \quad y_n = \varphi_n(x_1,\dots,x_m).$$

При этом если функции $f_1,\dots,f_n$ имеют в U непрерывные частные производные первого порядка по некоторой переменной x_j, то функции $\varphi_1,\dots,\varphi_n$ имеют непрерывные частные производные первого порядка по x_j.

Доказательство. Доказательство проведём индукцией по n – числу уравнений системы (13.2.1). При $n = 1$ утверждения теоремы 13.2.1 вытекают из теорем 13.1.1 и 13.1.2.

Предположим, что теорема 13.2.1 справедлива для системы из $n-1$, $n \geqslant 2$, уравнения, и докажем её для системы n уравнений.

В силу условий теоремы якобиан

$$\frac{\partial(f_1,\dots,f_n)}{\partial(y_1,\dots,y_n)}(\mathbf{x};\mathbf{y}) \tag{13.2.2}$$

является непрерывной функцией. Будем считать, что он не обращается в нуль не только в точке $(\mathbf{x}^0;\mathbf{y}^0)$, а во всей окрестности U.

Разложив определитель (13.2.2) по последнему столбцу, видим, что не равны нулю по крайней мере один элемент этого столбца и его алгебраическое дополнение. Уменьшив (в случае необходимости) окрестность U и изменив нумерацию функций $f_1,\dots,f_n$ и переменных $y_1,\dots,y_n$, будем, не теряя общности, считать, что в U не обращаются в нуль производная $\partial f_n/\partial y_n$ и её алгебраическое дополнение, равное якобиану

$$\frac{\partial(f_1,\dots,f_{n-1})}{\partial(y_1,\dots,y_{n-1})}(\mathbf{x};\mathbf{y}). \tag{13.2.3}$$

Рассмотрим систему первых $n-1$ уравнения системы (13.2.1)

$$f_p(x_1,\ldots,x_m;y_1,\ldots,y_n)=0, \qquad p=1,\ldots,n-1, \tag{13.2.4}$$

как систему относительно $y_1,\ldots,y_{n-1}$. Поскольку якобиан (13.2.3) не равен нулю, то по предположению индукции систему (13.2.4) можно разрешить относительно переменных $y_1,\ldots,y_{n-1}$, т.е. существуют непрерывные функции

$$y_1=\psi_1(x_1,\ldots,x_m,y_n), \quad \ldots, \quad y_{n-1}=\psi_{n-1}(x_1,\ldots,x_m,y_n) \tag{13.2.5}$$

и $(m+1)$-мерный куб

$$\{(x_1,\ldots,x_m,y_n): |x_k-x_k^0|<\delta,\ k=1,\ldots,m;\ |y_n-y_n^0|<\delta\} \tag{13.2.6}$$

такие, что в этом кубе выполняются тождества

$$f_p(x_1,\ldots,x_m;\psi_1(x_1,\ldots,x_m,y_n),\ldots,\psi_{n-1}(x_1,\ldots,x_m,y_n),y_n)\equiv 0, \\ p=1,\ldots,n-1. \tag{13.2.7}$$

При этом справедливы равенства

$$y_1^0=\psi_1(x_1^0,\ldots,x_m^0,y_n^0), \quad \ldots, \quad y_{n-1}^0=\psi_{n-1}(x_1^0,\ldots,x_m^0,y_n^0)$$

и функции $\psi_1,\ldots,\psi_{n-1}$ имеют в кубе (13.2.6) непрерывные частные производные по переменной y_n.

Последнее уравнение системы (13.2.1) означает, что в кубе (13.2.6) должно иметь место равенство

$$\begin{aligned}F(x_1,\ldots,x_m,y_n)&:=\\ :=f_n\bigl(x_1,\ldots,x_m;&\psi_1(x_1,\ldots,x_m,y_n),\ldots,\\ &\psi_{n-1}(x_1,\ldots,x_m,y_n),y_n\bigr)=0.\end{aligned}$$

Докажем, что уравнение $F(x_1,\ldots,x_m,y_n)=0$ можно разрешить относительно y_n.

Понятно, что в кубе (13.2.6) функция $F(x_1,\ldots,x_m,y_n)$ и её производная по y_n непрерывны и $F(x_1^0,\ldots,x_m^0,y_n^0)=0$. Чтобы воспользоваться теоремой 13.1.2, покажем, что

$$\frac{\partial F}{\partial y_n}(x_1^0,\ldots,x_m^0,y_n^0)\neq 0.$$

Частные производные функций из (13.2.7) по переменной y_n непрерывны в кубе $|x_k - x_k^0| < \delta$, $k = 1, \ldots, m$, и вычислив эти производные, находим, что

$$\frac{\partial f_p}{\partial y_1}\frac{\partial \psi_1}{\partial y_n} + \cdots + \frac{\partial f_p}{\partial y_{n-1}}\frac{\partial \psi_{n-1}}{\partial y_n} + \frac{\partial f_p}{\partial y_n} = 0, \qquad p = 1, \ldots, n-1. \tag{13.2.8}$$

Кроме того,

$$\frac{\partial F}{\partial y_n} = \frac{\partial f_n}{\partial y_1}\frac{\partial \psi_1}{\partial y_n} + \cdots + \frac{\partial f_n}{\partial y_{n-1}}\frac{\partial \psi_{n-1}}{\partial y_n} + \frac{\partial f_n}{\partial y_n}\,. \tag{13.2.9}$$

Если к последнему столбцу якобиана

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial y_1} & \ldots & \dfrac{\partial f_1}{\partial y_{n-1}} & \dfrac{\partial f_1}{\partial y_n} \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial f_n}{\partial y_1} & \ldots & \dfrac{\partial f_n}{\partial y_{n-1}} & \dfrac{\partial f_n}{\partial y_n} \end{vmatrix}$$

прибавить сумму первого столбца, умноженного на $\partial\psi_1/\partial y_n$, второго, умноженного на $\partial\psi_2/\partial y_n$, и т.д., $(n-1)$-го столбца, умноженного на $\partial\psi_{n-1}/\partial y_n$, то в силу (13.2.8) и (13.2.9) получим

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial y_1} & \ldots & \dfrac{\partial f_1}{\partial y_{n-1}} & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial f_{n-1}}{\partial y_1} & \ldots & \dfrac{\partial f_{n-1}}{\partial y_{n-1}} & 0 \\ \dfrac{\partial f_n}{\partial y_1} & \ldots & \dfrac{\partial f_n}{\partial y_{n-1}} & \dfrac{\partial F}{\partial y_n} \end{vmatrix} = \frac{\partial(f_1, \ldots, f_n)}{\partial(y_1, \ldots, y_n)}\,.$$

Так как этот якобиан не равен нулю, то $\partial F/\partial y_n \neq 0$.

Итак, для уравнения

$$F(x_1, \ldots, x_m, y_n) = 0 \tag{13.2.10}$$

выполнены все условия теоремы 13.1.2. Решив это уравнение относительно y_n, получим

$$y_n = \varphi_n(x_1, \ldots, x_m), \tag{13.2.11}$$

где функция φ_n непрерывна.

С помощью функции φ_n определим функции

$$\varphi_1(x_1,\ldots,x_m) := \psi_1(x_1,\ldots,x_m,\varphi_n(x_1,\ldots,x_m)),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\varphi_{n-1}(x_1,\ldots,x_m) := \psi_{n-1}(x_1,\ldots,x_m,\varphi_n(x_1,\ldots,x_m)),$$

непрерывные в достаточно малой окрестности точки $(x_1^0,\ldots,x_m^0)$.

Функции

$$y_1 = \varphi_1(x_1,\ldots,x_m), \quad \ldots,$$
$$y_{n-1} = \varphi_{n-1}(x_1,\ldots,x_m), \quad y_n = \varphi_n(x_1,\ldots,x_m)$$

в достаточно малой окрестности точки $(\mathbf{x}^0;\mathbf{y}^0)$ являются решением системы (13.2.1).

Осталось убедиться в том, что каждая функция $\varphi_1,\ldots,\varphi_n$ имеет непрерывную частную производную первого порядка по переменной x_j, если все функции $f_1,\ldots,f_n$ системы (13.2.1) имели непрерывные частные производные первого порядка по x_j.

Согласно предположению индукции таким свойством обладают функции $\psi_1,\ldots,\psi_{n-1}$ из (13.2.5). Отсюда в силу свойств сложных функций следует, что функция F имеет непрерывную частную производную по x_j. Значит, согласно теореме 13.1.2 функция φ_n имеет непрерывную частную производную по x_j. Поэтому непрерывные частные производные по переменной x_j имеют и функции $\varphi_1,\ldots,\varphi_{n-1}$.

Теорема доказана.

Глава 14. Экстремумы функций многих переменных

§ 14.1. Локальные экстремумы

Определение экстремумов функций многих переменных не отличается от соответствующих определений для функций одной переменной.

Определение. Пусть функция $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ задана в окрестности точки $\mathbf{x}^0$. Если для всех $\mathbf{x}$ из некоторой окрестности точки $\mathbf{x}^0$ справедлива оценка $f(\mathbf{x}) \leqslant f(\mathbf{x}^0)$, то говорят, что f имеет в точке $\mathbf{x}^0$ *локальный максимум*.

А если для всех $\mathbf{x}$ из некоторой проколотой окрестности точки $\mathbf{x}^0$ справедлива оценка $f(\mathbf{x}) < f(\mathbf{x}^0)$, то говорят: f имеет в точке $\mathbf{x}^0$ *строгий локальный максимум*.

Аналогично вводятся локальный минимум и строгий локальный минимум.

Точки, в которых функция имеет локальный максимум или локальный минимум, называют точками её *локального экстремума* и точками *строгого локального экстремума*, если максимум или минимум является строгим.

Слово "локальный" здесь показывает, что число $f(\mathbf{x}^0)$ сравнивается со значениями функции $f(\mathbf{x})$ только в некоторой окрестности точки $\mathbf{x}^0$, а не во всей области определения f.

В этой главе будут получены необходимые, а также достаточные условия существования локального экстремума функции в точке, в которой эта функция дифференцируема или имеет частные производные.

Сначала рассмотрим необходимые условия.

Пусть функция f имеет в точке $\mathbf{x}^0$ локальный экстремум. Зафиксируем значения всех переменных $x_1, \ldots, x_m$, кроме какой-либо одной, например x_k, равными их значениям в точке $\mathbf{x}^0$. Тогда полученная функция переменной x_k будет иметь в точке x_k^0 соответствующий локальный экстремум.

Поэтому из теоремы Ферма 6.1.2 для функций одной переменной вытекает следующее утверждение.

Теорема 14.1.1. *Если функция $f(\mathbf{x})$ имеет в точке локального экстремума $\mathbf{x}^0$ частные производные первого порядка по всем переменным, то все эти производные обращаются в точке $\mathbf{x}^0$ в нуль.*

Следовательно, для функций, дифференцируемых в точке, необходимым условием существования локального экстремума в этой точке является равенство первого дифференциала нулю.

Условие $df = 0$ можно записать так: $\operatorname{grad} f = \overline{0}$.

Определение. Точки, в которых функция f дифференцируема и $df = 0$, называют *стационарными точками* функции f.

Таким образом, точки локальных экстремумов дифференцируемых функций являются их стационарными точками.

Это условие не является достаточным для существования локального экстремума, что видно уже на функциях одной переменной.

Рассмотрим вопрос о локальных экстремумах функций, имеющих частные производные второго порядка.

В дальнейшем, до конца этого параграфа будем считать, что $\mathbf{x}^0$ является стационарной точкой функции $f(\mathbf{x})$ и в точке $\mathbf{x}^0$ непрерывны все частные производные второго порядка функции f.

Для сокращения записей введём обозначение

$$f_{kj} := \frac{\partial^2 f}{\partial x_k \partial x_j}(\mathbf{x}^0), \qquad k,\, j = 1, \ldots, m.$$

Если f дважды непрерывно дифференцируема в некоторой окрестности стационарной точки $\mathbf{x}^0$, то согласно формуле Тейлора с остаточным членом в форме Пеано (12.6.6)

$$f(\mathbf{x}) = f(\mathbf{x}^0) + \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m} f_{kj}(x_k - x_k^0)(x_j - x_j^0) + o(|\Delta\mathbf{x}|^2), \quad |\Delta\mathbf{x}| \to 0, \tag{14.1.1}$$

где как обычно,

$$|\Delta\mathbf{x}| = \sqrt{\sum_{k=1}^{m}(x_k - x_k^0)^2}.$$

Положим

$$a_k := \frac{x_k - x_k^0}{|\Delta\mathbf{x}|}, \qquad k = 1, \ldots, m.$$

В силу равенства

$$\sum_{k=1}^{m} a_k^2 = \sum_{k=1}^{m} \frac{(x_k - x_k^0)^2}{|\Delta\mathbf{x}|^2} = 1 \tag{14.1.2}$$

числа a_k представляют собой направляющие косинусы единичного вектора, задающего луч с вершиной в точке $\mathbf{x}^0$, на котором лежит точка $\mathbf{x}$. Обозначим этот вектор $\overline{a}$.

Из (14.1.1) получаем

$$f(\mathbf{x}) - f(\mathbf{x}^0) = \frac{1}{2}|\Delta\mathbf{x}|^2 \left\{ \sum_{k=1}^{m}\sum_{j=1}^{m} f_{kj} a_i k a_j + \varepsilon(\Delta\mathbf{x}) \right\}, \tag{14.1.3}$$

где

$$\lim_{|\Delta\mathbf{x}|\to 0} \varepsilon(\Delta\mathbf{x}) = 0. \tag{14.1.4}$$

Введём обозначение квадратичной формы из правой части (14.1.3):

$$\varPhi(\overline{a}) := \sum_{k=1}^{m}\sum_{j=1}^{m} f_{kj} a_k a_j.$$

Для заданной функции f значения квадратичной формы $\varPhi(\overline{a})$ зависят только от направляющих косинусов единичного вектора луча, о котором говорилось выше, и значит, во всех точках $\mathbf{x}$ этого луча функция $\varPhi(\overline{a})$ принимает одно и то же значение.

Чтобы выяснить, имеет ли функция $f(\mathbf{x})$ локальный экстремум в точке $\mathbf{x}^0$, нужно рассмотреть знак разности $f(\mathbf{x}) - f(\mathbf{x}^0)$ в достаточно малой окрестности точки $\mathbf{x}^0$. Согласно (14.1.3) знак этой разности тот же, что и знак выражения

$$\varPhi(\overline{a}) + \varepsilon(\Delta\mathbf{x}). \tag{14.1.5}$$

Поскольку для $\varepsilon(\Delta\mathbf{x})$ выполняется равенство (14.1.4), f не может иметь в точке $\mathbf{x}^0$ локальный максимум, если функция $\varPhi(\overline{a})$ для некоторого вектора $\overline{a}$ принимает положительное значение. А если $\varPhi(\overline{a})$ может принимать отрицательные значения, у f в точке $\mathbf{x}^0$ нет локального минимума.

Таким образом, имеет место следующее утверждение.

Теорема 14.1.2. *Если функция $f(\mathbf{x})$ дважды непрерывно дифференцируема в некоторой окрестности своей стационарной точки $\mathbf{x}^0$, то*

1°) *чтобы f имела в точке $\mathbf{x}^0$ локальный максимум, необходимо условие $\Phi(\overline{a}) \leqslant 0$;*
2°) *чтобы f имела в точке $\mathbf{x}^0$ локальный минимум, необходимо условие $\Phi(\overline{a}) \geqslant 0$;*
3°) *если квадратичная форма $\Phi(\overline{a})$ принимает и положительные и отрицательные значения, функция f в точке $\mathbf{x}^0$ локального экстремума не имеет.*

Получим теперь достаточные условия существования локального экстремума.

Пусть квадратичная форма $\Phi(\overline{a})$ положительно определена, т.е. принимает положительные значения для любого ненулевого вектора $\overline{a}$.

Сфера, заданная в $\mathbb{R}^m$ формулой (14.1.2), является замкнутым ограниченным множеством. Так как функция $\Phi(\overline{a})$ непрерывна, она принимает в некоторой точке сферы своё минимальное на ней значение.

Значит, существует положительное число α такое, что для всех $a_1, \ldots, a_m$, удовлетворяющих условию (14.1.2), справедлива оценка

$$\Phi(\overline{a}) \geqslant \alpha.$$

В силу (14.1.4) для достаточно малых $|\Delta\mathbf{x}|$

$$|\varepsilon(\Delta\mathbf{x})| < \frac{\alpha}{2}\,.$$

Поэтому при таких $\Delta\mathbf{x}$ знак выражения (14.1.5) определяется знаком $\Phi(\overline{a})$, т.е. это выражение принимает только положительные значения.

Таким образом, согласно (14.1.3) в достаточно малой проколотой окрестности точки $\mathbf{x}^0$ выполняется неравенство

$$f(\mathbf{x}) - f(\mathbf{x}^0) > 0$$

и, следовательно, функция $f(\mathbf{x})$ имеет в точке $\mathbf{x}^0$ строгий локальный минимум. Напомним, что $\mathbf{x}^0$ предполагается стационарной точкой.

Точно также устанавливается, что $f(\mathbf{x})$ имеет в точке $\mathbf{x}^0$ строгий локальный максимум, если квадратичная форма $\Phi(\overline{a})$ отрицательно определена, т.е. принимает отрицательные значения для всех ненулевых векторов $\overline{a}$.

Запишем полученные результаты в виде следующего утверждения.

Теорема 14.1.3. *Пусть функция $f(\mathbf{x})$ дважды непрерывно дифференцируема в некоторой окрестности своей стационарной точки $\mathbf{x}^0$. Тогда*

1°) *если квадратичная форма $\Phi(\overline{a})$ положительно определена, функция $f(\mathbf{x})$ имеет в точке $\mathbf{x}^0$ строгий локальный минимум;*
2°) *если квадратичная форма $\Phi(\overline{a})$ отрицательно определена, функция $f(\mathbf{x})$ имеет в точке $\mathbf{x}^0$ строгий локальный максимум.*

Если же квадратичная форма $\Phi(\overline{a})$ принимает и положительные и отрицательные значения, то f не имеет в точке $\mathbf{x}^0$ локального экстремума.

Обсудим теперь случай, когда квадратичная форма $\Phi(\overline{a})$ является полуположительно определённой (нестрого положительно определённой), т.е. $\Phi(\overline{a}) \geqslant 0$ для всех векторов $\overline{a}$ и существует ненулевой вектор $\overline{a}^0$, для которого $\Phi(\overline{a}^0) = 0$.

Тогда согласно (14.1.3) для точек $\mathbf{x}$, лежащих на луче с вершиной в $\mathbf{x}^0$ и направляющим вектором $\overline{a}^0$, имеем

$$f(\mathbf{x}) - f(\mathbf{x}^0) = \frac{1}{2}|\Delta\mathbf{x}|^2\varepsilon(\Delta\mathbf{x}).$$

Так как о знаке величины $\varepsilon(\Delta\mathbf{x})$ ничего не известно, нельзя сделать вывод ни о наличии, ни об отсутствии локального экстремума функции f в точке $\mathbf{x}^0$. Достаточно рассмотреть в качестве примера функции $x^4 + y^4$ и $x^4 - y^4$ в точке $(0, 0)$.

Случай, когда квадратичная форма $\Phi(\overline{a})$ является полузнакоопределённой, можно исследовать, используя частные производные более высокого порядка (если она существуют), наподобие того, как это делалось для функций одной переменной в § 6.6. Но не будем на этом останавливаться.

Заметим, что в доказательстве теоремы 14.1.3 непрерывность частных производных функции f сама по себе не использовалась. Нужно было только представление (14.1.1) приращения функции

f в точке $\mathbf{x}^0$. Числа f_{kj} в формуле (14.1.1) можно было бы считать частными производными функции f в смысле Пеано. О производных в смысле Пеано функций одной переменной говорилось в § 6.4.

§ 14.2. Условный локальный экстремум

В § 14.1 рассматривались локальные экстремумы функций, когда значение функции в точке $\mathbf{x}^0$ сравнивалось с её значениями во всех точках из достаточно малой окрестности $\mathbf{x}^0$.

Изучаются также локальные экстремумы, когда значение функции в точке сравнивается со значениями не во всей окрестности этой точки, а только в точках, удовлетворяющих каким-либо дополнительным условиям. В таких случаях говорят об *условном экстремуме*, или относительном экстремуме, или экстремуме при наличии связей.

Будем рассматривать вопрос об условном экстремуме функций многих переменных в следующей постановке.

В некоторой окрестности точки $\mathbf{x}^0 \in \mathbb{E}^m$ задана функция $f(\mathbf{x})$, которую будем исследовать на экстремум, и ещё n функций $g_1(\mathbf{x})$, $\ldots, g_n(\mathbf{x})$, где $n < m$. Пусть A – множество точек этой окрестности, в которых справедливы равенства

$$g_1(\mathbf{x}) = 0, \quad \ldots, \quad g_n(\mathbf{x}) = 0, \tag{14.2.1}$$

и точка $\mathbf{x}^0 \in A$.

Определение. Функция $f(\mathbf{x})$ имеет в точке $\mathbf{x}^0 \in A$ *локальный максимум при наличии связей* (14.2.1), если для всех точек $\mathbf{x}$ множества A, из некоторой окрестности точки $\mathbf{x}^0$ выполняется неравенство $f(\mathbf{x}) \leqslant f(\mathbf{x}^0)$.

Условный локальный максимум называют *строгим*, если для всех точек множества A из некоторой проколотой окрестности точки $\mathbf{x}^0$, справедливо неравенство $f(\mathbf{x}) < f(\mathbf{x}^0)$.

Определения условного локального минимума и строгого условного локального минимума аналогичны.

Когда ясно, какие связи имеются в виду, упоминание о связях обычно опускают и говорят просто об условном экстремуме.

Точки, в которых достигаются условный локальный максимум или минимум, называют точками *условного локального экстремума* и, соответственно, точками *строгого условного локального экстремума*.

Рассматривавшиеся в § 14.1 экстремумы называют безусловными.

Будем предполагать, что и функция $f(\mathbf{x})$ и функции $g_1(\mathbf{x}), \ldots, g_n(\mathbf{x})$ в некоторой окрестности точки $\mathbf{x}^0$ имеют непрерывные частные производные первого порядка по всем переменным $x_1, \ldots, x_m$ и ранг матрицы

$$\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \ldots & \dfrac{\partial g_1}{\partial x_m} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial g_n}{\partial x_1} & \ldots & \dfrac{\partial g_n}{\partial x_m} \end{pmatrix} \tag{14.2.2}$$

в точке $\mathbf{x}^0$ равен n. Тогда ранг этой матрицы равен n и в некоторой окрестности точки $\mathbf{x}^0$.

Для определённости будем считать, что в окрестности точки $\mathbf{x}^0$ отличен от нуля минор матрицы (14.2.2), построенный на последних n столбцах, т.е. что

$$\begin{vmatrix} \dfrac{\partial g_1}{\partial x_{m-n+1}} & \ldots & \dfrac{\partial g_1}{\partial x_m} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial g_n}{\partial x_{m-n+1}} & \ldots & \dfrac{\partial g_n}{\partial x_m} \end{vmatrix} \neq 0. \tag{14.2.3}$$

Согласно теореме 13.2.1 систему уравнений (14.2.1) в достаточно малой окрестности точки $\mathbf{x}^0$ можно разрешить относительно последних n переменных, т.е. существуют такие непрерывные функции $\varphi_{m-n+1}, \ldots, \varphi_m$, что

$$x_{m-n+1} = \varphi_{m-n+1}(x_1, \ldots, x_{m-n}), \quad \ldots, \quad x_m = \varphi_m(x_1, \ldots, x_{m-n}). \tag{14.2.4}$$

Соотношения (14.2.1) и (14.2.4) равносильны. При этом функции $\varphi_{m-n+1}, \ldots, \varphi_m$ в некоторой окрестности точки $(x_1^0, \ldots, x_{m-n}^0)$ имеют непрерывные частные производные первого порядка по всем переменным $x_1, \ldots, x_{m-n}$.

В формулах (14.2.4) переменные $x_1, \dots, x_{m-n}$ являются независимыми.

Существование экстремума функции $f(\mathbf{x})$ при наличии связей (14.2.1) равносильно существованию безусловного экстремума функции

$$h(x_1, \dots, x_{m-n}) := \\ := f\big(x_1, \dots, x_{m-n}, \varphi_{m-n+1}(x_1, \dots, x_{m-n}), \dots, \\ \varphi_m(x_1, \dots, x_{m-n})\big). \qquad (14.2.5)$$

Функция $h(x_1, \dots, x_{m-n})$ имеет непрерывные частные производные первого порядка по всем переменным. Согласно теореме 14.1.1 для существования локального экстремума функции h в точке $(x_1^0, \dots, x_{m-n}^0)$ необходимо условие $dh(x_1^0, \dots, x_{m-n}^0) = 0$ или

$$\frac{\partial h}{\partial x_k}(x_1^0, \dots, x_{m-n}^0) = 0, \qquad k = 1, \dots, m-n, \qquad (14.2.6)$$

так как переменные $x_1, \dots, x_{m-n}$ являются независимыми.

Таким образом, получено необходимое условие существования условного локального экстремума.

Однако, условие (14.2.6) не являются эффективным, так как в определении функции h участвуют функции $\varphi_{m-n+1}, \dots, \varphi_m$, полученные при переходе от уравнений связей (14.2.1) к равенствам (14.2.4). Такой переход в явном виде в большинстве случаев является трудноразрешимой задачей.

Поэтому будем искать условия, равносильные (14.2.6), но явным образом выраженные через функции f и $g_1, \dots, g_n$.

Условие $dh(x_1, \dots, x_{m-n}) = 0$ в силу инвариантности формы первого дифференциала можно записать в виде равенства

$$\sum_{k=1}^{m} \frac{\partial f}{\partial x_k}(x_1^0, \dots, x_m^0)\, dx_k = 0, \qquad (14.2.7)$$

в котором $dx_1, \dots, dx_{m-n}$ являются дифференциалами независимых переменных, а дифференциалы $dx_{m-n+1}, \dots, dx_m$ можно выразить через них, используя уравнения связей (14.2.1).

Подставим представления (14.2.4) переменных $x_{m-n+1}, \dots, x_m$ в (14.2.1). Получим n тождеств:

$$g_1\big(x_1, \dots, x_{m-n}, \varphi_{m-n+1}(x_1, \dots, x_{m-n}), \dots, \varphi_m(x_1, \dots, x_{m-n})\big) \equiv 0,$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$g_n\big(x_1, \dots, x_{m-n}, \varphi_{m-n+1}(x_1, \dots, x_{m-n}), \dots, \varphi_m(x_1, \dots, x_{m-n})\big) \equiv 0,$$

функции переменных $x_1, \dots, x_{m-n}$ в которых дифференцируемы. Поэтому

$$\sum_{k=1}^{m} \frac{\partial g_1}{\partial x_k}(\mathbf{x}^0)\, dx_k = 0, \qquad \dots, \qquad \sum_{k=1}^{m} \frac{\partial g_n}{\partial x_k}(\mathbf{x}^0)\, dx_k = 0. \qquad (14.2.8)$$

Таким образом, для того чтобы функция $f(\mathbf{x})$ имела в точке $\mathbf{x}^0$ условный экстремум при наличии связей (14.2.1), необходимо чтобы равенство (14.2.7) было справедливо для всех дифференциалов $dx_1, \dots, dx_m$, для которых выполняются соотношения (14.2.8).

Это необходимое условие можно выразить в терминах градиентов.

Рассмотрим две системы линейных уравнений относительно $dx_1, \dots, dx_m$. Первую систему составляют уравнения (14.2.8), а вторая получена добавлением к (14.2.8) уравнения (14.2.7).

Полученное выше необходимое условие означает, что каждое решение первой системы должно быть решением второй. Для этого последняя строка матрицы второй системы

$$\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}(\mathbf{x}) & \dots & \dfrac{\partial g_1}{\partial x_m}(\mathbf{x}) \\ \vdots & \vdots & \vdots \\ \dfrac{\partial g_n}{\partial x_1}(\mathbf{x}) & \dots & \dfrac{\partial g_n}{\partial x_m}(\mathbf{x}) \\ \dfrac{\partial f}{\partial x_1}(\mathbf{x}) & \dots & \dfrac{\partial f}{\partial x_m}(\mathbf{x}) \end{pmatrix}$$

должна быть линейной комбинацией первых n строк этой матрицы, т.е. должны существовать такие числа $\lambda_1, \dots, \lambda_n$, что при всех $j = 1, \dots, m$

$$\frac{\partial f}{\partial x_j}(\mathbf{x}^0) = \sum_{k=1}^{n} \lambda_k \frac{\partial g_k}{\partial x_j}(\mathbf{x}^0)$$

или в терминах градиентов

$$\operatorname{grad} f(\mathbf{x}^0) = \sum_{k=1}^{n} \lambda_k \operatorname{grad} g_k(\mathbf{x}^0). \qquad (14.2.9)$$

Таким образом, установленное выше необходимое условие относительного экстремума состоит в том, что градиент $\operatorname{grad} f(\mathbf{x}^0)$ можно представить в виде линейной комбинации градиентов $\operatorname{grad} g_1(\mathbf{x}^0), \ldots, \operatorname{grad} g_n(\mathbf{x}^0)$.

§ 14.3. Метод неопределённых множителей Лагранжа

При исследовании условных экстремумов часто бывает удобен следующий приём, восходящий к Лагранжу.

Будем пользоваться обозначениями из § 14.2 – требуется найти экстремум функции $f(x_1, \ldots, x_m)$ при наличии связей (14.2.1), для которых выполнено условие (14.2.3).

Возьмём n произвольных пока чисел $\lambda_1, \ldots, \lambda_n$ и составим с их помощью *функцию Лагранжа*

$$F(x_1, \ldots, x_m; \lambda_1, \ldots, \lambda_n) := f(x_1, \ldots, x_m) - \sum_{l=1}^{n} \lambda_l g_l(x_1, \ldots, x_m). \tag{14.3.1}$$

Числа $\lambda_1, \ldots, \lambda_n$ называют *неопределёнными множителями Лагранжа*.

Рассмотрим задачу на безусловный экстремум функции F как функции переменных $x_1, \ldots, x_m$ и проследим, что это даёт для задачи на условный экстремум для функции f.

Согласно теореме 14.1.1 для существования безусловного экстремума функции F в точке $\mathbf{x}^0$ необходимо, чтобы в этой точке выполнялись m равенств

$$\frac{\partial F}{\partial x_1}(\mathbf{x}^0) = 0, \quad \ldots, \quad \frac{\partial F}{\partial x_m}(\mathbf{x}^0) = 0. \tag{14.3.2}$$

Равенства (14.3.2) вместе с(14.2.1) представляют собой систему $m+n$ уравнений, с помощью которых нужно определить значения $m+n$ неизвестных величин: $x_1^0, \ldots, x_m^0$ и $\lambda_1^0, \ldots, \lambda_n^0$.

Заметим, что если искать экстремум функции F как функции $m+n$ переменных $x_1, \ldots, x_m$, $\lambda_1, \ldots, \lambda_n$, то наряду с (14.3.2) получим

$$\frac{\partial F}{\partial \lambda_1} = 0, \quad \ldots, \quad \frac{\partial F}{\partial \lambda_n} = 0,$$

т.е. уравнения связей (14.2.1).

Необходимое условие (14.3.2) на функцию Лагранжа равносильно необходимому условию (14.2.9). В самом деле, уравнения (14.3.2)

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^0)-\sum_{l=1}^{n}\lambda_l\frac{\partial g_l}{\partial x_1}(\mathbf{x}^0)=0,\quad \ldots,\quad \frac{\partial f}{\partial x_m}(\mathbf{x}^0)-\sum_{l=1}^{n}\lambda_l\frac{\partial g_l}{\partial x_m}(\mathbf{x}^0)=0$$

являются координатной записью векторного равенства

$$\operatorname{grad} f(\mathbf{x}^0)-\sum_{l=1}^{n}\lambda_l\operatorname{grad} g_l(\mathbf{x}^0)=\overline{0}, \tag{14.3.3}$$

т.е. равенства (14.2.9).

Итак, необходимые условия безусловного экстремума функции Лагранжа F привели к необходимому условию в задаче на экстремум исходной функции f при наличии связей (14.2.1).

Получим теперь достаточные условия относительного экстремума в терминах функции Лагранжа. Будем считать, что в точке $\mathbf{x}^0$ выполняется необходимое условие (14.3.3) и в качестве множителей Лагранжа $\lambda_1,\ldots,\lambda_n$ в F взяты те числа, при которых выполнено равенство (14.3.3). Поэтому функцию Лагранжа F можно считать зависящей только от $x_1,\ldots,x_m$.

Предположим, что все частные производные второго порядка функций $f(\mathbf{x})$, $g_1(\mathbf{x}),\ldots,g_n(\mathbf{x})$ непрерывны в некоторой окрестности точки $\mathbf{x}^0$.

Сейчас нас интересуют только те точки $\mathbf{x}$, в которых выполнены уравнения связей (14.2.1), а в этих точках

$$F(x_1,\ldots,x_m)=f(x_1,\ldots,x_m). \tag{14.3.4}$$

Будем считать $(x_1^0,\ldots,x_m^0)$ стационарной точкой функции F.

В § 14.2 показано, что задача об относительном экстремуме функции f при условии (14.2.3) на связи равносильна задаче о безусловном экстремуме функции $h(x_1,\ldots,x_{m-n})$, заданной формулой (14.2.5).

Достаточные условия экстремума функции h можно выразить в терминах квадратичной формы d^2h.

Покажем, что для точек $(x_1,\ldots,x_m)$, удовлетворяющих уравнениям связей (14.2.1), второй дифференциал функции h равен второму дифференциалу функции F, когда переменные x_{m-n+1}, $\ldots,x_m$ выражены через $x_1,\ldots,x_{m-n}$, а дифференциалы dx_{m-n+1},

$\ldots, dx_m$ выражены через $dx_1, \ldots, dx_{n-m}$ с помощью соотношений связи.

В самом деле,

$$\begin{aligned} d^2h(x_1,\ldots,x_{m-n}) &= \\ &= d^2f\big(x_1,\ldots,x_{m-n},\varphi_{m-n+1}(x_1,\ldots,x_{m-n}),\ldots, \\ &\qquad\qquad \varphi_m(x_1,\ldots,x_{m-n})\big) = \\ &= \sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 f}{\partial x_k\partial x_j}(x_1,\ldots,x_m)\,dx_k\,dx_j + \\ &\qquad + \sum_{j=1}^{m}\frac{\partial f}{\partial x_j}(x_1,\ldots,x_m)\,d^2x_j. \end{aligned} \tag{14.3.5}$$

В точках $(x_1, \ldots, x_m)$, удовлетворяющих уравнениям (14.2.1), имеем

$$\sum_{j=1}^{m}\frac{\partial g_l}{\partial x_j}(x_1,\ldots,x_m)dx_j = 0, \qquad l = 1,\ldots n,$$

и значит,

$$\sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 g_l}{\partial x_k\partial x_j}(x_1,\ldots,x_m)dx_kdx_k + \sum_{j=1}^{m}\frac{\partial g_l}{\partial x_j}(x_1,\ldots,x_m)d^2x_j = 0. \tag{14.3.6}$$

При каждом l умножим равенство из (14.2.6) на λ_l и сложим эти произведения по l от 1 до n. Вычитая получившееся равенство из правой части (14.3.5), находим

$$\begin{aligned} d^2h(x_1,\ldots,x_{m-n}) &= \\ &= \sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 F}{\partial x_k\partial x_j}(x_1,\ldots,x_m)\,dx_k\,dx_j + \\ &\qquad + \sum_{j=1}^{m}\frac{\partial F}{\partial x_j}(x_1,\ldots,x_m)\,d^2x_j. \end{aligned} \tag{14.3.7}$$

В стационарной точке $(x_1^0, \ldots, x_m^0)$ функции F последняя сумма в правой части (14.3.7) равна нулю. Следовательно,

$$d^2\,h(x_1^0,\ldots,x_{m-n}^0) = \sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 F}{\partial x_k\partial x_j}(x_1^0,\ldots,x_m^0)\,dx_k\,dx_j. \tag{14.3.8}$$

В тех случаях, когда квадратичная форма в правой части равенства (14.3.8)

$$\sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 F}{\partial x_k \partial x_j}(x_1^0,\ldots,x_m^0)\,dx_k\,dx_j \qquad (14.3.9)$$

является знакоопределённой, можно сразу сделать вывод о наличии у функции f строгого относительного экстремума и сказать, является этот экстремум максимумом или минимумом.

Если же квадратичная форма (14.3.9) не является знакоопределённой, то нужно учесть, что нужно рассматривать не все дифференциалы $dx_1,\ldots,dx_m$, а только те, для которых выполнены условия связей. Подставив в квадратичную форму (14.3.9) выражения $dx_{m-n+1},\ldots,dx_m$ через дифференциалы независимых переменных $dx_1,\ldots,dx_{m-n}$, получим квадратичную форму относительно $dx_1,\ldots,dx_{m-n}$.

Если эта новая квадратичная форма является знакоопределённой, то опираясь на теорему 14.1.3, можно сделать вывод о наличии строгого относительного экстремума функции f. А если полученная квадратичная форма может принимать значения разных знаков, то в точке $\mathbf{x}^0$ у функции f относительного экстремума нет.

Пример. Проиллюстрируем применение полученных результатов об условном экстремуме на задаче об экстремуме функции $f(x,y)=xy$ при условии связи $g(x,y)=x+y-2=0$.

Согласно необходимому условию (14.2.9) градиент функции f должен быть коллинеарен градиенту g, т.е. вектор (y,x) должен быть коллинеарен вектору $(1,1)$. Таким образом, должно выполняться равенство $y=x$.

Из уравнения связи находим $x=y=1$. Значит, условный экстремум f может иметь только в точке $(1,1)$.

Далее нетрудно убедиться, что f действительно имеет экстремум в точке $(1,1)$ и что это условный максимум.

Решим теперь эту же задачу, используя неопределённые множители Лагранжа.

Для функции Лагранжа $F(x,y;\lambda)=xy-\lambda(x+y-2)$ имеем

$$\frac{\partial F}{\partial x}=y-\lambda, \qquad \frac{\partial F}{\partial y}=x-\lambda.$$

В стационарной точке $y^0 - \lambda = 0$, $x^0 - \lambda = 0$. Из уравнения связи находим $x^0 + y^0 = 2$. Значит, $\lambda = 1$ и $x^0 = y^0 = 1$.

Запишем второй дифференциал функции F:

$$\frac{\partial^2 F}{\partial x \partial y} = 2\,dx\,dy.$$

Эта квадратичная форма не является знакоопределённой. Но согласно уравнению связи $dx + dy = 0$, т.е. $dy = -dx$. Значит,

$$\frac{\partial^2 F}{\partial x \partial y} = -2\,dx^2$$

и f имеет в точке $(1, 1)$ локальный относительный максимум.

Краткие сведения об ученых, упоминаемых в тексте

Бонне́ Пьер (Bonnet Pierrt, 1819 – 1892) – французский математик.

Боре́ль Эмиль (Borel Êmil, 1871 – 1956) – французкий математик.

Гамильто́н Вильям (Hamilton William, 1805 – 1865) – ирландский математик и астроном.

Дирихле́ Петер Густав Лежён (Dirichle Peter Gustav Lejeune, 1805 – 1859) – немецкий математик.

Ку́рцвейль Ярослав (Kurzweil Jaroslav, 1926 –) – чешский математик.

Лебе́г Анри (Lebesgue Henri Léon, 1875 – 1941) – французский математик.

Острогра́дский Михаил Васильевич (1801 – 1861) – русский математик, академик Петербургской Академии наук.

Пифаго́р Самосский (ок. 570 – ок. 500 до н.э.) – древнегреческий мыслитель, математик.

Ри́ман Бернхард (Riemann Bernhard, 1826 – 1866) – немецкий математик.

Рош Эдуард Альберт (Roche Édouard Albert, 1820 – 1883) – французский астроном и математик.

Стилтье́с Томас Иоаннес (Stieltijes Thomas Joannes, 1856 – 1894) – голландский математик.

Фрулла́ни Джулиано (Frullani Giuliano, 1795 – 1834) – итальянский математик.

Хе́нсток Ральф (Henstok Ralph, 1923 – 2007) – английский математик.

Шварц Герман Амандус (Schwarz Hermann Amandus, 1843 – 1921) – немецкий математик.

Шлёмильх Оскар (Schlömilch Oskar, 1823 – 1901) – немецкий математик.

Яко́би Карл Густав Якоб (Jacobi Karl Gustav Jacob, 1804 – 1851) – немецкий математик.

Сведения об ученых, упоминавшихся в семестре I настоящего курса, здесь не приводятся.

Научное издание

Лекционные курсы НОЦ

Выпуск 17

Сергей Александрович Теляковский

Курс лекций по по математическому анализу
Семестр II

Сдано в набор 28.01.2011. Подписано в печать 09.08.2011.
Формат 60×90/16. Усл. печ. л. 13,25. Тираж 200 экз.

Отпечатано в Математическом институте им. В. А. Стеклова РАН
Москва, 119991, ул. Губкина, 8.
`http://www.mi.ras.ru/noc/` e-mail: `pupyrev@mi.ras.ru`