ПРЕДИСЛОВИЕ

Эта книга представляет собой вторую часть курса математического анализа, читаемого на математико-механическом факультете Санкт-Петербургского государственного университета. Объем и содержание второй части примерно соответствуют материалу, традиционно входящему во второй семестр пятисеместрового курса (или четырехсеместрового, при условии, что теория функций комплексной переменной читается отдельно). Этот том включает в себя следующие разделы: определенный интеграл, предел и непрерывность в $\mathbb{R}^n$, дифференциальное исчисление функций нескольких вещественных переменных.

Про вторую часть, как и про первую, можно сказать, что она написана в жанре подробного конспекта лекций.

Интегральное исчисление (глава 4) начинается с определенного интеграла, потому что неопределенный интеграл (§ 1) вошел в первую часть курса. Набор фактов, содержащийся в этой главе, традиционный, чего нельзя сказать о ее структуре.

На математико-механическом факультете Санкт-Петербургского государственного университета принято рассказывать теорию меры и интеграла в общем курсе анализа (этот раздел начинается в третьем или четвертом семестре, в зависимости от потока). Таким образом, кратные и поверхностные интегралы, интегралы с параметром, интеграл Стилтьеса, ряды и интегралы Фурье излагаются на базе интеграла по мере. При этом подходе на первом курсе необходимо лишь предварительное знакомство с интегралом от непрерывной (или кусочно-непрерывной) функции по отрезку, а изложение интеграла Римана с полным доказательством всех свойств не является обязательным. Определенный интеграл можно вводить разными способами, и, вероятно, среди наших коллег-лекторов не найдется двоих, рассказывающих этот раздел одинаково. Последнее обстоятельство побудило нас включить в эту книгу три способа построения интегрального исчисления: интеграл Римана, интеграл Ньютона – Лейбница и аксиоматическое определение интеграла. Хотя внутри каждого из этих подходов возможна дальнейшая детализация, мы рассчитываем, что каждый студент сможет найти в нашей книге способ изложения, достаточно близкий к тому, который использует преподаватель.

Теория пределов и непрерывных отображений (глава 5) и дифференциальное исчисление (глава 6) излагаются лишь для пространств $\mathbb{R}^n$. В этой же ситуации рассматриваются открытые, замкнутые и компактные множества. Мы не выступаем против рассказа теории пределов и непрерывности в большей общности (нормированные, метрические и даже топологические пространства), но, следуя выбранному в первом томе "медленному" стилю, отложили его до следующей части.

Дифференциальное исчисление излагается в большей общности, чем обычно. Во-первых, там, где можно, рассматриваются не скалярные, а векторные функции нескольких переменных. Это не должно создать новых трудностей читателю, привыкшему к скалярным функциям, так как в подобных вопросах обобщение не требует дополнительных усилий. В тех же ситуациях, когда многомерность области значений принципиальна (например, в формуле Лагранжа или теореме об обратном отображении), нет и выбора, как поступать. Во-вторых, мы определяем кратную дифференцируемость в точке, а не только кратную непрерывную дифференцируемость на множестве. Это существенно и приводит к более тонким результатам. В частности, формула Тейлора – Пеано доказывается при тех же условиях, что и в одномерном случае. В-третьих, мы доказываем равносильность трех способов задания гладкой поверхности, что иногда относят на более позднее время.

К сожалению, полностью бескоординатное изложение дифференциального исчисления (и, тем более, изложение в нормированных пространствах) привело бы и к большему абстрагированию, и к дополнительным техническим трудностям, и к увеличению объема учебника, чего в этой части хотелось избежать. С другой стороны, изложение только на координатной основе, без использования понятия линейного оператора, представляется нам устаревшим и методически неудачным. Поэтому в книге мы выбрали средний, компромиссный вариант.

Нумерация теорем и лемм ведется отдельно в каждом параграфе; нумерация формул — отдельно в каждой главе; нумерация следствий и замечаний — отдельно к каждому утверждению или группе утверждений, к которым эти следствия и замечания относятся. Конец доказательства обозначается символом $\square$.

Авторы благодарны А. Н. Подкорытову, высказавшему множество ценных замечаний по тексту рукописи, а также всем коллегам по кафедре математического анализа Санкт-Петербургского государственного университета, чьи методические находки использовались в этой книге.

ГЛАВА 4. ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ВЕЩЕСТВЕННОЙ ПЕРЕМЕННОЙ

Существует несколько подходов к определению интеграла от непрерывной функции по отрезку. Нагляднее всего геометрический подход, при котором за интеграл от неотрицательной функции принимается площадь ее подграфика. Для того, чтобы реализовать этот подход напрямую, необходимо предварительное построение теории площади, что требует довольно трудоемкой подготовки.

Возможная модификация геометрического подхода заключается в том, что площадь подграфика приближается суммами площадей прямоугольников специального вида. Этот процесс приводит к конструкции *интеграла Римана.*

Другая модификация состоит в том, что интеграл определяется *аксиоматически:* переформулировки некоторых привычных свойств площади принимаются за аксиомы интеграла. При аксиоматическом определении существование интеграла, то есть объекта, задаваемого системой аксиом, остается недоказанным.

Возможен и аналитический подход, при котором за интеграл принимается разность значений первообразной на концах отрезка. Такой интеграл называют *интегралом Ньютона – Лейбница.* При этом подходе многие доказательства упрощаются, но связь с приложениями интеграла уже не столь ясна. Кроме того, остается недоказанным существование первообразной.

Связь между тремя способами построения интеграла устанавливается с помощью формулы Ньютона – Лейбница, которая при первых двух подходах является теоремой, а при третьем принимается за определение.

При втором и третьем способе изложения существование интеграла и, соответственно, первообразной должно быть доказано независимо. В основу доказательства обычно кладут какую-то геометрическую конструкцию: римановскую или более общую, лебеговскую, о которой пойдет речь в одной из следующих частей курса.

Интеграл Римана и его свойства рассматриваются в параграфах 2 и 3. Интеграл Ньютона – Лейбница определяется в § $2'$. Аксиоматический подход излагается в § $2''$, и там же доказывается формула Ньютона – Лейбница для интеграла, определенного аксиоматически. В § 3′ содержатся свойства интеграла, определенного в параграфах 2′ и 2″. Приложения интеграла излагаются двумя способами: с помощью конструкции Римана в § 7 и с использованием понятия плотности в § 7′. Остальные параграфы этой главы — общие для всех подходов к определению интеграла. В случае, когда трактовки того или иного утверждения различны для различных подходов, мы будем указывать на это в соответствующем месте.

Рекомендуем читателю придерживаться одного из трех следующих вариантов чтения этой главы. Интеграл Римана: параграфы 2, 3, 4, 5, 6, 7, 8; интеграл Ньютона – Лейбница: параграфы 2′, 3′, 4, 5, 6, 7′, 8; аксиоматически определенный интеграл: параграфы 2″, 3′, 4, 5, 6, 7′, 8.

§ 2. Определенный интеграл Римана и интегрируемые функции

Напомним, что через $[p:q]$ обозначается множество целых чисел из отрезка $[p,q]$, то есть $[p:q]=[p,q]\cap\mathbb{Z}$.

Определение 1. Пусть $[a,b]$ — невырожденный отрезок. Набор точек

$$\tau=\{x_k\}_{k=0}^{n}:\quad a=x_0<x_1<\ldots<x_n=b$$

называется *дроблением* или *разбиением* отрезка $[a,b]$. Отрезки $[x_k,x_{k+1}]$ $(k\in[0:n-1])$ называют *отрезками дробления,* через Δx_k обозначается длина k-го отрезка дробления: $\Delta x_k=x_{k+1}-x_k$. Наибольшая из длин отрезков дробления, то есть величина

$$\lambda=\lambda_\tau=\max_{0\leqslant k\leqslant n-1}\Delta x_k,$$

называется *рангом* или *мелкостью* дробления τ. Набор точек $\xi=\{\xi_k\}_{k=0}^{n-1}$, таких что $\xi_k\in[x_k,x_{k+1}]$ при всех $k\in[0:n-1]$, называется *оснащением* дробления. Дробление вместе с его оснащением, то есть пара (τ,ξ), называется *оснащенным дроблением.*

Эти обозначения, связанные с отрезком $[a,b]$, будут далее употребляться без дополнительных пояснений.

Определение 2. Суммы Римана. Пусть $f\colon [a,b] \to \mathbb{R}$. Суммы

$$\sigma = \sigma_\tau(f,\xi) = \sum_{k=0}^{n-1} f(\xi_k)\Delta x_k$$

называются *интегральными суммами* или *суммами Римана* функции f, отвечающими оснащенному дроблению (τ,ξ).

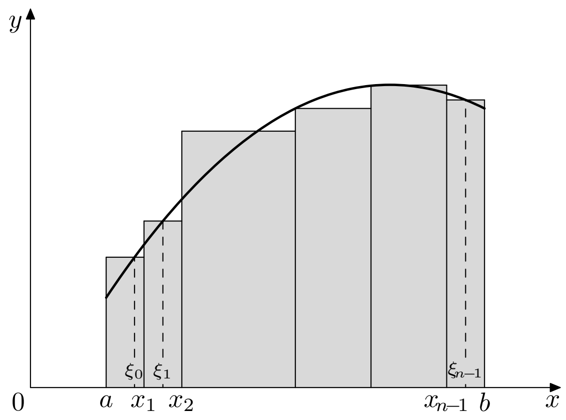

Рис. 1

На рисунке 1 изображен график неотрицательной функции f, а интегральная сумма равна сумме площадей прямоугольников с основаниями Δx_k и высотами $f(\xi_k)$. Естественно ожидать, что для "достаточно хороших" функций с измельчением дробления сумма площадей прямоугольников будет все меньше отличаться от площади подграфика (определение подграфика см. в § 7). Чтобы превратить эту идею в четкую формулировку, необходимо определение площади. Обсуждение понятия площади мы отложим до § 6, а сейчас дадим определение предела интегральных сумм и займемся его изучением.

Определение 3. Предел интегральных сумм. Пусть $f\colon [a,b] \to \mathbb{R}$. Число $I \in \mathbb{R}$ называют *пределом интегральных сумм* при ранге дробления, стремящемся к нулю, и пишут

$$I = \lim_{\lambda_\tau \to 0} \sigma_\tau(f,\xi) \quad \text{или} \quad I = \lim_{\lambda \to 0} \sigma,$$

если

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall \tau : \lambda_\tau < \delta \ \ \forall \xi \ \ |\sigma_\tau(f, \xi) - I| < \varepsilon, \tag{1}$$

то есть для любого положительного числа ε существует такое положительное число δ, что для любого оснащенного дробления (τ, ξ), ранг которого меньше δ, интегральная сумма отличается от числа I меньше чем на ε.

Замечание 1. Утверждение (1) допускает следующую равносильную переформулировку на языке последовательностей. Для любой последовательности оснащенных дроблений $\{(\tau^{(j)}, \xi^{(j)})\}$, такой что последовательность их рангов $\{\lambda^{(j)}\}$ стремится к нулю, соответствующая последовательность интегральных сумм стремится к числу I:

$$\forall \{(\tau^{(j)}, \xi^{(j)})\} : \lambda^{(j)} \to 0 \quad \sigma_{\tau^{(j)}}(f, \xi^{(j)}) \to I.$$

Равносильность определений доказывается так же, как и в случае предела функции. Предлагаем читателю проверить ее.

Замечание 2. Понятие предела интегральных сумм не является частным случаем понятия предела функции, так как интегральная сумма является функцией оснащенного дробления, а не его ранга. И, хотя можно дать определение предела в такой общей ситуации, которая охватит и предел функции, и предел интегральных сумм (так называемый *предел по базе*), мы не будем этого делать и ограничимся определениями по отдельности. Аналогично определяются пределы других функций, зависящих от дробления (возможно, оснащенного). Мы не всегда будем давать явную расшифровку определения предела, оставляя это читателю. Пределы такого типа по замечанию 1 можно свести к пределу последовательности, что позволяет распространить на них теоремы теории пределов из главы 2.

Определение 4. Интеграл Римана. Пусть $f\colon [a, b] \to \mathbb{R}$. Если существует предел интегральных сумм $\lim\limits_{\lambda \to 0} \sigma$, равный числу I, то функция f называется *интегрируемой по Риману* на $[a, b]$, а число I называется *интегралом* (*определенным интегралом, интегралом Римана*) от функции f по отрезку $[a, b]$ и обозначается $\int_a^b f$.

Множество интегрируемых по Риману на $[a,b]$ функций обозначается через $R[a,b]$.

Числа a и b в обозначении $\int_a^b f$ называют *пределами интегрирования,* а f — *подынтегральной функцией.* Часто бывает удобно явно указывать переменную интегрирования и писать $\int_a^b f(x)\,dx$. Переменная x здесь немая и может быть заменена другой буквой. Символ dx обсуждался в § 1 в первой части.

Слова "по Риману" мы будем обычно опускать и говорить просто "интегрируемая функция".

Итак, интеграл есть предел интегральных сумм:

$$\int_a^b f = \lim_{\lambda \to 0} \sigma.$$

После определения интеграла возникают следующие вопросы.

1. Какие функции интегрируемы?

2. Какими свойствами обладает интеграл?

3. Как найти интеграл?

Для рассмотрения первого вопроса нам потребуются интегральные суммы специального вида.

Определение 5. Суммы Дарбу.
Пусть $f\colon [a,b] \to \mathbb{R}$, $\tau = \{x_k\}_{k=0}^n$ — дробление $[a,b]$,

$$M_k = \sup_{x \in [x_k, x_{k+1}]} f(x), \quad m_k = \inf_{x \in [x_k, x_{k+1}]} f(x), \qquad k \in [\,0 : n-1].$$

Суммы

$$S = S_\tau(f) = \sum_{k=0}^{n-1} M_k \Delta x_k \quad \text{и} \quad s = s_\tau(f) = \sum_{k=0}^{n-1} m_k \Delta x_k$$

называются *верхней и нижней интегральными суммами* или *суммами Дарбу* функции f, отвечающими дроблению τ.

Как и для сумм σ, мы часто будем опускать аргументы у сумм S и s.

Если f непрерывна на $[a,b]$, то по теореме Вейерштрасса числа M_k и m_k являются наибольшим и наименьшим значениями f на

$[x_k, x_{k+1}]$. В общем же случае M_k и m_k не обязаны быть значениями функции, поэтому суммы Дарбу могут не быть суммами Римана. Тем не менее, суммы Дарбу устроены проще сумм Римана, так как в их определении не участвует оснащение дробления.

На рисунке 2 верхняя сумма есть сумма площадей бóльших, а нижняя — меньших прямоугольников.

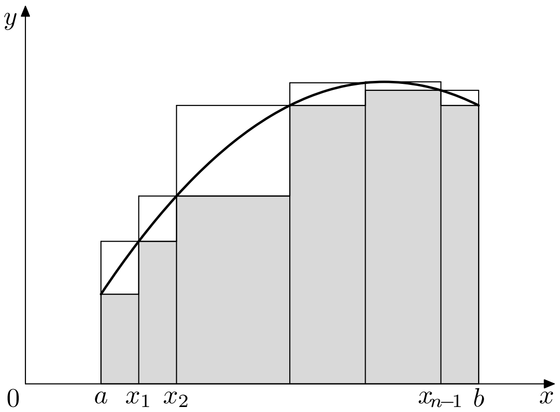

Рис. 2

Отметим, что ограниченность f сверху (снизу) равносильна конечности S (соответственно s). Действительно, если f ограничена сверху, то все $M_k < +\infty$, а тогда и $S < +\infty$; неравенство же $S > -\infty$ выполняется всегда. Обратно, если f не ограничена сверху на $[a, b]$, то f не ограничена сверху на $[x_k, x_{k+1}]$ при некотором k, а тогда соответствующее M_k равно $+\infty$ и $S = +\infty$. Аналогично рассматривается нижняя сумма.

Установим несколько свойств сумм Дарбу.

Д1. $S_\tau(f) = \sup\limits_{\xi} \sigma_\tau(f, \xi)$, $s_\tau(f) = \inf\limits_{\xi} \sigma_\tau(f, \xi)$ *(грани берутся по всевозможным оснащениям дробления τ).*

Доказательство. Для определенности докажем утверждение о верхних суммах. Очевидно, что $f(\xi_k) \leqslant M_k$ при всех $k \in [0 : n-1]$. Умножая эти неравенства на Δx_k и суммируя по k, мы получаем неравенство $\sigma \leqslant S$, то есть S — верхняя граница для интегральных сумм Римана. Докажем, что эта верхняя граница точная.

Пусть f ограничена сверху на $[a,b]$. Возьмем $\varepsilon > 0$ и для каждого k по определению верхней грани подберем такую точку $\xi_k^* \in [x_k, x_{k+1}]$, что $f(\xi_k^*) > M_k - \frac{\varepsilon}{b-a}$. Тогда

$$\sigma^* = \sum_{k=0}^{n-1} f(\xi_k^*)\Delta x_k > S - \frac{\varepsilon}{b-a}\sum_{k=0}^{n-1}\Delta x_k = S - \varepsilon.$$

Так как ε произвольно, S — точная верхняя граница.

Пусть f не ограничена сверху на $[a,b]$. Тогда существует такое ν, что f не ограничена сверху на $[x_\nu, x_{\nu+1}]$. Возьмем $A > 0$ и выберем точки ξ_k^* при $k \neq \nu$ произвольно, а ξ_ν^* — так, чтобы

$$f(\xi_\nu^*) > \frac{1}{\Delta x_\nu}\left(A - \sum_{k\neq\nu} f(\xi_k^*)\Delta x_k\right).$$

Тогда

$$\sigma^* = \sum_{k=0}^{n-1} f(\xi_k^*)\Delta x_k > A.$$

Так как A произвольно, $\sup\limits_{\xi} \sigma = +\infty = S$. $\square$

Д2. *При добавлении новых точек дробления верхняя сумма не увеличится, а нижняя — не уменьшится.*

Доказательство. Для определенности докажем утверждение о верхних суммах. В силу принципа математической индукции достаточно проверить, что верхняя сумма не увеличится при добавлении одной новой точки дробления. Пусть дробление T получено из дробления $\tau = \{x_k\}_{k=0}^n$ добавлением точки $c \in (x_\nu, x_{\nu+1})$. Тогда

$$S_\tau = \sum_{k=0}^{\nu-1} M_k\Delta x_k + M_\nu\Delta x_\nu + \sum_{k=\nu+1}^{n-1} M_k\Delta x_k,$$

$$S_T = \sum_{k=0}^{\nu-1} M_k\Delta x_k + M'(c - x_\nu) + M''(x_{\nu+1} - c) + \sum_{k=\nu+1}^{n-1} M_k\Delta x_k,$$

где $M' = \sup\limits_{x\in[x_\nu,c]} f(x)$, $M'' = \sup\limits_{x\in[c,x_{\nu+1}]} f(x)$. Поскольку при сужении множества его супремум не увеличивается, $M' \leqslant M_\nu$ и $M'' \leqslant M_\nu$. Поэтому

$$\begin{aligned}S_\tau - S_T = M_\nu\Delta x_\nu - M'(c-x_\nu) - M''(x_{\nu+1}-c) \geqslant \\ \geqslant M_\nu(x_{\nu+1} - x_\nu - c + x_\nu + c - x_{\nu+1}) = 0. \quad \square\end{aligned}$$

Д3. *Каждая нижняя сумма Дарбу не превосходит каждой верхней (даже отвечающей другому дроблению).*

Доказательство. Неравенство $s_\tau \leqslant S_\tau$ между суммами для одного и того же дробления τ тривиально. Пусть τ_1 и τ_2 — два дробления отрезка $[a,b]$. Докажем, что $s_{\tau_1} \leqslant S_{\tau_2}$. Положим $\tau = \tau_1 \cup \tau_2$. Тогда по свойству Д2

$$s_{\tau_1} \leqslant s_\tau \leqslant S_\tau \leqslant S_{\tau_2}. \quad \square$$

Лемма 1. *Интегрируемая на отрезке функция ограничена на нем.*

Доказательство. Пусть f не ограничена на $[a,b]$, например, сверху. Тогда для всякого дробления τ по свойству Д1 сумм Дарбу $\sup\limits_\xi \sigma_\tau(f,\xi) = +\infty$. Поэтому для любых числа I и дробления τ найдется такое оснащение ξ, что $\sigma_\tau(f,\xi) > I+1$. Значит, никакое число I не является пределом интегральных сумм. $\square$

Определение 6. Пусть $f\colon[a,b]\to\mathbb{R}$. Величины

$$I^* = \inf_\tau S_\tau, \quad \text{и} \quad I_* = \sup_\tau s_\tau$$

называются *верхним и нижним интегралами Дарбу* функции f.

Из свойства Д3 следует, что $I_* \leqslant I^*$. Как и для сумм Дарбу, ограниченность f сверху (снизу) равносильна соотношению $I^* < +\infty$ (соответственно $I^* > -\infty$).

Теорема 1. Критерий интегрируемости функции. *Пусть $f\colon[a,b]\to\mathbb{R}$. Тогда $f\in R[a,b]$ в том и только том случае, когда $S_\tau(f) - s_\tau(f) \underset{\lambda\to0}{\longrightarrow} 0$, то есть*

$$\forall\varepsilon>0 \;\; \exists\delta>0 \;\; \forall\tau : \lambda_\tau < \delta \;\; S_\tau(f) - s_\tau(f) < \varepsilon.$$

Доказательство. *Необходимость.* Пусть $f \in R[a,b]$. Обозначим $I = \int_a^b f$. По $\varepsilon > 0$ подберем такое $\delta > 0$ из определения предела интегральных сумм, что для любого оснащенного дробления (τ, ξ), ранг которого меньше δ,

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

Переходя к супремуму и инфимуму по ξ, в силу свойства Д1 мы получим

$$I - \frac{\varepsilon}{3} \leqslant s_\tau \leqslant S_\tau \leqslant I + \frac{\varepsilon}{3},$$

откуда $S_\tau - s_\tau \leqslant \frac{2\varepsilon}{3} < \varepsilon$.

Достаточность. Пусть $S_\tau - s_\tau \underset{\lambda \to 0}{\longrightarrow} 0$. Тогда все суммы S_τ и s_τ конечны. Для любого τ

$$s_\tau \leqslant I_* \leqslant I^* \leqslant S_\tau,$$

поэтому

$$0 \leqslant I^* - I_* \leqslant S_\tau - s_\tau.$$

Так как правая часть последнего неравенства принимает сколь угодно малые значения, $I_* = I^*$. Обозначим общее значение I_* и I^* через I и докажем, что $I = \lim\limits_{\lambda \to 0} \sigma$. Из неравенств

$$s_\tau \leqslant I \leqslant S_\tau, \qquad s_\tau \leqslant \sigma_\tau \leqslant S_\tau$$

следует, что

$$|\sigma_\tau - I| \leqslant S_\tau - s_\tau.$$

По $\varepsilon > 0$ можно подобрать такое $\delta > 0$, что для любого дробления τ, ранг которого меньше δ, будет $S_\tau - s_\tau < \varepsilon$, а тогда для любого оснащения ξ такого дробления $|\sigma_\tau(f, \xi) - I| < \varepsilon$. $\square$

Замечание 1. В процессе доказательства теоремы 1 установлено, что если $f \in R[a,b]$, то для любого дробления τ

$$s_\tau \leqslant \int_a^b f \leqslant S_\tau.$$

Определение 7. Пусть $f\colon D \subset \mathbb{R} \to \mathbb{R}$. Величина

$$\omega(f)_D = \sup_{x,y \in D} \bigl(f(x) - f(y)\bigr)$$

называется *колебанием* функции f на множестве D.

Из определений граней функции ясно, что

$$\omega(f)_D = \sup_{x \in D} f(x) - \inf_{y \in D} f(y).$$

Для заданного дробления отрезка $[a,b]$ точками x_k обозначим через $\omega_k(f)$ колебание f на $[x_k, x_{k+1}]$:

$$\omega_k(f) = \omega(f)_{[x_k, x_{k+1}]} = M_k - m_k.$$

Замечание 2. Теорема 1 может быть переформулирована следующим образом. *Пусть* $f\colon [a,b] \to \mathbb{R}$. *Тогда* $f \in R[a,b]$ *в том и только том случае, когда*

$$\lim_{\lambda \to 0} \sum_{k=0}^{n-1} \omega_k(f) \Delta x_k = 0.$$

Действительно,

$$S_\tau(f) - s_\tau(f) = \sum_{k=0}^{n-1} \omega_k(f) \Delta x_k.$$

Следствие 1. *Если* $f \in R[a,b]$, *то*

$$\lim_{\lambda \to 0} S_\tau(f) = \lim_{\lambda \to 0} s_\tau(f) = \int_a^b f.$$

Доказательство вытекает из замечания 1, поскольку

$$0 \leqslant S_\tau - I \leqslant S_\tau - s_\tau, \quad 0 \leqslant I - s_\tau \leqslant S_\tau - s_\tau,$$

где $I = \int_a^b f$. $\square$

Приведем без доказательства еще несколько утверждений.

Замечание 3. *Если* $f\colon[a,b]\to\mathbb{R}$, *то*

$$\lim_{\lambda\to 0} S_\tau(f) = I^*, \qquad \lim_{\lambda\to 0} s_\tau(f) = I_*.$$

Замечание 4. Критерий Дарбу интегрируемости функции. *Пусть* $f\colon[a,b]\to\mathbb{R}$. *Тогда* $f\in R[a,b]$ *в том и только том случае, когда* f *ограничена на* $[a,b]$ *и* $I_* = I^*$.

Замечание 5. Критерий Римана интегрируемости функции. *Пусть* $f\colon[a,b]\to\mathbb{R}$. *Тогда* $f\in R[a,b]$ *в том и только том случае, когда*

$$\forall\varepsilon>0\ \ \exists\tau:\ S_\tau(f)-s_\tau(f)<\varepsilon.$$

Критерий Римана усиливает теорему 1 в части достаточности: для установления интегрируемости функции достаточно по любому $\varepsilon>0$ найти хоть одно дробление, для которого $S_\tau - s_\tau < \varepsilon$, а не добиваться выполнения этого неравенства для всех дроблений достаточно малого ранга.

Геометрически разность $S_\tau - s_\tau$ есть сумма площадей незакрашенных прямоугольников на рисунке 2. Поэтому критерий Римана имеет наглядное геометрическое истолкование: $f\in R[a,b]$ в том и только том случае, когда график f можно заключить в объединение конечного набора прямоугольников указанного вида сколь угодно малой суммарной площади.

В дальнейших доказательствах интегрируемости мы будем опираться лишь на теорему 1.

Теорема 2. Интегрируемость непрерывной функции. *Непрерывная на отрезке функция интегрируема на нем.* Другими словами, *справедливо включение* $C[a,b]\subset R[a,b]$.

Доказательство. Пусть $f\in C[a,b]$. Тогда по теореме Кантора f равномерно непрерывна на $[a,b]$. Возьмем $\varepsilon>0$ и по определению равномерной непрерывности подберем такое $\delta>0$, что для любых $t',t''\in[a,b]$, удовлетворяющих условию $|t'-t''|<\delta$, верно неравенство $|f(t')-f(t'')|<\frac{\varepsilon}{b-a}$. По теореме Вейерштрасса функция f принимает на каждом отрезке наибольшее и наименьшее значение в некоторых точках t' и t''. Поэтому колебание f на всяком отрезке, длина которого меньше δ, будет меньше $\frac{\varepsilon}{b-a}$. Следовательно, для любого дробления τ, ранг которого меньше δ,

$$\sum_{k=0}^{n-1} \omega_k(f)\Delta x_k < \sum_{k=0}^{n-1} \frac{\varepsilon}{b-a}\,\Delta x_k = \varepsilon,$$

то есть для функции f выполнено условие интегрируемости. $\square$

Теорема 3. Интегрируемость монотонной функции. *Монотонная на отрезке функция интегрируема на нем.*

Доказательство. Пусть для определенности f возрастает на $[a,b]$. Если $f(a) = f(b)$, то f постоянна, и ее интегрируемость вытекает из теоремы 2. Пусть $f(a) < f(b)$. Для $\varepsilon > 0$ положим $\delta = \frac{\varepsilon}{f(b)-f(a)}$. Возьмем произвольное дробление τ, такое что $\lambda_\tau < \delta$. В силу возрастания функции f верны равенства $\omega_k(f) = f(x_{k+1}) - f(x_k)$. Поэтому

$$\sum_{k=0}^{n-1} \omega_k(f)\Delta x_k < \sum_{k=0}^{n-1} \bigl(f(x_{k+1}) - f(x_k)\bigr)\frac{\varepsilon}{f(b)-f(a)} = \varepsilon$$

(неравенство строгое ввиду того, что хотя бы одна из разностей $f(x_{k+1}) - f(x_k)$ положительна). Это и означает, что для f выполнено условие интегрируемости. $\square$

Замечание 1. *Если значения интегрируемой функции изменить на конечном множестве точек, то интегрируемость не нарушится и интеграл не изменится.*

Доказательство. Пусть $f \in R[a,b]$, а функция $\widetilde{f}$ отличается от f в точках $t_1, \ldots, t_m$. Тогда, поскольку $|f|$ ограничен некоторым числом A, $|\widetilde{f}|$ тоже ограничен некоторым числом $\widetilde{A}$ (например, можно положить $\widetilde{A} = \max\bigl\{A, \bigl|\widetilde{f}(t_1)\bigr|, \ldots, \bigl|\widetilde{f}(t_m)\bigr|\bigr\}$). В интегральных суммах для f и $\widetilde{f}$ отличаются не более $2m$ слагаемых, откуда

$$|\sigma_\tau(f,\xi) - \sigma_\tau(\widetilde{f},\xi)| \leqslant 2m(A+\widetilde{A})\lambda_\tau \underset{\lambda_\tau\to 0}{\longrightarrow} 0.$$

Поэтому предел $\sigma_\tau(\widetilde{f},\xi)$ существует и равен пределу $\sigma_\tau(f,\xi)$. $\square$

Замечание 1 позволяет определить интеграл для функций, заданных на отрезке всюду, за исключением конечного множества точек, и говорить об интегрируемости таких функций. Именно, если множество $E \subset [a,b]$ конечно, $f\colon [a,b] \setminus E \to \mathbb{R}$, то обозначим через $\widetilde{f}$ какое-нибудь продолжение f на $[a,b]$. Будем говорить, что f интегрируема, если интегрируема $\widetilde{f}$, и в этом случае положим $\int_a^b f = \int_a^b \widetilde{f}$. В силу замечания 1 такое определение корректно, так как не зависит от способа продолжения функции.

В следующей теореме нам будет удобно не различать в обозначениях функцию и ее сужение.

Теорема 4. Интегрируемость функции и ее сужения.
1. *Если $f \in R[a,b]$, $[\alpha,\beta] \subset [a,b]$, то $f \in R[\alpha,\beta]$.*
2. *Если $a < c < b$, $f\colon [a,b] \to \mathbb{R}$, f интегрируема на $[a,c]$ и на $[c,b]$, то $f \in R[a,b]$.*

Доказательство. 1. Проверим выполнение условия интегрируемости f на отрезке $[\alpha,\beta]$. Возьмем $\varepsilon > 0$ и подберем $\delta > 0$ из критерия интегрируемости f на $[a,b]$: если ранг дробления τ отрезка $[a,b]$ меньше δ, то $S_\tau - s_\tau < \varepsilon$. Покажем, что это δ подходит и для критерия интегрируемости f на $[\alpha,\beta]$. Пусть τ_0 — дробление $[\alpha,\beta]$, $\lambda_{\tau_0} < \delta$. Возьмем какие-нибудь дробления отрезков $[a,\alpha]$ и $[\beta,b]$ (если эти отрезки невырожденные) ранга, меньшего δ, и объединим их с τ_0. Получим дробление $\tau = \{x_k\}_{k=0}^n$ отрезка $[a,b]$:

$$a = x_0 < \ldots < x_\mu = \alpha < x_{\mu+1} < \ldots < x_\nu = \beta < x_{\nu+1} < \ldots < x_n = b,$$

причем $\lambda_\tau < \delta$. Тогда

$$S_{\tau_0} - s_{\tau_0} = \sum_{k=\mu}^{\nu-1} \omega_k(f)\Delta x_k \leqslant \sum_{k=0}^{n-1} \omega_k(f)\Delta x_k < \varepsilon.$$

2. Проверим выполнение условия интегрируемости f на отрезке $[a,b]$. Не умаляя общности, можно считать, что f не постоянна, то есть что $\omega = \omega(f)_{[a,b]} > 0$. Возьмем $\varepsilon > 0$. По критерию интегрируемости подберем такие $\delta_1 > 0$ и $\delta_2 > 0$, что для любых дроблений τ_1 отрезка $[a,c]$ и τ_2 отрезка $[c,b]$, удовлетворяющих условиям $\lambda_{\tau_1} < \delta_1$, $\lambda_{\tau_2} < \delta_2$, выполняются неравенства

$$S_{\tau_1} - s_{\tau_1} < \frac{\varepsilon}{3}, \qquad S_{\tau_2} - s_{\tau_2} < \frac{\varepsilon}{3}.$$

Положим $\delta = \min\left\{\delta_1, \delta_2, \frac{\varepsilon}{3\omega}\right\}$. Пусть τ — дробление $[a,b]$, $\lambda_\tau < \delta$. Точка c не обязана принадлежать τ; пусть $c \in [x_\nu, x_{\nu+1})$. Обозначим

$$\tau' = \tau \cup \{c\}, \quad \tau_1 = \tau' \cap [a,c], \quad \tau_2 = \tau' \cap [c,b].$$

Тогда по по выбору δ

$$S_\tau - s_\tau \leqslant S_{\tau_1} - s_{\tau_1} + S_{\tau_2} - s_{\tau_2} + \omega_\nu(f)\delta < \varepsilon. \quad \square$$

Второе утверждение теоремы 4 справедливо и тогда, когда отрезок $[a,b]$ разбит на несколько отрезков. Это проверяется по индукции.

Определение 8. Функция $f\colon [a,b] \to \mathbb{R}$ называется *кусочно-непрерывной* на $[a,b]$, если множество ее точек разрыва пусто или конечно, и все имеющиеся разрывы — первого рода.

Следствие 1. *Кусочно-непрерывная на отрезке функция интегрируема на нем.*

Доказательство. Пусть $a_1, \ldots, a_m$ — все точки разрыва f на (a,b), $a_1 < \ldots < a_m$. Функция f непрерывна во внутренних точках и имеет конечные односторонние пределы на концах каждого из отрезков $[a,a_1]$, $[a_1,a_2]$, $\ldots$, $[a_m,b]$. Поэтому на каждом таком отрезке f интегрируема по замечанию 1, так как отличается от непрерывной функции не более, чем в двух точках. Следовательно, $f \in R[a,b]$ по теореме 4. $\square$

Из теорем 2 и 3, а также из следствия 1 вытекает, что класс интегрируемых функций шире класса непрерывных функций. Тем не менее, оказывается, что интегрируемая функция не может быть "слишком разрывна". Следующий критерий очень удобен для вывода многих утверждений об интегрируемых функциях. Для его формулировки нам понадобится еще одно понятие.

Определение 9. Говорят, что множество $E \subset \mathbb{R}$ имеет *нулевую меру*, если для любого $\varepsilon > 0$ множество E можно заключить в не более чем счетное объединение интервалов, суммарная длина которых меньше ε. (Под суммой счетного семейства положительных чисел a_k понимается $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} a_k$; это понятие будет подробно обсуждаться в главе 7 о числовых рядах.)

В частности, несложно доказать, что любое не более чем счетное множество имеет нулевую меру.

Теорема 5. Критерий Лебега интегрируемости функции по Риману. *Пусть $f\colon [a,b] \to \mathbb{R}$. Тогда $f \in R[a,b]$ в том и только том случае, когда f ограничена на $[a,b]$ и множество ее точек разрыва имеет нулевую меру.*

Эта теорема будет доказана при изучении интеграла Лебега.

Теорема 6. Арифметические действия над интегрируемыми функциями. *Пусть $f, g \in R[a,b]$. Тогда*
1) $f + g \in R[a,b]$;
2) $fg \in R[a,b]$;
3) $\alpha f \in R[a,b]$ $(\alpha \in \mathbb{R})$;
4) $|f| \in R[a,b]$;
5) *если* $\inf\limits_{x\in[a,b]} |g(x)| > 0$, *то* $\frac{f}{g} \in R[a,b]$.

Доказательство. 1) Каковы бы ни были множество E и точки $x, y \in E$,

$$\big|(f+g)(x) - (f+g)(y)\big| \leqslant |f(x) - f(y)| + |g(x) - g(y)| \leqslant \omega(f)_E + \omega(g)_E.$$

Поэтому
$$\omega(f+g)_E \leqslant \omega(f)_E + \omega(g)_E.$$

В частности, для любого дробления $[a,b]$ точками x_k

$$\omega(f+g)_{[x_k,x_{k+1}]} \leqslant \omega(f)_{[x_k,x_{k+1}]} + \omega(g)_{[x_k,x_{k+1}]},$$

то есть
$$\omega_k(f+g) \leqslant \omega_k(f) + \omega_k(g).$$

Умножая эти неравенства на Δx_k, складывая их и пользуясь критерием интегрируемости для f и g, мы получаем

$$\sum_{k=0}^{n-1} \omega_k(f+g)\Delta x_k \leqslant \sum_{k=0}^{n-1} \omega_k(f)\Delta x_k + \sum_{k=0}^{n-1} \omega_k(g)\Delta x_k \underset{\lambda\to 0}{\longrightarrow} 0,$$

то есть для $f + g$ выполнено условие интегрируемости.

2) Поскольку f и g интегрируемы, они ограничены на $[a, b]$. Пусть $|f|$ ограничен числом K, а $|g|$ — числом L. Тогда

$$\begin{aligned}\big|(fg)(x) - (fg)(y)\big| = \big|f(x)g(x) - f(y)g(x) + f(y)g(x) - f(y)g(y)\big| \leqslant \\ \leqslant \big|(f(x) - f(y))g(x)\big| + \big|f(y)(g(x) - g(y))\big| \leqslant \\ \leqslant L|f(x) - f(y)| + K|g(x) - g(y)|.\end{aligned}$$

Следовательно,

$$\omega_k(f + g) \leqslant L\omega_k(f) + K\omega_k(g).$$

Оставшаяся часть доказательства проводится аналогично.

3) Утверждение для αf следует из доказанного утверждения для fg, если взять в качестве g функцию, тождественно равную α.

4) Утверждение для модуля доказывается тем же способом с помощью неравенства

$$\big||f(x)| - |f(y)|\big| \leqslant \big|f(x) - f(y)\big|.$$

5) Интегрируемость частного $\frac{f}{g} = f \cdot \frac{1}{g}$ будет вытекать из утверждения для произведения, если доказать интегрируемость $\frac{1}{g}$. Обозначим $m = \inf\limits_{x\in[a,b]} |g(x)|$. Тогда

$$\left|\frac{1}{g(x)} - \frac{1}{g(y)}\right| = \left|\frac{g(y) - g(x)}{g(x)g(y)}\right| \leqslant \frac{|g(x) - g(y)|}{m^2}$$

и, следовательно,

$$\omega_k\left(\frac{1}{g}\right) \leqslant \frac{\omega_k(g)}{m^2}.$$

Доказательство завершается аналогично. $\square$

Замечание 2. Утверждения 1) и 3) могут быть объединены в одно утверждение: *если* $f, g \in R[a, b]$, $\alpha, \beta \in \mathbb{R}$, *то* $\alpha f + \beta g \in R[a, b]$.

Приведем несколько примеров.

Пример 1. Пусть $b > 0$. Вычислим по определению $\int_0^b x^2\,dx$.

Существование интеграла очевидно из непрерывности подынтегральной функции $f(x) = x^2$, поэтому достаточно найти предел какой-нибудь одной последовательности интегральных сумм. Разобьем отрезок $[0, b]$ на n равных частей точками $x_k = \frac{kb}{n}$ $(k \in [\,0 : n])$. Тогда $\Delta x_k = \frac{b}{n}$. Положим еще $\xi_k = x_k$ $(k \in [\,0 : n-1])$. Для такого оснащенного дробления

$$\sum_{k=0}^{n-1} f(\xi_k)\Delta x_k = \sum_{k=0}^{n-1} \frac{k^2 b^2}{n^2} \cdot \frac{b}{n} = \frac{b^3}{n^3} \cdot \frac{(n-1)n(2n-1)}{6} \underset{n\to\infty}{\longrightarrow} \frac{b^3}{3}.$$

Мы воспользовались легко проверяемой по индукции формулой для суммы квадратов нескольких первых натуральных чисел. Таким образом,

$$\int_0^b x^2\,dx = \frac{b^3}{3}.$$

На практике находить интегралы как пределы интегральных сумм приходится редко; для этой цели гораздо удобнее формула Ньютона – Лейбница, доказываемая в следующем параграфе. Однако, интегральные суммы и их модификации используются для приближенного вычисления интегралов.

Пример 2. *Функция Дирихле*

$$\chi(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q} \end{cases}$$

не интегрируема ни на каком невырожденном отрезке $[a, b]$.

В самом деле, поскольку в каждом интервале есть как рациональное, так и иррациональное число, колебание χ на любом отрезке равно 1. Поэтому для всякого дробления $[a, b]$

$$\sum_{k=0}^{n-1} \omega_k(\chi)\Delta x_k = \sum_{k=0}^{n-1} \Delta x_k = b - a,$$

что не стремится к нулю при ранге дробления, стремящемся к нулю. Таким образом, для функции χ не выполнено условие интегрируемости. $\square$

Пример 3. *Функция Римана*

$$\psi(x) = \begin{cases} \frac{1}{q}, & x = \frac{p}{q} \in \mathbb{Q}, \quad p \in \mathbb{Z},\ q \in \mathbb{N},\ \text{дробь несократима}, \\ 0, & x \notin \mathbb{Q} \end{cases}$$

интегрируема на любом отрезке, и ее интеграл равен нулю.

Не умаляя общности, проведем рассуждение для отрезка $[0,1]$. Ясно, что $s_\tau(\psi) = 0$ для любого дробления τ отрезка $[0,1]$. Возьмем $\varepsilon > 0$ и подберем такое $N \in \mathbb{N}$, что $\frac{1}{N} < \frac{\varepsilon}{2}$. Множество рациональных чисел из $[0,1]$ со знаменателями, не большими N, конечно; пусть оно содержит C_N элементов. Положим $\delta = \frac{\varepsilon}{4C_N}$ и возьмем произвольное дробление τ с рангом, меньшим δ. Указанные точки попадут не более чем в $2C_N$ отрезков дробления; на остальных же отрезках все значения функции меньше $\frac{1}{N}$. Поэтому

$$S_\tau(\psi) = \sum_{k:\, M_k \geqslant \frac{1}{N}} M_k \Delta x_k + \sum_{k:\, M_k < \frac{1}{N}} M_k \Delta x_k \leqslant 2C_N\delta + \frac{1}{N} < \varepsilon.$$

Таким образом,

$$S_\tau(\psi) - s_\tau(\psi) = S_\tau(\psi) \underset{\lambda_\tau \to 0}{\longrightarrow} 0.$$

По критерию интегрируемости и следствию из него $\psi \in R[0,1]$ и $\int_0^1 \psi = 0$. $\square$

Пример 4. Пусть

$$f(y) = \begin{cases} 1, & 0 < y \leqslant 1, \\ 0, & y = 0. \end{cases}$$

Тогда $f \in R[0,1]$, а $f \circ \psi = \chi \notin R[0,1]$.

Этот пример показывает, что композиция двух интегрируемых функций не обязана быть интегрируемой. Если же дополнительно потребовать непрерывность внешней функции, то утверждение об интегрируемости композиции становится верным.

Замечание 3. Интегрируемость композиции. *Пусть* $\varphi\colon [\alpha,\beta] \to [a,b]$, $\varphi \in R[\alpha,\beta]$, $f \in C[a,b]$. *Тогда* $f \circ \varphi \in R[\alpha,\beta]$.

Читатель может вывести это утверждение, а также все утверждения параграфа, начиная с теоремы 3, из критерия Лебега.

§ 3. Свойства интеграла Римана

В определении интеграла предполагалось, что $a < b$. Примем следующее дополнительное соглашение. Если $b < a$, $f \in R[b, a]$, то положим

$$\int_a^b f = -\int_b^a f.$$

Будем также считать, что на вырожденном отрезке любая функция f интегрируема и $\int_a^a f = 0$.

Далее для любых $a, b \in \mathbb{R}$ символом $[a, b]$ мы будем обозначать отрезок с концами a и b. Если не оговорено противное, в формулировках следующих утверждений мы не предполагаем, что $a < b$. Однако, при доказательстве обычно достаточно ограничиться случаем $a < b$, так как общий случай получается переменой знака или принимает вид равенства $0 = 0$. В таких случаях этот шаг доказательства не будет упоминаться.

Если интегрируемость f известна, то для вычисления интеграла достаточно найти предел какой-нибудь одной последовательности интегральных сумм, когда последовательность рангов дроблений стремится к нулю. Например, можно дробить отрезок на n равных частей, а в качестве оснащения брать левые или правые концы или середины отрезков дробления. Можно также рассматривать последовательности верхних или нижних сумм Дарбу. Поэтому для интегральных сумм нет необходимости доказывать теоремы теории пределов, такие как теорема о пределе суммы или о предельном переходе в неравенстве, поскольку достаточно использовать уже известные утверждения о пределах последовательностей.

Установим несколько свойств интеграла.

И1. Аддитивность интеграла по отрезку. *Если* $a, b, c \in \mathbb{R}$, $f \in R\big[\min\{a, b, c\}, \max\{a, b, c\}\big]$, *то*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Доказательство. Пусть $a < c < b$, $f \in R[a,b]$. Тогда по теореме 4 § 2 $f \in R[a,c]$ и $f \in R[c,b]$. Пусть $\left\{\left(\bar\tau^{(n)}, \bar\xi^{(n)}\right)\right\}$, $\left\{\left(\bar{\bar\tau}^{(n)}, \bar{\bar\xi}^{(n)}\right)\right\}$ — последовательности оснащенных дроблений отрезков $[a,c]$ и $[c,b]$ на n равных частей, $\tau^{(n)} = \bar\tau^{(n)} \cup \bar{\bar\tau}^{(n)}$, $\xi^{(n)} = \bar\xi^{(n)} \cup \bar{\bar\xi}^{(n)}$, $\bar\sigma_n$, $\bar{\bar\sigma}_n$ и σ_n — соответствующие последовательности интегральных сумм. Тогда

$$\sigma_n = \bar\sigma_n + \bar{\bar\sigma}_n.$$

Остается перейти к пределу при $n \to \infty$.

Если $a < b < c$, то по доказанному

$$\int_a^b f = \int_a^c f - \int_b^c f = \int_a^c f + \int_c^b f.$$

Если $a = b$, то

$$\int_a^b f = 0 = \int_a^c f + \int_c^b f.$$

Остальные случаи разбираются аналогично. $\square$

Мы не различаем в обозначениях число K и функцию, тождественно равную K.

И2. *Если функция K постоянна на $[a,b]$, то $\int_a^b K = K(b-a)$.*

Доказательство. Поскольку все интегральные суммы равны $K(b-a)$, их предел также равен $K(b-a)$. $\square$

И3. Линейность интеграла. *Если $f, g \in R[a,b]$, $\alpha, \beta \in \mathbb{R}$, то*

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g.$$

Доказательство. Интегрируемость $\alpha f + \beta g$ следует из теоремы 6 § 2. Остается перейти к пределу в равенстве

$$\sigma_\tau(\alpha f + \beta g) = \alpha\sigma_\tau(f) + \beta\sigma_\tau(g). \quad \square$$

Замечание 1. По индукции свойство И1 распространяется на случай нескольких точек, а свойство И3 — на случай нескольких слагаемых.

Замечание 2. Линейность эквивалентна двум свойствам: *аддитивности по функции*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g$$

и *однородности*

$$\int_a^b \alpha f = \alpha \int_a^b f, \qquad \alpha \in \mathbb{R}.$$

Напомним, что запись $f \leqslant g$ на множестве E означает, что $f(x) \leqslant g(x)$ для всех $x \in E$. Если E — общая область определения f и g, то пишут просто $f \leqslant g$.

И4. Монотонность интеграла. *Если $a < b$, $f, g \in R[a, b]$, $f \leqslant g$, то*

$$\int_a^b f \leqslant \int_a^b g.$$

Другими словами, *неравенства можно интегрировать.*

Для доказательства надо перейти к пределу в неравенстве

$$\sigma_\tau(f) \leqslant \sigma_\tau(g).$$

Следствие 1. *Пусть $a < b$, $f \in R[a, b]$. Если $M \in \mathbb{R}$, $f \leqslant M$, то*

$$\int_a^b f \leqslant M(b-a),$$

а если $m \in \mathbb{R}$, $f \geqslant m$, то

$$\int_a^b f \geqslant m(b-a).$$

В частности, если $f \in R[a, b]$, $f \geqslant 0$, то

$$\int_a^b f \geqslant 0.$$

И5. *Пусть $a < b$, $f \in R[a,b]$, $f \geqslant 0$ и существует такая точка $x_0 \in [a,b]$, что $f(x_0) > 0$ и f непрерывна в точке x_0. Тогда*

$$\int_a^b f > 0.$$

Доказательство. Возьмем $\varepsilon = \frac{f(x_0)}{2} > 0$ и по определению непрерывности f в точке x_0 подберем такое $\delta > 0$, что

$$f(x) > f(x_0) - \varepsilon = \frac{f(x_0)}{2} \quad \text{для всех } x \in [x_0 - \delta, x_0 + \delta] \cap [a,b].$$

Обозначим $[\alpha, \beta] = [x_0 - \delta, x_0 + \delta] \cap [a,b]$. По следствию 1 из свойства И4

$$\int_a^b f = \int_a^\alpha f + \int_\alpha^\beta f + \int_\beta^b f \geqslant \int_\alpha^\beta f \geqslant (\beta - \alpha)\frac{f(x_0)}{2} > 0. \quad \square$$

Замечание 1. Без условия непрерывности f в точке x_0 утверждение неверно. Контрпримером служит функция, равная 0 всюду, кроме одной точки, в которой она положительна.

Замечание 2. Утверждение, аналогичное И5, справедливо и для двух функций. Сформулируем его.

Пусть $a < b$, $f, g \in R[a,b]$, $f \leqslant g$ и существует такая точка $x_0 \in [a,b]$, что $f(x_0) < g(x_0)$ и f, g непрерывны в точке x_0. Тогда

$$\int_a^b f < \int_a^b g.$$

Для доказательства достаточно применить И5 к функции $g - f$.

Замечание 3. *Пусть $a < b$, $f \in R[a,b]$, $f > 0$. Тогда*

$$\int_a^b f > 0.$$

Аналогичное утверждение верно и для двух функций.

Действительно, из критерия Лебега легко вытекает, что на $[a,b]$ есть точки непрерывности f.

И6. *Пусть* $a < b$, $f \in R[a,b]$. *Тогда*

$$\left| \int_a^b f \right| \leqslant \int_a^b |f|.$$

Доказательство. Интегрируя неравенство $-|f| \leqslant f \leqslant |f|$, получаем:

$$-\int_a^b |f| \leqslant \int_a^b f \leqslant \int_a^b |f|,$$

что равносильно доказываемому. $\square$

Замечание 4. Если отказаться от требования $a < b$, свойство И6 надо изменить так: *если* $f \in R[a,b]$, *то*

$$\left| \int_a^b f \right| \leqslant \left| \int_a^b |f| \right|.$$

Утверждения следующей серии объединяются названием "первая теорема о среднем интегрального исчисления".

Теорема 1. *Пусть* $f, g \in R[a,b]$, $g \geqslant 0$ *(или* $g \leqslant 0$*)*, $m, M \in \mathbb{R}$, $m \leqslant f \leqslant M$. *Тогда существует такое* $\mu \in [m, M]$, *что*

$$\int_a^b fg = \mu \int_a^b g.$$

Доказательство. Для определенности будем полагать, что $a < b$, $g \geqslant 0$. Тогда $\int_a^b g \geqslant 0$ и

$$mg \leqslant fg \leqslant Mg.$$

Проинтегрируем это неравенство и вынесем постоянные множители за знаки интегралов:

$$m \int_a^b g \leqslant \int_a^b fg \leqslant M \int_a^b g.$$

Отсюда если $\int_a^b g = 0$, то и $\int_a^b fg = 0$, а тогда подходит любое μ. Если же $\int_a^b g > 0$, то следует положить

$$\mu = \frac{\int_a^b fg}{\int_a^b g}.$$

Условия на μ, очевидно, выполнены. $\square$

Следствие 1. *Пусть* $f \in C[a,b]$, $g \in R[a,b]$, $g \geqslant 0$ *(или* $g \leqslant 0$*). Тогда существует такое* $c \in [a,b]$, *что*

$$\int_a^b fg = f(c) \int_a^b g.$$

Доказательство. По теореме Вейерштрасса о непрерывных функциях существуют

$$m = \min_{x \in [a,b]} f(x) \quad \text{и} \quad M = \max_{x \in [a,b]} f(x).$$

Подберем $\mu \in [m, M]$ из теоремы 1. По теореме Больцано – Коши о промежуточном значении найдется такое $c \in [a,b]$, что $\mu = f(c)$. $\square$

Следствие 2. *Пусть* $f \in R[a,b]$, $m, M \in \mathbb{R}$, $m \leqslant f \leqslant M$. *Тогда существует такое* $\mu \in [m, M]$, *что*

$$\int_a^b f = \mu(b-a).$$

Для доказательства надо положить $g \equiv 1$ в теореме 1.

Следствие 3. *Пусть* $f \in C[a,b]$. *Тогда существует такое* $c \in [a,b]$, *что*

$$\int_a^b f = f(c)(b-a).$$

Для доказательства надо положить $g \equiv 1$ в следствии 1.

Замечание 1. Можно доказать, что в условиях следствий 1 и 3 точка c найдется на интервале (a,b).

Замечание 2. Обычно следствия 2 и 3 называют *первой теоремой о среднем,* а теорему 1 и следствие 1 — *усиленной* или *обобщенной первой теоремой о среднем* интегрального исчисления. Эти теоремы находятся в тесной связи с теоремами о среднем дифференциального исчисления, что станет понятно после знакомства с формулой Ньютона – Лейбница.

Поясним еще термин “среднее” в названии теоремы.

Определение 1. Пусть $a < b$, $f \in R[a,b]$. Величина $\frac{1}{b-a}\int_a^b f$ называется *интегральным средним арифметическим* функции f на $[a,b]$.

Если разбить отрезок $[a,b]$ на равные части длины $\frac{b-a}{n}$ и составить интегральную сумму $\sigma_n = \sum\limits_{k=0}^{n-1} f(\xi_k)\frac{b-a}{n}$, то $\frac{\sigma_n}{b-a}$ будет средним арифметическим значений функции в точках оснащения дробления. При этом $\frac{\sigma_n}{b-a} \to \frac{1}{b-a}\int_a^b f$, поэтому и принимается такое определение среднего.

Если $f(t)$ означает мгновенную скорость материальной точки в момент времени t, то $\frac{1}{b-a}\int_a^b f$ есть средняя скорость точки за время от a до b. Это станет ясно из формулы Ньютона – Лейбница и трактовки производной как скорости.

Замечание 3. Следствие 2 утверждает, что среднее находится в тех же границах, что и подынтегральная функция, а следствие 3 — что среднее непрерывной функции равно ее значению в некоторой точке. Теореме 1 и следствию 1 можно придать тот же смысл, если рассматривать *взвешенное среднее* $\dfrac{\int_a^b fg}{\int_a^b g}$ $\left(g \geqslant 0,\ \int_a^b g > 0\right)$ функции f.

Замечание 4. Следствие 3 допускает и такую трактовку: интеграл от непрерывной функции равен некоторой римановой сумме с единственным слагаемым. Геометрически это означает, что площадь подграфика на рисунке 3 равна площади закрашенного прямоугольника.

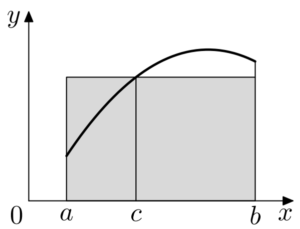

Рис. 3

Определение 2. Пусть $E \subset \mathbb{R}$ — невырожденный промежуток, $f\colon E \to \mathbb{R}$, f интегрируема на каждом отрезке, содержащемся в E, $a \in E$. Функция

$$\Phi(x) = \int_a^x f, \quad x \in E$$

называется *интегралом с переменным верхним пределом.*

Теорема 2. Об интеграле с переменным верхним пределом. *Пусть $E \subset \mathbb{R}$ — невырожденный промежуток, $f\colon E \to \mathbb{R}$, f интегрируема на каждом отрезке, содержащемся в E, $a \in E$, $\Phi(x) = \int_a^x f$ $(x \in E)$. Тогда справедливы следующие утверждения.*

1. *$\Phi \in C(E)$.*

2. *Если, кроме того, f непрерывна в точке $x_0 \in E$, то Φ дифференцируема в точке x_0 и $\Phi'(x_0) = f(x_0)$.*

Утверждение 2 часто называют *теоремой Барроу.*

Доказательство. 1. Возьмем $x_0 \in E$ и докажем непрерывность Φ в точке x_0. Выберем такое $\delta > 0$, что $[x_0 - \delta, x_0 + \delta] \cap E$ есть невырожденный отрезок $[A, B]$. Функция f ограничена на $[A, B]$ некоторым числом M. Пусть Δx таково, что $x_0 + \Delta x \in [A, B]$. Тогда по аддитивности интеграла

$$\Phi(x_0 + \Delta x) - \Phi(x_0) = \int_{x_0}^{x_0 + \Delta x} f,$$

и по свойствам И6 и И4

$$\bigl|\Phi(x_0 + \Delta x) - \Phi(x_0)\bigr| \leqslant \left|\int_{x_0}^{x_0 + \Delta x} |f|\right| \leqslant M\Delta x \underset{\Delta x \to 0}{\longrightarrow} 0.$$

Это и доказывает непрерывность Φ в точке x_0.

2. Проверим, что

$$\frac{\Phi(x_0 + \Delta x) - \Phi(x_0)}{\Delta x} \underset{\Delta x \to 0}{\longrightarrow} f(x_0). \tag{2}$$

Возьмем $\varepsilon > 0$ и по определению непрерывности подберем такое $\delta > 0$, что при всех $t \in E$, удовлетворяющих условию $|t - x_0| < \delta$, будет $|f(t) - f(x_0)| < \varepsilon$. Тогда для всех Δx, таких что $x_0 + \Delta x \in E$ и $0 < |\Delta x| < \delta$, по свойствам И6, И5 и замечаниям к ним

$$\left|\frac{\Phi(x_0 + \Delta x) - \Phi(x_0)}{\Delta x} - f(x_0)\right| = \left|\frac{1}{\Delta x}\int_{x_0}^{x_0+\Delta x} \big(f(t) - f(x_0)\big)\,dt\right| < \\ < \frac{1}{|\Delta x|}\cdot\varepsilon|\Delta x| = \varepsilon,$$

откуда и следует (2). $\square$

Следствие 1. *Функция, непрерывная на промежутке, имеет на нем первообразную.*

Согласно теореме Барроу, первообразной является интеграл с переменным верхним пределом.

Замечание 1. Следствие 1 было сформулировано без доказательства в первой части курса (теорема 2 § 1 главы 4); теперь оно доказано.

Замечание 2. Аналогично при $b \in E$ определяется *интеграл с переменным нижним пределом*

$$\Psi(x) = \int_x^b f, \quad x \in E.$$

Так как $\int_x^b f = -\int_b^x f$, из теоремы 2 следует, что Ψ непрерывна и

$$\Psi'(x_0) = -f(x_0)$$

во всех точках непрерывности f.

Следующая теорема — формула Ньютона – Лейбница — важнейшая в интегральном исчислении. Она устанавливает связь определенного интеграла с неопределенным и позволяет вычислять определенный интеграл от функции, первообразная которой известна.

Теорема 3. Формула Ньютона – Лейбница.
Пусть $f \in R[a,b]$, F — первообразная f на $[a,b]$. Тогда

$$\int_a^b f = F(b) - F(a).$$

Доказательство. При каждом $n \in \mathbb{N}$ положим $x_k = \frac{k(b-a)}{n}$. Тогда

$$F(b) - F(a) = \sum_{k=0}^{n-1} \big(F(x_{k+1}) - F(x_k)\big).$$

По теореме Лагранжа для каждого $k \in [0 : n-1]$ найдется такая точка $\xi_k^{(n)} \in (x_k, x_{k+1})$, что

$$F(x_{k+1}) - F(x_k) = F'\big(\xi_k^{(n)}\big)\Delta x_k = f\big(\xi_k^{(n)}\big)\Delta x_k.$$

В силу интегрируемости f

$$\int_a^b f = \lim_{n\to\infty} \sum_{k=0}^{n-1} f\big(\xi_k^{(n)}\big)\Delta x_k = \lim_{n\to\infty} \big(F(b) - F(a)\big) = F(b) - F(a). \quad \square$$

Замечание 1. Пусть $F\colon [a,b] \to \mathbb{R}$. Разность $F(b) - F(a)$ называется *двойной подстановкой* функции F на $[a,b]$ и обозначается $F|_a^b$, $F(x)|_a^b$, $F(x)|_{x=a}^b$ или $[F(x)]_{x=a}^b$.

Как обычно, переменная x здесь немая и может быть заменена другой буквой. Последняя запись позволяет отметить начало выражения для F, что бывает удобно, когда оно длинное.

Таким образом, формула Ньютона – Лейбница может быть записана в виде

$$\int_a^b f = F\big|_a^b.$$

Замечание 2. Формула Ньютона – Лейбница доказана для любой первообразной подынтегральной функции. То, что двойная подстановка не зависит от выбора первообразной, ясно и так. Действительно, если F и Φ — первообразные f на $[a,b]$, то они отличаются на константу: $\Phi = F + C$. Но тогда их двойные подстановки совпадают:

$$\Phi(b) - \Phi(a) = F(b) + C - (F(a) + C) = F(b) - F(a).$$

Интеграл, который был сосчитан в конце § 2 как предел интегральных сумм, с помощью формулы Ньютона – Лейбница считается моментально:

$$\int_0^b x^2\,dx = \frac{x^3}{3}\Big|_0^b = \frac{b^3}{3}.$$

Попробуем применить формулу Ньютона – Лейбница к интегралу $\int_{-1}^1 \frac{dx}{x^2}$:

$$\int_{-1}^1 \frac{dx}{x^2} = \left[-\frac{1}{x}\right]_{-1}^1 = -2.$$

Получается нелепый результат — интеграл от положительной функции отрицателен. В этом примере нарушены два условия теоремы 3. Во-первых, функция $f(x) = \frac{1}{x^2}$ не интегрируема на $[-1,1]$ (так как не ограничена). Во-вторых, равенство $\left(-\frac{1}{x}\right)' = \frac{1}{x^2}$ не имеет смысла в точке 0.

Тем не менее, формула Ньютона – Лейбница допускает некоторое обобщение.

Замечание 3. *Пусть $f \in R[a,b]$, $F \in C[a,b]$, F — первообразная f на $[a,b]$ за вычетом конечного множества точек. Тогда*

$$\int_a^b f = F(b) - F(a).$$

Доказательство. Пусть $\alpha_1 < \ldots < \alpha_{m-1}$ — все точки интервала (a,b), в которых нарушается равенство $F' = f$; положим также $\alpha_0 = a$, $\alpha_m = b$. Пользуясь последовательно непрерывностью интеграла с переменными пределами интегрирования, формулой Ньютона – Лейбница и непрерывностью F, находим:

$$\int_{\alpha_k}^{\alpha_{k+1}} f = \lim_{\varepsilon\to 0+} \int_{\alpha_k+\varepsilon}^{\alpha_{k+1}-\varepsilon} f = \\ = \lim_{\varepsilon\to 0+} \big(F(\alpha_{k+1}-\varepsilon) - F(\alpha_k+\varepsilon)\big) = F(\alpha_{k+1}) - F(\alpha_k).$$

Остается воспользоваться аддитивностью интеграла:

$$\int_a^b f = \sum_{k=0}^{m-1} \int_{\alpha_k}^{\alpha_{k+1}} f = \sum_{k=0}^{m-1} \big(F(\alpha_{k+1}) - F(\alpha_k)\big) = F(b) - F(a). \quad \square$$

Замечание 4. Условие $F \in C[a,b]$ в замечании 3 существенно. Для функций $f(x) = 0$, $F(x) = \operatorname{sign} x$ формула Ньютона – Лейбница на $[-1,1]$ неверна:

$$0 = \int_{-1}^{1} f \neq F\Big|_{-1}^{1} = 2.$$

Замечание 5. Теорему 3 можно переформулировать и так: *если F дифференцируема на $[a,b]$, $F' \in R[a,b]$, то*

$$\int_a^b F' = F(b) - F(a).$$

Аналогично замечанию 3, и в этой формулировке можно разрешить функции $F \in C[a,b]$ не иметь производной на конечном множестве точек.

Замечание 6. Условие $F' \in R[a,b]$ в замечании 5 опустить нельзя, так как производная может не быть интегрируемой, и тогда интеграл Римана от нее не имеет смысла. Примером служит функция

$$F(x) = \begin{cases} x^2 \sin \frac{1}{x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Для нее

$$F'(x) = \begin{cases} 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

(первая строчка получается по обычным правилам дифференцирования, а вторая — нахождением предела разностного отношения). Поскольку $F'\left(\frac{1}{\sqrt{2k\pi}}\right) \underset{k\to\infty}{\longrightarrow} -\infty$, F' не ограничена и, следовательно, не интегрируема на $[-1,1]$.

Этот пример показывает, что интеграл Римана не всегда решает задачу восстановления функции по ее производной. Полностью эта задача была решена французским математиком А. Данжуа (1912 г.) и немецким математиком О. Перроном (1914 г.). Как выяснилось позже, их конструкции приводят к одному и тому же результату, поэтому построенный ими интеграл стали называть *интегралом Данжуа – Перрона.*

Замечание 7. Приведенный в замечании 6 пример показывает, что из существования у функции f первообразной не следует интегрируемость f. С другой стороны, функция sign интегрируема на отрезке $[-1,1]$, но не имеет на нем первообразной. Таким образом, условия интегрируемости функции f и существования у нее первообразной независимы.

При вычислении определенных, как и неопределенных, интегралов полезны приемы *интегрирования по частям* и *замены переменной.*

Теорема 4. Интегрирование по частям в определенном интеграле. *Пусть f, g дифференцируемы на $[a,b]$, $f', g' \in R[a,b]$. Тогда*

$$\int_a^b fg' = fg\Big|_a^b - \int_a^b f'g.$$

Доказательство. Будучи дифференцируемыми, функции f и g непрерывны и, следовательно, интегрируемы. По теореме об арифметических действиях над интегрируемыми функциями $f'g, fg' \in R[a,b]$, а тогда и $(fg)' = f'g + fg' \in R[a,b]$. По формуле Ньютона – Лейбница

$$\int_a^b fg' + \int_a^b f'g = \int_a^b (fg)' = fg\Big|_a^b.$$

Остается перенести второе слагаемое из левой части в правую. $\square$

Замечание 1. Формулу интегрирования по частям записывают и в виде

$$\int_a^b f\,dg = fg\Big|_a^b - \int_a^b g\,df,$$

трактуя $f'(x)\,dx$ и $g'(x)\,dx$ как дифференциалы. Аналогичная форма записи для неопределенных интегралов обсуждалась в § 1.

Пример 1. Полагая $f(x) = \ln x$, $g'(x) = 1$, находим:

$$\int_1^2 \ln x\,dx = x\,\ln x\Big|_1^2 - \int_1^2 x\,\frac{dx}{x} = 2\ln 2 - 1.$$

Теорема 5. Замена переменной в определенном интеграле. *Пусть* $\varphi\colon [\alpha,\beta] \to [A,B]$, φ *дифференцируема на* $[\alpha,\beta]$, $\varphi' \in R[\alpha,\beta]$, $f \in C[A,B]$. *Тогда*

$$\int_\alpha^\beta (f\circ\varphi)\varphi' = \int_{\varphi(\alpha)}^{\varphi(\beta)} f.$$

Доказательство. Поскольку $f\circ\varphi \in C[\alpha,\beta] \subset R[\alpha,\beta]$, по теореме об арифметических действиях над интегрируемыми функциями $(f\circ\varphi)\varphi' \in R[\alpha,\beta]$. Пусть F — первообразная f на $[A,B]$. Тогда по правилу дифференцирования композиции $F\circ\varphi$ — первообразная $(f\circ\varphi)\varphi'$ на $[A,B]$. Применяя к обоим интегралам формулу Ньютона – Лейбница, получаем:

$$\int_\alpha^\beta (f\circ\varphi)\varphi' = F\circ\varphi\Big|_\alpha^\beta = F\Big|_{\varphi(\alpha)}^{\varphi(\beta)} = \int_{\varphi(\alpha)}^{\varphi(\beta)} f. \quad \Box$$

Замечание 2. Правило замены переменной может применяться как слева направо, так и справа налево. Допустим, что в интеграле $\int_a^b f(x)\,dx$ мы хотим сделать замену $x=\varphi(t)$, где $\varphi(\alpha)=a$, $\varphi(\beta)=b$. Тогда надо трактовать dx как дифференциал: $dx=\varphi'(t)\,dt$, и поменять пределы интегрирования: a на α и b на β. Получим:

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt.$$

В отличие от неопределенного интеграла, при вычислении определенного не надо возвращаться к старой переменной, но надо не забыть поменять пределы интегрирования.

Замечание 3. В условиях теоремы некоторые значения $\varphi(t)$ при $t\in[\alpha,\beta]$ могут не принадлежать отрезку $[\varphi(\alpha),\varphi(\beta)]$. Важно, что они принадлежат отрезку $[A,B]$, на котором определена функция f.

В формуле замены переменной особенно удобно соглашение о том, что нижний предел интегрирования не обязательно меньше верхнего. Например, если φ строго убывает, а $\alpha<\beta$, то $\varphi(\alpha)>\varphi(\beta)$.

Замечание 4. В формуле замены переменной на функции можно накладывать и другие условия. Сформулируем одно из таких утверждений. *Пусть функция* φ *дифференцируема, строго монотонна на* $[\alpha,\beta]$, $\varphi'\in R[\alpha,\beta]$, $f\in R[\varphi(\alpha),\varphi(\beta)]$. *Тогда*

$$\int_\alpha^\beta (f\circ\varphi)\varphi' = \int_{\varphi(\alpha)}^{\varphi(\beta)} f.$$

Как видно, здесь от функции φ требуется больше, а от f — меньше, чем в теореме 5.

Доказательство этого утверждения (вместе с доказательством существования интеграла в левой части) можно провести с помощью римановых сумм, и оно остается читателю в качестве задачи.

Пример 2. Вычислим интеграл $\int_0^a \sqrt{a^2-x^2}\,dx$ $(a>0)$ с помощью тригонометрической подстановки $x = a\sin t$. Функция $\varphi(t) = a\sin t$ отображает $\left[0,\frac{\pi}{2}\right]$ на $[0,a]$, $\varphi(0)=0$, $\varphi\left(\frac{\pi}{2}\right)=a$, $\varphi'(t)=a\cos t$. Поэтому

$$\int_0^a \sqrt{a^2-x^2}\,dx = \int_0^{\pi/2} a^2\cos^2 t\,dt = a^2\int_0^{\pi/2}\frac{1+\cos 2t}{2}\,dt =$$
$$= \frac{a^2}{2}\left[t+\frac{\sin 2t}{2}\right]_0^{\pi/2} = \frac{\pi a^2}{4}.$$

Трактуя интеграл как площадь, можно сразу сказать, что интеграл в этом примере равен площади четверти круга радиуса a.

Замечание 5. Если $f\in R[-a,0\,]$, то, полагая $x=-t$, находим

$$\int_{-a}^0 f(x)\,dx = \int_a^0 f(-t)(-1)\,dt = \int_0^a f(-t)\,dt.$$

Поэтому для $f\in R[-a,a]$

$$\int_{-a}^a f = \int_{-a}^0 f + \int_0^a f = \int_0^a \big(f(t)+f(-t)\big)\,dt.$$

Следовательно, если f четна, то

$$\int_{-a}^a f = 2\int_0^a f,$$

а если f нечетна, то

$$\int_{-a}^{a} f = 0.$$

Эти простые соображения часто облегчают вычисление интегралов.

§ 2′. Определенный интеграл Ньютона – Лейбница

Определение 1. Пусть $F\colon [a,b] \to \mathbb{R}$. Разность $F(b) - F(a)$ называется *двойной подстановкой* функции F на $[a,b]$ и обозначается $F|_a^b$, $F(x)|_a^b$, $F(x)|_{x=a}^b$ или $[F(x)]_{x=a}^b$.

Как обычно, переменная x здесь немая и может быть заменена другой буквой. Последняя запись позволяет отметить начало выражения для F, что бывает удобно, когда оно длинное.

Во многих задачах искомые величины выражаются в виде двойной подстановки первообразной известной функции. Приведем два примера.

1. Пусть $f(t)$ означает скорость материальной точки в момент времени t, а $s(t)$ — путь, пройденный точкой к моменту t. Тогда s есть первообразная f. Путь, пройденный точкой за отрезок времени $[a,b]$, равен двойной подстановке первообразной: $s(b) - s(a)$.

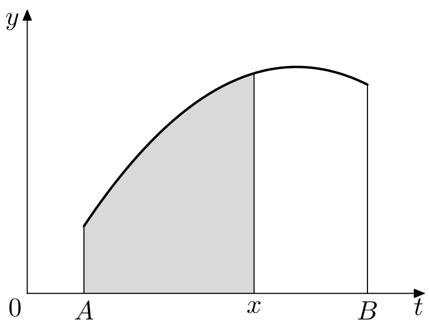

Рис. 4

2. На рисунке 4 изображен график неотрицательной функции $f \in C[A,B]$. Пусть $S(x)$ означает площадь подграфика сужения f на отрезок $[A,x]$ ($x \in [A,B]$). Используя естественные свойства площади (см. § 6; определение подграфика см. в § 7′), можно

доказать, что S есть первообразная f. Таким образом, площадь подграфика сужения f на некоторый отрезок $[a,b]$ равна двойной подстановке первообразной: $S(b)-S(a)$.

Определение 2. Интеграл Ньютона – Лейбница. Пусть $f\in C[a,b]$, F — первообразная f на $[a,b]$. Двойная подстановка функции F называется *определенным интегралом* от функции f по отрезку $[a,b]$ и обозначается $\int_a^b f$.

Итак,

$$\int_a^b f = F\Big|_a^b. \tag{3}$$

Числа a и b в обозначении $\int_a^b f$ называют *пределами интегрирования,* а f — *подынтегральной функцией.* Часто бывает удобно явно указывать переменную интегрирования и писать $\int_a^b f(x)\,dx$. Переменная x здесь также немая.

Равенство (3) называется *формулой Ньютона – Лейбница.*

Замечание 1. Определение интеграла не зависит от выбора первообразной. Действительно, если F и Φ — первообразные f на $[a,b]$, то они отличаются на константу: $\Phi=F+C$. Но тогда их двойные подстановки совпадают:

$$\Phi(b)-\Phi(a)=F(b)+C-(F(a)+C)=F(b)-F(a).$$

В определении интеграла предполагалось, что $a<b$. Примем следующее дополнительное соглашение. Если $b<a$, $f\in C[b,a]$, то положим

$$\int_a^b f = -\int_b^a f.$$

Положим также $\int_a^a f=0$ для любой функции f. При этом соглашении формула Ньютона – Лейбница остается верной для любых $a,b\in\mathbb{R}$: если $a>b$, то

$$\int_a^b f = -\int_b^a f = -\big(F(a)-F(b)\big)=F(b)-F(a).$$

Далее для любых $a,b\in\mathbb{R}$ символом $[a,b]$ мы будем обозначать отрезок с концами a и b. Если не оговорено противное, в формулировках следующих утверждений не предполагается, что $a<b$.

Определение 3. Пусть $E \subset \mathbb{R}$ — невырожденный промежуток, $f \in C(E)$, $a \in E$. Функция

$$\Phi(x) = \int_a^x f, \quad x \in E$$

называется *интегралом с переменным верхним пределом.*

Замечание 2. Функция Φ является первообразной функции f на E.

Действительно, если F — первообразная f, то по определению интеграла $\Phi(x) = F(x) - F(a)$, а тогда Φ — тоже первообразная f.

Замечание 2 иногда называют *теоремой Барроу.*

Замечание 3. В этом параграфе принято определение интеграла, при котором формула Ньютона – Лейбница и теорема Барроу выполняются по определению. Плата за такую простоту — отсутствие доказательства существования первообразной и, тем самым, существования интеграла у любой непрерывной на промежутке функции. Если определить интеграл независимо от понятия первообразной, то формула Ньютона – Лейбница и теорема Барроу становятся содержательными утверждениями, подлежащими доказательству. При этом из теоремы Барроу вытекает существование первообразной: таковой служит интеграл с переменным верхним пределом.

Установим несколько свойств интеграла.

И1. Аддитивность интеграла по отрезку.
Если $a, b, c \in \mathbb{R}$, $f \in C\big[\min\{a,b,c\}, \max\{a,b,c\}\big]$, *то*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

В самом деле,

$$\int_a^b f = F(b) - F(a) = F(c) - F(a) + F(b) - F(c) = \int_a^c f + \int_c^b f.$$

Подчеркнем, что свойство И1 верно при любом расположении точек a, b и c.

Замечание 1. По индукции свойство И1 распространяется на произвольный конечный набор чисел $x_0, \ldots, x_n$, где $x_0 = a$, $x_n = b$:

$$\int_a^b f = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f.$$

Мы не различаем в обозначениях число K и функцию, тождественно равную K.

И2. *Если функция K постоянна на $[a, b]$, то*

$$\int_a^b K = K(b-a).$$

Доказательство. По формуле Ньютона – Лейбница

$$\int_a^b K = Kx\big|_a^b = K(b-a). \quad \square$$

И3. Линейность интеграла. *Если $f, g \in C[a, b]$, $\alpha, \beta \in \mathbb{R}$, то*

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g.$$

Доказательство. Пусть F и G — первообразные для f и g. Тогда $\alpha F + \beta G$ — первообразная для $\alpha f + \beta g$. Поэтому

$$\int_a^b (\alpha f + \beta g) = (\alpha F + \beta G)(b) - (\alpha F + \beta G)(a) =$$

$$= \alpha(F(b) - F(a)) + \beta(G(b) - G(a)) = \alpha \int_a^b f + \beta \int_a^b g. \quad \square$$

Замечание 2. По индукции свойство И3 распространяется на случай нескольких слагаемых.

Замечание 3. Линейность эквивалентна двум свойствам: *аддитивности по функции*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g$$

и *однородности*

$$\int_a^b \alpha f = \alpha \int_a^b f, \qquad \alpha \in \mathbb{R}.$$

Напомним, что запись $f \leqslant g$ на множестве E означает, что $f(x) \leqslant g(x)$ для всех $x \in E$. Если E — общая область определения f и g, то пишут просто $f \leqslant g$.

И4. *Если* $a < b$, $f \in C[a,b]$, $f \geqslant 0$, *то*

$$\int_a^b f \geqslant 0.$$

Доказательство. Поскольку $f \geqslant 0$, ее первообразная F возрастает на $[a,b]$. Поэтому

$$\int_a^b f = F(b) - F(a) \geqslant 0. \quad \square$$

Следствие 1. Монотонность интеграла. *Если* $a < b$, $f, g \in C[a,b]$, $f \leqslant g$, *то*

$$\int_a^b f \leqslant \int_a^b g.$$

Другими словами, *неравенства можно интегрировать.*

Для доказательства надо применить свойство И4 к разности $g - f$ и воспользоваться свойством И3.

Следствие 2. *Пусть* $a < b$, $f \in C[a,b]$. *Если* $M \in \mathbb{R}$, $f \leqslant M$, *то*

$$\int_a^b f \leqslant M(b-a),$$

а если $m \in \mathbb{R}$, $f \geqslant m$, *то*

$$\int_a^b f \geqslant m(b-a).$$

Замечание 4. Если в условиях свойства И4 $f > 0$, то F строго возрастает, и поэтому $\int_a^b f > 0$. Аналогичное уточнение верно и для следствия 1.

Это замечание будет усилено в § 3′.

§ 2″. Аксиоматическое определение интеграла

Напомним, что запись $f \leqslant g$ на множестве E означает, что $f(x) \leqslant g(x)$ для всех $x \in E$. Если E — общая область определения f и g, то пишут просто $f \leqslant g$.

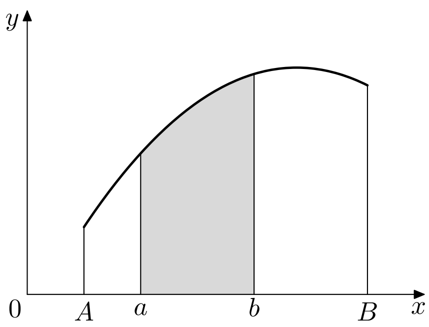

Рис. 5

К понятию интеграла приводит задача о нахождении площади подграфика функции. На рисунке 5 изображен график неотрицательной функции $f \in C[A, B]$. Рассмотрим площадь подграфика (определение подграфика см. в § 7′) как величину, зависящую от отрезка и функции, и обозначим ее через $S\bigl([a, b], f\bigr)$. Мы не будем давать определение площади, а примем без доказательства, что площадь подграфика существует и обладает следующими естественными свойствами.

1. Если $a < c < b$, $f \in C[a, b]$, $f \geqslant 0$, то

$$S\bigl([a, b], f\bigr) = S\bigl([a, c], f\bigr) + S\bigl([c, b], f\bigr).$$

2. Если $f, g \in C[a, b]$, $0 \leqslant f \leqslant g$, то

$$S\bigl([a, b], f\bigr) \leqslant S\bigl([a, b], g\bigr).$$

3. Если функция K постоянна на $[a,b]$, $K \geqslant 0$, то

$$S\big([a,b],K\big) = K(b-a).$$

Последнее означает, что площадь прямоугольника с основанием $[a,b]$ и высотой K равна $K(b-a)$. Нам удобно не различать в обозначениях функцию и ее сужение, а также число K и функцию, тождественно равную K.

Свойства площади наводят на мысль рассмотреть отображение с теми же свойствами, что и S, но отказавшись от требования неотрицательности функции и уже не прибегая к геометрической мотивировке.

Определение 1. Аксиоматическое определение интеграла. *Интегралом* (*определенным интегралом*) называется функционал I, заданный на множестве пар: первая компонента пары — отрезок, вторая — непрерывная на этом отрезке функция, удовлетворяющий следующим трем условиям (аксиомам).

I1. Аддитивность по отрезку. Если $a<c<b$, $f \in C[a,b]$, то

$$I\big([a,b],f\big) = I\big([a,c],f\big) + I\big([c,b],f\big).$$

I2. Монотоность. Если $f,g \in C[a,b]$, $f \leqslant g$, то

$$I\big([a,b],f\big) \leqslant I\big([a,b],g\big).$$

I3. Нормированность. Если функция K постоянна на $[a,b]$, то

$$I\big([a,b],K\big) = K(b-a).$$

Значение $I\big([a,b],f\big)$ называют интегралом от функции f по отрезку $[a,b]$ и обозначают $\int_a^b f$. Перепишем аксиомы I1–I3 с помощью этого обозначения.

I1. Аддитивность по отрезку. Если $a<c<b$, $f \in C[a,b]$, то

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

I2. Монотоность. Если $f, g \in C[a,b]$, $f \leqslant g$, то

$$\int_a^b f \leqslant \int_a^b g.$$

I3. Нормированность. Если функция K постоянна на $[a,b]$, то

$$\int_a^b K = K(b-a).$$

Числа a и b в обозначении $\int_a^b f$ называют *пределами интегрирования,* а f — *подынтегральной функцией.* Часто бывает удобно явно указывать переменную интегрирования и писать $\int_a^b f(x)\,dx$. Переменная x здесь немая и может быть заменена другой буквой.

После определения интеграла возникает три вопроса.

1. Существует ли интеграл?
2. Единственный ли интеграл?
3. Как найти интеграл?

Ответы на первые два вопроса утвердительные: интеграл существует и единственный. Можно догадаться, что существование интеграла равносильно существованию площади подграфика непрерывной функции. Существование интеграла в этом параграфе доказываться не будет. Единственность интеграла будет вскоре доказана с помощью формулы Ньютона – Лейбница, которая служит и основным приемом нахождения интегралов.

В определении интеграла $\int_a^b f$ предполагалось, что $a < b$. Примем следующее дополнительное соглашение. Если $b < a$, $f \in C[b,a]$, то положим

$$\int_a^b f = -\int_b^a f.$$

Условимся также, что $\int_a^a f = 0$ для любой функции f.

Далее для любых $a, b \in \mathbb{R}$ символом $[a,b]$ мы будем обозначать отрезок с концами a и b. При этом соглашении свойство аддитивности интеграла становится верным при любом расположении точек a, b и c.

Если не оговорено противное, в формулировках следующих утверждений не предполагается, что $a < b$.

И1. Аддитивность интеграла по отрезку. *Если* $a, b, c \in \mathbb{R}$, $f \in C\big[\min\{a,b,c\}, \max\{a,b,c\}\big]$, *то*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Доказательство. При $a < c < b$ это аксиома I2. Если $a < b < c$, то по аксиоме I2

$$\int_a^b f = \int_a^c f - \int_b^c f = \int_a^c f + \int_c^b f.$$

Если $a = b$, то

$$\int_a^b f = 0 = \int_a^c f + \int_c^b f.$$

Остальные случаи разбираются аналогично. $\square$

Замечание 1. По индукции свойство И1 распространяется на произвольный конечный набор чисел $x_0, \ldots, x_n$, где $x_0 = a$, $x_n = b$:

$$\int_a^b f = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f.$$

И2. *Пусть* $a < b$, $f \in C[a,b]$. *Если* $M \in \mathbb{R}$, $f \leqslant M$, *то*

$$\int_a^b f \leqslant M(b-a),$$

а если $m \in \mathbb{R}$, $f \geqslant m$, *то*

$$\int_a^b f \geqslant m(b-a).$$

Это свойство вытекает из монотонности и нормированности интеграла.

Определение 2. Пусть $E \subset \mathbb{R}$ — невырожденный промежуток, $f \in C(E)$, $a \in E$. Функция

$$\Phi(x) = \int_a^x f, \quad x \in E$$

называется *интегралом с переменным верхним пределом.*

Теорема 1 (И. Барроу). *Пусть $E \subset \mathbb{R}$ — невырожденный промежуток, $f \in C(E)$, $a \in E$, $\Phi(x) = \int_a^x f$ $(x \in E)$. Тогда Φ дифференцируема на E и $\Phi' = f$ на E.*

Доказательство. Возьмем $x \in E$ и проверим, что

$$\frac{\Phi(x+h) - \Phi(x)}{h} \underset{h \to 0}{\longrightarrow} f(x). \tag{4}$$

При $h \neq 0$, $x + h \in E$ по аддитивности интеграла

$$\frac{\Phi(x+h) - \Phi(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t)\, dt.$$

Возьмем $\varepsilon > 0$ и по определению непрерывности подберем такое $\delta > 0$, что при всех $t \in E$, удовлетворяющих условию $|t - x| < \delta$, будет $f(x) - \varepsilon < f(t) < f(x) + \varepsilon$. Тогда для всех h, таких что $x + h \in E$ и $0 < |h| < \delta$, по свойству И2

$$f(x) - \varepsilon \leqslant \frac{\Phi(x+h) - \Phi(x)}{h} \leqslant f(x) + \varepsilon,$$

откуда и следует (4). $\square$

Замечание 1. Теорема о существовании первообразной у любой непрерывной на промежутке функции была сформулирована без доказательства в первой части курса (теорема 2 § 1 главы 4). Согласно теореме Барроу, первообразной является интеграл с переменным верхним пределом. Таким образом, существование первообразной сводится к существованию интеграла и, значит, остается недоказанным.

Замечание 2. Аналогично при $b \in E$ определяется *интеграл с переменным нижним пределом*

$$\Psi(x) = \int_x^b f, \quad x \in E.$$

Так как $\int_x^b f = -\int_b^x f$, из теоремы Барроу следует, что Ψ дифференцируема и $\Psi' = -f$ на E.

Следующая теорема — формула Ньютона – Лейбница — важнейшая в интегральном исчислении. Она устанавливает связь определенного интеграла с неопределенным и позволяет вычислять определенный интеграл от функции, первообразная которой известна.

Теорема 2. Формула Ньютона – Лейбница.
Пусть $f \in C[a,b]$, F — первообразная f на $[a,b]$. Тогда

$$\int_a^b f = F(b) - F(a).$$

Доказательство. Пусть $\Phi(x) = \int_a^x f$. По теореме Барроу Φ — первообразная f. Тогда F и Φ — две первообразные f на $[a,b]$ — отличаются на константу: $F = \Phi + C$. Поэтому

$$F(b) - F(a) = \Phi(b) + C - (\Phi(a) + C) = \Phi(b) - \Phi(a) = \\ = \int_a^b f - \int_a^a f = \int_a^b f. \quad \square$$

Замечание 1. Пусть $F\colon [a,b] \to \mathbb{R}$. Разность $F(b) - F(a)$ называется *двойной подстановкой* функции F на $[a,b]$ и обозначается $F\big|_a^b$, $F(x)\big|_a^b$, $F(x)\big|_{x=a}^b$ или $[F(x)]_{x=a}^b$.

Как обычно, переменная x здесь немая и может быть заменена другой буквой. Последняя запись позволяет отметить начало выражения для F, что бывает удобно, когда оно длинное.

Таким образом, формула Ньютона – Лейбница может быть записана в виде

$$\int_a^b f = F\big|_a^b.$$

Замечание 2. Двойная подстановка не зависит от выбора первообразной. Поэтому из формулы Ньютона – Лейбница следует единственность интеграла. В самом деле, интеграл (то есть функционал, удовлетворяющий аксиомам I1–I3) обязан равняться двойной подстановке первообразной.

Замечание 3. Теорему 2 можно переформулировать и так: *если $F \in C^1[a,b]$, то*

$$\int_a^b F' = F(b) - F(a).$$

И3. Линейность интеграла. *Если $f, g \in C[a,b]$, $\alpha, \beta \in \mathbb{R}$, то*

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g.$$

Доказательство. Пусть F и G — первообразные для f и g. Тогда $\alpha F + \beta G$ — первообразная для $\alpha f + \beta g$. Поэтому

$$\int_a^b (\alpha f + \beta g) = (\alpha F + \beta G)(b) - (\alpha F + \beta G)(a) =$$
$$= \alpha\bigl(F(b) - F(a)\bigr) + \beta\bigl(G(b) - G(a)\bigr) = \alpha \int_a^b f + \beta \int_a^b g. \quad \square$$

Замечание 1. По индукции свойство И3 распространяется на случай нескольких слагаемых.

Замечание 2. Линейность эквивалентна двум свойствам: *аддитивности по функции*

$$\int_a^b (f + g) = \int_a^b f + \int_a^b g$$

и *однородности*

$$\int_a^b \alpha f = \alpha \int_a^b f, \qquad \alpha \in \mathbb{R}.$$

§ 3′. Свойства интеграла Ньютона – Лейбница

В этом параграфе устанавливаются свойства интеграла, определенного в параграфах 2′ и 2″. Мы выяснили, что эти интегралы совпадают: для аксиоматически определенного интеграла верна формула Ньютона – Лейбница, а для интеграла Ньютона – Лейбница — аксиомы I1–I3. Далее для краткости мы будем использовать название "интеграл Ньютона – Лейбница".

По свойству монотонности интеграл от неотрицательной функции неотрицателен. Следующее свойство уточняет это замечание.

J1. *Пусть* $a < b$, $f \in C[a,b]$, $f \geqslant 0$, *и существует такая точка* $x_0 \in [a,b]$, *что* $f(x_0) > 0$. *Тогда*

$$\int_a^b f > 0.$$

Доказательство. Возьмем $\varepsilon = \frac{f(x_0)}{2} > 0$ и по определению непрерывности f в точке x_0 подберем такое число $\delta > 0$, что $f(x) > f(x_0) - \varepsilon = \frac{f(x_0)}{2}$ для всех $x \in [x_0 - \delta, x_0 + \delta] \cap [a,b]$; обозначим этот отрезок через $[\alpha, \beta]$. По свойствам аддитивности и монотонности

$$\int_a^b f = \int_a^\alpha f + \int_\alpha^\beta f + \int_\beta^b f \geqslant \int_\alpha^\beta f \geqslant (\beta - \alpha)\frac{f(x_0)}{2} > 0. \quad \square$$

Замечание 1. Утверждение, аналогичное J1, справедливо и для двух функций. *Пусть* $a < b$, $f, g \in C[a,b]$, $f \leqslant g$ *и существует такая точка* $x_0 \in [a,b]$, *что* $f(x_0) < g(x_0)$. *Тогда*

$$\int_a^b f < \int_a^b g.$$

Для доказательства достаточно применить J1 к функции $g - f$.

J2. *Пусть* $a < b$, $f \in C[a,b]$. *Тогда*

$$\left|\int_a^b f\right| \leqslant \int_a^b |f|.$$

Доказательство. Проинтегрируем по отрезку $[a, b]$ неравенство $-|f| \leqslant f \leqslant |f|$. Учитывая, что по свойству однородности $\int_a^b(-|f|) = -\int_a^b |f|$, имеем:

$$-\int_a^b |f| \leqslant \int_a^b f \leqslant \int_a^b |f|,$$

что равносильно доказываемому. □

Замечание 2. Если отказаться от требования $a < b$, то утверждение J2 можно записать так:

$$\left|\int_a^b f\right| \leqslant \left|\int_a^b |f|\right|.$$

Следующие два утверждения объединяются названием "первая теорема о среднем интегрального исчисления".

Теорема 1. Обобщенная теорема о среднем.
Пусть $f, g \in C[a, b]$, $g \geqslant 0$ (или $g \leqslant 0$). Тогда существует такое $c \in [a, b]$, что

$$\int_a^b fg = f(c) \int_a^b g.$$

Доказательство. Для определенности будем полагать, что $a < b$, $g \geqslant 0$; тогда $\int_a^b g \geqslant 0$. По теореме Вейерштрасса о непрерывных функциях существуют

$$m = \min_{x \in [a,b]} f(x) \quad \text{и} \quad M = \max_{x \in [a,b]} f(x).$$

Тогда

$$mg \leqslant fg \leqslant Mg.$$

Проинтегрируем это неравенство и вынесем постоянные множители за знаки интегралов:

$$m \int_a^b g \leqslant \int_a^b fg \leqslant M \int_a^b g.$$

Отсюда, если $\int_a^b g = 0$, то и $\int_a^b fg = 0$, а тогда подходит любое c. Если же $\int_a^b g > 0$, то положим

$$\mu = \frac{\int_a^b fg}{\int_a^b g}.$$

Так как $\mu \in [m, M]$, по теореме Больцано – Коши о промежуточном значении найдется такое $c \in [a, b]$, что $\mu = f(c)$. $\square$

Применив теорему 1 к функции $g \equiv 1$, мы получим теорему о среднем (без слова "обобщенная").

Следствие 1. Теорема о среднем. *Пусть* $f \in C[a, b]$. *Тогда существует такое* $c \in [a, b]$, *что*

$$\int_a^b f = f(c)(b - a).$$

Поясним еще термин "среднее" в названии теоремы.

Определение 1. Пусть $a < b$, $f \in C[a, b]$. Величина $\frac{1}{b-a} \int_a^b f$ называется *интегральным средним арифметическим* функции f на $[a, b]$.

Если $f(t)$ означает мгновенную скорость материальной точки в момент времени t, то $\frac{1}{b-a} \int_a^b f$ есть средняя скорость точки за время от a до b.

Замечание 1. Теорема о среднем утверждает, что среднее значение непрерывной функции на отрезке равно ее значению в некоторой точке. Обобщенной теореме о среднем можно придать тот же смысл, если рассматривать *взвешенное среднее* $\dfrac{\int_a^b fg}{\int_a^b g}$ ($g \geqslant 0$, $\int_a^b g > 0$) функции f.

Геометрически теорема о среднем означает, что площадь подграфика на рисунке 6 равна площади закрашенного прямоугольника.

Замечание 2. Можно доказать, что в условиях теоремы 1 и следствия 1 точка c найдется на интервале (a, b).

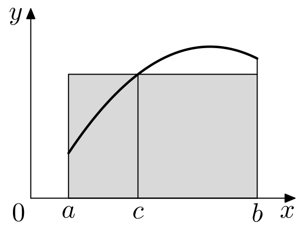

Рис. 6

Замечание 3. Теорема о среднем интегрального исчисления (для непрерывной подынтегральной функции) вытекает из формулы Ньютона – Лейбница и теоремы Лагранжа о среднем (см. § 4 главы 3), если применить последнюю к первообразной:

$$\int_a^b f = F(b) - F(a) = F'(c)(b-a) = f(c)(b-a).$$

Обобщенная теорема о среднем интегрального исчисления (для непрерывных подынтегральных функций) при дополнительном условии, что g не обращается в нуль, вытекает из формулы Ньютона – Лейбница и теоремы Коши о среднем. Если обозначить через H и G первообразные fg и g, то

$$\frac{\int_a^b fg}{\int_a^b g} = \frac{H(b) - H(a)}{G(b) - G(a)} = \frac{H'(c)}{G'(c)} = \frac{f(c)g(c)}{g(c)} = f(c).$$

При вычислении определенных, как и неопределенных, интегралов полезны приемы *интегрирования по частям* и *замены переменной.*

Теорема 2. Интегрирование по частям в определенном интеграле. *Пусть* $f, g \in C^1[a,b]$. *Тогда*

$$\int_a^b fg' = fg\Big|_a^b - \int_a^b f'g.$$

Доказательство. По формуле Ньютона – Лейбница

$$\int_a^b fg' + \int_a^b f'g = \int_a^b (fg)' = fg\Big|_a^b.$$

Остается перенести второе слагаемое из левой части в правую. □

Замечание 1. Формулу интегрирования по частям записывают и в виде

$$\int_a^b f\,dg = fg\Big|_a^b - \int_a^b g\,df,$$

трактуя $f'(x)\,dx$ и $g'(x)\,dx$ как дифференциалы. Аналогичная форма записи для неопределенных интегралов обсуждалась в § 1.

Пример 1. Полагая $f(x) = \ln x$, $g'(x) = 1$, находим:

$$\int_1^2 \ln x\,dx = x\,\ln x\Big|_1^2 - \int_1^2 x\,\frac{dx}{x} = 2\ln 2 - 1.$$

Теорема 3. Замена переменной в определенном интеграле. *Пусть* $\varphi \in C^1\big([\alpha,\beta] \to [A,B]\big)$, $f \in C[A,B]$. *Тогда*

$$\int_\alpha^\beta (f\circ\varphi)\varphi' = \int_{\varphi(\alpha)}^{\varphi(\beta)} f.$$

Доказательство. Пусть F — первообразная f на $[A,B]$. Тогда по правилу дифференцирования композиции $F\circ\varphi$ — первообразная $(f\circ\varphi)\varphi'$ на $[A,B]$. Применяя к обоим интегралам формулу Ньютона – Лейбница, получаем:

$$\int_\alpha^\beta (f\circ\varphi)\varphi' = F\circ\varphi\Big|_\alpha^\beta = F\Big|_{\varphi(\alpha)}^{\varphi(\beta)} = \int_{\varphi(\alpha)}^{\varphi(\beta)} f. \quad \square$$

Замечание 2. Правило замены переменной может применяться как слева направо, так и справа налево. Допустим, что в интеграле $\int_a^b f(x)\,dx$ мы хотим сделать замену $x = \varphi(t)$, где $\varphi(\alpha) = a$, $\varphi(\beta) = b$.

Тогда надо трактовать dx как дифференциал: $dx = \varphi'(t)\,dt$, и поменять пределы интегрирования: a на α и b на β. Получим:

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt.$$

В отличие от неопределенного интеграла, при вычислении определенного не надо возвращаться к старой переменной, но надо не забыть поменять пределы интегрирования.

Замечание 3. В условиях теоремы некоторые значения $\varphi(t)$ при $t \in [\alpha, \beta]$ могут не принадлежать отрезку $[\varphi(\alpha), \varphi(\beta)]$. Важно, что они принадлежат отрезку $[A, B]$, на котором определена функция f.

В формуле замены переменной особенно удобно соглашение о том, что нижний предел интегрирования не обязательно меньше верхнего. Например, если φ строго убывает, а $\alpha < \beta$, то $\varphi(\alpha) > \varphi(\beta)$.

Пример 2. Вычислим интеграл $\int_0^a \sqrt{a^2 - x^2}\,dx$ $(a > 0)$ с помощью тригонометрической подстановки $x = a\sin t$. Функция $\varphi(t) = a\sin t$ отображает $\left[0, \frac{\pi}{2}\right]$ на $[0, a]$, $\varphi(0) = 0$, $\varphi\left(\frac{\pi}{2}\right) = a$, $\varphi'(t) = a\cos t$. Поэтому

$$\int_0^a \sqrt{a^2 - x^2}\,dx = \int_0^{\pi/2} a^2\cos^2 t\,dt = a^2\int_0^{\pi/2} \frac{1 + \cos 2t}{2}\,dt =$$
$$= \frac{a^2}{2}\left[t + \frac{\sin 2t}{2}\right]_0^{\pi/2} = \frac{\pi a^2}{4}.$$

Трактуя интеграл как площадь, можно сразу сказать, что интеграл в этом примере равен площади четверти круга радиуса a.

Замечание 4. Если $f \in C[-a, 0]$, то, полагая $x = -t$, находим

$$\int_{-a}^0 f(x)\,dx = \int_a^0 f(-t)(-1)\,dt = \int_0^a f(-t)\,dt.$$

Поэтому для $f \in C[-a, a]$

$$\int_{-a}^a f = \int_{-a}^0 f + \int_0^a f = \int_0^a \big(f(t) + f(-t)\big)\,dt.$$

Следовательно, если f четна, то

$$\int_{-a}^{a} f = 2\int_{0}^{a} f,$$

а если f нечетна, то

$$\int_{-a}^{a} f = 0.$$

Эти простые соображения часто облегчают вычисление интегралов.

По аддитивности определение интеграла можно распространить на кусочно-непрерывные функции. Определим их.

Определение 2. Функция $f\colon [a,b] \to \mathbb{R}$ называется *кусочно-непрерывной* на $[a,b]$, если множество ее точек разрыва пусто или конечно, и все имеющиеся разрывы — первого рода.

Интеграл от кусочно-непрерывной функции определяется так. Если f непрерывна на (a,b), а в точках a и b имеет конечные односторонние пределы, то полагают $\int_a^b f = \int_a^b \widetilde{f}$, где $\widetilde{f}$ — функция, доопределенная по непрерывности: $\widetilde{f} = f$ на (a,b), $\widetilde{f}(a) = f(a+)$, $\widetilde{f}(b) = f(b-)$.

Если f кусочно-непрерывна на $[a,b]$, $x_1, \ldots, x_{n-1}$ — все ее точки разрыва на (a,b), $a = x_0 < x_1 < \ldots < x_n = b$, то для каждого отрезка $[x_k, x_{k+1}]$ функция f непрерывна во внутренних точках и имеет конечные односторонние пределы на концах. Поэтому $\int_{x_k}^{x_{k+1}} f$ уже определен. Теперь можно положить

$$\int_a^b f = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f.$$

Замечание 1. Из определения и свойства аддитивности ясно, что если значения кусочно-непрерывной функции изменить на конечном множестве точек, то функция останется кусочно-непрерывной и ее интеграл не изменится. Поэтому можно в определении кусочно-непрерывной функции считать, что функция задана на $[a,b]$, за исключением, быть может, конечного множества точек, и определять интеграл от такой функции.

Замечание 2. Свойства аддитивности по отрезку, линейности, монотонности и оценки интеграла очевидным образом переносятся на интеграл от кусочно-непрерывной функции.

В утверждение J1 об интегрировании строгих неравенств надо внести небольшое изменение.

J1′. *Пусть $a < b$, f кусочно-непрерывна на $[a, b]$, $f \geqslant 0$ и существует такая точка $x_0 \in [a, b]$, что $f(x_0) > 0$ и f непрерывна в точке x_0. Тогда*

$$\int_a^b f > 0.$$

Аналогичное уточнение верно и для замечания 1 к свойству J1. Доказательства сохраняют силу.

Замечание 3. Без условия непрерывности f в точке x_0 утверждение неверно. Контрпримером служит функция, равная 0 всюду, кроме одной точки, в которой она положительна.

Замечание 4. Теперь интеграл с переменным верхним пределом $\Phi(x) = \int_a^x f$ определен для функций $f\colon E \to \mathbb{R}$, кусочно-непрерывных на каждом отрезке, содержащемся в промежутке E. Из теоремы Барроу следует, что Φ дифференцируема и $\Phi' = f$ во всех точках непрерывности f. В точках же разрыва f функция Φ имеет конечные односторонние производные, которые равны односторонним пределам f. Аналогичными свойствами ($\Psi' = -f$) обладает и интеграл Ψ с переменным нижним пределом.

Формула Ньютона – Лейбница также допускает некоторое обобщение.

Замечание 5. *Пусть функция f кусочно-непрерывна на $[a, b]$, $F \in C[a, b]$, F — первообразная f на $[a, b]$ за вычетом конечного множества точек. Тогда*

$$\int_a^b f = F(b) - F(a).$$

Доказательство. Пусть $\alpha_1 < \ldots < \alpha_{m-1}$ — все точки интервала (a, b), в которых нарушается равенство $F' = f$; положим также $\alpha_0 = a$, $\alpha_m = b$. По следствию из теоремы Дарбу (см. § 7 главы 3) функция F' не имеет скачков на своей области определения. Тогда функция f может иметь разрывы (а они первого рода) только в точках, где она не совпадает с F', каковыми могут быть лишь точки α_k. Пользуясь последовательно непрерывностью интеграла с переменными пределами интегрирования, формулой Ньютона – Лейбница и непрерывностью F, находим:

$$\int_{\alpha_k}^{\alpha_{k+1}} f = \lim_{\varepsilon\to 0+} \int_{\alpha_k+\varepsilon}^{\alpha_{k+1}-\varepsilon} f = \\ = \lim_{\varepsilon\to 0+} \big(F(\alpha_{k+1}-\varepsilon) - F(\alpha_k+\varepsilon)\big) = F(\alpha_{k+1}) - F(\alpha_k).$$

Остается воспользоваться аддитивностью интеграла:

$$\int_a^b f = \sum_{k=0}^{m-1} \int_{\alpha_k}^{\alpha_{k+1}} f = \sum_{k=0}^{m-1} \big(F(\alpha_{k+1}) - F(\alpha_k)\big) = F(b) - F(a). \quad \square$$

Замечание 6. Условие $F \in C[a,b]$ в замечании 5 существенно. Для функций $f(x) = 0$, $F(x) = \operatorname{sign} x$ формула Ньютона – Лейбница на $[-1,1]$ неверна:

$$0 = \int_{-1}^{1} f \neq F\Big|_{-1}^{1} = 2.$$

Замечание 7. В равенстве

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

можно разрешить функции $F \in C[a,b]$ иметь кусочно-непрерывную производную, определенную на $[a,b]$ за вычетом конечного множества точек.

Обобщенная теорема о среднем (теорема 1 этого параграфа) в такой формулировке сохраняет силу и для кусочно-непрерывной функции g. Если отказаться и от непрерывности f, теорема о среднем будет формулироваться так.

Теорема 1′. *Пусть функции f, g кусочно-непрерывны на $[a,b]$, $g \geqslant 0$ (или $g \leqslant 0$), $m, M \in \mathbb{R}$, $m \leqslant f \leqslant M$. Тогда существует такое $\mu \in [m, M]$, что*

$$\int_a^b fg = \mu \int_a^b g.$$

Доказательство. Для определенности мы будем полагать, что $a < b$, $g \geqslant 0$. Тогда $\int_a^b g \geqslant 0$ и

$$mg \leqslant fg \leqslant Mg.$$

Проинтегрируем это неравенство и вынесем постоянные множители за знаки интегралов:

$$m \int_a^b g \leqslant \int_a^b fg \leqslant M \int_a^b g.$$

Отсюда, если $\int_a^b g = 0$, то и $\int_a^b fg = 0$, а тогда подходит любое μ. Если же $\int_a^b g > 0$, то следует положить

$$\mu = \frac{\int_a^b fg}{\int_a^b g}.$$

Условия на μ, очевидно, выполнены. $\square$

Следствие 1$'$. *Пусть функция f кусочно-непрерывна на $[a, b]$, $m, M \in \mathbb{R}$, $m \leqslant f \leqslant M$. Тогда существует такое $\mu \in [m, M]$, что*

$$\int_a^b f = \mu(b - a).$$

Интеграл тесно связан с суммами специального вида, к описанию которых мы и переходим.

Определение 3. Набор точек

$$\tau = \{x_k\}_{k=0}^n : \quad a = x_0 < x_1 < \ldots < x_n = b$$

называется *дроблением* или *разбиением* отрезка $[a, b]$. Отрезки $[x_k, x_{k+1}]$ ($k \in [0 : n-1]$) называют отрезками дробления, через Δx_k обозначается длина k-го отрезка дробления: $\Delta x_k = x_{k+1} - x_k$. Величина

$$\lambda = \lambda_\tau = \max_{0 \leqslant k \leqslant n-1} \Delta x_k,$$

то есть наибольшая из длин отрезков дробления, называется *рангом* или *мелкостью* дробления τ. Набор точек $\xi = \{\xi_k\}_{k=0}^{n-1}$, таких что $\xi_k \in [x_k, x_{k+1}]$ при всех $k \in [0 : n-1]$, называется *оснащением* дробления. Дробление вместе с его оснащением, то есть пара (τ, ξ), называется *оснащенным дроблением*.

Пусть $f \in C[a,b]$, $\tau = \{x_k\}_{k=0}^{n}$ — дробление $[a,b]$. По теореме о среднем найдутся такие точки $c_k \in [x_k, x_{k+1}]$, что

$$\int_a^b f = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f = \sum_{k=0}^{n-1} f(c_k)\Delta x_k. \tag{5}$$

Если заменить c_k на произвольные точки $\xi_k \in [x_k, x_{k+1}]$, то получившаяся сумма уже не обязана равняться интегралу. Однако, если ранг дробления мал, то точки c_k и ξ_k близки, а тогда в силу непрерывности f близки и значения $f(c_k)$ и $f(\xi_k)$. Можно предположить, что и сумма будет мало отличаться от интеграла. Дадим точные определения.

Определение 4. Суммы Римана. Пусть $f\colon [a,b] \to \mathbb{R}$. Суммы

$$\sigma = \sigma_\tau(f, \xi) = \sum_{k=0}^{n-1} f(\xi_k)\Delta x_k$$

называются *интегральными суммами* или *суммами Римана* функции f, отвечающими оснащенному дроблению (τ, ξ).

Определение 5. Предел интегральных сумм. Пусть $f\colon [a,b] \to \mathbb{R}$. Число $I \in \mathbb{R}$ называют *пределом интегральных сумм* при ранге дробления, стремящемся к нулю, и пишут

$$I = \lim_{\lambda_\tau \to 0} \sigma_\tau(f, \xi) \quad \text{или} \quad I = \lim_{\lambda \to 0} \sigma,$$

если

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall \tau : \lambda_\tau < \delta \ \ \forall \xi \ \ |\sigma_\tau(f, \xi) - I| < \varepsilon,$$

то есть для любого положительного числа ε существует такое положительное число δ, что для любого оснащенного дробления (τ, ξ) ранга, меньшего чем δ, интегральная сумма отличается от числа I меньше чем на ε.

Теорема 4. Интеграл как предел интегральных сумм. *Если* $f \in C[a,b]$, *то*

$$\lim_{\lambda\to 0} \sigma = \int_a^b f.$$

Доказательство. Возьмем $\varepsilon > 0$. По теореме Кантора функция f равномерно непрерывна на $[a,b]$. Поэтому найдется такое $\delta > 0$, что если $s, t \in [a,b]$, $|s-t| < \delta$, то $|f(s) - f(t)| < \frac{\varepsilon}{b-a}$. Покажем, что это δ — требуемое. Если $\lambda_\tau < \delta$, то по теореме о среднем и формуле (5) при любом оснащении ξ

$$\left|\sigma - \int_a^b f\right| = \left|\sum_{k=0}^{n-1} f(\xi_k)\Delta x_k - \sum_{k=0}^{n-1} f(c_k)\Delta x_k\right| \leqslant$$
$$\leqslant \sum_{k=0}^{n-1} |f(\xi_k) - f(c_k)|\Delta x_k < \sum_{k=0}^{n-1} \frac{\varepsilon}{b-a}\Delta x_k = \varepsilon.$$

Это и означает, что $\lim\limits_{\lambda\to 0} \sigma = \int_a^b f$. $\square$

Замечание 1. Легко доказать, что теорема 4 остается верной и для кусочно-непрерывных функций.

Замечание 2. Свойство, выраженное теоремой 4, можно принять за определение интеграла.

Определение 6. Интеграл Римана. Пусть $f\colon [a,b] \to \mathbb{R}$. Если существует предел интегральных сумм $\lim\limits_{\lambda\to 0} \sigma$, равный числу I, то функция f называется *интегрируемой по Риману* на $[a,b]$, а число I — *интегралом* (*определенным интегралом, интегралом Римана*) от функции f по отрезку $[a,b]$.

Множество интегрируемых по Риману на $[a,b]$ функций обозначается через $R[a,b]$.

Класс интегрируемых по Риману функций оказывается шире класса кусочно-непрерывных функций. В частности, он содержит все монотонные функции, а такие функции могут иметь счетное число скачков.

Интеграл Римана рассматривается в параграфах 2 и 3.

Замечание 3. Интегральные суммы используются для приближенного вычисления интегралов, которое может понадобиться в тех случаях, когда первообразная неэлементарна или громоздка.

§ 4. Формулы Тейлора и Валлиса и интегральные неравенства

Теорема 1. Формула Тейлора с остатком в интегральной форме. *Пусть* $n \in \mathbb{Z}_+$, $f \in C^{n+1}\langle A, B\rangle$, $a, x \in \langle A, B\rangle$. *Тогда*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt.$$

Доказательство проведем по индукции. База индукции (случай $n = 0$) представляет собой формулу Ньютона – Лейбница:

$$f(x) = f(a) + \int_a^x f'(t)\,dt.$$

Пусть утверждение верно для некоторого $n - 1 \in \mathbb{Z}_+$. Докажем его для номера n. Для этого проинтегрируем по частям в остаточном члене:

$$\int_a^x f^{(n)}(t)\frac{(x-t)^{n-1}}{(n-1)!}\,dt = \int_a^x f^{(n)}(t)\,d\left(-\frac{(x-t)^n}{n!}\right) =$$
$$= -\frac{1}{n!}\left[f^{(n)}(t)(x-t)^n\right]_{t=a}^{x} + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt =$$
$$= \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt.$$

Первое слагаемое в правой части есть слагаемое с номером n в многочлене Тейлора, а второе — новый остаточный член:

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k +$$
$$+ \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt =$$
$$= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt. \quad \square$$

Замечание 1. Интегральную форму остатка иногда называют *формой К. Якоби.*

Замечание 2. По первой теореме о среднем в условиях теоремы 1 найдется такая точка $c \in [a, x]$, что

$$\frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n\,dt = \frac{f^{(n+1)}(c)}{n!}\int_a^x (x-t)^n\,dt = \\ = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

По замечанию 1 к теореме о среднем при $x \neq a$ точку c можно выбрать на (a, x). Таким образом, лагранжева форма остатка следует из интегральной (правда, при более ограничительных условиях на функцию). Интегральная форма остатка имеет то преимущество, что она не содержит неизвестной точки c.

Далее мы выведем формулу Валлиса, которая выражает число π в виде предела последовательности рациональных чисел.

Введем стандартное обозначение $m!!$ — *двойной факториал* числа m. При $m \in \mathbb{N}$ это произведение всех натуральных чисел, не превосходящих m, одной четности с m; кроме того, положим $0!! = (-1)!! = 1$.

Лемма 1. *Если* $m \in \mathbb{Z}_+$, *то*

$$\int_0^{\pi/2} \sin^m x\,dx = \frac{(m-1)!!}{m!!}\cdot\begin{cases} \dfrac{\pi}{2}, & m \text{ четно}, \\ 1, & m \text{ нечетно}. \end{cases}$$

Доказательство. Обозначим $J_m = \displaystyle\int_0^{\pi/2} \sin^m t\,dt$. Легко проверить, что $J_0 = \frac{\pi}{2}$, $J_1 = 1$. При $m - 1 \in \mathbb{N}$ проинтегрируем по частям:

$$J_m = \int_0^{\pi/2} \sin^{m-1} x\,d(-\cos x) = -\sin^{m-1} x\,\cos x\Big|_0^{\pi/2} + \\ + (m-1)\int_0^{\pi/2} \sin^{m-2} x\cos^2 x\,dx = (m-1)(J_{m-2} - J_m)$$

(в последнем равенстве мы учли, что двойная подстановка обнулилась, и применили формулу $\cos^2 x = 1 - \sin^2 x$). Выражая J_m, получаем рекуррентное соотношение

$$J_m = \frac{m-1}{m} J_{m-2}.$$

Остается применить его несколько раз и выразить J_m через J_0 или J_1 в зависимости от четности m. $\square$

Теорема 2. Формула Валлиса.

$$\pi = \lim_{n\to\infty} \frac{1}{n} \left(\frac{(2n)!!}{(2n-1)!!} \right)^2 .$$

Доказательство. При всех $x \in \left(0, \frac{\pi}{2}\right)$ выполняется неравенство $0 < \sin x < 1$, поэтому для любого $n \in \mathbb{N}$

$$\sin^{2n+1} x < \sin^{2n} x < \sin^{2n-1} x,$$

а тогда и

$$J_{2n+1} < J_{2n} < J_{2n-1}.$$

Применяя лемму 1, получаем двойное неравенство

$$\frac{(2n)!!}{(2n+1)!!} < \frac{(2n-1)!!}{(2n)!!} \cdot \frac{\pi}{2} < \frac{(2n-2)!!}{(2n-1)!!},$$

что равносильно

$$\frac{1}{2n+1} \left(\frac{(2n)!!}{(2n-1)!!} \right)^2 < \frac{\pi}{2} < \frac{1}{2n} \left(\frac{(2n)!!}{(2n-1)!!} \right)^2 .$$

Обозначим $x_n = \frac{1}{n} \left(\frac{(2n)!!}{(2n-1)!!} \right)^2$. Двойное неравенство можно преобразовать к виду

$$\pi < x_n < \frac{2n+1}{2n} \pi,$$

откуда $x_n \to \pi$. $\square$

Теорема 3 (О. Бонне). Вторая теорема о среднем интегрального исчисления. *Пусть* $f \in C[a,b]$, $g \in C^1[a,b]$, g *монотонна на* $[a,b]$. *Тогда существует такое* $c \in [a,b]$, *что*

$$\int_a^b fg = g(a)\int_a^c f + g(b)\int_c^b f.$$

Доказательство. Положим $F(x) = \int_a^x f$. Тогда $F' = f$, $F(a) = 0$. Проинтегрируем по частям:

$$\int_a^b fg = \int_a^b gF' = Fg\Big|_a^b - \int_a^b Fg' = g(b)\int_a^b f - \int_a^b Fg'.$$

Поскольку g монотонна, g' сохраняет знак на $[a,b]$. По обобщенной первой теореме о среднем найдется такое $c \in [a,b]$, что

$$\int_a^b Fg' = F(c)\int_a^b g' = \big(g(b) - g(a)\big)\int_a^c f.$$

Группируя слагаемые с множителем $g(b)$ и пользуясь аддитивностью интеграла, получаем требуемое:

$$\int_a^b fg = g(b)\int_a^b f - \big(g(b)-g(a)\big)\int_a^c f = g(a)\int_a^c f + g(b)\int_c^b f. \quad \square$$

Замечание 1. Отметим без доказательства, что теорема Бонне верна при более слабых предположениях: $f \in R[a,b]$, g монотонна на $[a,b]$.

Замечание 2. Если в условиях замечания 1 функция g убывает и $g \geqslant 0$, то существует такое $c \in [a,b]$, что

$$\int_a^b fg = g(a)\int_a^c f.$$

а если g возрастает и $g \geqslant 0$, то существует такое $c \in [a,b]$, что

$$\int_a^b fg = g(b)\int_c^b f.$$

Для доказательства надо применить замечание 1 к функциям f и соответственно

$$g_1(x) = \begin{cases} g(x), & x \in [a, b), \\ 0, & x = b, \end{cases} \qquad g_2(x) = \begin{cases} g(x), & x \in (a, b], \\ 0, & x = a, \end{cases}$$

которые по-прежнему монотонны.

Вторую теорему о среднем бывает удобно применять для оценки интегралов от колеблющихся функций.

Пример. Оценим интеграл

$$I = \int_{100\pi}^{200\pi} \frac{\sin x}{x}\, dx.$$

Заметим, что интеграл неберущийся, поэтому вычислить его явно по формуле Ньютона – Лейбница не удается. Оценка модуля подынтегральной функции (или первая теорема о среднем) дает неравенство

$$|I| \leqslant \int_{100\pi}^{200\pi} \frac{dx}{x} = \ln 200\pi - \ln 100\pi = \ln 2. \tag{6}$$

Применим вторую теорему о среднем, положив $f(x) = \sin x$, $g(x) = \frac{1}{x}$:

$$\begin{aligned} I &= \frac{1}{100\pi} \int_{100\pi}^{c} \sin x\, dx + \frac{1}{200\pi} \int_{c}^{200\pi} \sin x\, dx = \\ &= \frac{1 - \cos c}{100\pi} + \frac{\cos c - 1}{200\pi} = \frac{1 - \cos c}{200\pi}. \end{aligned}$$

Учитывая, что $\cos c \in [-1, 1]$, получаем оценку

$$0 \leqslant I \leqslant \frac{1}{100\pi}, \tag{7}$$

что гораздо точнее, чем (6).

Замечание 3. Вторая теорема о среднем обычно не дает новых результатов по сравнению с теми, которые можно получить интегрированием по частям. В приведенном примере

$$I = \int_{100\pi}^{200\pi} \frac{d(1 - \cos x)}{x} = \int_{100\pi}^{200\pi} \frac{1 - \cos x}{x^2} dx$$

(двойная подстановка обнуляется). Поэтому

$$0 < I < 2\int_{100\pi}^{200\pi} \frac{dx}{x^2} = \frac{1}{100\pi},$$

и мы снова получили неравенство (7), даже строгое.

Далее мы установим несколько интегральных неравенств, аналогичных доказанным в главе 3 для сумм. В их формулировках мы будем подразумевать, что $a < b$. Доказательство этих неравенств может быть проведено или повторением рассуждений для сумм, или предельным переходом из неравенств для сумм. Мы проиллюстрируем оба способа.

Участвующие в этих неравенствах функции будут предполагаться непрерывными, что обеспечит существование интегралов. В ряде случаев требование непрерывности может быть ослаблено, но мы не будем на этом останавливаться.

Теорема 4. Неравенство Иенсена для интегралов. *Пусть f выпукла и непрерывна на $\langle A, B\rangle$, $\varphi \in C\big([a,b] \to \langle A, B\rangle\big)$, $\lambda \in C\big([a,b] \to [\,0,+\infty)\big)$, $\int_a^b \lambda = 1$. Тогда*

$$f\left(\int_a^b \lambda\varphi\right) \leqslant \int_a^b \lambda\cdot(f\circ\varphi).$$

Доказательство. Обозначим

$$c = \int_a^b \lambda\varphi, \qquad E = \big\{x \in [a,b] : \lambda(x) > 0\big\},$$

$$m = \inf_E \varphi, \qquad M = \sup_E \varphi$$

(m и M конечны по теореме Вейерштрасса). Если $m = M$, то есть φ постоянна на E, то $c = m$ и обе части неравенства Иенсена равны $f(m)$.

Пусть $m < M$. Тогда $c \in (m, M)$ и, следовательно, $c \in (A, B)$. Функция f имеет в точке c опорную прямую (см. § 8 главы 3); пусть она задается уравнением $y = \alpha x + \beta$. По определению опорной прямой $f(c) = \alpha c + \beta$ и $f(t) \geqslant \alpha t + \beta$ при всех $t \in \langle A, B\rangle$. Поэтому

$$f(c) = \alpha c + \beta = \alpha \int_a^b \lambda\varphi + \beta \int_a^b \lambda = \\ = \int_a^b \lambda \cdot (\alpha\varphi + \beta) \leqslant \int_a^b \lambda \cdot (f \circ \varphi). \quad \square$$

Замечание 1. Если φ не постоянна на E, а f строго выпукла, то неравенство Иенсена строгое.

Действительно, в этом случае опорная прямая в точке c будет строго опорной, то есть при всех $t \in \langle A, B\rangle \setminus \{c\}$ будет выполняться строгое неравенство $f(t) > \alpha t + \beta$. Остается воспользоваться свойством И5 § 3 (или свойством J1 § 3′), согласно которому интеграл от неотрицательной функции, положительной хотя бы в одной точке непрерывности, положителен.

Замечание 2. Для вогнутой функции f неравенство Иенсена выполняется с противоположным знаком.

Напомним, что *сопряженными показателями* называются числа p и q из $(1, +\infty)$, связанные соотношением $\frac{1}{p} + \frac{1}{q} = 1$.

Теорема 5. Неравенство Гёльдера для интегралов. *Пусть $f, g \in C[a, b]$, p и q — сопряженные показатели. Тогда*

$$\left|\int_a^b fg\right| \leqslant \left(\int_a^b |f|^p\right)^{1/p} \left(\int_a^b |g|^q\right)^{1/q}.$$

Доказательство. Положим $x_k = a + \frac{k(b-a)}{n}$ $(k \in [0 : n])$, $a_k = f(x_k)(\Delta x_k)^{1/p}$, $b_k = g(x_k)(\Delta x_k)^{1/q}$ $(k \in [0 : n-1])$. Тогда $a_k b_k = f(x_k) g(x_k) \Delta x_k$ в силу равенства $\frac{1}{p} + \frac{1}{q} = 1$. Воспользуемся неравенством Гёльдера для сумм:

$$\left|\sum_{k=0}^{n-1} a_k b_k\right| \leqslant \left(\sum_{k=0}^{n-1} |a_k|^p\right)^{1/p} \left(\sum_{k=0}^{n-1} |b_k|^q\right)^{1/q},$$

которое принимает вид

$$\left|\sum_{k=0}^{n-1} f(x_k) g(x_k) \Delta x_k\right| \leqslant \left(\sum_{k=0}^{n-1} |f(x_k)|^p \Delta x_k\right)^{1/p} \left(\sum_{k=0}^{n-1} |g(x_k)|^q \Delta x_k\right)^{1/q}.$$

В последнем неравенстве участвуют суммы Римана для непрерывных функций fg, $|f|^p$ и $|g|^q$. При $n \to \infty$ суммы стремятся к интегралам от этих функций. Остается сделать предельный переход в неравенстве и воспользоваться непрерывностью модуля и степенных функций. $\square$

Замечание 3. Другой способ доказать неравенство Гёльдера — воспользоваться неравенством Юнга (теорема 8 § 8 главы 3), как это было сделано для сумм.

Следствие. Неравенство Коши – Буняковского для интегралов. *Пусть* $f, g \in C[a, b]$. *Тогда*

$$\left|\int_a^b fg\right| \leqslant \sqrt{\int_a^b f^2} \cdot \sqrt{\int_a^b g^2}.$$

Для доказательства следствия надо положить в неравенстве Гёльдера $p = q = 2$.

Теорема 6. Неравенство Минковского для интегралов. *Пусть* $f, g \in C[a, b]$, $p \geqslant 1$. *Тогда*

$$\left(\int_a^b |f+g|^p\right)^{1/p} \leqslant \left(\int_a^b |f|^p\right)^{1/p} + \left(\int_a^b |g|^p\right)^{1/p}.$$

Для доказательства неравенства Минковского можно сделать предельный переход в неравенстве для сумм. Подробности мы оставляем читателю.

Следствие. *Пусть* $f, g \in C[a, b]$. *Тогда*

$$\sqrt{\int_a^b (f+g)^2} \leqslant \sqrt{\int_a^b f^2} + \sqrt{\int_a^b g^2}.$$

Для доказательства следствия надо положить в неравенстве Минковского $p = 2$.

Неравенство Коши между средним арифметическим и средним геометрическим также имеет интегральный аналог. Определение среднего арифметического функции уже было дано при обсуждении теоремы о среднем. Напомним его, а также определим среднее геометрическое.

Определение 1. Пусть $f \in C[a,b]$.

1. Величина

$$\frac{1}{b-a}\int_a^b f$$

называется *интегральным средним арифметическим* функции f на $[a,b]$.

2. Если $f > 0$, то величина

$$\exp\left(\frac{1}{b-a}\int_a^b \ln f\right)$$

называется *интегральным средним геометрическим* функции f на $[a,b]$.

Эти величины суть пределы при $n \to \infty$ последовательностей средних арифметических

$$\frac{1}{n}\sum_{k=0}^{n-1} f(x_k) = \frac{1}{b-a}\sum_{k=0}^{n-1} f(x_k)\Delta x_k$$

и средних геометрических

$$\sqrt[n]{\prod_{k=0}^{n-1} f(x_k)} = \exp\left(\frac{1}{n}\sum_{k=0}^{n-1} \ln f(x_k)\right) = \exp\left(\frac{1}{b-a}\sum_{k=0}^{n-1} \ln f(x_k)\,\Delta x_k\right)$$

значений функции f в точках $x_k = a + \frac{k(b-a)}{n}$ $(k \in [\,0 : n-1])$.

Теорема 7. Неравенство для интегральных средних. *Пусть* $f \in C[a,b]$, $f > 0$. *Тогда*

$$\exp\left(\frac{1}{b-a}\int_a^b \ln f\right) \leqslant \frac{1}{b-a}\int_a^b f.$$

Доказательство проводится предельным переходом или применением интегрального неравенства Иенсена к вогнутой функции ln.

Теорема 8. Неравенство Чебышёва для интегралов. *Пусть f возрастает, а g убывает на $[a,b]$. Тогда*

$$\frac{1}{b-a}\int_a^b fg \leqslant \left(\frac{1}{b-a}\int_a^b f\right)\cdot\left(\frac{1}{b-a}\int_a^b g\right).$$

Другими словами, *среднее арифметическое от произведения разноименно монотонных функций не превосходит произведения средних.*

Замечание 1. Если читатель знаком с определением интеграла только от кусочно-непрерывных функций, то он может считать и здесь их кусочно-непрерывными.

Доказательство. Обозначим

$$A = \frac{1}{b-a}\int_a^b f, \qquad E = \{x \in [a,b] : f(x) \leqslant A\}.$$

Ясно, что $E \neq \varnothing$, так как в противном случае $f > A$ на $[a,b]$, что приводит к абсурдному неравенству $A > A$. Положим $c = \sup E$. Тогда $A - f \geqslant 0$, $g \geqslant g(c)$ на $[a,c)$ и $A - f \leqslant 0$, $g \leqslant g(c)$ на $(c,b]$. Поэтому

$$\int_a^b (A-f)g = \int_a^c (A-f)g + \int_c^b (A-f)g \geqslant$$
$$\geqslant g(c)\int_a^c (A-f) + g(c)\int_c^b (A-f) = g(c)\int_a^b (A-f) = 0,$$

что равносильно доказываемому. □

Покажем на примере неравенства Чебышёва, что иногда неравенства для сумм могут быть получены, в свою очередь, как частные случаи интегральных неравенств.

Следствие. Неравенство Чебышёва для сумм. *Пусть $n \in \mathbb{N}$, $a, b \in \mathbb{R}^n$, $a_1 \leqslant \ldots \leqslant a_n$, $b_1 \geqslant \ldots \geqslant b_n$. Тогда*

$$\frac{1}{n}\sum_{k=1}^n a_k b_k \leqslant \left(\frac{1}{n}\sum_{k=1}^n a_k\right)\cdot\left(\frac{1}{n}\sum_{k=1}^n b_k\right).$$

Для доказательства следует записать интегральное неравенство Чебышёва для кусочно-постоянных функций $f, g\colon [0,1] \to \mathbb{R}$, равных соответственно a_k и b_k на промежутках $\left(\frac{k-1}{n}, \frac{k}{n}\right)$ (значения функций на конечном множестве точек несущественны).

Замечание 2. Для одноименно монотонных функций f и g неравенство Чебышёва выполняется с противоположным знаком.

Этот случай сводится к разобранному рассмотрением функций f и $-g$.

§ 5. Несобственные интегралы

Задача нахождения площадей неограниченных фигур требует расширения понятия интеграла.

Определение 1. Функция f называется *локально интегрируемой* (по Риману) на промежутке E, если f интегрируема (по Риману) на каждом отрезке, содержащемся в E. Множество функций, локально интегрируемых на E, обозначается через $R_{loc}(E)$.

Из теоремы 2 § 2 ясно, что $C(E) \subset R_{loc}(E)$. Как правило, мы будем формулировать утверждения для локально интегрируемых функций. Читатель, не изучавший свойства функций, интегрируемых по Риману, может считать, что подынтегральные функции непрерывны.

Определение 2. Несобственный интеграл.
Пусть $-\infty < a < b \leqslant +\infty$, $f \in R_{loc}[a,b)$. Символ $\int_a^{\to b} f$ называется *несобственным интегралом.* Интегралы $\int_a^A f$ при $A \in [a,b)$ называются *частными* или *частичными.* Если существует предел $\lim\limits_{A\to b-} \int_a^A f$ в $\overline{\mathbb{R}}$, равный I, то символу $\int_a^{\to b} f$ приписывают значение I. В противном случае символу $\int_a^{\to b} f$ не приписывают никакого значения. Если $I \in \mathbb{R}$, то говорят, что несобственный интеграл *сходится*; в противном случае говорят, что он *расходится.*

Итак, по определению

$$\int_a^{\to b} f = \lim_{A\to b-} \int_a^A f,$$

если предел существует в $\overline{\mathbb{R}}$.

Аналогично определяется несобственный интеграл в симметричной ситуации: для $-\infty \leqslant a < b < +\infty$, $f \in R_{loc}(a, b]$ полагают

$$\int_{\to a}^{b} f = \lim_{B \to a+} \int_{B}^{b} f,$$

если предел существует в $\overline{\mathbb{R}}$, и называют интеграл сходящимся, если предел конечен. В конце параграфа определение несобственного интеграла будет дано в более общем случае.

Для определенности далее мы будем формулировать утверждения в первой ситуации.

Интеграл Римана от функции по отрезку называют еще *собственным.*

Пусть $b < +\infty$, $f \in R[a, b]$ (напомним, что мы определяли интеграл от функций, заданных на отрезке всюду, за исключением конечного множества точек, так что неважно, задана функция в точке b или нет). Тогда по непрерывности интеграла с переменным верхним пределом в точке b несобственный интеграл существует и равен собственному:

$$\int_{a}^{\to b} f = \lim_{A \to b-} \int_{a}^{A} f = \int_{a}^{b} f.$$

Поэтому несобственый интеграл — обобщение собственного, и для него можно использовать тот же символ $\int_a^b f$. Обозначение $\int_a^{\to b} f$ удобно, когда нужно подчеркнуть, что переходить к пределу следует именно в точке b.

Новая ситуация возникает в двух случаях:

а) если $b = +\infty$;

б) если $b < +\infty$, но $f \notin R[a, b]$.

Из теоремы 1 § 2 или критерия Лебега легко вывести, что если $b < +\infty$, f ограничена на $[a, b)$ и $f \in R_{loc}[a, b)$, то $f \in R[a, b]$. С другой стороны, если $f \in R[a, b]$, то f ограничена. Поэтому вариант б) реализуется в том и только том случае, когда f не ограничена на $[a, b)$. В условиях определения последнее равносильно тому, что f не ограничена ни в какой левой окрестности точки b.

Например, функция $f(x) = \sin \frac{1}{x}$ интегрируема по Риману на $[-1, 0]$, несмотря на то, что в точке 0 она имеет разрыв второго рода.

Теорема 1. Критерий Больцано – Коши сходимости интегралов. *Пусть* $-\infty < a < b \leqslant +\infty$, $f \in R_{loc}[a, b)$. *Тогда сходимость интеграла* $\int_a^b f$ *равносильна условию*

$$\forall \varepsilon > 0 \ \ \exists \Delta \in (a, b): \ \ \forall A, B \in (\Delta, b) \ \ \left| \int_A^B f \right| < \varepsilon.$$

Доказательство. Положим $\Phi(A) = \int_a^A f$. По определению сходимость интеграла $\int_a^b f$ равносильна существованию конечного предела $\Phi(A)$ при $A \to b-$. Остается воспользоваться критерием Больцано – Коши существования предела функции (теорема 9 § 3 главы 2):

$$\forall \varepsilon > 0 \ \ \exists \Delta \in (a, b): \ \ \forall A, B \in (\Delta, b) \ \ |\Phi(B) - \Phi(A)| < \varepsilon$$

и учесть, что по аддитивности интеграла $\Phi(B) - \Phi(A) = \int_A^B f$. $\square$

Замечание 1. Критерий Больцано – Коши чаще используется для установления *расходимости* интегралов. Если существуют последовательности точек A_n и B_n из $[a, b)$, стремящиеся к b, для которых $\int_{A_n}^{B_n} f \not\to 0$, то интеграл $\int_a^b f$ расходится.

Замечание 2. Пусть функция f имеет первообразную F на $[a, b)$. Тогда по формуле Ньютона – Лейбница

$$\int_a^b f = \lim_{A \to b-} \int_a^A f = \lim_{A \to b-} (F(A) - F(a)) = F(b-) - F(a).$$

Таким образом, сходимость несобственного интеграла равносильна существованию конечного предела первообразной. Двойную подстановку $F(b-) - F(a)$ также удобно обозначать через $F|_a^b$, понимая под $F(b)$ предел, то есть $F(b-)$.

Пример 1. Исследуем сходимость интеграла $\displaystyle\int_1^{+\infty} \frac{dx}{x^\alpha}$, $\alpha \in \mathbb{R}$. Имеем:

$$\int_1^{+\infty} \frac{dx}{x^\alpha} = \left\{ \begin{array}{ll} \left.\dfrac{x^{1-\alpha}}{1-\alpha}\right|_1^{+\infty}, & \alpha \neq 1 \\ \ln x\big|_1^{+\infty}, & \alpha = 1 \end{array} \right\} = \left\{ \begin{array}{ll} \dfrac{1}{\alpha - 1}, & \alpha > 1 \\ +\infty, & \alpha \leqslant 1 \end{array} \right\}.$$

Таким образом, интеграл $\displaystyle\int_1^{+\infty} \frac{dx}{x^\alpha}$ сходится при $\alpha > 1$ и расходится при $\alpha \leqslant 1$.

Пример 2. Исследуем сходимость интеграла $\displaystyle\int_0^1 \frac{dx}{x^\alpha}$, $\alpha \in \mathbb{R}$. Имеем:

$$\int_0^1 \frac{dx}{x^\alpha} = \left\{ \begin{array}{ll} \left.\dfrac{x^{1-\alpha}}{1-\alpha}\right|_0^1, & \alpha \neq 1 \\ \ln x\big|_0^1, & \alpha = 1 \end{array} \right\} = \left\{ \begin{array}{ll} \dfrac{1}{1-\alpha}, & \alpha < 1 \\ +\infty, & \alpha \geqslant 1 \end{array} \right\}.$$

Таким образом, интеграл $\displaystyle\int_0^1 \frac{dx}{x^\alpha}$ сходится при $\alpha < 1$ (при $\alpha \leqslant 0$ он даже собственный) и расходится при $\alpha \geqslant 1$.

С помощью доказываемого далее признака сравнения сходимость многих интегралов более общего вида сводится к сходимости интегралов от степенных функций.

Установим несколько свойств несобственных интегралов. В их формулировках мы предполагаем, что выполнены условия из определения несобственного интеграла, то есть функции локально интегрируемы на соответствующих промежутках.

Н1. Аддитивность несобственного интеграла по промежутку. *Если интеграл $\int_a^b f$ сходится, то для любой точки $c \in (a, b)$ интеграл $\int_c^b f$ тоже сходится и*

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{8}$$

Обратно, если при некотором $c \in (a, b)$ интеграл $\int_c^b f$ сходится, то сходится и интеграл $\int_a^b f$.

Доказательство. При всех $A \in (c, b)$ по свойству аддитивности интеграла

$$\int_a^A f = \int_a^c f + \int_c^A f. \tag{9}$$

При $A \to b-$ предел обеих частей равенства (9) существует или нет одновременно, то есть сходимость $\int_c^b f$ эквивалентна сходимости $\int_a^b f$. Равенство (8) получается переходом к пределу в (9). $\square$

Определение 3. Несобственный интеграл $\int_A^{\to b} f$ называется *остатком* интеграла $\int_a^{\to b} f$.

Свойство Н1 утверждает, что интеграл и любой его остаток сходятся или расходятся одновременно.

Н2. *Если интеграл $\int_a^b f$ сходится, то $\int_A^b f \underset{A\to b-}{\longrightarrow} 0$.* Другими словами, *остаток сходящегося интеграла стремится к нулю.*

Действительно,

$$\int_A^b f = \int_a^b f - \int_a^A f \underset{A\to b-}{\longrightarrow} \int_a^b f - \int_a^b f = 0.$$

Н3. Линейность несобственного интеграла. *Если интегралы $\int_a^b f$, $\int_a^b g$ сходятся, $\alpha, \beta \in \mathbb{R}$, то интеграл $\int_a^b(\alpha f + \beta g)$ сходится и*

$$\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g.$$

Для доказательства надо перейти к пределу в равенстве для частичных интегралов

$$\int_a^A (\alpha f + \beta g) = \alpha \int_a^A f + \beta \int_a^A g.$$

Замечание 1. Если интеграл $\int_a^b f$ расходится, а интеграл $\int_a^b g$ сходится, то интеграл $\int_a^b (f+g)$ расходится.

В самом деле, если бы интеграл от $f+g$ сходился, то сходился бы и интеграл от $f = (f+g) - g$, что неверно.

Н4. Монотонность несобственного интеграла. *Если интегралы $\int_a^b f$, $\int_a^b g$ существуют в $\overline{\mathbb{R}}$, $f \leqslant g$ на $[a,b)$, то*

$$\int_a^b f \leqslant \int_a^b g.$$

Для доказательства надо перейти к пределу в неравенстве для частичных интегралов

$$\int_a^A f \leqslant \int_a^A g.$$

Замечание 2. Аналогично, с помощью предельного перехода, на несобственные интегралы переносятся неравенства Иенсена, Гёльдера, Минковского.

Н5. Интегрирование по частям в несобственном интеграле. *Пусть f, g дифференцируемы на $[a,b)$, $f', g' \in R_{loc}[a,b)$. Тогда*

$$\int_a^b fg' = fg\Big|_a^b - \int_a^b f'g. \tag{10}$$

Заключение надо понимать так: *если существуют два конечных предела из трех, то третий предел также существует и конечен, причем имеет место равенство* (10).

Для доказательства надо устремить A к b слева в равенстве

$$\int_a^A fg' = fg\Big|_a^A - \int_a^A f'g.$$

Для несобственных интегралов, аналогично собственным, принимается соглашение $\int_a^b f = -\int_b^a f$. Как и в случае отрезка, удобно считать, что при $a > b$ символы $[a,b)$ и (a,b) означают соответственно $(b,a]$ и (b,a).

Н6. Замена переменной в несобственном интеграле. *Пусть $\varphi\colon [\alpha,\beta) \to [A,B)$, функция φ дифференцируема на $[\alpha,\beta)$, $\varphi' \in R_{loc}[\alpha,\beta)$, существует $\varphi(\beta-) \in \overline{\mathbb{R}}$, $f \in C[A,B)$. Тогда*

$$\int_\alpha^\beta (f\circ\varphi)\varphi' = \int_{\varphi(\alpha)}^{\varphi(\beta-)} f. \tag{11}$$

Заключение надо понимать так: *если существует один из интегралов, то существует и другой, и имеет место равенство* (11).

Доказательство. Обозначим

$$\Phi(\gamma) = \int_\alpha^\gamma (f\circ\varphi)\varphi', \quad F(C) = \int_{\varphi(\alpha)}^C f.$$

По формуле замены переменной в собственном интеграле

$$\Phi(\gamma) = F(\varphi(\gamma)).$$

1. Пусть существует интеграл $\int_{\varphi(\alpha)}^{\varphi(\beta-)} f = I \in \overline{\mathbb{R}}$. Докажем, что интеграл $\int_\alpha^\beta (f \circ \varphi)\varphi'$ также существует и равен I, то есть $\Phi(\gamma) \to I$ при $\gamma \to \beta-$. Возьмем последовательность $\{\gamma_n\}$ из промежутка $[\alpha, \beta)$, такую что $\gamma_n \to \beta$. Тогда $\varphi(\gamma_n) \in [A, B)$, $\varphi(\gamma_n) \to \varphi(\beta-)$. Поэтому $\Phi(\gamma_n) = F(\varphi(\gamma_n)) \to I$. Ввиду произвольности последовательности $\{\gamma_n\}$ отсюда вытекает, что $\Phi(\gamma) \to I$ при $\gamma \to \beta-$.

2. Пусть существует интеграл $\int_\alpha^\beta (f \circ \varphi)\varphi' = J \in \overline{\mathbb{R}}$. Докажем, что интеграл $\int_{\varphi(\alpha)}^{\varphi(\beta-)} f$ существует; отсюда уже по пункту 1 будет следовать, что он равен J. Если $\varphi(\beta-) \in [A, B)$, то доказывать нечего: интеграл существует в собственном смысле. Пусть $\varphi(\beta-) = B$. Возьмем последовательность $\{C_n\}$ из промежутка $[A, B)$, такую что $C_n \to B$. Не уменьшая общности, можно считать, что $C_n \in [\varphi(\alpha), B)$ при всех n. Тогда по теореме Больцано – Коши о промежуточном значении найдутся такие точки $\gamma_n \in [\alpha, \beta)$, что $\varphi(\gamma_n) = C_n$.

Докажем, что $\gamma_n \to \beta$. Пусть $\beta' \in [\alpha, \beta)$. Так как $\max\limits_{[\alpha,\beta']} \varphi < B$, а $\varphi(\gamma_n) \to B$, то, начиная с некоторого номера, $\gamma_n \in (\beta', \beta)$. Поэтому $\gamma_n \to \beta$, откуда $F(C_n) = \Phi(\gamma_n) \to J$. $\square$

Замечание 1. Если φ строго монотонна, то вторая часть доказательства упрощается, так как можно положить $\gamma_n = \varphi^{-1}(C_n)$.

Замечание 2. При доказательстве мы воспользовались теоремой 5 § 3 (или теоремой 3 § 3$'$) о замене переменной. Используя замечание 4 к теореме 5 § 3, сформулируем другой вариант Н6.

Пусть функция φ дифференцируема, строго монотонна на промежутке $[\alpha, \beta)$, $\varphi(\alpha) = a$, $\varphi(\beta-) = b$, $\varphi' \in R_{loc}[\alpha, \beta)$, $f \in R_{loc}[a, b)$. Тогда

$$\int_\alpha^\beta (f \circ \varphi)\varphi' = \int_a^b f.$$

Замечание 3. Возможны случаи, когда оба интеграла в равенстве (11) собственные, оба несобственные или один собственный, а другой несобственный. Таким образом, замена переменной в собственном интеграле может привести к несобственному, и обратно. Например, это бывает, если функция φ не ограничена. Среди стандартных подстановок, используемых при нахождении первообразных, такие встречаются довольно часто.

Пример 3. Найдем $\int_0^\pi \frac{dx}{2+\cos x}$. Этот интеграл собственный (подынтегральная функция непрерывна), и его удобно вычислить с помощью подстановки $t = \operatorname{tg}\frac{x}{2}$. Функция $\varphi(t) = 2\operatorname{arctg} t$ отображает $[\,0, +\infty)$ на $[\,0, \pi)$,

$$\varphi(0) = 0, \quad \varphi(+\infty) = \pi, \quad \varphi'(t) = \frac{2}{1+t^2}, \quad \cos x = \frac{1-t^2}{1+t^2}.$$

Поэтому

$$\int_0^\pi \frac{dx}{2+\cos x} = \int_0^{+\infty} \frac{1}{2+\frac{1-t^2}{1+t^2}} \cdot \frac{2dt}{1+t^2} =$$
$$= 2\int_0^{+\infty} \frac{dt}{3+t^2} = \frac{2}{\sqrt{3}} \operatorname{arctg} \frac{t}{\sqrt{3}}\Big|_0^{+\infty} = \frac{\pi}{\sqrt{3}}\,.$$

Замечание 4. Интеграл по конечному промежутку $\int_a^b f(x)\,dx$ заменой $x = b - \frac{1}{t}$ можно свести к интегралу с бесконечным верхним пределом:

$$\int_a^b f(x)\,dx = \int_{\frac{1}{b-a}}^{+\infty} f\Big(b - \frac{1}{t}\Big)\frac{dt}{t^2}.$$

Поэтому, не уменьшая общности, можно ограничиться изучением несобственных интегралов по бесконечному промежутку.

Далее мы выведем несколько признаков сходимости несобственных интегралов. Сначала рассмотрим интегралы от неотрицательных функций.

Лемма 1. *Пусть $f \in R_{loc}[a,b)$, $f \geqslant 0$. Тогда сходимость интеграла $\int_a^b f$ равносильна ограниченности функции $F(A) = \int_a^A f$ на $[a,b)$ сверху.*

Доказательство. Функция F возрастает на $[a,b)$, так как при $a \leqslant A < B < b$ по свойству аддитивности интеграла

$$F(B) - F(A) = \int_A^B f \geqslant 0.$$

Сходимость интеграла $\int_a^b f$ по определению означает существование конечного предела $F(A)$ при $A \to b-$, которое по теореме о пределе монотонной функции равносильно ограниченности F сверху. $\square$

Замечание 1. Из теоремы о пределе монотонной функции также следует, что для $f \geqslant 0$ интеграл $\int_a^b f$ либо сходится, либо расходится к $+\infty$, причем

$$\int_a^b f = \sup_{A \in [a,b)} \int_a^A f.$$

Замечание 2. Для ограниченности возрастающей функции F сверху достаточно ограниченности сверху некоторой последовательности $\{F(A_n)\}$, где $A_n \in [a,b)$, $\{A_n\}$ возрастает, $A_n \to b$.

Действительно,

$$\sup_{n \in \mathbb{N}} F(A_n) = \lim_{n \to \infty} F(A_n) = \lim_{A \to b-} F(A) = \sup_{A \in [a,b)} F(A).$$

Теорема 2. Признак сравнения сходимости несобственных интегралов. *Пусть $f, g \in R_{loc}[a,b)$, $f, g \geqslant 0$,*

$$f(x) = O(g(x)) \qquad \textit{при} \ \ x \to b-.$$

1. *Если интеграл $\int_a^b g$ сходится, то и интеграл $\int_a^b f$ сходится.*

2. *Если интеграл $\int_a^b f$ расходится, то и интеграл $\int_a^b g$ расходится.*

Доказательство. 1. По определению символа O найдутся такие $\Delta \in (a,b)$ и $K > 0$, что $f(x) \leqslant Kg(x)$ при всех $x \in [\Delta, b)$. Следовательно,

$$\int_\Delta^b f \leqslant K \int_\Delta^b g < +\infty,$$

то есть остаток интеграла $\int_a^b f$ сходится, а тогда и сам интеграл $\int_a^b f$ сходится.

2. Если бы интеграл $\int_a^b g$ сходился, то по пункту 1 сходился бы и интеграл $\int_a^b f$, что неверно. $\square$

Следствие 1. Признак сравнения в предельной форме. *Пусть $f, g \in R_{loc}[a,b)$, $f \geqslant 0$, $g > 0$ и существует предел $\lim\limits_{x \to b-} \frac{f(x)}{g(x)} = \ell \in [0, +\infty]$.*

1. *Если $\ell \in [0, +\infty)$, а интеграл $\int_a^b g$ сходится, то интеграл $\int_a^b f$ сходится.*

2. *Если $\ell \in (0, +\infty]$, а интеграл $\int_a^b f$ сходится, то интеграл $\int_a^b g$ сходится.*

3. *Если $\ell \in (0, +\infty)$, то интегралы $\int_a^b f$ и $\int_a^b g$ сходятся или расходятся одновременно.*

Доказательство. Очевидно, что третий пункт вытекает из первых двух. Перейдем к их доказательству.

1. Из конечности ℓ следует, что частное $\frac{f}{g}$ ограничено в некоторой левой окрестности точки b, то есть $f(x) = O(g(x))$ при $x \to b-$. Остается воспользоваться теоремой 2.

2. Так как $\ell > 0$, то и $f > 0$ в некоторой левой окрестности b. Остается поменять f и g ролями и свести утверждение к первому пункту. $\square$

Следствие 2. *Интегралы от неотрицательных эквивалентных в точке b функций сходятся или расходятся одновременно.*

Пример 4. Исследуем сходимость интеграла $\displaystyle\int_2^{+\infty} \frac{dx}{x^\alpha \ln^\beta x}$ при $\alpha, \beta \in \mathbb{R}$.

В § 4 главы 3 было доказано, что $\ln x = o(x^q)$ при $x \to +\infty$, $q > 0$. Отсюда для любых $p \in \mathbb{R}$ и $q > 0$ верно соотношение $\ln^p x = o(x^q)$ при $x \to +\infty$. Действительно, при $p \leqslant 0$ это очевидно, а при $p > 0$

$$\frac{\ln^p x}{x^q} = \left(\frac{\ln x}{x^{q/p}}\right)^p \underset{x \to +\infty}{\longrightarrow} 0.$$

Если $\alpha > 1$, то

$$\frac{1}{x^\alpha \ln^\beta x} = O\left(\frac{1}{x^{\frac{\alpha+1}{2}}}\right), \quad x \to +\infty,$$

поскольку $\lim\limits_{x\to+\infty}\dfrac{\ln^{-\beta}x}{x^{\frac{\alpha-1}{2}}}=0$. Так как $\frac{\alpha+1}{2}>1$, интеграл $\displaystyle\int_2^{+\infty}\frac{dx}{x^{\frac{\alpha+1}{2}}}$ сходится, а тогда исходный интеграл сходится по признаку сравнения при любом β.

Если $\alpha<1$, то аналогично

$$\frac{1}{x^{\frac{1+\alpha}{2}}}=O\left(\frac{1}{x^\alpha\ln^\beta x}\right),\quad x\to+\infty.$$

Так как $\frac{1+\alpha}{2}<1$, интеграл $\displaystyle\int_2^{+\infty}\frac{dx}{x^{\frac{1+\alpha}{2}}}$ расходится, а тогда исходный интеграл расходится по признаку сравнения при любом β.

Если $\alpha=1$, то сделаем замену $\ln x=t$:

$$\int_2^{+\infty}\frac{dx}{x\ln^\beta x}=\int_{\ln 2}^{+\infty}\frac{dt}{t^\beta}.$$

Последний интеграл сходится при $\beta>1$ и расходится при $\beta\leqslant 1$.

Итак, интеграл $\displaystyle\int_2^{+\infty}\frac{dx}{x^\alpha\ln^\beta x}$ сходится ровно в двух случаях: при $\alpha>1$ и произвольном β или при $\alpha=1$ и $\beta>1$.

Замечание 3. Из сходимости интеграла $\int_a^{+\infty}f$ не вытекает, что $f(x)\underset{x\to+\infty}{\longrightarrow}0$, даже если $f\geqslant 0$ и f непрерывна.

Пример 5. Пусть $E=\bigcup\limits_{k=1}^{\infty}\left(k-\frac{1}{k^2(k+1)},k+\frac{1}{k^2(k+1)}\right)$, $f(k)=k$ при всех $k\in\mathbb{N}$, $f(x)=0$ при всех $x\in[0,+\infty)\setminus E$, f линейна на промежутках $\left[k-\frac{1}{k^2(k+1)},k\right]$ и $\left[k,k+\frac{1}{k^2(k+1)}\right]$. График f изображен на рисунке 7. Тогда $f\in C[0,+\infty)$. Если $N\in\mathbb{N}$, то

$$\int_0^{N+1/2}f=\sum_{k=1}^N\int_{k-\frac{1}{k^2(k+1)}}^{k+\frac{1}{k^2(k+1)}}f=\sum_{k=1}^N\frac{1}{2}\cdot\frac{2}{k^2(k+1)}\cdot k=$$
$$=\sum_{k=1}^N\frac{1}{k(k+1)}=\sum_{k=1}^N\left(\frac{1}{k}-\frac{1}{k+1}\right)=1-\frac{1}{N+1}\underset{N\to\infty}{\longrightarrow}1$$

(при вычислении интегралов мы применили формулу для площади треугольника). Поэтому $\int_0^{+\infty}f=1$. Вместе с тем, f не только не стремится к нулю, но даже не ограничена ни в какой окрестности $+\infty$.

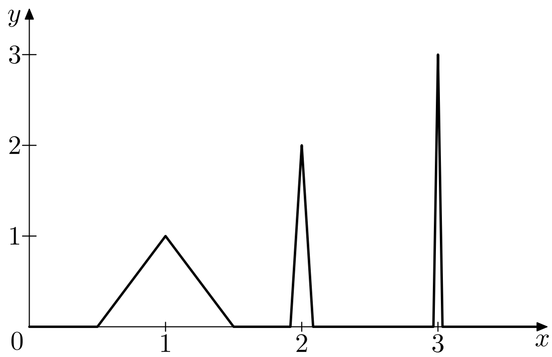

Рис. 7

Теперь рассмотрим несобственные интегралы от функций произвольного знака.

Замечание 1. Ограниченность частичных интегралов является необходимым, но не достаточным условием сходимости. Например, интеграл $\int_0^{+\infty} \cos x\, dx$ расходится, так как частичные интегралы $\int_0^A \cos x\, dx = \sin A$ не имеют предела при $A \to +\infty$.

Определение 4. Пусть $-\infty < a < b \leqslant +\infty$, $f \in R_{loc}[a, b)$. Говорят, что интеграл $\int_a^b f$ *сходится абсолютно,* если сходится интеграл $\int_a^b |f|$.

Замечание 2. Если интегралы $\int_a^b f$, $\int_a^b g$ сходятся абсолютно, $\alpha, \beta \in \mathbb{R}$, то интеграл $\int_a^b (\alpha f + \beta g)$ сходится абсолютно.

Это утверждение следует из неравенства

$$|\alpha f + \beta g| \leqslant |\alpha||f| + |\beta||g|$$

и признака сравнения.

Замечание 3. Если интеграл $\int_a^b f$ существует в $\overline{\mathbb{R}}$, то

$$\left| \int_a^b f \right| \leqslant \int_a^b |f|.$$

Доказательство получается переходом к пределу в неравенстве для частичных интегралов.

Лемма 2. *Если интеграл сходится абсолютно, то он сходится.*

Эту лемму мы докажем двумя способами.

Первое доказательство. Возьмем $\varepsilon > 0$ и по критерию Больцано – Коши сходимости интеграла $\int_a^b |f|$ подберем $\Delta \in (a, b)$ так, что для любых $A, B \in (\Delta, b)$ $(A < B)$ будет $\int_A^B |f| < \varepsilon$. Но тогда тем более

$$\left|\int_A^B f\right| \leqslant \int_A^B |f| < \varepsilon.$$

Следовательно, интеграл $\int_a^b f$ сходится по критерию Больцано – Коши. $\square$

Прежде чем дать другое доказательство, введем новые понятия, важные и сами по себе.

Для $x \in \mathbb{R}$ положим

$$x_+ = \max\{x, 0\} = \begin{cases} x, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$
$$x_- = \max\{-x, 0\} = \begin{cases} 0, & x > 0, \\ -x, & x \leqslant 0. \end{cases}$$

Функции x_+ и x_- называют *положительной* и *отрицательной частями* числа x. Графики x_+ и x_- изображены на рисунках 8a и 8b.

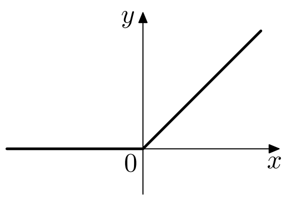

Рис. 8a

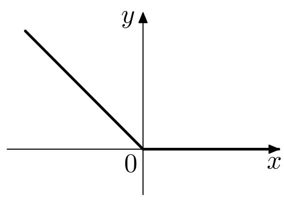

Рис. 8b

Непосредственно из определений следуют соотношения

$$x_+ - x_- = x, \quad x_+ + x_- = |x|, \quad 0 \leqslant x_\pm \leqslant |x|,$$
$$x_+ = \frac{|x| + x}{2}, \quad x_- = \frac{|x| - x}{2}.$$

Если задана функция f, то функции f_+ и f_-, определяемые равенствами $f_\pm(x) = \big(f(x)\big)_\pm$, называются *положительной* и *отрицательной частями* функции f.

Из теорем об арифметических действиях над непрерывными (интегрируемыми) функциями следует, что если f непрерывна (интегрируема) на $[a, b]$, то таковыми будут и $f_\pm$.

Второе доказательство леммы 2. Поскольку интеграл $\int_a^b |f|$ сходится, по признаку сравнения сходятся и интегралы $\int_a^b f_\pm$, а тогда сходится и интеграл $\int_a^b f$ как разность двух сходящихся интегралов. □

Замечание 4. Утверждение, обратное к лемме 2, неверно: интеграл может сходиться, но не абсолютно. Примеры будут приведены после теоремы 3.

Если интеграл сходится, но не абсолютно, то говорят, что он *сходится условно* или *неабсолютно.*

Замечание 5. Если интеграл $\int_a^b f$ сходится условно, а интеграл $\int_a^b g$ сходится абсолютно, то интеграл $\int_a^b (f + g)$ сходится условно.

В самом деле, если бы интеграл от $f + g$ сходился абсолютно, то по замечанию 2 абсолютно сходился бы и интеграл от $f = (f+g)-g$, что неверно.

Теорема 3. Признаки Дирихле и Абеля сходимости несобственных интегралов. *Пусть $f \in C[a, b)$, $g \in C^1[a, b)$, g монотонна.*

1. Признак Дирихле. *Если функция $F(A) = \int_a^A f$ ограничена, а $g(x) \underset{x\to b-}{\longrightarrow} 0$, то интеграл $\int_a^b fg$ сходится.*

2. Признак Абеля. *Если интеграл $\int_a^b f$ сходится, а g ограничена, то интеграл $\int_a^b fg$ сходится.*

Доказательство. 1. Проинтегрируем по частям:

$$\int_a^b fg = \int_a^b F'g = Fg\Big|_a^b - \int_a^b Fg' = -\int_a^b Fg'.$$

Двойная подстановка обнуляется, поэтому сходимость исходного интеграла равносильна сходимости интеграла $\int_a^b Fg'$. Докажем, что последний сходится абсолютно, по признаку сравнения. Пусть K таково, что $|F(x)| \leqslant K$ при всех $x \geqslant a$. Поскольку g монотонна, g' не меняет знака на $[a,b)$. Следовательно,

$$\int_a^b |Fg'| \leqslant K\int_a^b |g'| = K\left|\int_a^b g'\right| = K\left|[g]_a^b\right| = K|g(a)|.$$

2. Так как g монотонна и ограничена, существует конечный предел $\lim\limits_{x\to b-} g(x) = \alpha$. Функции f и $g-\alpha$ удовлетворяют условиям признака Дирихле. Поэтому интеграл $\int_a^b f(g-\alpha)$ сходится, а тогда и интеграл $\int_a^b fg$ сходится как сумма двух сходящихся:

$$\int_a^b fg = \int_a^b f(g-\alpha) + \alpha\int_a^b f. \quad \square$$

Замечание 1. Теорема 3 верна при менее ограничительных условиях: $f \in R_{loc}[a,b)$, g монотонна на $[a,b)$.

Доказательство может быть проведено с помощью критерия Больцано – Коши и второй теоремы о среднем. Однако, поскольку вторая теорема о среднем не доказывалась в полной общности, мы и здесь ограничились более слабой формулировкой.

Пример 6. Исследуем сходимость и абсолютную сходимость интегралов $\displaystyle\int_1^{+\infty} g(x)\sin\lambda x\,dx$ и $\displaystyle\int_1^{+\infty} g(x)\cos\lambda x\,dx$, где функция g монотонна, $g \geqslant 0$, $\lambda \neq 0$.

Для определенности рассмотрим интеграл с синусом; интеграл с косинусом исследуется аналогично. Достаточно считать, что $\lambda = 1$, так как случай $\lambda > 0$ сводится к этому заменой $\lambda x = t$, а случай $\lambda < 0$ — по нечетности синуса.

Если интеграл $\int_1^{+\infty} g$ сходится, то в силу очевидного неравенства $|g(x)\sin x| \leqslant g(x)$ интеграл $\int_1^{+\infty} g(x)\sin x\,dx$ сходится абсолютно по признаку сравнения.

Если интеграл $\int_1^{+\infty} g$ расходится, то признак сравнения не позволяет сделать вывод о сходимости исходного интеграла, и необходимы более тонкие рассуждения.

Обозначим $\ell = \lim\limits_{x\to+\infty} g(x) \in [\,0, +\infty]$.

Если $\ell = 0$, то интеграл $\int_1^{+\infty} g(x)\sin x\,dx$ сходится по признаку Дирихле, так как g монотонна, а $\left|\int_1^A \sin x\,dx\right| = |\cos 1 - \cos A| \leqslant 2$.

Докажем, что если $\ell = 0$, но интеграл $\int_1^{+\infty} g$ расходится, то сходимость исходного интеграла не абсолютна, то есть что интеграл $\int_1^{+\infty} g(x)|\sin x|\,dx$ расходится. Воспользуемся оценкой

$$|\sin x| \geqslant \sin^2 x = \frac{1}{2} - \frac{\cos 2x}{2}.$$

Аналогично предыдущему, интеграл $\int_1^{+\infty} \frac{1}{2}\,g(x)\cos 2x\,dx$ сходится по признаку Дирихле, так как $\frac{1}{2}\,g(x) \underset{x\to+\infty}{\longrightarrow} 0$ монотонно, а

$$\left|\int_1^A \cos 2x\,dx\right| = \tfrac{1}{2}\bigl|\sin 2A - \sin 2\bigr| \leqslant 1.$$

Следовательно, интеграл $\int_1^{+\infty} \frac{1}{2}\,g(x)(1-\cos 2x)\,dx$ расходится как разность расходящегося и сходящегося интегралов (см. замечание к свойству Н3). По признаку сравнения интеграл $\int_1^{+\infty} g(x)|\sin x|\,dx$ расходится.

Наконец, докажем, что при $\ell > 0$ интеграл расходится. Так как при $k \in \mathbb{N}$ и $x \in \left[2k\pi + \frac{\pi}{6}, 2k\pi + \frac{5\pi}{6}\right]$ верно неравенство $\sin x \geqslant \frac{1}{2}$,

$$\int_{2k\pi+\frac{\pi}{6}}^{2k\pi+\frac{5\pi}{6}} g(x)\sin x\,dx \geqslant$$
$$\geqslant \frac{1}{2}\cdot\frac{2\pi}{3}\cdot\min\Bigl\{g\Bigl(2k\pi+\frac{\pi}{6}\Bigr), g\Bigl(2k\pi+\frac{5\pi}{6}\Bigr)\Bigr\} \underset{k\to\infty}{\longrightarrow} \frac{\pi\ell}{3} > 0,$$

и интеграл расходится по замечанию 1 к критерию Больцано – Коши.

Таким образом, можно сделать следующие выводы об интегралах $\int_1^{+\infty} g(x)\sin\lambda x\,dx$ и $\int_1^{+\infty} g(x)\cos\lambda x\,dx$. При $g(x) \underset{x\to+\infty}{\longrightarrow} 0$ интегралы сходятся, причем абсолютно, если интеграл $\int_1^{+\infty} g$ сходится, и условно, если он расходится. При $g(x) \underset{x\to+\infty}{\not\longrightarrow} 0$ интегралы расходятся.

В частности, если $\lambda \neq 0$, то оба интеграла $\int_1^{+\infty} \frac{\sin\lambda x}{x^\alpha}\,dx$ и $\int_1^{+\infty} \frac{\cos\lambda x}{x^\alpha}\,dx$ сходятся при $\alpha > 0$, причем абсолютно при $\alpha > 1$ и условно при $\alpha \in (0,1]$, и расходятся при $\alpha \leqslant 0$.

До сих пор мы рассматривали случаи, когда интегрируемость нарушается вблизи одного из концов промежутка, на котором определена функция. Теперь изучим более общую ситуацию.

Пусть сначала $-\infty \leqslant a < b \leqslant +\infty$, $f \in R_{loc}(a,b)$. Тогда полагают

$$\int_{\to a}^{\to b} f = \int_{\to a}^{c} f + \int_{c}^{\to b} f,$$

где $c \in (a,b)$, если интегралы в правой части существуют в $\overline{\mathbb{R}}$ и не равны бесконечностям разных знаков. При этом интеграл $\int_{\to a}^{\to b} f$ называют сходящимся, если сходятся оба интеграла в правой части.

В силу аддитивности несобственного интеграла это определение корректно, то есть не зависит от выбора точки c. Действительно, если $a < c < d < b$, то

$$\int_{\to a}^{c} f + \int_{c}^{\to b} f = \int_{\to a}^{c} f + \int_{c}^{d} f + \int_{d}^{c} f + \int_{c}^{\to b} f = \int_{\to a}^{d} f + \int_{d}^{\to b} f,$$

причем обе части существуют или нет одновременно.

Пусть теперь $-\infty \leqslant a < b \leqslant +\infty$, функция f задана на (a,b), за исключением, быть может, конечного множества точек. Будем называть точку $c \in (a,b)$ *особой* точкой функции f, если для любых A и B, удовлетворяющих неравенствам $a < A < c < B < b$, $f \notin R[A,B]$. Точку a будем называть особой, если $a = -\infty$ или $a \in \mathbb{R}$, но для любого $B \in (a,b)$ функция f не интегрируема на $[a,B]$. Точку b будем называть особой, если $b = +\infty$ или $b \in \mathbb{R}$, но для любого $A \in (a,b)$ функция f не интегрируема на $[A,b]$.

Предположим, что множество особых точек функции f на (a,b) конечно. Пусть $c_1 < \ldots < c_{n-1}$ — все особые точки f на (a,b), $c_0 = a$, $c_n = b$. С помощью теоремы Гейне – Бореля (см. § 3 главы 5) можно доказать, что $f \in R_{loc}(c_k, c_{k+1})$ при всех $k \in [\,0 : n-1]$. Тогда интегралы $\int_{c_k}^{c_{k+1}} f$ уже определены, и мы можем положить

$$\int_a^b f = \sum_{k=0}^{n-1} \int_{c_k}^{c_{k+1}} f,$$

если все слагаемые в правой части и их сумма имеют смысл в $\overline{\mathbb{R}}$.

Таким образом, общий случай сводится к первоначальному, в котором единственной особой точкой является конец промежутка. Ясно, что свойство аддитивности несобственного интеграла сохраняется и в новой ситуации.

Пример 7. Из примеров 1 и 2 следует, что интеграл $\displaystyle\int_0^{+\infty} \frac{dx}{x^\alpha}$ расходится при всех α, так как при $\alpha \leqslant 1$ расходится интеграл $\displaystyle\int_1^{+\infty} \frac{dx}{x^\alpha}$, а при $\alpha \geqslant 1$ — интеграл $\displaystyle\int_0^1 \frac{dx}{x^\alpha}$.

Пример 8. Поскольку $\sin x \sim x$ при $x \to 0$, интеграл от положительной функции $\displaystyle\int_0^1 \frac{\sin x}{x^\alpha}\,dx$ ведет себя так же, как и $\displaystyle\int_0^1 \frac{dx}{x^{\alpha-1}}$, то есть расходится при $\alpha \geqslant 2$ и сходится (абсолютно) при $\alpha < 2$ (при $\alpha \leqslant 1$ он даже собственный). Учитывая результат примера 6, заключаем, что интеграл $\displaystyle\int_0^{+\infty} \frac{\sin x}{x^\alpha}\,dx$ сходится при $\alpha \in (0,2)$, причем абсолютно при $\alpha \in (1,2)$ и условно при $\alpha \in (0,1]$, и расходится при $\alpha \notin (0,2)$.

Иногда оказывается полезным еще одно обобщение несобственного интеграла.

Определение 5. Главное значение несобственного интеграла. Пусть $-\infty < a < b < +\infty$, $c \in (a,b)$ — единственная особая точка функции f на $[a,b]$. Предел

$$\lim_{\varepsilon \to 0+} \left(\int_a^{c-\varepsilon} f + \int_{c+\varepsilon}^b f \right)$$

называется *главным значением* несобственного интеграла $\int_a^b f$ и обозначается v.p. $\int_a^b f$ (от французского “valeur principale”) или p.v. $\int_a^b f$ (от английского “principal value”).

В обычном смысле несобственный интеграл определялся равенством

$$\int_a^b f = \int_a^c f + \int_c^b f = \lim_{\varepsilon_1 \to 0+} \int_a^{c-\varepsilon_1} f + \lim_{\varepsilon_2 \to 0+} \int_{c+\varepsilon_2}^b f.$$

Поэтому ясно, что если несобственный интеграл существует в обычном смысле, то его главное значение также существует и совпадает с обычным. Обратное неверно, как показывает следующий пример.

Пример 9. Пусть $-\infty < a < c < b < +\infty$. Интеграл $\displaystyle\int_a^b \frac{dx}{x-c}$ расходится, так как расходятся интегралы $\displaystyle\int_a^c \frac{dx}{x-c}$ и $\displaystyle\int_c^b \frac{dx}{x-c}$ (см. пример 2). Главное же значение этого интеграла существует, так как

$$\left(\int_a^{c-\varepsilon} \frac{dx}{x-c} + \int_{c+\varepsilon}^b \frac{dx}{x-c}\right) =$$

$$= \Big[\ln|x-c|\Big]_{x=a}^{c-\varepsilon} + \Big[\ln|x-c|\Big]_{x=c+\varepsilon}^{b} = \ln\frac{b-c}{c-a}.$$

Если $f\colon \mathbb{R} \to \mathbb{R}$ не имеет особых точек на $\mathbb{R}$, то главное значение интеграла от f по $\mathbb{R}$ определяется равенством

$$\text{v.p.} \int_{-\infty}^{+\infty} f = \lim_{A \to +\infty} \int_{-A}^{A} f.$$

Например, v.p. $\int_{-\infty}^{+\infty} x\,dx = 0$, а v.p. $\int_{-\infty}^{+\infty} x^2\,dx = +\infty$.

В случае, когда особых точек несколько, возможны различные обобщения (например, можно удалять одинаковые симметричные окрестности всех точек, а можно разные). Мы не будем останавливаться на этих обобщениях.

§ 6. Длина, площадь, объем

В этом вспомогательном параграфе приводятся краткие сведения о геометрических характеристиках множеств: площади, объеме и длине. Эти факты используются в следующих параграфах о геометрических приложениях интеграла. Систематически площади и объемы будут изучаться в главе, посвященной мере.

1. Площадь. Понятие площади некоторых геометрических фигур известно из школьного курса геометрии. Определение площади для более широкого класса множеств мы дадим лишь частично. Начнем с определения движения. Здесь и далее символом $\|x\|$ будет обозначаться длина вектора $x \in \mathbb{R}^n$. Тогда для любых $A, B \in \mathbb{R}^n$ длина отрезка AB в $\mathbb{R}^n$ выражается формулой

$$\|A - B\| = \sqrt{\sum_{i=1}^{n} (A_i - B_i)^2}.$$

Определение 1. Движение. Отображение $U: \mathbb{R}^n \to \mathbb{R}^n$ называется *движением* пространства $\mathbb{R}^n$, если оно сохраняет расстояние между точками, то есть $\|A - B\| = \|U(A) - U(B)\|$ для любых $A, B \in \mathbb{R}^n$.

Определение 2. Площадь. *Площадью* называется функционал $S: \{P\} \to [0, +\infty)$, заданный на некотором классе $\{P\}$ подмножеств плоскости, называемых *квадрируемыми фигурами,* и обладающий следующими тремя свойствами.

1. *Аддитивность.* Если P_1 и P_2 — квадрируемые фигуры, причем $P_1 \cap P_2 = \varnothing$, то $P_1 \cup P_2$ — квадрируемая фигура и

$$S(P_1 \cup P_2) = S(P_1) + S(P_2).$$

2. *Нормированность на прямоугольниках.* Площадь прямоугольника со сторонами a и b равна ab.

3. *Инвариантность относительно движений.* Если P — квадрируемая фигура, U — движение плоскости, то $U(P)$ — квадрируемая фигура и $S\big(U(P)\big) = S(P)$.

Свойство 3 в школьном курсе геометрии обычно формулируется так: равные фигуры имеют равные площади, а равными как раз называются фигуры, получающиеся друг из друга движением (наложением). В свойстве 2 точки, лежащие на границе прямоугольника, могут как принадлежать, так и не принадлежать ему. В силу инвариантности площади относительно движений определение не изменится, если требовать выполнение свойства 2 лишь для прямоугольников со сторонами, параллельными осям координат.

Это определение нельзя считать завершенным, так как не определено множество квадрируемых фигур, то есть фигур, имеющих площадь. Определять квадрируемые фигуры мы не будем, как не будем и заниматься вопросами о существовании и единственности площади. Ответы на эти вопросы будут даны позже, при изучении меры Лебега. В связи с этим мы во многих случаях лишены возможности доказать, что та или иная фигура имеет площадь, и принимаем существование площади на веру. Так мы поступаем и в следующих свойствах 4 – 6, и при выводе других формул для площади.

Отметим еще три свойства площади.

4. *Монотонность.* Если P и P_1 — квадрируемые фигуры, $P_1 \subset P$, то $S(P_1) \leqslant S(P)$.

Для доказательства запишем, что $P = P_1 \cup (P \setminus P_1)$, причем $P_1 \cap (P \setminus P_1) = \varnothing$. По аддитивности и неотрицательности площади

$$S(P) = S(P_1) + S(P \setminus P_1) \geqslant S(P_1)$$

(квадрируемость $P \setminus P_1$ мы не доказываем).

5. Если P содержится в некотором отрезке, то $S(P) = 0$.

Действительно, P можно поместить в прямоугольник сколь угодно малой площади, а тогда в силу монотонности $S(P)$ меньше любого положительного числа, то есть равна нулю (квадрируемость P мы не доказываем).

6. *Усиленная аддитивность.* Если квадрируемые фигуры P_1 и P_2 пересекаются по множеству нулевой площади (в частности, по отрезку), то $S(P_1 \cup P_2) = S(P_1) + S(P_2)$.

Для доказательства обозначим $P = P_1 \cap P_2$. Тогда по аддитивности

$$S(P_1) = S(P_1 \setminus P) + S(P) = S(P_1 \setminus P),$$
$$S(P_1 \cup P_2) = S(P_1 \setminus P) + S(P_2) = S(P_1) + S(P_2).$$

2. Объем. Далее слово "тело" будет означать то же, что "подмножество $\mathbb{R}^3$". Понятие объема некоторых геометрических тел известно из школьного курса геометрии. Определение объема для более широкого класса тел мы дадим лишь частично, полностью аналогично определению площади.

Определение 3. Объем. *Объемом* называется функционал $V:\{T\}\to[0,+\infty)$, заданный на некотором классе $\{T\}$ подмножеств трехмерного пространства, называемых *кубируемыми телами,* и обладающий следующими свойствами.

1. *Аддитивность.* Если T_1 и T_2 — кубируемые тела, причем $T_1\cap T_2=\varnothing$, то $T_1\cup T_2$ — кубируемое тело и

$$V(T_1\cup T_2)=V(T_1)+V(T_2).$$

2. *Нормированность на прямоугольных параллелепипедах.* Объем прямоугольного параллелепипеда с ребрами a, b и c равен abc.

3. *Инвариантность относительно движений.* Если T — кубируемое тело, U — движение пространства, то $U(T)$ — кубируемое тело и $V\big(U(T)\big)=V(T)$.

Множество кубируемых тел, то есть тел, имеющих объем, мы определять не будем, и поэтому существование объема в утверждениях этого пункта примем на веру.

Доказательство следующих свойств объема полностью аналогично плоскому случаю.

4. *Монотонность.* Если T и T_1 — кубируемые тела, $T_1\subset T$, то $V(T_1)\leqslant V(T)$.

5. Если тело T содержится в некотором прямоугольнике, то $V(T)=0$.

6. *Усиленная аддитивность.* Если кубируемые тела T_1 и T_2 пересекаются по множеству нулевого объема (в частности, по части прямоугольника), то $V(T_1\cup T_2)=V(T_1)+V(T_2)$.

Определение 4. Предположим, что $P\subset\mathbb{R}^2$, $h\geqslant 0$. Множество $Q=P\times[0,h]$, а также всякий образ Q при движении называется *прямым цилиндром* с основанием P и высотой h.

Примем без доказательства, что если P — квадрируемая фигура, то цилиндр Q кубируем и $V(Q)=S(P)h$. Идея доказательства этого факта такая же, как в знакомом из школьного курса частном случае прямого кругового цилиндра. Она состоит в приближении с любой точностью изнутри и снаружи множества P многоугольниками и, как следствие, цилиндра Q — цилиндрами с многоугольными основаниями (прямыми призмами).

Определение 5. Сечение. Пусть $T \subset \mathbb{R}^3$, $x \in \mathbb{R}$. Множество

$$T(x) = \big\{(y, z) \in \mathbb{R}^2 : (x, y, z) \in T\big\}$$

называется *сечением* множества T первой координатой x.

Мы будем рассматривать сечения только первой координатой; если же сечения производятся разными координатами, то удобно обозначать номер координаты индексом, например: $T_1(3)$, $T_2(a)$ или $T_3(\pi)$. Читателю может быть привычнее называть сечением подмножество $\mathbb{R}^3$, полученное пересечением тела с плоскостью, проходящей через точку оси OX с абсциссой x перпендикулярно OX. В приведенном определении мы “забываем” про первую координату и проектируем это пересечение на плоскость OYZ.

3. Длина пути. Дадим определение пути в $\mathbb{R}^m$ и введем некоторые понятия, связанные с путями. Пусть $\gamma\colon [a, b] \to \mathbb{R}^m$. Через $\gamma_i(t)$ ($i \in [1 : m]$) будут обозначаться координаты вектора $\gamma(t)$. Функция $\gamma_i\colon [a, b] \to \mathbb{R}$ называется *i-й координатной функцией* отображения γ. Если $\gamma_i \in C[a, b]$ при всех $i \in [1 : m]$, то будем говорить, что отображение γ непрерывно на $[a, b]$. Подробнее отображения со значениями в $\mathbb{R}^m$ изучаются в главах 5 и 6.

Определение 6. Путь. *Путем* в $\mathbb{R}^m$ называется непрерывное отображение отрезка в $\mathbb{R}^m$:

$$\gamma = (\gamma_1, \ldots, \gamma_m)\colon [a, b] \to \mathbb{R}^m.$$

Точка $\gamma(a)$ называется *началом,* $\gamma(b)$ — *концом* пути. Множество

$$\gamma^* = \gamma\big([a, b]\big),$$

то есть образ отрезка $[a, b]$, называется *носителем пути* γ.

Если $\gamma(a) = \gamma(b)$, то путь γ называется *замкнутым.* Если равенство $\gamma(t_1) = \gamma(t_2)$ имеет место лишь при $t_1 = t_2$ или $t_1, t_2 \in \{a, b\}$, то путь γ называется *простым* или *несамопересекающимся.*

Если $\gamma_i \in C^r[a,b]$ при всех $i \in [1:m]$ $\bigl(r \in \mathbb{N} \cup \{\infty\}\bigr)$, то путь γ называют *r раз непрерывно дифференцируемым* или *r-гладким* и пишут $\gamma \in C^r[a,b]$. Путь гладкости 1 называют просто *гладким.*

Эти определения — частные случаи понятия гладкой вектор-функции, которое будет введено в главе 6.

Если существует такое дробление $\{t_k\}_{k=0}^n$ отрезка $[a,b]$, что сужение γ на каждый отрезок дробления $[t_k, t_{k+1}]$ $(k \in [0:n-1])$ — гладкий путь, то путь γ называется *кусочно-гладким.*

Путь γ^-, задаваемый формулой

$$\gamma^-(t) = \gamma(a+b-t), \qquad t \in [a,b],$$

называется *противоположным* γ.

Термин “кривая” употребляется в математике в разных значениях. Часто говорят, что “кривая на плоскости задается уравнениями $x = \gamma_1(t)$, $y = \gamma_2(t)$”, то есть понимают под *кривой* носитель пути.

Однако, такое определение хотя и возможно, но охватывает множества, не похожие на линию в привычном представлении. Например, существует путь, носитель которого — m-мерный куб $[0,1]^m$. Такие отображения получили название *кривых Пеано.*

Желая исключить из рассмотрения подобные патологические примеры, на путь накладывают дополнительные требования. Например, гладкой (r-гладкой, кусочно-гладкой) кривой называют носитель гладкого (r-гладкого, кусочно-гладкого) пути, а *простой* или *жордановой кривой* — носитель простого пути. Упомянутые кривые Пеано не удовлетворяют ни одному из перечисленных ограничений, то есть имеют самопересечения и не являются кусочно-гладкими.

Возможен и другой подход к понятию кривой, при котором кривая вообще определяется не как множество точек.

Разные пути могут иметь одинаковые носители. Например, полуокружность $\{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 1,\ y \geqslant 0\}$ является носителем путей

$$\begin{aligned}
\gamma^1(t) &= \left(t, \sqrt{1-t^2}\right), && t \in [-1,1],\\
\gamma^2(t) &= (-\cos t, \sin t), && t \in [0,\pi],\\
\gamma^3(t) &= (\cos t, \sin t), && t \in [0,\pi],\\
\gamma^4(t) &= (\cos t, |\sin t|), && t \in [-\pi,\pi].
\end{aligned}$$

В этом примере $\gamma^3 = (\gamma^2)^-$, а путь γ^4 описывает "дважды пробегаемую" полуокружность. Пути γ^1 и γ^2 можно, в определенном смысле, не различать.

Определение 7. Два пути $\gamma\colon [a,b] \to \mathbb{R}^m$ и $\widetilde{\gamma}\colon [\alpha,\beta] \to \mathbb{R}^m$ называются *эквивалентными,* если существует строго возрастающая функция $u\colon [a,b] \xrightarrow{\text{на}} [\alpha,\beta]$, такая что $\gamma = \widetilde{\gamma} \circ u$.

Аргумент пути (точку отрезка) часто называют *параметром,* а функцию u со свойствами из определения эквивалентных путей — *допустимым преобразованием параметра.*

Замечание 1. В условиях определения 7 функция u непрерывна по теореме 8 § 4 главы 2 о разрывах и непрерывности монотонной функции.

Замечание 2. Введенное отношение, действительно, является отношением эквивалентности на множестве путей.

Доказательство. Для доказательства соотношения $\gamma \sim \gamma$ надо положить $u = \mathrm{id}_{[a,b]}$. Если $\gamma \sim \widetilde{\gamma}$, u — преобразование параметра для этой эквивалентности, то u^{-1} — преобразование параметра для эквивалентности $\widetilde{\gamma} \sim \gamma$ (функция, обратная к непрерывной строго возрастающей биекции — непрерывная строго возрастающая биекция). Наконец, если $\gamma_1 \sim \gamma_2$, $\gamma_2 \sim \gamma_3$, u_1 и u_2 — соответствующие преобразования параметра, то $u_2 \circ u_1$ — преобразование параметра для эквивалентности $\gamma_1 \sim \gamma_3$ (композиция непрерывных строго возрастающих биекций — непрерывная строго возрастающая биекция). $\square$

Класс эквивалентных путей называется *кривой,* а каждый представитель класса — *параметризацией* кривой. Кривую обозначают $\{\gamma\}$, где γ — какая-то ее параметризация.

Из определения ясно, что носители эквивалентных путей совпадают.

Носителем кривой называется общий носитель всех ее параметризаций.

Кривую $\{\gamma^-\}$ называют *ориентированной противоположно* $\{\gamma\}$.

Ясно, что носители противоположных путей совпадают и, следовательно, носители противоположно ориентированных кривых совпадают.

Кривая называется *гладкой* (*r-гладкой, кусочно-гладкой*), если у нее есть гладкая (r-гладкая, кусочно-гладкая) параметризация.

Замечание 3. В определении гладкой кривой требуется существование хотя бы одной гладкой параметризации и не запрещается существование негладких параметризаций.

Иногда эквивалентность путей определяют с учетом гладкости, а именно, дополнительно накладывают условия $u, u^{-1} \in C^r$. При таком определении всякий путь, эквивалентный r-гладкому, также будет r-гладким, и следовательно, всякая параметризация r-гладкой кривой также будет r-гладкой. Далее будет использоваться первоначальное определение эквивалентности путей.

Замечание 4. Имеет смысл рассматривать отображения (в том числе разрывные) промежутков другого типа, однако многие важные свойства (например, теорема Вейерштрасса) справедливы лишь в случае путей — непрерывных отображений отрезка. Поэтому, если не оговорено противное, будут рассматриваться именно пути.

В отличие от площади и объема, определение длины пути (кривой) мы дадим полностью.

Пусть $\gamma \in C\big([a,b] \to \mathbb{R}^m\big)$ — путь в $\mathbb{R}^m$. Постараемся дать определение длины s_γ пути γ так, чтобы удовлетворить нескольким естественным требованиям. Во-первых, длина пути, соединяющего точки A и B, должна быть не меньше длины отрезка AB. Во вторых, длина пути должна быть аддитивной функцией отрезка: если $a < c < b$, $\gamma^1 = \gamma|_{[a,c]}$, $\gamma^2 = \gamma|_{[c,b]}$, то

$$s_\gamma = s_{\gamma^1} + s_{\gamma^2}.$$

Пусть $\tau = \{t_k\}_{k=0}^n$ — дробление отрезка $[a,b]$. Семейство отрезков, соединяющих точки $\gamma(t_k)$ и $\gamma(t_{k+1})$ ($k \in [0:n-1]$), называется *ломаной*, вписанной в путь γ (рисунок 9). Длиной ℓ_τ ломаной, отвечающей дроблению τ, называют сумму длин составляющих ее отрезков.

Из первых двух требований вытекает, что длина пути должна быть не меньше длины любой вписанной в этот путь ломаной. Наконец, потребуем, чтобы длина пути могла быть приближена с любой точностью длинами вписанных ломаных. Эти условия приводят к следующему определению.

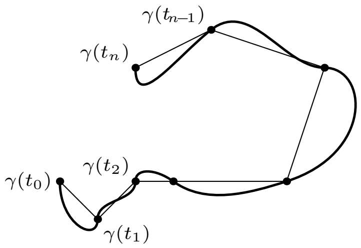

Рис. 9

Определение 8. Длина пути. Пусть γ — путь в $\mathbb{R}^m$. *Длиной пути* γ называется величина

$$s_\gamma = \sup_\tau \ell_\tau.$$

Хотя длина определяется для любого пути, из определения не следует, что она конечна.

Определение 9. Если $s_\gamma < +\infty$, то путь γ называется *спрямляемым.*

Пример неспрямляемого пути будет приведен в § 8.

Лемма 1. *Длины эквивалентных путей равны.*

Доказательство. Пусть $\gamma = \widetilde{\gamma} \circ u$, функция $u\colon [a,b] \xrightarrow{\text{на}} [\alpha,\beta]$ строго возрастает. Возьмем дробление $\tau = \{t_k\}_{k=0}^n$ отрезка $[a,b]$ и положим $\widetilde{t}_k = u(t_k)$. Тогда $\widetilde{\tau} = \{\widetilde{t}_k\}$ — дробление $[\alpha,\beta]$, и

$$\ell_\tau = \sum_{k=0}^{n-1} \bigl\|\gamma(t_{k+1}) - \gamma(t_k)\bigr\| = \sum_{k=0}^{n-1} \bigl\|\widetilde{\gamma}(\widetilde{t}_{k+1}) - \widetilde{\gamma}(\widetilde{t}_k)\bigr\| = \ell_{\widetilde{\tau}} \leqslant s_{\widetilde{\gamma}}.$$

В силу произвольности дробления τ имеем $s_\gamma \leqslant s_{\widetilde{\gamma}}$. Меняя γ и $\widetilde{\gamma}$ ролями, получаем противоположное неравенство $s_{\widetilde{\gamma}} \leqslant s_\gamma$. $\square$

Замечание 5. Аналогично доказывается, что длины противоположных путей равны.

Лемма 1 обеспечивает корректность следующего определения длины кривой: *длиной кривой* называют длину любой ее параметризации. Также имеет смысл говорить о спрямляемых кривых.

Лемма 2. Аддитивность длины пути. *Если* $\gamma\colon [a,b] \to \mathbb{R}^m$, $c \in (a,b)$, $\gamma^1 = \gamma|_{[a,c]}$, $\gamma^2 = \gamma|_{[c,b]}$, *то*

$$s_\gamma = s_{\gamma^1} + s_{\gamma^2}.$$

Доказательство. Обозначим $s_1 = s_{\gamma^1}$, $s_2 = s_{\gamma^2}$. Возьмем дробления τ_1 и τ_2 отрезков $[a,c]$ и $[c,b]$; тогда $\tau = \tau_1 \cup \tau_2$ — дробление $[a,b]$. Построим по τ_1 и τ_2 ломаные, вписанные в γ^1 и γ^2, и обозначим через ℓ_1 и ℓ_2 их длины. Тогда $\ell_1 + \ell_2 = \ell_\tau \leqslant s_\gamma$. Последовательно переходя в левой части к супремуму по всевозможным дроблениям τ_1 и τ_2, получаем:

$$s_1 + \ell_2 \leqslant s_\gamma,$$
$$s_1 + s_2 \leqslant s_\gamma.$$

Докажем противоположное неравенство

$$s_\gamma \leqslant s_1 + s_2.$$

Возьмем дробление $\tau = \{t_k\}_{k=0}^n$ отрезка $[a,b]$ и докажем, что $\ell_\tau \leqslant s_1 + s_2$; отсюда и будет следовать требуемое. Если $c \in \tau$, то τ представляется в виде $\tau_1 \cup \tau_2$, где τ_1 и τ_2 — дробления $[a,c]$ и $[c,b]$. Поэтому

$$\ell_\tau = \ell_1 + \ell_2 \leqslant s_1 + s_2.$$

Если $c \notin \tau$, то добавим c в число точек дробления, то есть положим $\tau^* = \tau \cup \{c\}$. Пусть $c \in (t_\nu, t_{\nu+1})$. По неравенству треугольника (см. § 8 главы 3)

$$\ell_\tau = \sum_{k=0}^{\nu-1} \big\|\gamma(t_{k+1}) - \gamma(t_k)\big\| + \big\|\gamma(t_{\nu+1}) - \gamma(t_\nu)\big\| +$$
$$+ \sum_{k=\nu+1}^{n-1} \big\|\gamma(t_{k+1}) - \gamma(t_k)\big\| \leqslant \sum_{k=0}^{\nu-1} \big\|\gamma(t_{k+1}) - \gamma(t_k)\big\| + \big\|\gamma(c) - \gamma(t_\nu)\big\| +$$
$$+ \big\|\gamma(t_{\nu+1}) - \gamma(c)\big\| + \sum_{k=\nu+1}^{n-1} \big\|\gamma(t_{k+1}) - \gamma(t_k)\big\| = \ell_{\tau^*}.$$

По доказанному
$$\ell_\tau \leqslant \ell_{\tau^*} \leqslant s_1 + s_2. \quad \square$$

Замечание 6. В определении длины пути и леммах 1 и 2 можно отказаться от непрерывности отображения γ (она не использовалась в доказательствах).

Это замечание будет использовано в § 8 о функциях ограниченной вариации.

§ 7. Приложения интеграла Римана

В этом параграфе используется следующая схема применения интеграла, которую мы не будем формализовать. Сначала искомая величина приближается с любой точностью суммами специального вида. Эти суммы оказываются интегральными для некоторой функции. Отсюда можно заключить, что искомая величина выражается интегралом.

1. Вычисление площадей.

Определение 1. Пусть $f\colon [a,b] \to \mathbb{R}$, $f \geqslant 0$. Множество
$$Q_f = \big\{(x,y) \in \mathbb{R}^2 : x \in [a,b],\ 0 \leqslant y \leqslant f(x)\big\}$$
называется *подграфиком* функции f. Если f непрерывна, то подграфик называют еще *криволинейной трапецией.*

Пусть $f \in R[a,b]$. Примем без доказательства, что подграфик f имеет площадь, и найдем ее. Для этого мы повторим рассуждения, которыми мотивировалась конструкция интеграла Римана. Возьмем дробление $\tau = \{x_k\}_{k=0}^n$ отрезка $[a,b]$, обозначим
$$m_k = \inf_{x\in[x_k,x_{k+1}]} f(x), \qquad M_k = \sup_{x\in[x_k,x_{k+1}]} f(x)$$
и составим суммы
$$s_\tau = \sum_{k=0}^{n-1} m_k \Delta x_k, \qquad S_\tau = \sum_{k=0}^{n-1} M_k \Delta x_k.$$

Геометрически s_τ есть сумма площадей меньших, а S_τ — бо́льших прямоугольников на рисунке 10, что по усиленной аддитивности совпадает с площадями объединений указанных прямоугольников. Поскольку Q_f содержит объединение меньших и содержится в объединении бо́льших прямоугольников,

$$s_\tau \leqslant S(Q_f) \leqslant S_\tau. \tag{12}$$

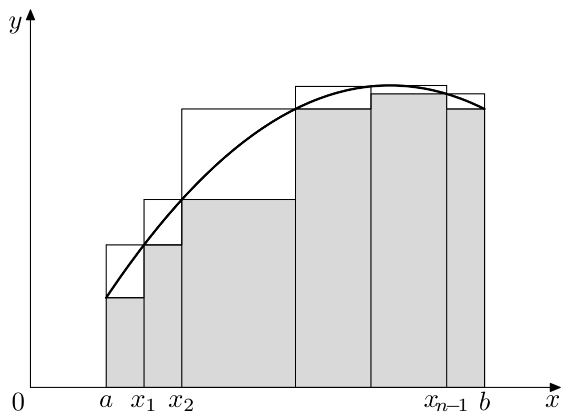

Рис. 10

С другой стороны, s_τ и S_τ — суммы Дарбу функции f. Так как f интегрируема,

$$\sup_\tau s_\tau = \inf_\tau S_\tau = \int_a^b f,$$

то есть неравенству (12) одновременно для всех τ удовлетворяет только одно число, а именно $\int_a^b f$. Значит,

$$S(Q_f) = \int_a^b f.$$

Замечание 1. Если определить подграфик f соотношениями $x \in [a, b]$, $0 \leqslant y < f(x)$, то его площадь будет выражаться той же самой формулой. Чтобы обосновать это, достаточно в приведенном доказательстве не присоединять границу к меньшим прямоугольникам. По аддитивности из сказанного следует, что площадь графика интегрируемой функции равна нулю. Далее при вычислении площади мы будем присоединять или не присоединять график к подграфику, в зависимости от ситуации.

Замечание 2. Доказанная формула допускает несколько очевидных обобщений.

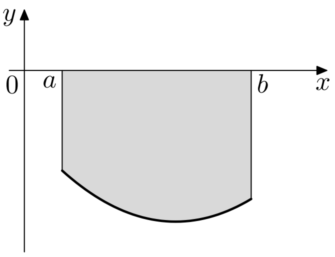

Рис. 11a

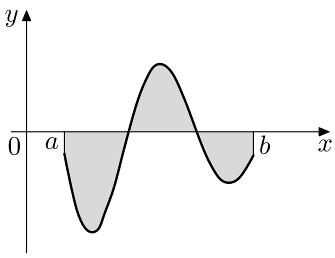

Рис. 11b

Если $f \in R[a,b]$, $f \leqslant 0$, то площадь закрашенной фигуры на рисунке 11a в силу инвариантности относительно движений совпадает с $S(Q_{-f})$ и потому равна $-\int_a^b f$.

В общем случае, если $f \in R[a,b]$, то площадь закрашенной фигуры на рисунке 11b равна $\int_a^b |f|$. Действительно, отразив часть фигуры, которая находится ниже оси абсцисс, относительно этой оси, мы получим, что площадь исходной фигуры равна $S(Q_{|f|})$.

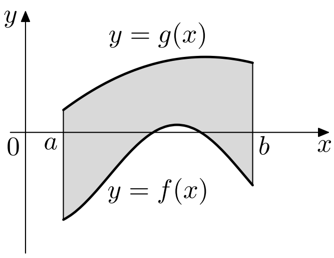

Рис. 11c

Если $f, g \in R[a, b]$, $f \leqslant g$, то площадь закрашенной фигуры на рисунке 11c (в случае непрерывных f и g эта фигура тоже называется *криволинейной трапецией*) равна $\int_a^b (g - f)$. Для доказательства следует перенести фигуру выше оси абсцисс (то есть добавить к f и g такую постоянную c, что $f + c \geqslant 0$) и представить ее в виде разности двух подграфиков. Получим:

$$S = \int_a^b \big((g + c) - (f + c)\big) = \int_a^b (g - f).$$

Пример 1. Найдем площадь S_E *эллипса*

$$E = \left\{(x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}, \qquad a, b > 0$$

(рисунок 12). Числа a и b называются *полуосями* эллипса.

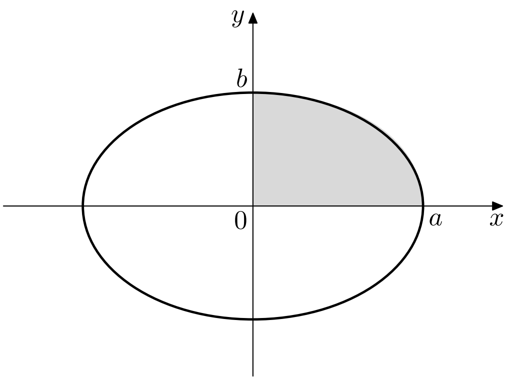

Рис. 12

Закрашенная четверть эллипса есть подграфик функции

$$y = b\sqrt{1 - \frac{x^2}{a^2}}, \qquad x \in [0, a].$$

Из соображений симметрии

$$S_E = 4\int_0^a b\sqrt{1 - \frac{x^2}{a^2}}\, dx = \frac{4b}{a}\int_0^a \sqrt{a^2 - x^2}\, dx = \pi ab$$

(последний интеграл был сосчитан в § 3 с помощью тригонометрической подстановки). При $b = a$ получается знакомая формула πa^2 для площади круга радиуса a.

Выведем теперь формулу площади в полярных координатах. Напомним, что полярные координаты r, φ связаны с декартовыми равенствами $x = r\cos\varphi$, $y = r\sin\varphi$.

Пусть $0 < \beta - \alpha \leqslant 2\pi$, $f\colon [\alpha, \beta] \to \mathbb{R}$, $f \geqslant 0$,

$$\widetilde{Q}_f = \bigl\{(r\cos\varphi, r\sin\varphi) \in \mathbb{R}^2 : \varphi \in [\alpha, \beta],\ 0 \leqslant r \leqslant f(\varphi)\bigr\}.$$

Если f непрерывна, то множество $\widetilde{Q}_f$ называют *криволинейным сектором.*

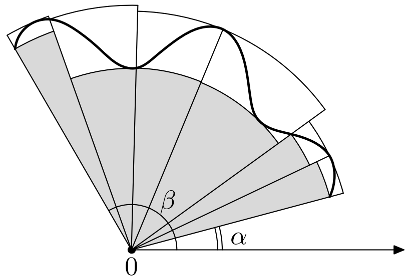

Рис. 13

Пусть $f \in R[\alpha, \beta]$. Примем без доказательства, что $\widetilde{Q}_f$ имеет площадь, и найдем ее. Напомним, что площадь кругового сектора с радиусом r и углом φ равна $\frac{1}{2}r^2\varphi$. Возьмем дробление $\tau = \{\varphi_k\}_{k=0}^n$ отрезка $[\alpha, \beta]$, обозначим

$$m_k = \inf_{\varphi \in [\varphi_k, \varphi_{k+1}]} f(\varphi), \qquad M_k = \sup_{\varphi \in [\varphi_k, \varphi_{k+1}]} f(\varphi)$$

и составим суммы

$$s_\tau = \frac{1}{2}\sum_{k=0}^{n-1} m_k^2\,\Delta\varphi_k, \qquad S_\tau = \frac{1}{2}\sum_{k=0}^{n-1} M_k^2\,\Delta\varphi_k.$$

Геометрически s_τ есть сумма площадей меньших, а S_τ — бóльших круговых секторов на рисунке 13, что по усиленной аддитивности совпадает с площадями объединений указанных секторов. Поскольку $\widetilde{Q}_f$ содержит объединение меньших и содержится в объединении бóльших секторов,

$$s_\tau \leqslant S(\widetilde{Q}_f) \leqslant S_\tau. \tag{13}$$

С другой стороны, s_τ и S_τ — суммы Дарбу функции $\frac{1}{2}f^2$. Так как эта функция интегрируема,

$$\sup_\tau s_\tau = \inf_\tau S_\tau = \frac{1}{2}\int_\alpha^\beta f^2,$$

то есть неравенству (13) одновременно для всех τ удовлетворяет только одно число, а именно $\frac{1}{2}\int_\alpha^\beta f^2$. Значит,

$$S(\widetilde{Q}_f) = \frac{1}{2}\int_\alpha^\beta f^2.$$

Пример 2. Найдем площадь $S_{\text{Л}}$ криволинейного сектора, ограниченного правым лепестком *лемнискаты Я. Бернулли*

$$r = a\sqrt{2\cos 2\varphi}, \qquad \varphi \in \left[-\frac{\pi}{4}, \frac{\pi}{4}\right],$$

где $a > 0$ (рисунок 14).

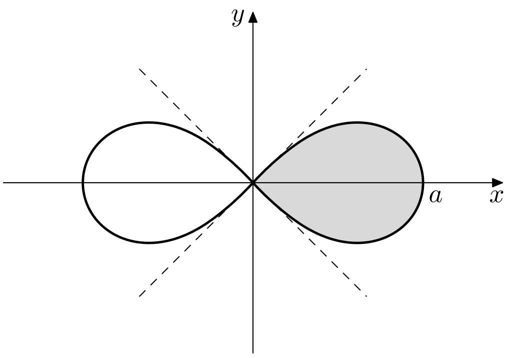

Рис. 14

Имеем:

$$S_{\text{Л}} = \frac{1}{2}\int_{-\pi/4}^{\pi/4} a^2 \cdot 2\cos 2\varphi\, d\varphi = a^2\, \frac{\sin 2\varphi}{2}\bigg|_{-\pi/4}^{\pi/4} = a^2.$$

2. Вычисление объемов. Предположим, что тело T удовлетворяет следующим условиям.

1. Существует такой отрезок $[a,b]$, что $T(x) = \varnothing$ для всех $x \notin [a,b]$.

2. При всех $x \in [a,b]$ сечение $T(x)$ — квадрируемая фигура с площадью $\mathcal{S}(x)$, причем $\mathcal{S} \in C[a,b]$.

3. Для любого отрезка $\Delta \subset [a,b]$ существуют такие $\xi_\Delta^*, \xi_\Delta^{**} \in \Delta$, что для всех $x \in \Delta$

$$T(\xi_\Delta^*) \subset T(x) \subset T(\xi_\Delta^{**}).$$

Примем без доказательства, что тело T имеет объем, и найдем его. Возьмем дробление $\tau = \{x_k\}_{k=0}^n$ отрезка $[a,b]$ и обозначим $\xi_k^* = \xi^*_{[x_k,x_{k+1}]}$, $\xi_k^{**} = \xi^{**}_{[x_k,x_{k+1}]}$,

$$m_k = \min_{x\in[x_k,x_{k+1}]} \mathcal{S}(x), \qquad M_k = \max_{x\in[x_k,x_{k+1}]} \mathcal{S}(x).$$

По монотонности площади

$$S\big(T(\xi_k^*)\big) = m_k, \qquad S\big(T(\xi_k^{**})\big) = M_k,$$

Обозначим через q_k и Q_k цилиндры, построенные на сечениях $T(\xi_k^*)$ и $T(\xi_k^{**})$ наименьшей и наибольшей площади (см. рисунок 15),

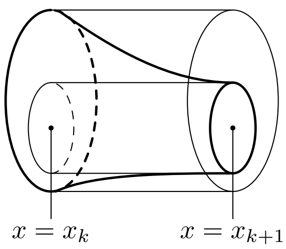

Рис. 15

то есть положим

$$q_k = [x_k, x_{k+1}] \times T(\xi_k^*), \qquad Q_k = [x_k, x_{k+1}] \times T(\xi_k^{**}).$$

В силу условия 3 для тела T верны включения

$$q_k \subset T_k \subset Q_k,$$

где множество

$$T_k = \{(x, y, z) \in T : x \in [x_k, x_{k+1}]\}$$

есть “слой” тела T между плоскостями $x = x_k$ и $x = x_{k+1}$ (рисунок 15). Поэтому

$$\bigcup_{k=0}^{n-1} q_k \subset T \subset \bigcup_{k=0}^{n-1} Q_k.$$

Составим суммы

$$w_\tau = \sum_{k=0}^{n-1} m_k \, \Delta x_k, \qquad W_\tau = \sum_{k=0}^{n-1} M_k \, \Delta x_k.$$

По усиленной аддитивности объема (свойству 6 объемов)

$$V\left(\bigcup_{k=0}^{n-1} q_k\right) = \sum_{k=0}^{n-1} V(q_k) = w_\tau, \qquad V\left(\bigcup_{k=0}^{n-1} Q_k\right) = \sum_{k=0}^{n-1} V(Q_k) = W_\tau.$$

По монотонности объема

$$w_\tau \leqslant V(T) \leqslant W_\tau. \tag{14}$$

С другой стороны, w_τ и W_τ — суммы Дарбу функции $\mathcal{S}$. Так как $\mathcal{S}$ интегрируема,

$$\sup_\tau w_\tau = \inf_\tau W_\tau = \int_a^b \mathcal{S},$$

то есть неравенству (14) одновременно для всех τ удовлетворяет только одно число, а именно $\int_a^b \mathcal{S}$. Значит,

$$V(T) = \int_a^b \mathcal{S}. \tag{15}$$

Пример 3. Найдем объем V_D *эллипсоида*

$$D = \left\{(x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1\right\}, \qquad a, b, c > 0.$$

Если $x \notin [-a, a]$, то $D(x) = \varnothing$; если $x = \pm a$, то $D(x) = \{(0, 0)\}$; если $x \in (-a, a)$, то

$$D(x) = \left\{(x, y) \in \mathbb{R}^2 : \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1 - \frac{x^2}{a^2}\right\}$$

есть эллипс с полуосями $b\sqrt{1 - \frac{x^2}{a^2}}$ и $c\sqrt{1 - \frac{x^2}{a^2}}$. Площадь эллипса вычислена в примере 1: $\mathcal{S}(x) = \pi bc\big(1 - \frac{x^2}{a^2}\big)$. Условие 3, накладывавшееся на тело, также выполнено, так как сечения расширяются с уменьшением $|x|$. Поэтому

$$V_D = \int_{-a}^{a} \pi bc \left(1 - \frac{x^2}{a^2}\right) dx = 2\pi bc \left[x - \frac{x^3}{3a^2}\right]_{x=0}^{a} = \frac{4}{3}\pi abc.$$

В частности, при $a = b = c$ получается формула $\frac{4}{3}\pi a^3$ для объема шара радиуса a.

Частным случаем равенства (15) является формула для объема *тела вращения*.

Пусть $f \colon [a, b] \to [\,0, +\infty)$, T_f — тело, получающееся вращением подграфика функции f вокруг оси OX (рисунок 16). Аналитически тело T_f задается равенством

$$T_f = \big\{(x, y, z) \in \mathbb{R}^3 : x \in [a, b],\ y^2 + z^2 \leqslant f^2(x)\big\}.$$

Аналогично можно определить тело вращения произвольной фигуры вокруг любой прямой.

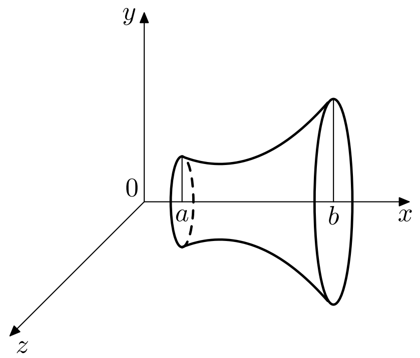

Рис. 16

Пусть $f \in C[a,b]$, $f \geqslant 0$. Для тела вращения T_f при каждом $x \in [a,b]$ сечение есть круг радиуса $f(x)$, поэтому $\mathcal{S}(x) = \pi f^2(x)$. Условия 1 – 3 выполнены и, значит,

$$V(T_f) = \pi \int_a^b f^2.$$

Пример 4. Найдем объем V_T *тора* — тела, образованного вращением круга $\{(x,y) : x^2 + (y-R)^2 \leqslant r^2\}$ $(0 < r < R)$ вокруг оси OX (рисунок 17).

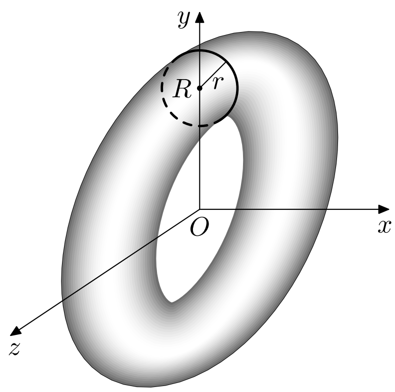

Рис. 17

Тор представляется в виде разности тел вращения подграфиков функций, графики которых — верхняя и нижняя полуокружности, то есть функций

$$f_1(x) = R + \sqrt{r^2 - x^2}, \quad f_2(x) = R - \sqrt{r^2 - x^2}, \qquad x \in [-r, r].$$

Поэтому

$$\begin{aligned} V_T &= \pi \int_{-r}^{r} f_1^2 - \pi \int_{-r}^{r} f_2^2 = \\ &= \pi \int_{-r}^{r} \left(\left(R + \sqrt{r^2 - x^2} \right)^2 - \left(R - \sqrt{r^2 - x^2} \right)^2 \right) dx = \\ &= 4\pi R \int_{-r}^{r} \sqrt{r^2 - x^2}\, dx = 2\pi^2 R r^2. \end{aligned}$$

Приведем без доказательства еще две формулы вычисления объемов.

Замечание 1. Пусть $0 \leqslant a < b$, T'_f — тело вращения подграфика непрерывной функции $f\colon [a, b] \to [\,0, +\infty)$ вокруг оси OY. Тогда

$$V(T'_f) = 2\pi \int_a^b x f(x)\, dx.$$

Замечание 2. Пусть $-\frac{\pi}{2} \leqslant \alpha < \beta \leqslant \frac{\pi}{2}$, $\widetilde{T}_f$ — тело вращения криволинейного сектора, определяемого непрерывной функцией $f\colon [\alpha, \beta] \to [\,0, +\infty)$, вокруг оси OY. Тогда

$$V(\widetilde{T}_f) = \frac{2\pi}{3} \int_\alpha^\beta f^3(\varphi) \cos\varphi\, d\varphi.$$

Замечание 3. Для неограниченных множеств формулы площадей и объемов остаются верными, но выражающие их интегралы будут несобственными.

3. Вычисление длин. Если $\gamma = (\gamma_1, \ldots, \gamma_m)$ — путь в $\mathbb{R}^m$, γ_i — дифференцируемые функции, то полагаем $\gamma' = (\gamma'_1, \ldots, \gamma'_m)$. Напомним, что по определению евклидовой длины

$$\|\gamma'\| = \sqrt{\sum_{i=1}^{m} {\gamma'_i}^2}.$$

Теорема 1. Длина гладкого пути. *Пусть* $\gamma : [a, b] \to \mathbb{R}^m$ — *гладкий путь. Тогда* γ *спрямляем и*

$$s_\gamma = \int_a^b \|\gamma'\|.$$

Доказательство. 1. Пусть $\Delta = [\alpha, \beta] \subset [a, b]$. Возьмем дробление $\eta = \{u_k\}_{k=0}^n$ отрезка Δ. Тогда по определению евклидовой длины

$$\ell_\eta = \sum_{k=0}^{n-1} \big\|\gamma(u_{k+1}) - \gamma(u_k)\big\| = \sum_{k=0}^{n-1} \sqrt{\sum_{i=1}^{m} \big(\gamma_i(u_{k+1}) - \gamma_i(u_k)\big)^2}.$$

По формуле Лагранжа при каждых i и k найдется такая точка $c_{ik} \in (u_k, u_{k+1})$, что

$$\gamma_i(u_{k+1}) - \gamma_i(u_k) = \gamma_i'(c_{ik})\Delta u_k.$$

Поэтому

$$\ell_\eta = \sum_{k=0}^{n-1} \sqrt{\sum_{i=1}^{m} \gamma_i'^2(c_{ik})} \cdot \Delta u_k.$$

Обозначим

$$M_\Delta^{(i)} = \max_{t \in \Delta} |\gamma_i'(t)|, \qquad m_\Delta^{(i)} = \min_{t \in \Delta} |\gamma_i'(t)|,$$
$$M_\Delta = \sqrt{\sum_{i=1}^{m} \big(M_\Delta^{(i)}\big)^2}, \qquad m_\Delta = \sqrt{\sum_{i=1}^{m} \big(m_\Delta^{(i)}\big)^2}$$

($M_\Delta^{(i)}$ и $m_\Delta^{(i)}$ существуют по теореме Вейерштрасса о непрерывных функциях). Тогда

$$m_\Delta(\beta - \alpha) \leqslant \ell_\eta \leqslant M_\Delta(\beta - \alpha).$$

Переходя к супремуму по всем дроблениям, мы получим

$$m_\Delta(\beta - \alpha) \leqslant s_{\gamma|_\Delta} \leqslant M_\Delta(\beta - \alpha).$$

В частности, при $\Delta = [a,b]$ отсюда следует, что путь γ спрямляем.

2. Возьмем дробление $\tau = \{t_k\}_{k=0}^n$ отрезка $[a,b]$ и обозначим

$$m_k = m_{[t_k,t_{k+1}]}, \qquad M_k = M_{[t_k,t_{k+1}]}.$$

По доказанному

$$m_k\Delta t_k \leqslant s_{\gamma|_{[t_k,t_{k+1}]}} \leqslant M_k\Delta t_k.$$

Кроме того, при всех $t \in [t_k, t_{k+1}]$

$$m_k \leqslant \|\gamma'(t)\| \leqslant M_k,$$

и поэтому

$$m_k\Delta t_k \leqslant \int_{t_k}^{t_{k+1}} \|\gamma'\| \leqslant M_k\Delta t_k.$$

Складывая неравенства и пользуясь аддитивностью длины пути и интеграла, получаем:

$$\begin{gathered}\sum_{k=0}^{n-1} m_k\Delta t_k \leqslant s_\gamma \leqslant \sum_{k=0}^{n-1} M_k\Delta t_k,\\ \sum_{k=0}^{n-1} m_k\Delta t_k \leqslant \int_a^b \|\gamma'\| \leqslant \sum_{k=0}^{n-1} M_k\Delta t_k.\end{gathered}\tag{16}$$

Осталось доказать, что между левыми и правыми частями неравенств (16) одновременно для всех дроблений лежит лишь одно число; из этого и будет следовать, что $s_\gamma = \int_a^b \|\gamma'\|$. Суммы в левой и правой части (16) не обязаны быть интегральными для $\|\gamma'\|$, поэтому оценим разность между ними непосредственно. Если $M_\Delta + m_\Delta \neq 0$, то

$$\begin{aligned}M_\Delta - m_\Delta = \frac{M_\Delta^2 - m_\Delta^2}{M_\Delta + m_\Delta} &= \frac{\sum\limits_{i=1}^m \Big(\big(M_\Delta^{(i)}\big)^2 - \big(m_\Delta^{(i)}\big)^2\Big)}{M_\Delta + m_\Delta} = \\ &= \sum_{i=1}^m \big(M_\Delta^{(i)} - m_\Delta^{(i)}\big)\frac{M_\Delta^{(i)} + m_\Delta^{(i)}}{M_\Delta + m_\Delta} \leqslant \sum_{i=1}^m \big(M_\Delta^{(i)} - m_\Delta^{(i)}\big).\end{aligned}$$

Если же $M_\Delta = m_\Delta = 0$, то доказанное неравенство очевидно.

Возьмем $\varepsilon > 0$. По теореме Кантора все функции $|\gamma_i'|$ равномерно непрерывны на $[a,b]$. Поэтому для каждого $i \in [1:m]$ найдется такое $\delta_i > 0$, что

$$\text{если} \quad x, y \in [a,b],\ |x-y| < \delta_i, \quad \text{то} \qquad \big||\gamma_i'(x)| - |\gamma_i'(y)|\big| < \frac{\varepsilon}{m(b-a)}.$$

Положим $\delta = \min\limits_{1 \leqslant i \leqslant m} \delta_i$. Если $\Delta = [\alpha, \beta] \subset [a,b]$, $\beta - \alpha < \delta$, то по теореме Вейерштрасса $M_\Delta^{(i)}$ и $m_\Delta^{(i)}$ суть значения $|\gamma_i'|$ в некоторых точках Δ, откуда $M_\Delta^{(i)} - m_\Delta^{(i)} < \frac{\varepsilon}{m(b-a)}$ и $M_\Delta - m_\Delta < \frac{\varepsilon}{b-a}$. Тогда для любого дробления τ, ранг которого меньше δ, при всех k будет $M_k - m_k < \frac{\varepsilon}{b-a}$. Поэтому

$$\left| s_\gamma - \int_a^b \|\gamma'\| \right| \leqslant \sum_{k=0}^{n-1} M_k \Delta t_k - \sum_{k=0}^{n-1} m_k \Delta t_k < \frac{\varepsilon}{b-a} \sum_{k=0}^{n-1} \Delta t_k = \varepsilon.$$

Так как ε произвольно, $s_\gamma = \int_a^b \|\gamma'\|$. $\square$

Замечание 1. По аддитивности теорема 1 распространяется на кусочно-гладкие пути.

Замечание 2. Запишем частный случай теоремы 1 при $m=2$. *Пусть* $\gamma = (\varphi, \psi) \in C^1\big([a,b] \to \mathbb{R}^2\big)$. *Тогда*

$$s_\gamma = \int_a^b \sqrt{\varphi'^2 + \psi'^2}.$$

Следствие 1. Длина графика. *График* Γ_f *функции* f *класса* $C^1[a,b]$ *спрямляем и*

$$s_{\Gamma_f} = \int_a^b \sqrt{1 + f'^2}.$$

Здесь под графиком f понимается путь

$$\Gamma_f(t) = \big(t, f(t)\big), \quad t \in [a,b].$$

Следствие 2. Длина пути в полярных координатах. *Пусть* $f \in C^1[\alpha, \beta]$, $f \geqslant 0$, *путь* γ *задается в полярных координатах равенством* $r = f(\theta)$, $\theta \in [\alpha, \beta]$. *Тогда*

$$s_\gamma = \int_a^b \sqrt{f^2 + f'^2}.$$

Поясним, что означает фраза “путь задается в полярных координатах равенством”. Подставляя $r = f(\theta)$ в формулы, связывающие декартовы и полярные координаты, получаем:

$$x = r\cos\theta = f(\theta)\cos\theta, \qquad y = r\sin\theta = f(\theta)\sin\theta.$$

Таким образом, речь идет о пути

$$\gamma(\theta) = \big(\varphi(\theta), \psi(\theta)\big) = \big(f(\theta)\cos\theta, f(\theta)\sin\theta\big), \qquad \theta \in [\alpha, \beta].$$

Доказательство. Дифференцируя, находим:

$$\varphi'^2(\theta) + \psi'^2(\theta) =$$
$$= (f'(\theta)\cos\theta - f(\theta)\sin\theta)^2 + (f'(\theta)\sin\theta + f(\theta)\cos\theta)^2 =$$
$$= \big(f'^2(\theta) + f^2(\theta)\big)(\cos^2\theta + \sin^2\theta) = f'^2(\theta) + f^2(\theta). \quad \square$$

Пример 5. Пусть $a > 0$. Найдем длину $s_\Pi(a)$ части *параболы* $y = x^2$ от точки $(0, 0)$ до точки (a, a^2) (рисунок 18).

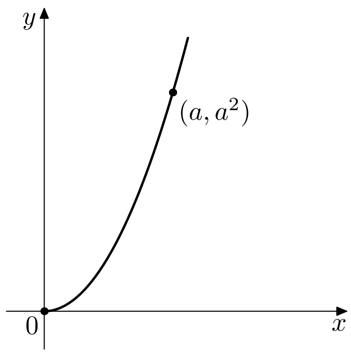

Рис. 18

Так как $\left(x^2\right)' = 2x$, по формуле для длины графика

$$s_{\Pi}(a) = \int_0^a \sqrt{1+4x^2}\,dx.$$

Проинтегрируем по частям и сведем интеграл к самому себе:

$$\begin{aligned} s_{\Pi}(a) &= x\sqrt{1+4x^2}\Big|_0^a - \int_0^a \frac{4x^2}{\sqrt{1+4x^2}}\,dx = \\ &= a\sqrt{1+4a^2} - \int_0^a \frac{1+4x^2}{\sqrt{1+4x^2}}\,dx + \int_0^a \frac{dx}{\sqrt{1+4x^2}} = \\ &= a\sqrt{1+4a^2} - s_{\Pi}(a) + \frac{1}{2}\ln\left(2a+\sqrt{1+4a^2}\right), \end{aligned}$$

откуда

$$s_{\Pi}(a) = \frac{a}{2}\sqrt{1+4a^2} + \frac{1}{4}\ln\left(2a+\sqrt{1+4a^2}\right).$$

Пример 6. Пусть $\beta \in [\,0, 2\pi]$. Найдем длину $s_{Э}(\beta)$ дуги *эллипса*

$$x = a\cos t, \quad y = b\sin t, \qquad t \in [\,0, \beta].$$

Для определенности будем считать, что $0 < a \leqslant b$ (рисунок 19).

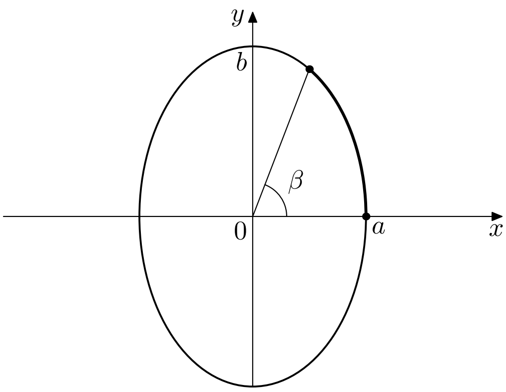

Рис. 19

Имеем:

$$s_Э(\beta) = \int_0^\beta \sqrt{a^2 \sin^2 t + b^2 \cos^2 t}\, dt =$$
$$= \int_0^\beta \sqrt{b^2 - (b^2 - a^2) \sin^2 t}\, dt = b \int_0^\beta \sqrt{1 - \varepsilon^2 \sin^2 t}\, dt.$$

Величина $\varepsilon = \frac{\sqrt{b^2 - a^2}}{b}$ называется *эксцентриситетом* эллипса. Эксцентриситет характеризует меру "сжатости" эллипса. Очевидно, что $\varepsilon \in [0, 1)$; если $\varepsilon = 0$, то $a = b$, то есть эллипс есть окружность. В предельном случае $\varepsilon = 1$ мы получаем $a = 0$, то есть эллипс вырождается в дважды пробегаемый отрезок.

Интеграл

$$E(\varepsilon, \beta) = \int_0^\beta \sqrt{1 - \varepsilon^2 \sin^2 t}\, dt$$

называется *эллиптическим интегралом второго рода.* При $\beta = \frac{\pi}{2}$ эллиптический интеграл называется *полным.*

Таким образом, длина всего эллипса равна

$$s_Э = 4bE\Big(\varepsilon, \frac{\pi}{2}\Big).$$

Эллиптическим интегралом первого рода называется интеграл

$$K(\varepsilon, \beta) = \int_0^\beta \frac{1}{\sqrt{1 - \varepsilon^2 \sin^2 t}}\, dt.$$

Этот интеграл возникает при вычислении длины дуги лемнискаты.

Эллиптические интегралы, вообще говоря, неберущиеся, то есть первообразная подынтегральной функции не является элементарной.

4. Работа силы. В качестве одного из физических приложений интеграла разберем вычисление работы силы.

Рассмотрим движение тела (материальной точки) по прямой, которую примем за числовую ось. Пусть на тело действует сила F, зависящая лишь от положения (координаты) тела. Будем считать, что вектор силы параллелен оси. Мы не будем различать вектор силы и его единственную координату, которую также обозначим F. Таким образом, F есть функция одной переменной, заданная на некотором промежутке E. Ясно, что $F(x) > 0$, если в точке x вектор силы сонаправлен с осью, и $F(x) < 0$, если он направлен противоположно оси.

Обозначим через $A_F\bigl([a,b]\bigr)$ работу силы F при перемещении тела из точки a в точку b. Если сила F постоянна на $[a,b]$, то $A_F\bigl([a,b]\bigr) = F\cdot(b-a)$ (это равенство, верное при любом расположении точек a и b, часто принимают за определение работы постоянной силы). Задача состоит в том, чтобы найти работу силы в общем случае, то есть когда сила переменная. Мы не будем обсуждать физическое содержание понятия работы (работой в физике называют количество энергии, затраченное при перемещении тела под действием силы), а сформулируем два свойства работы. Для определенности будем считать, что $a < b$.

1. Функция A_F аддитивна по отрезку.
2. Если $m \leqslant F \leqslant M$ на отрезке $[a,b]$, то

$$m(b-a) \leqslant A_F\bigl([a,b]\bigr) \leqslant M(b-a).$$

Эти свойства будем считать известными из физических соображений.

Предположим еще, что функция F непрерывна.

Возьмем дробление $\tau = \{x_k\}_{k=0}^n$ отрезка $[a,b]$ и положим

$$m_k = \min_{x\in[x_k,x_{k+1}]} F(x), \qquad M_k = \max_{x\in[x_k,x_{k+1}]} F(x),$$
$$s_\tau = \sum_{k=0}^{n-1} m_k\Delta x_k, \qquad S_\tau = \sum_{k=0}^{n-1} M_k\Delta x_k.$$

Из свойств 1 и 2 следует, что

$$s_\tau \leqslant A_F\bigl([a,b]\bigr) \leqslant S_\tau.$$

С другой стороны, s_τ и S_τ — суммы Дарбу функции F, а между множествами всех нижних и всех верхних сумм лежит лишь одно число, а именно $\int_a^b F$. Значит,

$$A_F\bigl([a,b]\bigr) = \int_a^b F.$$

Если сила направлена произвольно, то работа силы равна работе ее составляющей, параллельной оси движения.

§ 7′. Приложения интеграла как аддитивной функции отрезка

1. Аддитивные функции отрезка. Сначала опишем общую схему применения интеграла, позволяющую заключать, что при выполнении ряда условий искомая величина выражается интегралом от некоторой функции.

Далее E — невырожденный промежуток в $\mathbb{R}$, J_E — множество всех невырожденных отрезков, содержащихся в E.

Определение 1. Отображение $I\colon J_E \to \mathbb{R}$ называется *функцией отрезка.* Функция отрезка I называется *аддитивной,* если для любых a, b и c, таких что $a, b \in E$, $a < c < b$,

$$I\bigl([a,b]\bigr) = I\bigl([a,c]\bigr) + I\bigl([c,b]\bigr). \tag{17}$$

По индукции свойство аддитивности распространяется на случай нескольких слагаемых: если I аддитивна, $x_0 < x_1 < \ldots < x_n$, $x_0, x_n \in E$, то

$$I\bigl([x_0,x_n]\bigr) = \sum_{k=0}^{n-1} I\bigl([x_k,x_{k+1}]\bigr).$$

Интеграл $I\bigl([a,b]\bigr) = \int_a^b f$ является примером аддитивной функции отрезка. Здесь f кусочно-непрерывна на любом отрезке, содержащемся в E. В частности, если $f \equiv 1$, то I есть длина отрезка. Длину отрезка Δ мы будем обозначать через $|\Delta|$. Многие важные функции отрезка представляются в виде интеграла; конкретные примеры приводятся далее.

Мы будем рассматривать два вида пределов функций отрезка, когда длина отрезка стремится к нулю.

Определение 2. Пусть $\Phi\colon J_E \to \mathbb{R}$, $A \in \mathbb{R}$.

1. Число A называют *пределом* функции Φ при $|\Delta| \to 0$ и пишут $A = \lim\limits_{|\Delta|\to 0} \Phi(\Delta)$ или $\Phi(\Delta) \underset{|\Delta|\to 0}{\longrightarrow} A$, если

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall \Delta \in J_E : \; |\Delta| < \delta \;\; |\Phi(\Delta) - A| < \varepsilon.$$

2. Пусть $x_0 \in E$. Число A называют *пределом* функции Φ при $\Delta \to \{x_0\}$ и пишут $A = \lim\limits_{\Delta \to \{x_0\}} \Phi(\Delta)$ или $\Phi(\Delta) \underset{\Delta \to \{x_0\}}{\longrightarrow} A$, если

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall \Delta \in J_E : |\Delta| < \delta, \, x_0 \in \Delta \ \ |\Phi(\Delta) - A| < \varepsilon.$$

Пункт 2 отличается от 1 тем, что все отрезки Δ содержат фиксированную точку x_0.

Определение 3. Пусть функция $I\colon J_E \to \mathbb{R}$ аддитивна, $x_0 \in E$. Предел $\lim\limits_{\Delta \to \{x_0\}} \dfrac{I(\Delta)}{|\Delta|}$ называется *плотностью* или *производной* функции I в точке x_0 и обозначается $I'(x_0)$.

Если $I'(x_0)$ существует в каждой точке $x_0 \in E$, то функция $x \mapsto I'(x)$ $(x \in E)$ называется *плотностью* или *производной* (*производной функцией*) функции I.

Поясним оба термина. Рассмотрим тонкий стержень (не обязательно однородный), который будем отождествлять с отрезком E, и обозначим через $M(\Delta)$ массу части стержня $\Delta \in J_E$. Из физических соображений функция M аддитивна. Плотностью массы стержня в точке x_0 в физике как раз называют предел $\frac{M(\Delta)}{|\Delta|}$ при $\Delta \to \{x_0\}$. Использование термина “производная” объясняется следующей далее леммой 1.

Примем следующее соглашение: если $a, b \in E$, $b < a$, то положим $I\big([a,b]\big) = -I\big([b,a]\big)$; также положим $I\big([a,a]\big) = 0$. Этим соглашением функция I распространяется на направленные отрезки с сохранением свойства аддитивности: для любых $a, b, c \in E$ верно равенство (17). При этом отношение $\dfrac{I\big([a,b]\big)}{b-a}$ не меняется при перестановке a и b.

Лемма 1. *Пусть функция $I\colon J_E \to \mathbb{R}$ аддитивна, $a \in E$, $\Phi(x) = I\big([a,x]\big)$ $(x \in E)$, $x_0 \in E$. Тогда $I'(x_0)$ и $\Phi'(x_0)$ существуют или нет одновременно и, если существуют, то равны.*

Доказательство. 1. Пусть существует $I'(x_0) = \rho$. Возьмем $\varepsilon > 0$ и подберем $\delta > 0$ из определения $I'(x_0)$. Тогда для любого $x \in E$, такого что $0 < |x - x_0| < \delta$,

$$\left| \frac{\Phi(x) - \Phi(x_0)}{x - x_0} - \rho \right| = \left| \frac{I\big([x_0, x]\big)}{x - x_0} - \rho \right| < \varepsilon.$$

Это и значит, что $\rho = \Phi'(x_0)$.

2. Пусть существует $\Phi'(x_0) = \rho$. Возьмем $\varepsilon > 0$ и по определению $\Phi'(x_0)$ подберем такое $\delta > 0$, что для любого $x \in E$, удовлетворяющего неравенству $0 < |x - x_0| < \delta$,

$$\rho - \varepsilon < \frac{\Phi(x) - \Phi(x_0)}{x - x_0} < \rho + \varepsilon.$$

Пусть $[a, b] \subset E$, $x_0 \in [a, b]$, $0 < b - a < \delta$. Тогда

$$I\big([a,b]\big) = I\big([a,x_0]\big) + I\big([x_0,b]\big) = (\Phi(x_0) - \Phi(a)) + (\Phi(b) - \Phi(x_0)) < \\ < (\rho+\varepsilon)(x_0 - a) + (\rho+\varepsilon)(b - x_0) = (\rho+\varepsilon)(b-a),$$

и, аналогично,

$$I\big([a,b]\big) > (\rho - \varepsilon)(b-a).$$

Это и значит, что $\rho = I'(x_0)$. □

Следствие. Восстановление аддитивной функции отрезка по ее плотности. *Пусть функция* $I\colon J_E \to \mathbb{R}$ *аддитивна,* $f \in C(E)$. *Тогда следующие утверждения равносильны.*

1. f *— плотность* I.
2. *Для любого* $[a, b] \in J_E$ *верно равенство* $I\big([a,b]\big) = \int_a^b f$.

Доказательство. Если $f = I'$, то по лемме 1 и формуле Ньютона – Лейбница для любого отрезка $[a, b] \subset E$

$$\int_a^b f = \Phi(b) - \Phi(a) = I\big([a,b]\big).$$

Обратное верно по теореме Барроу. □

Докажем утверждение, позволяющее проверять, что некоторая функция является плотностью I.

Теорема 1. Признак плотности. *Пусть функция* $I\colon J_E \to \mathbb{R}$ *аддитивна,* $f\colon E \to \mathbb{R}$. *Предположим, что существуют функции* $m, M\colon J_E \to \mathbb{R}$, *удовлетворяющие следующим условиям:*

1) $m_\Delta|\Delta| \leqslant I(\Delta) \leqslant M_\Delta|\Delta|$ *для любого* $\Delta \in J_E$;
2) $m_\Delta \leqslant f(x) \leqslant M_\Delta$ *для любых* $\Delta \in J_E$, $x \in \Delta$;
3) $M_\Delta - m_\Delta \to 0$ *при* $|\Delta| \to 0$.

Тогда f — плотность I.

Здесь и далее значения функций m и M на отрезке Δ обозначаются m_Δ и M_Δ вместо $m(\Delta)$ и $M(\Delta)$.

Доказательство. Пусть $\Delta \in J_E$ и $x_0 \in \Delta$. Так как числа $\frac{I(\Delta)}{|\Delta|}$ и $f(x_0)$ лежат на $[m_\Delta, M_\Delta]$, в силу условий 1) и 2) верно неравенство

$$\left|\frac{I(\Delta)}{|\Delta|} - f(x_0)\right| \leqslant M_\Delta - m_\Delta.$$

По условию 3) для любого $\varepsilon > 0$ можно подобрать такое $\delta > 0$, что если $|\Delta| < \delta$, то $M_\Delta - m_\Delta < \varepsilon$. Тогда и $\left|\frac{I(\Delta)}{|\Delta|} - f(x_0)\right| < \varepsilon$. $\square$

Замечание 1. Если в условиях теоремы $f \in C(E)$, то для любого $[a,b] \in J_E \quad I\big([a,b]\big) = \int_a^b f$.

Следствие. *Пусть функция $I\colon J_E \to \mathbb{R}$ аддитивна, $f \in C(E)$,*

$$m_\Delta = \min_{x\in\Delta} f(x), \qquad M_\Delta = \max_{x\in\Delta} f(x). \tag{18}$$

Если для любого $\Delta \in J_E$

$$m_\Delta|\Delta| \leqslant I(\Delta) \leqslant M_\Delta|\Delta|,$$

то для всех $[a,b] \in J_E$ верно равенство $I\big([a,b]\big) = \int_a^b f$.

Отметим, что минимум и максимум в (18) существуют по теореме Вейерштрасса.

Доказательство. Применим теорему 1 для $E = [a,b]$ и функций (18). Условия 1) и 2) теоремы 1 очевидны, поэтому достаточно проверить 3). Пусть $\varepsilon > 0$. По теореме Кантора существует такое $\delta > 0$, что для любых $x, y \in [a,b]$: $|x-y| < \delta$ верно $|f(x) - f(y)| < \varepsilon$. По теореме Вейерштрасса для любого $\Delta \in J_E$ числа m_Δ и M_Δ суть значения f в некоторых точках $x_*, x^* \in \Delta$. Если $|\Delta| < \delta$, то $|x^* - x_*| < \delta$, откуда

$$M_\Delta - m_\Delta = |f(x^*) - f(x_*)| < \varepsilon. \quad \square$$

Замечание 2. Проверка условия следствия для m и M, определенных формулами (18), иногда бывает затруднительна, и тогда приходится использовать теорему 1 с другими функциями m и M.

Замечание 3. Если $f \in C(E)$ — плотность I, m и M определены формулами (18), то условия 1) и 2) выполняются автоматически, а 3) равносильно равномерной непрерывности f на E.

Таким образом, для равномерно непрерывных на E функций f теорему 1 можно обратить: если f есть плотность I, то существуют функции m и M (определяемые формулами (18)), для которых выполняются условия 1) – 3).

2. Вычисление площадей.

Определение 4. Пусть $f\colon [a,b] \to \mathbb{R}$, $f \geqslant 0$. Множество

$$Q_f = \bigl\{(x,y) \in \mathbb{R}^2 : x \in [a,b],\, 0 \leqslant y \leqslant f(x)\bigr\}$$

называется *подграфиком* функции f. Если f непрерывна, то подграфик называют еще *криволинейной трапецией.*

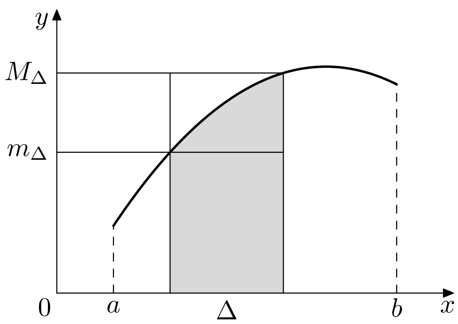

Рис. 20

Пусть $f \in C[a,b]$. Примем без доказательства, что подграфик f имеет площадь, и найдем ее. Для каждого отрезка $\Delta \subset [a,b]$ положим $I(\Delta) = S(Q_{f|_\Delta})$. Из усиленной аддитивности площади (см. § 6) следует, что функция I аддитивна. Определим m и M равенствами (18). Тогда условие 1) в теореме 1 выполняется ввиду монотонности площади (рисунок 20). По следствию теоремы 1 функция f является плотностью I и, следовательно,

$$S(Q_f) = \int_a^b f.$$

Замечание 1. Если определить подграфик f соотношениями $x \in [a, b]$, $0 \leqslant y < f(x)$, то его площадь будет выражаться той же самой формулой. Доказательство сохраняет силу. По аддитивности из сказанного следует, что площадь графика непрерывной функции равна нулю. Далее при вычислении площади мы будем присоединять или не присоединять график к подграфику, в зависимости от ситуации.

Замечание 2. Доказанная формула допускает несколько очевидных обобщений.

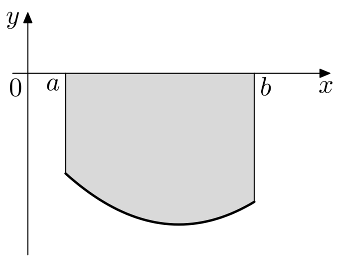

Рис. 21a

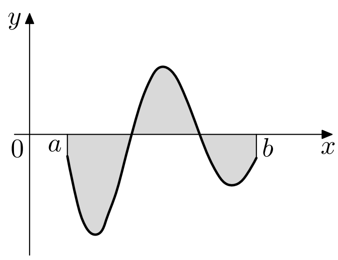

Рис. 21b

Если $f \in C[a, b]$, $f \leqslant 0$, то площадь закрашенной фигуры на рисунке 21a в силу инвариантности относительно движений совпадает с $S(Q_{-f})$ и потому равна $-\int_a^b f$.

В общем случае, если $f \in C[a, b]$, то площадь закрашенной фигуры на рисунке 21b равна $\int_a^b |f|$. Действительно, отразив часть фигуры, которая находится ниже оси абсцисс, относительно этой оси, получим, что площадь исходной фигуры равна $S(Q_{|f|})$.

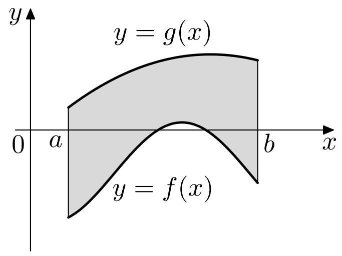

Рис. 21c

Если $f, g \in C[a,b]$, $f \leqslant g$, то площадь закрашенной фигуры на рисунке 21c (эта фигура тоже называется *криволинейной трапецией*) равна $\int_a^b (g-f)$. Для доказательства следует перенести фигуру выше оси абсцисс (то есть добавить к f и g такую постоянную c, что $f + c \geqslant 0$) и представить ее в виде разности двух подграфиков. Получим:

$$S = \int_a^b \big((g+c) - (f+c)\big) = \int_a^b (g-f).$$

Пример 1. Найдем площадь S_E *эллипса*

$$E = \left\{(x,y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leqslant 1\right\}, \qquad a, b > 0$$

(рисунок 22). Числа a и b называются *полуосями* эллипса.

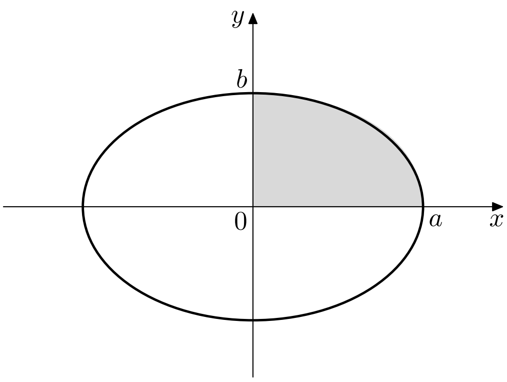

Рис. 22

Закрашенная четверть эллипса есть подграфик функции

$$y = b\sqrt{1 - \frac{x^2}{a^2}}, \qquad x \in [0, a].$$

Из соображений симметрии

$$S_E = 4\int_0^a b\sqrt{1 - \frac{x^2}{a^2}}\, dx = \frac{4b}{a}\int_0^a \sqrt{a^2 - x^2}\, dx = \pi ab$$

(последний интеграл был сосчитан в § $3'$ с помощью тригонометрической подстановки). При $b = a$ получается знакомая формула πa^2 для площади круга радиуса a.

Выведем теперь формулу площади в полярных координатах. Напомним, что полярные координаты r, φ связаны с декартовыми равенствами $x = r\cos\varphi$, $y = r\sin\varphi$.

Пусть $0 < \beta - \alpha \leqslant 2\pi$, $f\colon [\alpha, \beta] \to \mathbb{R}$, $f \geqslant 0$,

$$\widetilde{Q}_f = \bigl\{(r\cos\varphi, r\sin\varphi) \in \mathbb{R}^2 : \varphi \in [\alpha, \beta],\ 0 \leqslant r \leqslant f(\varphi)\bigr\}.$$

Если f непрерывна, то множество $\widetilde{Q}_f$ называют *криволинейным сектором*.

Пусть $f \in C[\alpha, \beta]$. Примем без доказательства, что $\widetilde{Q}_f$ имеет площадь, и найдем ее. Напомним, что площадь кругового сектора с радиусом r и углом φ равна $\frac{1}{2}r^2\varphi$.

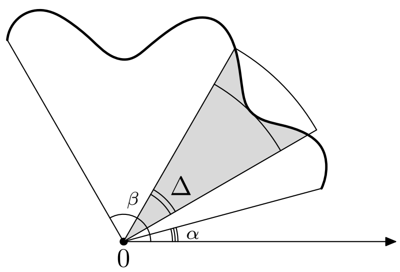

Рис. 23

Для каждого отрезка $\Delta \subset [\alpha, \beta]$ положим $I(\Delta) = S\big(\widetilde{Q}_{f|_\Delta}\big)$. Из усиленной аддитивности площади следует, что функция I аддитивна. Определим m и M равенствами

$$m_\Delta = \frac{1}{2}\min_{x\in\Delta} f^2(x), \qquad M_\Delta = \frac{1}{2}\max_{x\in\Delta} f^2(x).$$

Тогда двойное неравенство

$$m_\Delta|\Delta| \leqslant I(\Delta) \leqslant M_\Delta|\Delta|$$

выполняется ввиду монотонности площади (рисунок 23). По следствию теоремы 1 функция $\frac{1}{2}f^2$ — плотность I и, следовательно,

$$S\big(\widetilde{Q}_f\big) = \frac{1}{2}\int_\alpha^\beta f^2.$$

Пример 2. Найдем площадь $S_{\text{Л}}$ криволинейного сектора, образованного правым лепестком *лемнискаты Я. Бернулли*

$$r = a\sqrt{2\cos 2\varphi}, \qquad \varphi \in \Big[-\frac{\pi}{4}, \frac{\pi}{4}\Big],$$

где $a > 0$ (рисунок 24).

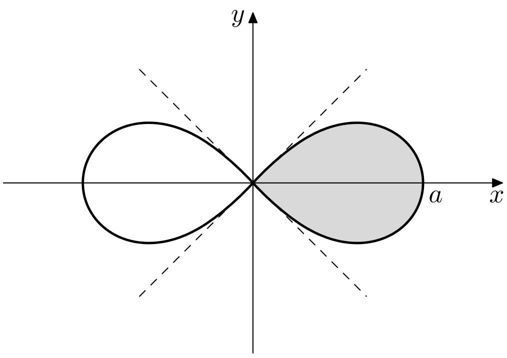

Рис. 24

Имеем:

$$S_{\text{Л}} = \frac{1}{2}\int_{-\pi/4}^{\pi/4} a^2 \cdot 2\cos 2\varphi \, d\varphi = a^2\,\frac{\sin 2\varphi}{2}\bigg|_{-\pi/4}^{\pi/4} = a^2.$$

3. Вычисление объемов. Предположим, что тело T удовлетворяет следующим условиям.

1. Существует такой отрезок $[a,b]$, что $T(x) = \varnothing$ для всех $x \notin [a,b]$.

2. При всех $x \in [a,b]$ сечение $T(x)$ — квадрируемая фигура с площадью $\mathcal{S}(x)$, причем $\mathcal{S} \in C[a,b]$.

3. Для любого отрезка $\Delta \subset [a,b]$ существуют такие $\xi^*_\Delta, \xi^{**}_\Delta \in \Delta$, что для всех $x \in \Delta$

$$T(\xi^*_\Delta) \subset T(x) \subset T(\xi^{**}_\Delta).$$

Примем без доказательства, что тело T имеет объем, и найдем его. Для каждого отрезка $\Delta \subset [a,b]$ положим

$$I(\Delta) = V(T_\Delta),$$

где множество

$$T_\Delta = \{(x,y,z) \in T : x \in \Delta\}$$

есть "слой" тела T (рисунок 25). Из усиленной аддитивности объема (см. § 6) следует, что функция I аддитивна. Определим m и M равенствами

$$m_\Delta = \min_{x\in\Delta} \mathcal{S}(x), \qquad M_\Delta = \max_{x\in\Delta} \mathcal{S}(x).$$

По монотонности площади

$$S\big(T(\xi^*_\Delta)\big) = m_\Delta, \qquad S\big(T(\xi^{**}_\Delta)\big) = M_\Delta.$$

Обозначим через q_Δ и Q_Δ цилиндры, построенные на сечениях $T(\xi^*_\Delta)$ и $T(\xi^{**}_\Delta)$ наименьшей и наибольшей площади (рисунок 25).

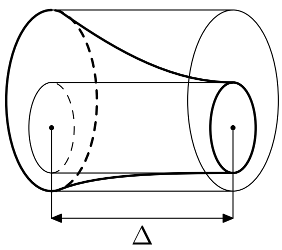

Рис. 25

Иначе говоря, положим

$$q_\Delta = \Delta \times T(\xi_\Delta^*), \qquad Q_\Delta = \Delta \times T(\xi_\Delta^{**}).$$

В силу условия 3 для тела T верны включения

$$q_\Delta \subset T_\Delta \subset Q_\Delta.$$

Поэтому двойное неравенство

$$m_\Delta |\Delta| \leqslant I(\Delta) \leqslant M_\Delta |\Delta|$$

выполняется ввиду монотонности объема. По следствию теоремы 1 функция $\mathcal{S}$ — плотность I и, значит,

$$V(T) = \int_a^b \mathcal{S}. \tag{19}$$

Пример 3. Найдем объем V_D *эллипсоида*

$$D = \left\{ (x,y,z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1 \right\}, \qquad a,b,c > 0.$$

Если $x \notin [-a,a]$, то $D(x) = \varnothing$; если $x = \pm a$, то $D(x) = \{(0,0)\}$; если $x \in (-a,a)$, то

$$D(x) = \left\{ (x,y) \in \mathbb{R}^2 : \frac{y^2}{b^2} + \frac{z^2}{c^2} \leqslant 1 - \frac{x^2}{a^2} \right\}$$

есть эллипс с полуосями $b\sqrt{1-\frac{x^2}{a^2}}$ и $c\sqrt{1-\frac{x^2}{a^2}}$. Площадь эллипса вычислена в примере 1: $\mathcal{S}(x) = \pi bc\bigl(1-\frac{x^2}{a^2}\bigr)$. Условие 3, накладывавшееся на тело, также выполнено, так как сечения расширяются с уменьшением $|x|$. Поэтому

$$V_D = \int_{-a}^a \pi bc \left(1 - \frac{x^2}{a^2}\right)\, dx = 2\pi bc \left[x - \frac{x^3}{3a^2} \right]_{x=0}^a = \frac{4}{3}\pi abc.$$

В частности, при $a = b = c$ получается формула $\frac{4}{3}\pi a^3$ для объема шара радиуса a.

Частным случаем формулы (19) является формула для объема *тела вращения*.

Пусть $f\colon[a,b]\to[0,+\infty)$, T_f — тело, получающееся вращением подграфика функции f вокруг оси OX (рисунок 26). Аналитически тело T_f задается равенством

$$T_f=\bigl\{(x,y,z)\in\mathbb{R}^3 : x\in[a,b],\, y^2+z^2\leqslant f^2(x)\bigr\}.$$

Аналогично можно определить тело вращения произвольной фигуры вокруг любой прямой.

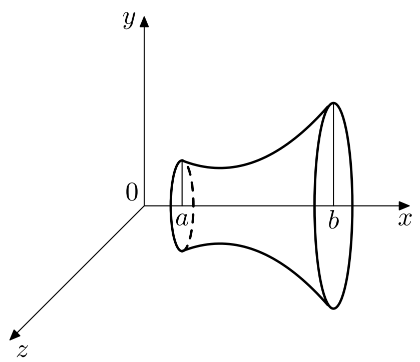

Рис. 26

Пусть $f\in C[a,b]$, $f\geqslant 0$. Для тела вращения T_f при каждом $x\in[a,b]$ сечение есть круг радиуса $f(x)$, поэтому $\mathcal{S}(x)=\pi f^2(x)$. Условия 1 – 3 выполнены и, значит,

$$V(T_f)=\pi\int_a^b f^2.$$

Пример 4. Найдем объем V_T *тора* — тела, образованного вращением круга $\bigl\{(x,y) : x^2+(y-R)^2\leqslant r^2\bigr\}$ $(0<r<R)$ вокруг оси OX (рисунок 27).

Тор представляется в виде разности тел вращения подграфиков функций, графики которых — верхняя и нижняя полуокружности, то есть функций

$$f_1(x)=R+\sqrt{r^2-x^2},\quad f_2(x)=R-\sqrt{r^2-x^2},\qquad x\in[-r,r].$$

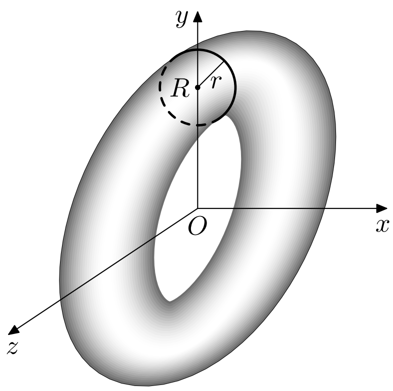

Рис. 27

Поэтому

$$\begin{aligned} V_T &= \pi \int_{-r}^{r} f_1^2 - \pi \int_{-r}^{r} f_2^2 = \\ &= \pi \int_{-r}^{r} \left(\left(R + \sqrt{r^2 - x^2} \right)^2 - \left(R - \sqrt{r^2 - x^2} \right)^2 \right) dx = \\ &= 4\pi R \int_{-r}^{r} \sqrt{r^2 - x^2}\, dx = 2\pi^2 R r^2. \end{aligned}$$

Приведем без доказательства еще две формулы вычисления объемов.

Замечание 1. Пусть $0 \leqslant a < b$, T_f' — тело вращения подграфика непрерывной функции $f\colon [a,b] \to [0, +\infty)$ вокруг оси OY. Тогда

$$V(T_f') = 2\pi \int_a^b x f(x)\, dx.$$

Замечание 2. Пусть $-\frac{\pi}{2} \leqslant \alpha < \beta \leqslant \frac{\pi}{2}$, $\widetilde{T}_f$ — тело вращения криволинейного сектора, определяемого непрерывной функцией $f\colon [\alpha, \beta] \to [0, +\infty)$, вокруг оси OY. Тогда

$$V\big(\widetilde{T}_f\big) = \frac{2\pi}{3} \int_\alpha^\beta f^3(\varphi) \cos\varphi\, d\varphi.$$

Замечание 3. Для неограниченных множеств формулы площадей и объемов остаются верными, но выражающие их интегралы будут несобственными.

4. Вычисление длин. Если $\gamma = (\gamma_1, \ldots, \gamma_m)$ — путь в $\mathbb{R}^m$, γ_i — дифференцируемые функции, то полагаем $\gamma' = (\gamma_1', \ldots, \gamma_m')$. Напомним, что по определению евклидовой длины

$$\|\gamma'\| = \sqrt{\sum_{i=1}^{m} \gamma_i'^2}.$$

Теорема 2. Длина гладкого пути. *Пусть $\gamma : [a, b] \to \mathbb{R}^m$ — гладкий путь. Тогда γ спрямляем и*

$$s_\gamma = \int_a^b \|\gamma'\|.$$

Доказательство. Докажем, что функция $\|\gamma'\|$ — плотность s.

Пусть $\Delta \subset [a, b]$. Возьмем дробление $\eta = \{u_k\}_{k=0}^{n}$ отрезка Δ. Тогда по определению евклидовой длины

$$\ell_\eta = \sum_{k=0}^{n-1} \big\|\gamma(u_{k+1}) - \gamma(u_k)\big\| = \sum_{k=0}^{n-1} \sqrt{\sum_{i=1}^{m} \big(\gamma_i(u_{k+1}) - \gamma_i(u_k)\big)^2}.$$

По формуле Лагранжа при каждых i и k найдется такая точка $c_{ik} \in (u_k, u_{k+1})$, что

$$\gamma_i(u_{k+1}) - \gamma_i(u_k) = \gamma_i'(c_{ik}) \Delta u_k.$$

Поэтому

$$\ell_\eta = \sum_{k=0}^{n-1} \sqrt{\sum_{i=1}^{m} \gamma_i'^2(c_{ik})} \cdot \Delta u_k.$$

Обозначим

$$M_\Delta^{(i)} = \max_{t \in \Delta} |\gamma_i'(t)|, \qquad m_\Delta^{(i)} = \min_{t \in \Delta} |\gamma_i'(t)|,$$

$$M_\Delta = \sqrt{\sum_{i=1}^{m} \big(M_\Delta^{(i)}\big)^2}, \qquad m_\Delta = \sqrt{\sum_{i=1}^{m} \big(m_\Delta^{(i)}\big)^2}$$

($M_\Delta^{(i)}$ и $m_\Delta^{(i)}$ существуют по теореме Вейерштрасса о непрерывных функциях) и покажем, что m и M удовлетворяют условиям теоремы 1. Для любого дробления η отрезка Δ

$$m_\Delta|\Delta| \leqslant \ell_\eta \leqslant M_\Delta|\Delta|.$$

Переходя к супремуму по всем дроблениям, мы получим

$$m_\Delta|\Delta| \leqslant s(\Delta) \leqslant M_\Delta|\Delta|.$$

Тем самым проверено условие 1) теоремы 1.

В частности, при $\Delta = [a, b]$ отсюда следует, что путь γ спрямляем.

Условие 2) теоремы 1, очевидно, следует из определения m и M: при всех $t \in \Delta$

$$m_\Delta \leqslant \|\gamma'(t)\| \leqslant M_\Delta.$$

Остается проверить выполнение условия 3). Если $M_\Delta + m_\Delta \neq 0$, то

$$M_\Delta - m_\Delta = \frac{M_\Delta^2 - m_\Delta^2}{M_\Delta + m_\Delta} = \frac{\sum\limits_{i=1}^{m}\Big(\big(M_\Delta^{(i)}\big)^2 - \big(m_\Delta^{(i)}\big)^2\Big)}{M_\Delta + m_\Delta} =$$

$$= \sum_{i=1}^{m}\big(M_\Delta^{(i)} - m_\Delta^{(i)}\big)\frac{M_\Delta^{(i)} + m_\Delta^{(i)}}{M_\Delta + m_\Delta} \leqslant \sum_{i=1}^{m}\big(M_\Delta^{(i)} - m_\Delta^{(i)}\big).$$

Если же $M_\Delta = m_\Delta = 0$, то доказанное неравенство очевидно.

Возьмем $\varepsilon > 0$. По теореме Кантора все функции $|\gamma_i'|$ равномерно непрерывны на $[a, b]$. Поэтому для каждого $i \in [1 : m]$ найдется такое $\delta_i > 0$, что

$$\text{если} \quad x, y \in [a, b],\ |x - y| < \delta_i, \quad \text{то} \qquad \big||\gamma_i'(x)| - |\gamma_i'(y)|\big| < \frac{\varepsilon}{m}.$$

Положим $\delta = \min\limits_{1 \leqslant i \leqslant m} \delta_i$. Если Δ — отрезок в $[a, b]$, $|\Delta| < \delta$, то по теореме Вейерштрасса $M_\Delta^{(i)}$ и $m_\Delta^{(i)}$ суть значения $|\gamma_i'|$ в некоторых точках Δ, откуда $M_\Delta^{(i)} - m_\Delta^{(i)} < \frac{\varepsilon}{m}$ и $M_\Delta - m_\Delta < \varepsilon$. Это и означает, что $M_\Delta - m_\Delta \underset{|\Delta| \to 0}{\longrightarrow} 0$. $\square$

Замечание 1. По аддитивности теорема 1 распространяется на кусочно-гладкие пути.

Замечание 2. Запишем частный случай теоремы 1 при $m=2$. *Пусть* $\gamma = (\varphi, \psi) \in C^1\big([a,b] \to \mathbb{R}^2\big)$. *Тогда*

$$s_\gamma = \int_a^b \sqrt{\varphi'^2 + \psi'^2}.$$

Следствие 1. Длина графика. *График* Γ_f *функции* f *класса* $C^1[a,b]$ *спрямляем и*

$$s_{\Gamma_f} = \int_a^b \sqrt{1 + f'^2}.$$

Здесь под графиком f понимается путь

$$\Gamma_f(t) = \big(t, f(t)\big), \quad t \in [a,b].$$

Следствие 2. Длина пути в полярных координатах. *Пусть* $f \in C^1[\alpha,\beta]$, $f \geqslant 0$, *путь* γ *задается в полярных координатах равенством* $r = f(\theta)$, $\theta \in [\alpha,\beta]$. *Тогда*

$$s_\gamma = \int_a^b \sqrt{f^2 + f'^2}.$$

Поясним, что означает фраза “путь задается в полярных координатах равенством”. Подставляя $r = f(\theta)$ в формулы, связывающие декартовы и полярные координаты, получаем:

$$x = r\cos\theta = f(\theta)\cos\theta, \qquad y = r\sin\theta = f(\theta)\sin\theta.$$

Таким образом, речь идет о пути

$$\gamma(\theta) = \big(\varphi(\theta), \psi(\theta)\big) = \big(f(\theta)\cos\theta, f(\theta)\sin\theta\big), \quad \theta \in [\alpha,\beta].$$

Доказательство. Дифференцируя, находим:

$$\begin{aligned} &\varphi'^2(\theta) + \psi'^2(\theta) = \\ &\quad = (f'(\theta)\cos\theta - f(\theta)\sin\theta)^2 + (f'(\theta)\sin\theta + f(\theta)\cos\theta)^2 = \\ &\qquad = \big(f'^2(\theta) + f^2(\theta)\big)(\cos^2\theta + \sin^2\theta) = f'^2(\theta) + f^2(\theta). \quad \square \end{aligned}$$

Пример 5. Пусть $a > 0$. Найдем длину $s_\Pi(a)$ части *параболы* $y = x^2$ от точки $(0,0)$ до точки (a, a^2) (рисунок 28).

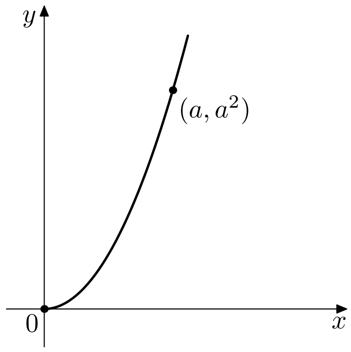

Рис. 28

Так как $\left(x^2\right)' = 2x$, по формуле для длины графика

$$s_\Pi(a) = \int_0^a \sqrt{1 + 4x^2}\, dx.$$

Проинтегрируем по частям и сведем интеграл к самому себе:

$$\begin{aligned} s_\Pi(a) &= x\sqrt{1+4x^2}\Big|_0^a - \int_0^a \frac{4x^2}{\sqrt{1+4x^2}}\, dx = \\ &= a\sqrt{1+4a^2} - \int_0^a \frac{1+4x^2}{\sqrt{1+4x^2}}\, dx + \int_0^a \frac{dx}{\sqrt{1+4x^2}} = \\ &= a\sqrt{1+4a^2} - s_\Pi(a) + \frac{1}{2}\ln\left(2a + \sqrt{1+4a^2}\right), \end{aligned}$$

откуда

$$s_\Pi(a) = \frac{a}{2}\sqrt{1+4a^2} + \frac{1}{4}\ln\left(2a + \sqrt{1+4a^2}\right).$$

Пример 6. Пусть $\beta \in [\,0, 2\pi]$. Найдем длину $s_Э(\beta)$ дуги *эллипса*

$$x = a\cos t, \quad y = b\sin t, \qquad t \in [\,0, \beta].$$

Для определенности будем считать, что $0 < a \leqslant b$ (рисунок 29).

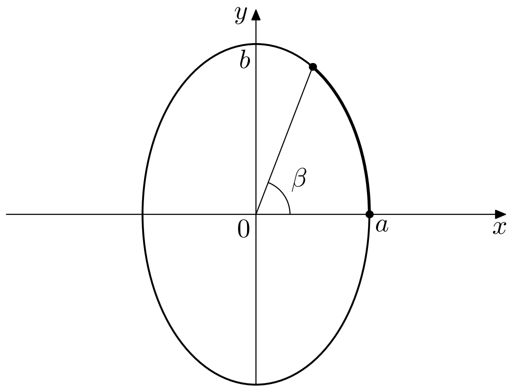

Рис. 29

Имеем:

$$s_Э(\beta) = \int_0^\beta \sqrt{a^2 \sin^2 t + b^2 \cos^2 t}\,dt =$$
$$= \int_0^\beta \sqrt{b^2 - (b^2 - a^2)\sin^2 t}\,dt = b\int_0^\beta \sqrt{1 - \varepsilon^2 \sin^2 t}\,dt.$$

Величина $\varepsilon = \frac{\sqrt{b^2-a^2}}{b}$ называется *эксцентриситетом* эллипса. Эксцентриситет характеризует меру "сжатости" эллипса. Очевидно, что $\varepsilon \in [0, 1)$; если $\varepsilon = 0$, то $a = b$, то есть эллипс есть окружность. В предельном случае $\varepsilon = 1$ мы получаем $a = 0$, то есть эллипс вырождается в дважды пробегаемый отрезок.

Интеграл

$$E(\varepsilon, \beta) = \int_0^\beta \sqrt{1 - \varepsilon^2 \sin^2 t}\,dt$$

называется *эллиптическим интегралом второго рода.* При $\beta = \frac{\pi}{2}$ эллиптический интеграл называется *полным.*

Таким образом, длина всего эллипса равна

$$s_Э = 4bE\Big(\varepsilon, \frac{\pi}{2}\Big).$$

Эллиптическим интегралом первого рода называется интеграл

$$K(\varepsilon, \beta) = \int_0^\beta \frac{1}{\sqrt{1-\varepsilon^2 \sin^2 t}}\, dt.$$

Этот интеграл возникает при вычислении длины дуги лемнискаты.

Эллиптические интегралы, вообще говоря, неберущиеся, то есть первообразная подынтегральной функции не является элементарной.

5. Работа силы. В качестве одного из физических приложений интеграла разберем вычисление работы силы.

Рассмотрим движение тела (материальной точки) по прямой, которую примем за числовую ось. Пусть на тело действует сила F, зависящая лишь от положения (координаты) тела. Будем считать, что вектор силы параллелен оси. Мы не будем различать вектор силы и его единственную координату, которую также обозначим F. Таким образом, F есть функция одной переменной, заданная на некотором промежутке E. Ясно, что $F(x) > 0$, если в точке x вектор силы сонаправлен с осью, и $F(x) < 0$, если он направлен противоположно оси.

Обозначим через $A_F\big([a,b]\big)$ работу силы F при перемещении тела из точки a в точку b. Если сила F постоянна на $[a,b]$, то $A_F\big([a,b]\big) = F \cdot (b-a)$ (это равенство, верное при любом расположении точек a и b, часто принимают за определение работы постоянной силы). Задача состоит в том, чтобы найти работу силы в общем случае, то есть когда сила переменная. Мы не будем обсуждать физическое содержание понятия работы (работой в физике называют количество энергии, затраченное при перемещении тела под действием силы), а сформулируем два свойства функции отрезка A_F.

1. Функция A_F аддитивна.
2. Если $m \leqslant F \leqslant M$ на отрезке Δ, то

$$m|\Delta| \leqslant A_F(\Delta) \leqslant M|\Delta|.$$

Эти свойства будем считать известными из физических соображений.

Предположим еще, что функция F непрерывна.

Для каждого отрезка $\Delta \subset E$ положим

$$m_\Delta = \min_{x\in\Delta} F(x), \qquad M_\Delta = \max_{x\in\Delta} F(x).$$

Тогда выполняются условия следствия теоремы 1. Поэтому F — плотность A_F, и работа выражается интегралом:

$$A_F\bigl([a,b]\bigr) = \int_a^b F.$$

Если сила направлена произвольно, то работа силы равна работе ее составляющей, параллельной оси движения.

§ 8. Функции ограниченной вариации

Понятие длины пути в пространстве $\mathbb{R}^m$ оказывается содержательным и при $m = 1$, однако от требования непрерывности имеет смысл отказаться.

Определение 1. Вариация функции. Пусть $f\colon [a,b] \to \mathbb{R}$. Величина

$$\bigvee_a^b f = \sup_\tau \sum_{k=0}^{n-1} \bigl|f(x_{k+1}) - f(x_k)\bigr|,$$

где верхняя грань берется по всем дроблениям $\tau = \{x_k\}_{k=0}^n$ отрезка $[a,b]$, называется *вариацией* (*полной вариацией*) или *полным изменением* функции f на отрезке $[a,b]$.

Если $\bigvee\limits_a^b f < +\infty$, то f называется функцией *ограниченной вариации* или функцией *с ограниченным изменением* на отрезке $[a,b]$. Множество всех функций ограниченной вариации на $[a,b]$ обозначается $V[a,b]$.

Исторически устоялся термин "функция ограниченной вариации", хотя было бы правильней говорить "функция конечной вариации".

Сравнив это определение с определением длины пути, можно сказать, что вариация — это "длина одномерного отображения", а функция ограниченной вариации — "спрямляемое одномерное отображение".

Рассмотрим свойства функций, связаные с понятием вариации.

V1. *Вариация аддитивна: если* $f\colon [a,b]\to\mathbb{R}$, $a<c<b$, *то*

$$\bigvee_a^b f = \bigvee_a^c f + \bigvee_c^b f.$$

V2. *Если* f *является кусочно-гладкой на* $[a,b]$, *то*

$$\bigvee_a^b f = \int_a^b |f'|.$$

Свойство V1 — это утверждение об аддитивности длины пути, а свойство V2 — формула для длины кусочно-гладкого пути.

V3. *Вариация монотонна: если* $f\colon [a,b]\to\mathbb{R}$, $[\alpha,\beta]\subset[a,b]$, *то*

$$\bigvee_\alpha^\beta f \leqslant \bigvee_a^b f.$$

Доказательство. По аддитивности

$$\bigvee_a^b f = \bigvee_a^\alpha f + \bigvee_\alpha^\beta f + \bigvee_\beta^b f \geqslant \bigvee_\alpha^\beta f. \quad \square$$

Вариацию можно определить и для функций, заданных на промежутке произвольного типа. Если $f\colon\langle a,b\rangle\to\mathbb{R}$, то полагают

$$\bigvee_a^b f = \sup_{[\alpha,\beta]\subset\langle a,b\rangle} \bigvee_\alpha^\beta f.$$

Определение корректно, поскольку для отрезка $[a,b]$ левая и правая части этого равенства в силу свойства монотонности совпадают.

V4. *Пусть* $\gamma=(\gamma_1,\ldots,\gamma_m)\colon[a,b]\to\mathbb{R}^m$. *Тогда* γ *спрямляем в том и только том случае, когда* $\gamma_i\in V[a,b]$ *при всех* $i\in[1:m]$.

Доказательство следует из двусторонней оценки

$$\bigl|\gamma_i(t_{k+1})-\gamma_i(t_k)\bigr| \leqslant \bigl\|\gamma(t_{k+1})-\gamma(t_k)\bigr\| \leqslant \sum_{j=1}^m \bigl|\gamma_j(t_{k+1})-\gamma_j(t_k)\bigr|. \quad \square$$

Из свойства V4 вытекает, что $f\in V[a,b]$ тогда и только тогда, когда график f спрямляем.

V5. *Если f монотонна на $[a,b]$, то $f \in V[a,b]$ и*

$$\bigvee_a^b f = |f(b) - f(a)|.$$

Доказательство. Для любого дробления

$$\sum_{k=0}^{n-1} \bigl|f(x_{k+1}) - f(x_k)\bigr| = \left|\sum_{k=0}^{n-1} \bigl(f(x_{k+1}) - f(x_k)\bigr)\right| = |f(b) - f(a)|. \quad \square$$

V6. *Если $f \in V[a,b]$, то f ограничена на $[a,b]$.*

Доказательство. При всех $x \in [a,b]$

$$|f(x)| \leqslant |f(a)| + |f(x) - f(a)| + |f(b) - f(x)| \leqslant |f(a)| + \bigvee_a^b f. \quad \square$$

Следующее утверждение доказывается аналогично теореме 6 § 2 об арифметических действиях над интегрируемыми функциями.

Теорема 1. Арифметические действия над функциями ограниченной вариации. *Пусть $f, g \in V[a,b]$. Тогда*

1) $f + g \in V[a,b]$;
2) $fg \in V[a,b]$;
3) $\alpha f \in V[a,b]$ *(*$\alpha \in \mathbb{R}$*)*;
4) $|f| \in V[a,b]$;
5) *если* $\inf\limits_{x \in [a,b]} |g(x)| > 0$*, то* $\frac{f}{g} \in V[a,b]$.

Доказательство. Обозначим $\Delta_k f = f(x_{k+1}) - f(x_k)$.

1. Складывая по всем k неравенства

$$|\Delta_k(f+g)| \leqslant |\Delta_k f| + |\Delta_k g|,$$

находим:

$$\sum_{k=0}^{n-1} |\Delta_k(f+g)| \leqslant \sum_{k=0}^{n-1} |\Delta_k f| + \sum_{k=0}^{n-1} |\Delta_k g| \leqslant \bigvee_a^b f + \bigvee_a^b g.$$

Переходя в левой части к супремуму по всем дроблениям, получаем, что

$$\bigvee_a^b (f+g) \leqslant \bigvee_a^b f + \bigvee_a^b g.$$

2. По свойству V6 функции f и g ограничены; пусть $|f|$ ограничен числом K, а $|g|$ — числом L. Тогда

$$|\Delta_k(fg)| \leqslant L|\Delta_k f| + K|\Delta_k g|$$

(подробно это неравенство доказано в теореме 6 § 2). Складывая эти неравенства и переходя к супремуму, мы получим

$$\bigvee_a^b fg \leqslant L\bigvee_a^b f + K\bigvee_a^b g.$$

3. Утверждение для αf следует из 2, если взять в качестве g функцию, тождественно равную α.

4. Утверждение для модуля доказывается тем же способом: из неравенств

$$\big|\Delta_k|f|\big| \leqslant |\Delta_k f|$$

вытекает, что

$$\bigvee_a^b |f| \leqslant \bigvee_a^b f.$$

5. Достаточно доказать, что $\frac{1}{g} \in V[a,b]$, после чего воспользоваться утверждением 2. Положим $m = \inf\limits_{x\in[a,b]} |g(x)|$. Тогда

$$\left|\Delta_k \frac{1}{g}\right| = \left|\frac{\Delta_k g}{g(x_k)g(x_{k+1})}\right| \leqslant \frac{|\Delta_k g|}{m^2}$$

и, следовательно,

$$\bigvee_a^b \frac{1}{g} \leqslant \frac{1}{m^2}\bigvee_a^b g. \quad \square$$

Теорема 2. Характеристика функций ограниченной вариации. *Пусть $f\colon[a,b]\to\mathbb{R}$. Тогда $f\in V[a,b]$ в том и только том случае, когда f представляется в виде разности двух возрастающих на $[a,b]$ функций.*

Доказательство. Достаточность очевидна из свойства V5 и теоремы 1. Докажем необходимость. Положим

$$g(x)=\bigvee_a^x f,\quad x\in[a,b];\qquad h=g-f.$$

Если $a\leqslant x_1<x_2\leqslant b$, то по аддитивности

$$\begin{aligned} g(x_2)-g(x_1)&=\bigvee_{x_1}^{x_2} f\geqslant 0,\\ h(x_2)-h(x_1)&=\bigvee_{x_1}^{x_2} f-\big(f(x_2)-f(x_1)\big)\geqslant 0,\end{aligned}$$

то есть функции g и h возрастают. $\square$

Выведем два следствия из доказанной теоремы.

V7. $V[a,b]\subset R[a,b]$.

Действительно, монотонная функция интегрируема и разность интегрируемых функций интегрируема.

V8. *Функция ограниченной вариации не может иметь разрывов второго рода.*

Свойство V8 следует из теоремы 2 и из того, что монотонная на отрезке функция не может иметь разрывов второго рода.

V9. *Ни один из классов $V[a,b]$ и $C[a,b]$ не содержится в другом.*

Доказательство. Так как существуют разрывные монотонные функции, $V[a,b]\not\subset C[a,b]$.

Приведем пример непрерывной функции, вариация которой бесконечна. Пусть

$$f(x)=\begin{cases} x\cos\frac{\pi}{x}, & x\in(0,1],\\ 0, & x=0.\end{cases}$$

Тогда $f \in C[0,1]$. Обозначим $x_k = \frac{1}{k}$ $(k \in \mathbb{N})$; при этом

$$f(x_k) = \frac{(-1)^k}{k}, \qquad \bigl|f(x_k) - f(x_{k+1})\bigr| = \frac{1}{k} + \frac{1}{k+1}.$$

Возьмем $n \in \mathbb{N}$ и рассмотрим дробление: $0 < x_n < \ldots < x_1 = 1$ (для удобства точки дробления занумерованы в порядке убывания, что несущественно). Сумма из определения вариации равна

$$\sum_{k=1}^{n-1} \bigl|f(x_{k+1}) - f(x_k)\bigr| + \bigl|f(x_n) - f(0)\bigr| = -1 + 2\sum_{k=1}^{n} \frac{1}{k}.$$

Докажем, что последовательность сумм

$$H_n = \sum_{k=1}^{n} \frac{1}{k},$$

называемых ***гармоническими,*** не ограничена сверху. При $m \in \mathbb{N}$ оценим сумму с номером 2^m снизу. Для этого сгруппируем слагаемые, а затем оценим сумму в каждой группе как количество слагаемых, умноженное на самое малое слагаемое:

$$\begin{aligned} H_{2^m} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) &+ \\ + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) &+ \ldots + \left(\frac{1}{2^{m-1}+1} + \ldots + \frac{1}{2^m}\right) \geqslant \\ &\geqslant 1 + \frac{1}{2} + 2\cdot\frac{1}{4} + 4\cdot\frac{1}{8} + \ldots + 2^{m-1}\cdot\frac{1}{2^m} = 1 + \frac{m}{2} \end{aligned}$$

(при $m \geqslant 2$ неравенство строгое).

Таким образом, последовательность $\{H_n\}$ не ограничена сверху. Поэтому $f \notin V[0,1]$. $\square$

Построенная функция f служит примером неспрямляемого пути в $\mathbb{R}$, а ее график — примером неспрямляемого пути в $\mathbb{R}^2$.

ГЛАВА 5. ПРЕДЕЛ И НЕПРЕРЫВНОСТЬ ОТОБРАЖЕНИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 1. Введение

В двух ближайших главах мы будем рассматривать отображения вида

$$f: E \subset \mathbb{R}^n \to \mathbb{R}^m, \quad \text{где } \ n, m \in \mathbb{N}.$$

Поскольку элементы $\mathbb{R}^n$ имеют n координат, f называют *отображением n вещественных переменных*, а в случае $m = 1$ — *функцией n вещественных переменных*. Цель главы 5 — ввести для таких отображений понятия предела и непрерывности и обобщить на многомерный случай основные результаты главы 2. При этом мы определим и изучим *открытые, замкнутые* и *компактные множества в* $\mathbb{R}^n$. Дифференциальное исчисление для отображений нескольких переменных будет изложено в следующей главе.

Введем некоторые стандартные соглашения относительно обозначений, применяемых в главах 5 и 6. Для нумерации координат векторов из $\mathbb{R}^n$ мы будем использовать *нижние индексы*. Так, k-я координата точки $x \in \mathbb{R}^n$ всегда будет обозначаться символом x_k. *Верхними индексами* нумеруются элементы семейств векторов и, в частности, последовательностей векторов, о которых речь пойдет в § 2. Аналогично, для отображений f со значениями в $\mathbb{R}^m$ через $f_1(x), \ldots, f_m(x)$ мы будем обозначать координаты вектора $f(x)$. При $k \in \{1, \ldots, m\}$ функция $x \mapsto f_k(x)$ называется *k-й координатной функцией отображения f*.

Напомним теперь основные операции над элементами $\mathbb{R}^n$, введенные в § 8 главы 3. Пусть $x, y \in \mathbb{R}^n$.

1) *Сложение*: $x + y = (x_1 + y_1, \ldots, x_n + y_n)$.

2) *Умножение на число*: $\lambda x = (\lambda x_1, \ldots, \lambda x_n) \quad (\lambda \in \mathbb{R})$.

3) *Скалярное произведение*: $x \cdot y = x_1 y_1 + \ldots + x_n y_n$.

Сложение и умножение на числа называют *линейными операциями* в $\mathbb{R}^n$.

Вектор $(0, \ldots, 0) \in \mathbb{R}^n$ называется *нулевым элементом* $\mathbb{R}^n$. Мы будем обозначать его символом 0, если из контекста ясно, что речь идет о векторе, а не о числе. Очевидно, нулевой элемент обладает следующими свойствами:

$$x + 0 = x \text{ при всех } x \in \mathbb{R}^n; \quad \lambda \cdot 0 = 0 \text{ для любых } \lambda \in \mathbb{R}.$$

Операции, определенные для векторов, можно распространить и на отображения. Пусть $E \subset \mathbb{R}^n$, $f, g\colon E \to \mathbb{R}^m$, $\lambda\colon E \to \mathbb{R}$. Положим

1) $(f+g)(x) = f(x) + g(x) \quad (x \in E)$;

2) $(\lambda f)(x) = \lambda(x) f(x) \quad (x \in E)$;

3) $(f \cdot g)(x) = \sum\limits_{k=1}^{m} f_k(x)\, g_k(x) \quad (x \in E)$.

Частными случаями 2) являются умножение отображения на число и умножение функции на вектор: для $\alpha \in \mathbb{R}$ и $v \in \mathbb{R}^m$

2$'$) $(\alpha f)(x) = \alpha f(x) \quad (x \in E)$;

2$''$) $(v \cdot \lambda)(x) = v \cdot \lambda(x) \quad (x \in E)$.

Правила 1) и 2$'$) называют *линейными операциями над отображениями*. Равенство 3) определяет *скалярное произведение отображений*.

Введем теперь норму вектора, которая является многомерным аналогом модуля числа.

Определение 1. *Нормой* или *евклидовой нормой* в $\mathbb{R}^n$ называется функция, которая каждому $x \in \mathbb{R}^n$ сопоставляет величину

$$\|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}.$$

Замечание 1. При $n = 1$ элементами $\mathbb{R}^n$ являются вещественные числа, а норма совпадает с модулем. Если n равно 2 или 3, то $\|x\|$ есть *длина вектора* x. Для произвольного $n \in \mathbb{N}$ норма является естественным обобщением длины.

Замечание 2. В § 8 главы 3 рассматривались *p-нормы* $\|\cdot\|_p$ при $p \geqslant 1$, а евклидова норма соответствовала случаю $p = 2$. Так как другие p нас сейчас не интересуют, мы используем обозначение $\|\cdot\|$ вместо $\|\cdot\|_2$.

Замечание 3. Евклидова норма обладает следующими тремя свойствами.

1) *Неотрицательность*:

$$\|x\| \geqslant 0 \quad \text{при всех } \ x \in \mathbb{R}^n,$$

причем равенство достигается только для $x = 0$.

2) *Положительная однородность*:

$$\|\lambda \cdot x\| = |\lambda| \cdot \|x\| \quad \text{при всех } \ x \in \mathbb{R}^n \ \text{ и } \ \lambda \in \mathbb{R}.$$

3) *Неравенство треугольника*:

$$\|x + y\| \leqslant \|x\| + \|y\| \quad \text{при всех } \ x, y \in \mathbb{R}^n.$$

Действительно, неравенство треугольника было доказано в § 8 главы 3, а два других утверждения очевидны. В дальнейшем мы определим норму *аксиоматически*, как функцию на векторном пространстве, удовлетворяющую условиям 1) – 3).

Замечание 4. Для любых $x, y \in \mathbb{R}^n$ справедливо неравенство

$$\|x - y\| \geqslant \big|\|x\| - \|y\|\big|$$

Действительно, по неравенству треугольника

$$\|x\| = \|x - y + y\| \leqslant \|x - y\| + \|y\|,$$

откуда

$$\|x\| - \|y\| \leqslant \|x - y\|.$$

Поменяв местами x и y, мы получим

$$-\big(\|x\| - \|y\|\big) \leqslant \|x - y\|.$$

Из этих двух неравенств и вытекает требуемая оценка.

Замечание 5. Напомним читателю также *неравенство Коши*, доказанное в § 8 главы 3:

$$|x \cdot y| \leqslant \|x\| \cdot \|y\| \quad \text{при всех } \ x, y \in \mathbb{R}^n.$$

Замечание 6. Непосредственно из определения 1 вытекает двусторонняя оценка нормы вектора: для любого $k \in \{1, \ldots, n\}$

$$|x_k| \leqslant \|x\| \leqslant \sqrt{n} \max_{1 \leqslant i \leqslant n} |x_i|. \tag{1}$$

При изучении предела и непрерывности функций одной переменной важное значение имели *окрестности* и *проколотые окрестности* точек в $\overline{\mathbb{R}}$. Теперь мы определим многомерные аналоги этих понятий. Положим $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup \{\infty\}$. Напомним, что в главе 2 было введено обозначение $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Таким образом, символы $\overline{\mathbb{R}^1}$ и $\overline{\mathbb{R}}$ имеют разный смысл. Это не создаст читателю неудобств, поскольку основным объектом изучения в главах 5 и 6 будет случай $n > 1$. Указанное различие объясняется тем, что при $n \geqslant 2$ на $\mathbb{R}^n$ не введен порядок, поэтому символы $\pm\infty$ теряют смысл.

Определение 2. Окрестности точек в $\overline{\mathbb{R}^n}$. Пусть $\delta > 0$.

1) Если $a \in \mathbb{R}^n$, то множества

$$V_a(\delta) = \{x \in \mathbb{R}^n : \|x - a\| < \delta\}, \quad \dot{V}_a(\delta) = V_a(\delta) \setminus \{a\}$$

называют соответственно *δ-окрестностью* и *проколотой δ-окрестностью* точки a.

2) Множество

$$V_\infty(\delta) = \{x \in \mathbb{R}^n : \|x\| > \delta\}$$

называется *δ-окрестностью в $\mathbb{R}^n$ бесконечно удаленной точки*. Положим также $\dot{V}_\infty(\delta) = V_\infty(\delta)$. Это будет удобно для единообразия записи окрестностей точек $\overline{\mathbb{R}^n}$.

Замечание 1. При $n = 1$ множество $V_a(\delta)$ представляет собой интервал $(a-\delta, a+\delta)$, что согласуется с определением главы 2. Если n равно 2 или 3, то $V_a(\delta)$ будет соответственно кругом или шаром радиуса δ с центром в точке a. Поэтому в общем случае мы будем называть окрестности точек $\mathbb{R}^n$ *открытыми n-мерными шарами*.

Замечание 2. Множества вида $V_\infty(\delta) \cup \{\infty\}$ называют *окрестностями бесконечно удаленной точки в $\overline{\mathbb{R}^n}$*. Они оказываются полезными в теории функций комплексной переменной. Отметим, что проколотые окрестности ∞ в $\mathbb{R}^n$ и в $\overline{\mathbb{R}^n}$ одинаковы.

Замечание 3. В ситуациях, когда радиус шара не имеет значения, мы будем его опускать и писать просто V_a. Это соответствует соглашениям главы 2.

В заключение сформулируем два простых, но важных утверждения, касающихся окрестностей. Их доказательство оставляется читателю в качестве упражнения.

Замечание 4. *Пусть* $a \in \overline{\mathbb{R}^n}$, $\delta_1, \delta_2 > 0$. *Тогда*

$$V_a(\delta_1) \cap V_a(\delta_2) = V_a(\delta),$$

где $\delta = \min\{\delta_1, \delta_2\}$ *в случае* $a \in \mathbb{R}^n$ *и* $\delta = \max\{\delta_1, \delta_2\}$ *при* $a = \infty$.

Таким образом, пересечение двух окрестностей одной и той же точки из $\overline{\mathbb{R}^n}$ снова будет окрестностью этой точки.

Замечание 5. *Пусть* $a, b \in \overline{\mathbb{R}^n}$, $a \neq b$. *Тогда существуют такие окрестности* V_a *и* V_b, *что* $V_a \cap V_b = \varnothing$.

§ 2. Предел последовательности векторов

В этом параграфе мы дадим определение предела последовательности элементов $\mathbb{R}^n$ и обобщим на многомерный случай большинство утверждений, доказанных в главе 2 для числовых последовательностей. Напомним, что для нумерации векторов используются верхние индексы, а нижними обозначаются координаты векторов.

Определение 1. Предел последовательности в $\mathbb{R}^n$. Пусть $\{x^k\}_{k=1}^{\infty}$ — последовательность в $\mathbb{R}^n$, $a \in \overline{\mathbb{R}^n}$. Будем говорить, что *предел последовательности* $\{x^k\}$ *равен* a, если

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall k \in \mathbb{N}\colon \ k > N \ \ x^k \in V_a(\varepsilon).$$

Мы будем использовать те же обозначения, что и в главе 2:

$$\lim_{k\to\infty} x^k = a \quad \text{или} \quad x^k \to a \ \ (k \to \infty).$$

Если последовательность $\{x^k\}$ имеет предел $a \in \mathbb{R}^n$, то говорят, что она *сходится к* a. В противном случае последовательность $\{x^k\}$ называется *расходящейся.*

Замечание 1. Определение предела последовательности в $\mathbb{R}^n$ можно переписать на языке неравенств, а также свести его к пределу числовой последовательности. Рассмотрим два случая.

1) Пусть $a \in \mathbb{R}^n$. Соотношение $\lim\limits_{k\to\infty} x^k = a$ означает, что

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall k \in \mathbb{N}\colon \; k > N \quad \left\| x^k - a \right\| < \varepsilon$$

или

$$\lim_{k\to\infty} \left\| x^k - a \right\| = 0.$$

2) Равенство $\lim\limits_{k\to\infty} x^k = \infty$ эквивалентно

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall k \in \mathbb{N}\colon \; k > N \quad \left\| x^k \right\| > \varepsilon$$

или

$$\lim_{k\to\infty} \left\| x^k \right\| = +\infty.$$

Замечание 2. Единственность предела последовательности. *Если $a, b \in \overline{\mathbb{R}^n}$ одновременно являются пределами последовательности $\left\{x^k\right\}$, то $a = b$.*

Действительно, пусть $a \neq b$. Тогда у a и b существуют непересекающиеся окрестности V_a и V_b соответственно (см. § 1). По определению предела найдутся такие $N_1, N_2 \in \mathbb{N}$, что $x^k \in V_a$ при $k > N_1$ и $x^k \in V_b$ при $k > N_2$. Значит, $x^k \in V_a \cap V_b$ при $k > \max\{N_1, N_2\}$, что невозможно.

Замечание 3. *Если $a \in \mathbb{R}^n$, $\lim\limits_{k\to\infty} x^k = a$, то $\lim\limits_{k\to\infty} \left\| x^k \right\| = \|a\|$.*

Действительно, в силу замечания 4 § 1

$$\left| \left\| x^k \right\| - \|a\| \right| \leqslant \left\| x^k - a \right\| \to 0 \quad (k \to \infty).$$

Замечание 4. Определение 1 включает в себя также понятие *предела последовательности комплексных чисел*, поскольку $\mathbb{C}$ как множество совпадает с $\mathbb{R}^2$. Пусть $\left\{x^k\right\}$ — последовательность в $\mathbb{C}$. Равенство $\lim\limits_{k\to\infty} x^k = a$ означает, что

$$\forall \varepsilon > 0 \;\; \exists N \in \mathbb{N} \;\; \forall k \in \mathbb{N}\colon \; k > N \;\; |x^k - a| < \varepsilon,$$

так как модуль комплексного числа совпадает с нормой этого числа как элемента $\mathbb{R}^2$.

Докажем, что сходимость в $\mathbb{R}^n$ последовательности $\{x^k\}$ равносильна сходимости всех ее *координатных последовательностей* $\{x_i^k\}_{k=1}^{\infty}$, где $i \in \{1, \dots, n\}$.

Теорема 1. Сходимость и покоординатная сходимость. *Пусть $\{x^k\}_{k=1}^{\infty}$ — последовательность в $\mathbb{R}^n$, $a \in \mathbb{R}^n$. Тогда равносильны следующие утверждения.*

1) *$\{x^k\}_{k=1}^{\infty}$ сходится к a.*

2) *Для любого $i \in \{1, \dots, n\}$ последовательность $\{x_i^k\}_{k=1}^{\infty}$ сходится к a_i.*

Доказательство. 1) $\Rightarrow$ 2) Пусть $i \in \{1, \dots, n\}$. В силу (1) и замечания 1

$$\left|x_i^k - a_i\right| \leqslant \left\|x^k - a\right\| \to 0 \quad (k \to \infty),$$

откуда $\lim\limits_{k\to\infty} x_i^k = a_i$.

2) $\Rightarrow$ 1) Если $x_i^k \to a_i$ при всех $i \in \{1, \dots, n\}$, то

$$\left\|x^k - a\right\| = \sqrt{\sum_{i=1}^{n} \left(x_i^k - a_i\right)^2} \to 0 \quad (k \to \infty),$$

откуда $\lim\limits_{k\to\infty} x^k = a$. Мы воспользовались замечанием 1, теоремой об арифметических операциях над пределами и непрерывностью квадратного корня. $\square$

Замечание 1. Если $\lim\limits_{k\to\infty} x^k = \infty$, то координатные последовательности $\{x^k\}$ могут и не иметь предела. Пусть, например, последовательность в $\mathbb{R}^2$ определяется формулой

$$x^k = \left(k \cos\frac{\pi k}{2},\, k \sin\frac{\pi k}{2}\right), \quad k \in \mathbb{N}.$$

Тогда

$$\left\|x^k\right\| = \sqrt{k^2\, \cos^2\frac{\pi k}{2} + k^2\, \sin^2\frac{\pi k}{2}} = k \to +\infty,$$

то есть $x^k \to \infty$. Тем не менее, последовательности $x_1^k = k \cos \dfrac{\pi k}{2}$ и $x_2^k = k \sin \dfrac{\pi k}{2}$ предела не имеют.

Замечание 2. Если $\{x^k\}$ — последовательность комплексных чисел, $a \in \mathbb{C}$, то равенство $\lim\limits_{k\to\infty} x^k = a$ эквивалентно условиям

$$\lim_{k\to\infty} \operatorname{Re} x^k = \operatorname{Re} a \quad \text{и} \quad \lim_{k\to\infty} \operatorname{Im} x^k = \operatorname{Im} a.$$

Обобщим теперь на многомерный случай некоторые утверждения, справедливые для числовых последовательностей.

Теорема 2. Предел и арифметические операции. *Пусть $\{x^k\}$, $\{y^k\}$ — сходящиеся последовательности в $\mathbb{R}^n$, $a = \lim\limits_{k\to\infty} x^k$, $b = \lim\limits_{k\to\infty} y^k$. Тогда справедливы следующие утверждения.*

1) $\lim\limits_{k\to\infty} (x^k + y^k) = a + b$.

2) *Если $\{\lambda_k\}$ — последовательность в $\mathbb{R}$, сходящаяся к λ, то*

$$\lim_{k\to\infty} \lambda_k\, x^k = \lambda\, a;$$

в частности,

$$\lim_{k\to\infty} \lambda\, x^k = \lambda\, a \quad \textit{для любого } \lambda \in \mathbb{R}.$$

3) $\lim\limits_{k\to\infty} x^k \cdot y^k = a \cdot b$.

Доказательство. Ограничимся проверкой первого утверждения. По теореме 1 для любого $i \in \{1, \dots, n\}$

$$\lim_{k\to\infty} x_i^k = a_i \quad \text{и} \quad \lim_{k\to\infty} y_i^k = b_i.$$

Тогда $\lim\limits_{k\to\infty} (x_i^k + y_i^k) = a_i + b_i$ (см. § 1 главы 2). Применяя теорему 1 еще раз, мы получим равенство $\lim\limits_{k\to\infty} (x^k + y^k) = a + b$. $\square$

Замечание. Комбинируя первые два утверждения теоремы 2, мы получим равенство

$$\lim_{k\to\infty} (\lambda x^k + \mu y^k) = \lambda a + \mu b \quad (\lambda, \mu \in \mathbb{R}).$$

Его называют *линейностью предела последовательности.*

Теорема 3. Предел подпоследовательности.
Пусть $\{x^k\}_{k=1}^{\infty}$ *— последовательность в* $\mathbb{R}^n$, $\{x^{k_i}\}_{i=1}^{\infty}$ *— ее подпоследовательность,* $a \in \overline{\mathbb{R}^n}$. *Если* $\lim\limits_{k\to\infty} x^k = a$, *то* $\lim\limits_{i\to\infty} x^{k_i}$ *также существует и равен* a.

Доказательство. Пусть $a \in \mathbb{R}^n$. По лемме 2 § 2 главы 2

$$\lim_{i\to\infty} \left\|x^{k_i} - a\right\| = \lim_{k\to\infty} \left\|x^k - a\right\| = 0,$$

откуда $\lim\limits_{i\to\infty} x^{k_i} = a$. Случай $a = \infty$ разбирается аналогично. $\square$

Для дальнейшего изложения нам потребуются два важных понятия — *ограниченные множества* и *ограниченные отображения* в $\mathbb{R}^n$. Введем их.

Определение 2. Ограниченность в $\mathbb{R}^n$**.**

1) Пусть E — подмножество $\mathbb{R}^n$. Оно называется *ограниченным*, если существует такое $C > 0$, что $E \subset V_0(C)$ (то есть $\|x\| < C$ для любого $x \in E$).

2) Пусть $f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m$. Отображение f называется *ограниченным*, если множество его значений ограничено в $\mathbb{R}^m$.

Замечание 1. Ограниченность множества $E \subset \mathbb{R}^n$ эквивалентна условию

$$\sup_{x\in E} \|x\| < +\infty,$$

а ограниченность отображения $f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m$ — условию

$$\sup_{x\in E} \|f(x)\| < +\infty.$$

В частности, ограниченность в $\mathbb{R}^n$ последовательности $\{x^k\}_{k=1}^{\infty}$ означает, что $\sup\limits_{k\in\mathbb{N}} \left\|x^k\right\| < +\infty$.

Замечание 2. Пусть $E \subset \mathbb{R}^n$, $k \in \{1, \ldots, n\}$. Назовем *проекцией* E *на* k*-ю ось* множество

$$E_k = \{x_k : x \in E\}.$$

Тогда ограниченность E равносильна ограниченности всех его проекций E_k $(k = 1, \dots, n)$. Это вытекает непосредственно из (1).

Замечание 3. В силу замечания 2 ограниченность отображения $f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m$ эквивалентна ограниченности всех его координатных функций. В частности, последовательность $\{x^k\}$ в $\mathbb{R}^n$ ограничена тогда и только тогда, когда все ее координатные последовательности ограничены.

Замечание 4. *Любая сходящаяся в $\mathbb{R}^n$ последовательность $\{x^k\}$ ограничена.* Действительно, по теореме 1 все координатные последовательности $\{x^k\}$ сходятся и потому ограничены (см. теорему 2 § 1 главы 2). Тогда по предыдущему замечанию ограничена и последовательность $\{x^k\}$.

Обобщим теперь принцип выбора Больцано – Вейерштрасса на многомерный случай.

Теорема 4. Принцип выбора Больцано – Вейерштрасса. *Пусть $\{x^k\}_{k=1}^{\infty}$ — последовательность в $\mathbb{R}^n$. Тогда справедливы следующие утверждения.*

1) *Если последовательность $\{x^k\}$ ограничена, то из нее можно выбрать сходящуюся подпоследовательность.*

2) *Если последовательность $\{x^k\}$ не ограничена, то у нее есть подпоследовательность $\{x^{k_i}\}$, стремящаяся к бесконечности.*

Доказательство. 1) Пусть последовательность $\{x^k\}$ ограничена в $\mathbb{R}^n$. В силу замечания 3 все ее координатные последовательности также ограничены, и для них можно воспользоваться одномерным принципом Больцано – Вейерштрасса. Выберем вначале из $\{x_1^k\}$ сходящуюся подпоследовательность $\{x_1^{r_i}\}$. Затем из последовательности $\{x_2^{r_i}\}$ выберем сходящуюся подпоследовательность $\{x_2^{s_i}\}$. Заметим, что последовательность $\{x_1^{s_i}\}$ тоже сходится как подпоследовательность $\{x_1^{r_i}\}$. Продолжая эти рассуждения для других координатных последовательностей $\{x^k\}$, мы в конце концов получим подпоследовательность $\{x^{k_i}\}$, у которой все координатные последовательности сходятся. Тогда $\{x^{k_i}\}$ сходится в $\mathbb{R}^n$ по теореме 1.

2) Так как числовая последовательность $\left\{\left\|x^k\right\|\right\}_{k=1}^{\infty}$ не ограничена сверху, у нее есть подпоследовательность $\left\{\left\|x^{k_i}\right\|\right\}_{i=1}^{\infty}$, стремящаяся к $+\infty$ (см. § 2 главы 2). Тогда $\lim\limits_{i\to\infty} x^{k_i} = \infty$. $\square$

Докажем теперь многомерный вариант критерия сходимости последовательности.

Теорема 5. Критерий Больцано – Коши. *Пусть* $\left\{x^k\right\}_{k=1}^{\infty}$ — *последовательность в* $\mathbb{R}^n$. *Тогда равносильны следующие утверждения.*

1) *Последовательность* $\left\{x^k\right\}$ *сходится в* $\mathbb{R}^n$.

2) $\forall \varepsilon > 0\ \ \exists N \in \mathbb{N}\ \ \forall k, m \in \mathbb{N}\colon\ k, m > N\ \ \left\|x^k - x^m\right\| < \varepsilon$.

Замечание. Последовательность в $\mathbb{R}^n$, удовлетворяющую условию 2), называют *сходящейся в себе* или *фундаментальной.*

Доказательство. 1) $\Rightarrow$ 2) По теореме 1 все координатные последовательности $\left\{x^k\right\}$ сходятся, и в силу одномерного критерия Больцано – Коши они фундаментальны. Тогда по $\varepsilon > 0$ можно подобрать такое $N \in \mathbb{N}$, что для любого $i \in \{1, \ldots, n\}$

$$\left|x_i^k - x_i^m\right| < \frac{\varepsilon}{\sqrt{n}} \quad \text{при всех } k, m > N.$$

Отсюда в силу (1) мы получим $\left\|x^k - x^m\right\| < \varepsilon$, что и дает 2).

2) $\Rightarrow$ 1) В силу (1) для любых $k, m \in \mathbb{N}$ и $i \in \{1, \ldots, n\}$ справедливо неравенство

$$\left|x_i^k - x_i^m\right| \leqslant \left\|x^k - x^m\right\|.$$

Поэтому условие 2) влечет сходимость в себе координатных последовательностей $\left\{x^k\right\}$. По одномерному критерию Больцано – Коши все они сходятся в $\mathbb{R}$, а тогда по теореме 1 последовательность $\left\{x^k\right\}$ сходится в $\mathbb{R}^n$. $\square$

В заключение этого параграфа введем понятие предельной точки подмножества $\mathbb{R}^n$ и дадим характеристику предельных точек в терминах последовательностей.

Определение 3. Предельная и изолированная точки множества. Пусть $E \subset \mathbb{R}^n$.

1) Точка $a \in \overline{\mathbb{R}^n}$ называется *предельной* для E, если для любого $\varepsilon > 0$ множество $E \cap \dot{V}_a(\varepsilon)$ непусто.

2) Точка $a \in E$ называется *изолированной* для E, если она не является предельной точкой E.

Теорема 6. Характеристика предельных точек. *Пусть $E \subset \mathbb{R}^n$, $a \in \overline{\mathbb{R}^n}$. Тогда равносильны следующие утверждения.*

1) *Точка a является предельной для E.*

2) *Существует последовательность $\{x^k\}_{k=1}^{\infty}$ в $E \setminus \{a\}$, для которой $\lim\limits_{k\to\infty} x^k = a$.*

Доказательство проведем только для $a \in \mathbb{R}^n$, оставляя читателю случай $a = \infty$ в качестве упражнения.

1) $\Rightarrow$ 2) Пусть a — предельная точка E. Тогда для любого $k \in \mathbb{N}$ найдется точка $x^k \in E \setminus \{a\}$, для которой $\|x^k - a\| < \frac{1}{k}$. Отсюда $\lim\limits_{k\to\infty} \|x^k - a\| = 0$, то есть $\lim\limits_{k\to\infty} x^k = a$.

2) $\Rightarrow$ 1) Пусть $\{x^k\}_{k=1}^{\infty}$ — последовательность в $E \setminus \{a\}$, сходящаяся к a. Для любого $\varepsilon > 0$ найдется такое $k \in \mathbb{N}$, что $\|x^k - a\| < \varepsilon$. Поэтому $x^k \in V_a(\varepsilon) \cap \big(E \setminus \{a\}\big)$, то есть $x^k \in \dot{V}_a(\varepsilon) \cap E$. Тогда при любом $\varepsilon > 0$ множество $\dot{V}_a(\varepsilon) \cap E$ непусто и, значит, точка a является предельной для E. $\square$

§ 3. Открытые, замкнутые и компактные множества.

Некоторые утверждения, доказанные в главах 2 и 3, были верны только при определенных условиях на область определения функции. Например, в теоремах Вейерштрасса и Кантора (см. § 4 главы 2) функция должна быть задана *на отрезке*, а дифференцируемость в главе 3 часто предполагалась *на интервале*. Для обобщения таких результатов на многомерный случай необходимо найти в $\mathbb{R}^n$ разумные аналоги открытых и замкнутых промежутков вещественной прямой. Перейдем к решению этой задачи.

Определение 1. Открытые и замкнутые множества в $\mathbb{R}^n$. Пусть $E \subset \mathbb{R}^n$.

1) Точка $a \in E$ называется *внутренней для* E, если у нее есть окрестность, содержащаяся в E.

2) Множество E называется *открытым в* $\mathbb{R}^n$, если любая его точка является внутренней для E.

3) Множество E называется *замкнутым в* $\mathbb{R}^n$, если его дополнение $\mathbb{R}^n \setminus E$ открыто в $\mathbb{R}^n$.

Замечание. Дополнение любого открытого в $\mathbb{R}^n$ множества, в свою очередь, замкнуто. Таким образом, замкнутые и открытые множества взаимно дополняют друг друга.

Для практической проверки замкнутости множества определение 1 не всегда удобно. Приведем явное описание замкнутых множеств в терминах последовательностей и предельных точек.

Теорема 1. Характеристика замкнутых множеств. *Пусть* $E \subset \mathbb{R}^n$. *Тогда равносильны следующие утверждения.*

1) E *замкнуто в* $\mathbb{R}^n$.

2) E *содержит все свои предельные точки в* $\mathbb{R}^n$.

3) *Если* $\{x^k\}_{k=1}^{\infty}$ *— сходящаяся последовательность в* E, *то* $\lim\limits_{k\to\infty} x^k \in E$.

Доказательство проведем по схеме 1) $\Rightarrow$ 3) $\Rightarrow$ 2) $\Rightarrow$ 1).

1) $\Rightarrow$ 3) Пусть $\{x^k\}_{k=1}^{\infty}$ — последовательность в E, $a \in \mathbb{R}^n$, $\lim\limits_{k\to\infty} x_k = a$. Покажем, что $a \in E$. Если это не так, то $a \in \mathbb{R}^n \setminus E$, и в силу открытости $\mathbb{R}^n \setminus E$ найдется окрестность V_a точки a, лежащая в $\mathbb{R}^n \setminus E$. По определению предела при всех достаточно больших $k \in \mathbb{N}$ справедливо включение $x^k \in V_a \subset \mathbb{R}^n \setminus E$. С другой стороны, $x^k \in E$ для любых $k \in \mathbb{N}$, и мы получаем противоречие.

3) $\Rightarrow$ 2) Пусть $a \in \mathbb{R}^n$ — предельная точка E. По теореме 6 § 2 найдется последовательность $\{x^k\}$ в E, сходящаяся к a. Тогда в силу 3) $a \in E$.

2) $\Rightarrow$ 1) Докажем, что $\mathbb{R}^n \setminus E$ открыто. Пусть $a \in \mathbb{R}^n \setminus E$. Проверим, что некоторая окрестность V_a точки a содержится в $\mathbb{R}^n \setminus E$. Если это не так, то для любого $\varepsilon > 0$ окрестность $\dot{V}_a(\varepsilon)$ содержит точку множества E, то есть $\dot{V}_a(\varepsilon) \cap E \neq \varnothing$. Тогда a является предельной точкой E, откуда в силу 2) $a \in E$, что невозможно. $\square$

Рассмотрим теперь теоретико-множественные операции над открытыми и замкнутыми множествами. Что можно сказать о результатах таких операций? Изучим этот вопрос более подробно.

Теорема 2. *Пусть* A *— множество,* $\{E_\alpha\}_{\alpha\in A}$ *и* $\{F_\alpha\}_{\alpha\in A}$ *— семейства открытых и замкнутых подмножеств* $\mathbb{R}^n$ *соответственно. Тогда справедливы следующие утверждения.*

1) $\bigcup\limits_{\alpha\in A} E_\alpha$ *открыто,* $\bigcap\limits_{\alpha\in A} F_\alpha$ *замкнуто в* $\mathbb{R}^n$.

2) *Если множество индексов* A *конечно, то* $\bigcap\limits_{\alpha\in A} E_\alpha$ *открыто,* $\bigcup\limits_{\alpha\in A} F_\alpha$ *замкнуто в* $\mathbb{R}^n$.

Доказательство. 1) Положим

$$E = \bigcup_{\alpha\in A} E_\alpha, \quad F = \bigcap_{\alpha\in A} F_\alpha.$$

Если $a \in E$, то существует такое $\alpha \in A$, что $a \in E_\alpha$. В силу открытости E_α у точки a есть окрестность $V_a \subset E_\alpha$. Тогда $V_a \subset E$, то есть E открыто. Для доказательства замкнутости F заметим, что в силу правила де Моргана (см. § 1 главы 1)

$$\mathbb{R}^n \setminus F = \mathbb{R}^n \setminus \bigcap_{\alpha\in A} F_\alpha = \bigcup_{\alpha\in A} \bigl(\mathbb{R}^n \setminus F_\alpha\bigr).$$

Так как множества $\mathbb{R}^n \setminus F_\alpha$ открыты при всех $\alpha \in A$, из первой части утверждения 1) вытекает открытость $\mathbb{R}^n \setminus F$. Поэтому множество F замкнуто.

2) Пусть множество A конечно. Положим $E = \bigcap\limits_{\alpha\in A} E_\alpha$. Если $a \in E$, то $a \in E_\alpha$ при всех $\alpha \in A$. Тогда по любому $\alpha \in A$ найдется $\delta_\alpha > 0$, для которого $V_a\bigl(\delta_\alpha\bigr) \subset E_\alpha$. Положим

$$\delta = \min\{\delta_\alpha\colon \alpha \in A\}.$$

Поскольку множество A конечно, этот минимум существует и положителен. Заметим, что $V_a(\delta) \subset E_\alpha$ при всех $\alpha \in A$, то есть $V_a(\delta) \subset E$. Поэтому множество E открыто. Замкнутость $\bigcup\limits_{\alpha\in A} F_\alpha$ выводится из открытости E с помощью правила де Моргана так же, как в доказательстве первого утверждения. Предлагаем читателю сделать это самостоятельно. □

Следствие. *Пусть* E *открыто, а* F *замкнуто в* $\mathbb{R}^n$. *Тогда* $E \setminus F$ *открыто, а* $F \setminus E$ *замкнуто в* $\mathbb{R}^n$.

Действительно, так как $E \setminus F = E \cap \bigl(\mathbb{R}^n \setminus F\bigr)$, а $\mathbb{R}^n \setminus F$ открыто, то по теореме 2 $E \setminus F$ тоже открыто. Второе утверждение проверяется аналогично. □

Рассмотрим несколько примеров открытых и замкнутых множеств.

Пример 1. *Множества* $\mathbb{R}^n$ *и* $\varnothing$ *одновременно и открыты, и замкнуты в* $\mathbb{R}^n$. Предлагаем читателю проверить это самостоятельно. Отметим без доказательства, что верно и обратное: *если подмножество* $\mathbb{R}^n$ *открыто и замкнуто, то оно либо пусто, либо совпадает с* $\mathbb{R}^n$.

Пример 2. *Любое конечное подмножество* $\mathbb{R}^n$ *замкнуто*. Действительно, одноточечное множество не имеет предельных точек, и по теореме 1 оно замкнуто. Общий случай вытекает из теоремы 2.

Пример 3. Шары в $\mathbb{R}^n$. *Пусть* $a \in \mathbb{R}^n$, $\delta > 0$. *Положим*

$$\overline{V_a(\delta)} = \{x \in \mathbb{R}^n : \|x - a\| \leqslant \delta\}.$$

Тогда справедливы следующие утверждения.

1) *Множество* $V_a(\delta)$ *открыто в* $\mathbb{R}^n$.

2) *Множество* $\overline{V_a(\delta)}$ *замкнуто в* $\mathbb{R}^n$.

Доказательство. 1) Пусть $b \in V_a(\delta)$. Нам нужно проверить, что b — внутренняя точка $V_a(\delta)$. Обозначим через σ положительное число $\delta - \|b - a\|$. Для любого $x \in V_b(\sigma)$

$$\|x - a\| = \|x - b + b - a\| \leqslant \|x - b\| + \|b - a\| < \sigma + \|b - a\| = \delta,$$

откуда $x \in V_a(\delta)$. Таким образом, $V_b(\sigma) \subset V_a(\delta)$.

2) Пусть $\{x^k\}$ — последовательность в $\overline{V_a(\delta)}$, сходящаяся к $x \in \mathbb{R}^n$. Так как

$$\|x - a\| = \lim_{k\to\infty} \left\|x^k - a\right\| \leqslant \delta,$$

то $x \in \overline{V_a(\delta)}$, что и дает замкнутость $\overline{V_a(\delta)}$. $\square$

Замечание. Множество $\overline{V_a(\delta)}$ называют *замкнутым шаром* или *замкнутой окрестностью* a. Можно определить также *замкнутую окрестность бесконечно удаленной точки*:

$$\overline{V_\infty(\delta)} = \{x \in \mathbb{R}^n : \|x - a\| \geqslant \delta\}.$$

Множество $\overline{V_\infty(\delta)}$ является дополнением $V_0(\delta)$ и потому замкнуто.

Пример 4. Сферы в $\mathbb{R}^n$. *Пусть $a \in \mathbb{R}^n$, $\delta > 0$. Положим*

$$S_a(\delta) = \{x \in \mathbb{R}^n : \|x - a\| = \delta\}.$$

Тогда $S_a(\delta)$ замкнуто в $\mathbb{R}^n$.

Действительно, $S_a(\delta) = \overline{V_a(\delta)} \setminus V_a(\delta)$, поэтому требуемое утверждение вытекает из предыдущего примера и следствия теоремы 2.

Заметим, что при $n = 3$ множество $S_a(\delta)$ представляет собой сферу радиуса δ с центром в точке a. Поэтому $S_a(\delta)$ называют *n-мерной сферой*.

Пример 5. Параллелепипеды в $\mathbb{R}^n$. Пусть $a, b \in \mathbb{R}^n$, причем $a_k \leqslant b_k$ при всех $k \in \{1, \dots, n\}$. Множества

$$(a, b) = (a_1, b_1) \times \ldots \times (a_n, b_n), \quad [a, b] = [a_1, b_1] \times \ldots \times [a_n, b_n].$$

называются соответственно *открытым* и *замкнутым параллелепипедами* в $\mathbb{R}^n$. Аналогичным образом определяются параллелепипеды $(a, b]$ и $[a, b)$. Промежутки вида $\langle a_k, b_k \rangle$ называют *ребрами параллелепипеда* соответствующего вида. Если все ребра параллелепипеда имеют одинаковую длину, то его называют *кубом*.

Заметим, что если $a_k < b_k$ при всех $k \in \{1, \dots, n\}$, то $(a, b) \neq \varnothing$, а параллелепипед $[a, b]$ в этом случае называют *невырожденным*. Докажем утверждение, поясняющее термины "открытый" и "замкнутый" в определении параллелепипедов.

Пусть $a, b \in \mathbb{R}^n$, причем $a_k \leqslant b_k$ при всех $k \in \{1, \dots, n\}$. Тогда справедливы следующие утверждения.

1) *Множество (a, b) открыто в $\mathbb{R}^n$.*
2) *Множество $[a, b]$ замкнуто в $\mathbb{R}^n$.*

Доказательство. 1) Пусть $c \in (a, b)$. Это равносильно включениям $c_k \in (a_k, b_k)$ при всех $k \in \{1, \dots, n\}$. Положим

$$\delta_k = \min\{c_k - a_k, b_k - c_k\}, \ \ k \in \{1, \dots, n\}; \quad \delta = \min\{\delta_1, \dots, \delta_n\}.$$

Тогда $(c_k - \delta, c_k + \delta) \subset (a_k, b_k)$ при всех $k \in \{1, \dots, n\}$. Проверим, что $V_c(\delta) \subset (a, b)$. Если $x \in V_c(\delta)$, то в силу (1)

$$|x_k - c_k| \leqslant \|x - c\| < \delta \quad \text{для любых} \ \ k \in \{1, \dots, n\},$$

откуда $x \in (a, b)$. Таким образом, c является внутренней точкой множества (a, b), поэтому (a, b) открыто.

2) Пусть $\{x^k\}$ — последовательность в $[a, b]$, сходящаяся к точке $c \in \mathbb{R}^n$. Если $i \in \{1, \ldots, n\}$, то по теореме 1 § 2 $\lim\limits_{k\to\infty} x_i^k = c_i$. Переходя к пределу при $k \to \infty$ в неравенствах $a_i \leqslant x_i^k \leqslant b_i$, мы получим $a_i \leqslant c_i \leqslant b_i$ при всех $i \in \{1, \ldots, n\}$, то есть $c \in [a, b]$. Тогда по теореме 1 множество $[a, b]$ замкнуто. $\square$

Замечание. Не следует думать, что любое подмножество $\mathbb{R}^n$ открыто или замкнуто. Например, множество $[0, 1)$ не будет ни открытым, ни замкнутым в $\mathbb{R}$. Действительно, точка 0 не является внутренней для $[0, 1)$, а точка 1 — для $\mathbb{R} \setminus [0, 1)$. Аналогичные утверждения верны для параллелепипедов в $\mathbb{R}^n$ вида $[a, b)$ и $(a, b]$.

Введем следующие важные понятия, касающиеся множеств в $\mathbb{R}^n$: *замыкание*, *внутренность* и *граница*. Разберем их по отдельности. Иногда бывает полезно расширить произвольное множество до замкнутого. Это приводит нас к следующему определению.

Определение 2. Замыкание множества. Пусть $E \subset \mathbb{R}^n$. *Замыканием* E называют замкнутое множество $F \subset \mathbb{R}^n$, удовлетворяющее двум условиям:

1) $F \supset E$;

2) если A — другое замкнутое подмножество $\mathbb{R}^n$, содержащее E, то $A \supset F$.

Замыкание E принято обозначать символом $\mathrm{Cl}\, E$ (используется также запись $\overline{E}$).

Замечание. $\mathrm{Cl}\, E$ является *наименьшим по включению замкнутым множеством, содержащим* E. Более того, верно равенство

$$\mathrm{Cl}\, E = \bigcap \bigl\{F \subset \mathbb{R}^n : \ F \ \text{замкнуто}, \ F \supset E\bigr\}.$$

Действительно, замкнутость множества, стоящего в правой части, вытекает из теоремы 2, а условия 1) и 2) для него, очевидно, выполнены. Тем самым доказаны *существование* и *единственность* замыкания. Понятно также, что если E замкнуто, то $\mathrm{Cl}\, E = E$.

Приведенное выше описание замыкания множества неудобно для практических целей. Сформулируем характеристику замыкания в терминах последовательностей и предельных точек.

Теорема 3. Описание замыкания. *Пусть* $E \subset \mathbb{R}^n$, $a \in \mathbb{R}^n$. *Тогда равносильны следующие утверждения.*

1) $a \in \operatorname{Cl} E$.

2) *Точка* a *лежит в* E *или является предельной для* E.

3) *В* E *существует последовательность, сходящаяся к* a.

Доказательство проведем по схеме 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 1).

1) $\Rightarrow$ 2) Если $a \notin E$ и a не является предельной точкой E, то найдется такое $\delta > 0$, что $V_a(\delta) \subset \mathbb{R}^n \setminus E$. Перепишем это включение в виде $E \subset \mathbb{R}^n \setminus V_a(\delta)$. Отсюда в силу замкнутости $\mathbb{R}^n \setminus V_a(\delta)$ вытекает, что $\operatorname{Cl} E \subset \mathbb{R}^n \setminus V_a(\delta)$. Но это невозможно, поскольку $a \in \operatorname{Cl} E$.

2) $\Rightarrow$ 3) Если a является предельной точкой E, то искомая последовательность существует по теореме 6 § 2. В случае $a \in E$ нам подойдет постоянная последовательность, все члены которой равны a.

3) $\Rightarrow$ 1) Пусть $\{x^k\}$ — последовательность в E, сходящаяся к a. Тогда $\{x^k\}$ лежит и в $\operatorname{Cl} E$, откуда по теореме 1 $a \in \operatorname{Cl} E$. $\square$

Замечание. Теорема 3 дает два различных способа описания замыкания E: как объединение E с множеством его предельных точек или как множество пределов всех сходящихся последовательностей, лежащих в E.

Введем теперь другое важное понятие — *внутренность множества*.

Определение 3. Внутренность множества. Пусть $E \subset \mathbb{R}^n$. Множество всех внутренних точек E называется *внутренностью* E и обозначается $\operatorname{Int} E$ или $\overset{\circ}{E}$.

Если множество E открыто, то, очевидно, $\operatorname{Int} E = E$. Сформулируем еще несколько простых свойств $\operatorname{Int} E$.

Замечание 1. *Для любого* $E \subset \mathbb{R}^n$ *множество* $\operatorname{Int} E$ *открыто в* $\mathbb{R}^n$. Действительно, если $a \in \operatorname{Int} E$, то существует такое $\delta > 0$, что $V_a(\delta) \subset E$. Из открытости $V_a(\delta)$ вытекает, что любая точка $x \in V_a(\delta)$ является внутренней для $V_a(\delta)$ и, тем более, для E. Таким образом, $V_a(\delta) \subset \operatorname{Int} E$.

Замечание 2. Пусть $E \subset \mathbb{R}^n$. Тогда $\operatorname{Int} E$ — наибольшее по включению открытое множество, содержащееся в E. Иными словами, *если множество G открыто в $\mathbb{R}^n$ и $G \subset E$, то $G \subset \operatorname{Int} E$.* Действительно, пусть $a \in G$. Тогда a является внутренней точкой для G и, тем более, для E. Поэтому $a \in \operatorname{Int} E$.

Замечание 3. Отметим взаимосвязь между замыканием и внутренностью: *для любого $E \subset \mathbb{R}^n$ справедливы равенства*

$$\mathbb{R}^n \setminus \operatorname{Cl} E = \operatorname{Int}(\mathbb{R}^n \setminus E), \quad \mathbb{R}^n \setminus \operatorname{Int} E = \operatorname{Cl}(\mathbb{R}^n \setminus E).$$

Предлагаем читателю проверить их самостоятельно.

Определение 4. Граница множества. Пусть $E \subset \mathbb{R}^n$. Множество $\operatorname{Cl} E \setminus \operatorname{Int} E$ называется *границей* E и обозначается через ∂E.

Замечание. *Граница любого множества $E \subset \mathbb{R}^n$ замкнута.* Это вытекает из следствия теоремы 2, так как ∂E есть разность замкнутого и открытого множеств.

Проиллюстрируем введенные понятия на примере шаров и параллелепипедов.

Пример 1. *Пусть $a \in \mathbb{R}^n$, $\delta > 0$. Тогда справедливы следующие утверждения.*

1) $\overline{V_a(\delta)} = \operatorname{Cl} V_a(\delta)$.
2) $\operatorname{Int} \overline{V_a(\delta)} = V_a(\delta)$.
3) $\partial \overline{V_a(\delta)} = \partial V_a(\delta) = S_a(\delta)$.

Доказательство. 1) Из замкнутости множества $\overline{V_a(\delta)}$ вытекает, что $\operatorname{Cl} V_a(\delta) \subset \overline{V_a(\delta)}$. Проверим обратное включение. Пусть $x \in \overline{V_a(\delta)}$. Положим

$$x^k = a + \left(1 - \tfrac{1}{k}\right)(x - a) \quad (k \in \mathbb{N}).$$

По теореме 2 § 2 $\lim\limits_{k\to\infty} x^k = x$. Кроме того,

$$\left\|x^k - a\right\| = \left(1 - \tfrac{1}{k}\right)\|x - a\| \leqslant \left(1 - \tfrac{1}{k}\right)\delta < \delta \quad (k \in \mathbb{N}),$$

то есть $x^k \in V_a(\delta)$ при всех $k \in \mathbb{N}$. Тогда по теореме 3 $x \in \operatorname{Cl} V_a(\delta)$.

2) Из открытости $V_a(\delta)$ вытекает, что $\operatorname{Int} \overline{V_a(\delta)} \supset V_a(\delta)$. Проверим обратное включение. Так как

$$\overline{V_a(\delta)} \setminus V_a(\delta) = S_a(\delta),$$

нам достаточно показать, что никакая точка $x \in S_a(\delta)$ не является внутренней для $\overline{V_a(\delta)}$. Положим

$$x^k = a + \left(1 + \tfrac{1}{k}\right)(x - a) \quad (k \in \mathbb{N}).$$

Тогда $x^k \notin \overline{V_a(\delta)}$ при всех $k \in \mathbb{N}$, поскольку

$$\left\|x^k - a\right\| = \left(1 + \tfrac{1}{k}\right)\|x - a\| = (1 + \tfrac{1}{k})\,\delta > \delta.$$

Так как $\lim\limits_{k\to\infty} x^k = x$, любая окрестность x содержит точки последовательности $\left\{x^k\right\}$. Поэтому $x \notin \operatorname{Int}\overline{V_a(\delta)}$.

3) В силу двух предыдущих утверждений у множеств $V_a(\delta)$ и $\overline{V_a(\delta)}$ замыкание равно $\overline{V_a(\delta)}$, а внутренность — $V_a(\delta)$. □

Замечание. Аналогичные утверждения справедливы и для окрестностей бесконечно удаленной точки: *если $\delta > 0$, то*

$$\operatorname{Cl} V_\infty(\delta) = \overline{V_\infty(\delta)}, \quad \operatorname{Int}\overline{V_\infty(\delta)} = V_\infty(\delta).$$

Действительно, в силу замечания 3 к определению 3

$$\operatorname{Cl} V_\infty(\delta) = \operatorname{Cl}\big(\mathbb{R}^n \setminus \overline{V_0(\delta)}\big) = \mathbb{R}^n \setminus \operatorname{Int}\overline{V_0(\delta)} = \mathbb{R}^n \setminus V_0(\delta) = \overline{V_\infty(\delta)}.$$

Второе равенство доказывается аналогично.

Пример 2. *Пусть $a, b \in \mathbb{R}^n$, причем при всех $k \in \{1, \ldots, n\}$ $a_k < b_k$. Тогда справедливы следующие утверждения.*

1) $[a, b] = \operatorname{Cl}\langle a, b\rangle$.

2) $\operatorname{Int}\langle a, b\rangle = (a, b)$.

3) $\partial\langle a, b\rangle = [a, b] \setminus (a, b)$.

Здесь $\langle a, b\rangle$ — параллелепипед произвольного вида.

Доказательство мы проведем только для первого утверждения, оставляя два других читателю в качестве упражнения. Из замкнутости $[a, b]$ вытекает, что $\operatorname{Cl}\langle a, b\rangle \subset [a, b]$. Проверим обратное включение. Пусть $x \in [a, b]$. Для любого $i \in \{1, \ldots, n\}$ можно подобрать последовательность $\left\{x_i^k\right\}_{k=1}^\infty$ в (a_i, b_i), сходящуюся к x_i (см. предыдущий пример). Поэтому последовательность $\left\{x^k\right\}$ лежит в $\langle a, b\rangle$ и сходится к x. Тогда по теореме 3 $x \in \operatorname{Cl}\langle a, b\rangle$. □

Замечание. Множество $[a,b]\setminus(a,b)$ можно описать более конструктивно. Для любых $k\in\{1,\ldots,n\}$ и $x\in[a_k,b_k]$ обозначим через $P_k(x)$ параллелепипед, у которого k-е ребро есть $\{x\}$, а все остальные ребра такие же, как у $[a,b]$. Таким образом,

$$[a,b]\setminus(a,b)=\bigcup_{k=1}^{n}\big(P_k(a_k)\cup P_k(b_k)\big).$$

Параллелепипеды $P_k(a_k)$ и $P_k(b_k)$ называют *гранями* $[a,b]$. В дальнейшем будет удобно отождествить $P_k(x)$ с параллелепипедом в $\mathbb{R}^{n-1}$ вида

$$[a_1,b_1]\times\ldots\times[a_{k-1},b_{k-1}]\times[a_{k+1},b_{k+1}]\times\ldots\times[a_n,b_n].$$

До сих пор мы говорили об окрестностях точек в $\mathbb{R}^n$. Это понятие можно обобщить и ввести окрестности точек в произвольном подмножестве $\mathbb{R}^n$. Пусть $E\subset\mathbb{R}^n$. Для любой точки $a\in E$ *окрестностями a в E* называются множества вида

$$V_a^E(\delta)=V_a(\delta)\cap E\quad(\delta>0).$$

Теперь мы можем определить открытые и замкнутые подмножества E.

Определение 5. Пусть $E\subset\mathbb{R}^n$, $F\subset E$.

1) Точка $a\in F$ называется *внутренней для F в E*, если у нее есть окрестность V_a^E, содержащаяся в F.

2) Множество F называется *открытым в E*, если любая его точка является внутренней для F в E.

3) Множество F называется *замкнутым в E*, если его дополнение $E\setminus F$ открыто в E.

Докажем теперь утверждение, которое окажется полезным в § 5 при изучении непрерывных отображений.

Теорема 4. *Пусть $F\subset E\subset\mathbb{R}^n$. Тогда равносильны следующие утверждения.*

1) *Множество F открыто в E.*

2) *Существует открытое в $\mathbb{R}^n$ множество G, для которого $F=E\cap G$.*

Доказательство. 1) $\Rightarrow$ 2) Любая точка $a \in F$ является внутренней в E, то есть существует такое $\delta_a > 0$, что

$$V_a^E(\delta_a) \subset F.$$

Положим $G = \bigcup\limits_{a \in F} V_a(\delta_a)$. Множество G открыто в $\mathbb{R}^n$ по теореме 2. Кроме того,

$$G \cap E = \bigcup_{a \in F} \big(V_a(\delta_a) \cap E\big) = \bigcup_{a \in F} V_a^E(\delta_a) \subset F.$$

Обратное включение очевидно, поскольку $F \subset G$ и $F \subset E$.

2) $\Rightarrow$ 1) Пусть $a \in F$. В силу открытости G найдется такое $\delta > 0$, что $V_a(\delta) \subset G$. Тогда

$$V_a^E(\delta) = V_a(\delta) \cap E \subset G \cap E = F,$$

то есть a является внутренней точкой F в E. $\square$

Замечание 1. Пусть $F \subset E \subset \mathbb{R}^n$. Если E открыто в $\mathbb{R}^n$, то по теоремам 2 и 4 открытость множества F в E эквивалентна его открытости в $\mathbb{R}^n$. Для произвольного E это уже не так. Например, множество $(0, 1) \cap \mathbb{Q}$ открыто в $\mathbb{Q}$, но не в $\mathbb{R}$.

Замечание 2. Утверждение, аналогичное теореме 4, справедливо и для замкнутых в E множеств. Предлагаем читателю самостоятельно его сформулировать и вывести из теоремы 4.

Перейдем теперь к рассмотрению еще одного фундаментального математического понятия — *компактности*. Компактные множества будут играть важную роль при изучении непрерывных отображений. Дадим вначале следующее вспомогательное определение.

Определение 6. Покрытие множества. Пусть $E \subset \mathbb{R}^n$.

1) Если Ω — такое семейство подмножеств $\mathbb{R}^n$, что $\bigcup\limits_{A \in \Omega} A \supset E$, то говорят, что Ω *является покрытием* E (или Ω *покрывает* E). Если при этом все элементы Ω открыты, то Ω называют *открытым покрытием* E.

2) Пусть $\widetilde{\Omega}$ — подсемейство Ω, которое также покрывает E. Тогда $\widetilde{\Omega}$ называют *подпокрытием* Ω.

Определение 7. Компактные множества в $\mathbb{R}^n$. Множество $E \subset \mathbb{R}^n$ называется *компактным* (или *компактом*), если из любого открытого покрытия E можно выбрать конечное подпокрытие.

Замечание. *Если множество $E \subset \mathbb{R}^n$ компактно, то любое замкнутое подмножество E также компактно.*

Доказательство. Пусть F — замкнутое подмножество E, Ω — открытое покрытие F. Покажем, что из Ω можно выбрать конечное подпокрытие. Добавляя к Ω множество $\mathbb{R}^n \setminus F$, мы получим открытое покрытие компакта E. Выберем из этого покрытия конечное подсемейство $\widetilde{\Omega}$, которое также покрывает E. Если множество $\mathbb{R}^n \setminus F$ входит в $\widetilde{\Omega}$, удалим его оттуда. Мы получим конечное подпокрытие Ω множества F. $\square$

Определение 7 имеет важное теоретическое значение, но для практических целей оно неудобно. Далее мы получим описание компактности в более привычных терминах, чем покрытия. Для этого нам потребуется некоторая подготовка. Обобщим вначале на многомерный случай теорему о стягивающихся отрезках, доказанную в § 2 главы 2.

Определение 8. Стягивающиеся множества.

1) Для множества $E \subset \mathbb{R}^n$ назовем *диаметром* E величину

$$\operatorname{diam} E = \sup\big\{\|x-y\| :\ x, y \in E\big\}.$$

2) Непустые множества $\big\{E_k\big\}_{k=1}^{\infty}$ в $\mathbb{R}^n$ называются *стягивающимися*, если

$$E_k \supset E_{k+1} \ \text{ при всех } \ k \in \mathbb{N} \quad \text{и} \quad \operatorname{diam} E_k \to 0 \ \ (k \to \infty).$$

Замечание. Если $P = [a, b]$ — замкнутый параллелепипед в $\mathbb{R}^n$, то $\operatorname{diam} P = \|b - a\|$. Действительно, для любых $x, y \in P$

$$\|x-y\| = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \leqslant \sqrt{\sum_{i=1}^{n}(b_i - a_i)^2} = \|b-a\|\,,$$

откуда $\operatorname{diam} P \leqslant \|b - a\|$. Обратное неравенство очевидно. Предлагаем читателю проверить, что $\operatorname{diam}\langle a, b\rangle = \|b - a\|$.

Лемма 1. О стягивающихся параллелепипедах. *Пусть* $\{[a^k, b^k]\}_{k=1}^{\infty}$ *— последовательность стягивающихся параллелепипедов. Тогда* $\bigcap\limits_{k=1}^{\infty} [a^k, b^k]$ *состоит из единственной точки.*

Доказательство. Положим $P_k = [a^k, b^k]$ $(k \in \mathbb{N})$, $P = \bigcap\limits_{k=1}^{\infty} P_k$. Покажем вначале, что $P \neq \varnothing$. Для любого $i \in \{1, \ldots, n\}$ отрезки $[a_i^k, b_i^k]$ являются вложенными, поэтому их пересечение содержит некоторую точку c_i. Тогда $(c_1, \ldots, c_n) \in P$, то есть $P \neq \varnothing$. Проверим теперь, что P состоит из *единственной* точки. Если $x, y \in P$, то $x, y \in P_k$ для любого $k \in \mathbb{N}$. Поэтому

$$\|x - y\| \leqslant \operatorname{diam} P_k \to 0 \quad (k \to \infty).$$

Отсюда $\|x - y\| = 0$, то есть $x = y$. $\square$

Приведем теперь пример компактного множества, который нам потребуется для дальнейших рассуждений.

Лемма 2. *Любой замкнутый куб* $P \subset \mathbb{R}^n$ *компактен.*

Доказательство. Опишем вначале процедуру дробления кубов. Пусть $[a, b]$ — замкнутый куб в $\mathbb{R}^n$. Назовем *полукубами* $[a, b]$ множества вида $\Delta_1 \times \ldots \times \Delta_n$, где при любом $i \in \{1, \ldots, n\}$

$$\Delta_i = \left[a_i, \frac{a_i + b_i}{2}\right] \quad \text{или} \quad \Delta_i = \left[\frac{a_i + b_i}{2}, b_i\right].$$

Каждый полукуб является кубом вдвое меньшего диаметра, чем $[a, b]$. Очевидно также, что $[a, b]$ имеет ровно 2^n полукубов, объединение которых дает $[a, b]$. Для $n = 3$ эта процедура имеет следующий геометрический смысл: мы пересекаем куб плоскостями, проходящими через середины всех его ребер перпендикулярно этим ребрам, и получаем полукубы.

Доказательство леммы проведем от противного. Пусть Ω — открытое покрытие P, из которого нельзя выбрать конечного подпокрытия. Тогда у P существует хотя бы один полукуб P_1, который не покрывается никаким конечным подсемейством Ω. Аналогично, у куба P_1 найдется полукуб P_2, который нельзя покрыть конечным числом элементов Ω. Продолжая этот процесс до бесконечности, мы получим последовательность замкнутых кубов $\left\{P_k\right\}_{k=1}^{\infty}$, удовлетворяющих следующим условиям:

1) $P \supset P_1 \supset P_2 \supset \ldots$;

2) $\operatorname{diam} P_k = 2^{-k} \cdot \operatorname{diam} P$ при всех $k \in \mathbb{N}$;

3) ни при каком $k \in \mathbb{N}$ не существует конечного подпокрытия Ω куба P_k.

В силу 1) и 2) кубы P_k являются стягивающимися, поэтому их пересечение содержит некоторую точку $c \in P$. Так как Ω покрывает P, найдется такое множество G из Ω, что $c \in G$. В силу открытости G существует $\delta > 0$, для которого $V_c(\delta) \subset G$. Поскольку $\operatorname{diam} P_k \to 0$ при $k \to \infty$, мы можем подобрать такой номер $k \in \mathbb{N}$, что $\operatorname{diam} P_k < \delta$. Тогда для любого $x \in P_k$

$$\|x - c\| \leqslant \operatorname{diam} P_k < \delta, \quad \text{то есть} \quad x \in V_c(\delta) \subset G.$$

Таким образом, $P_k \subset G$, что противоречит условию 3). $\square$

Теперь мы можем доказать важную теорему, дающую характеристику компактности в различных терминах.

Теорема 5. Критерий компактности в $\mathbb{R}^n$. *Пусть $E \subset \mathbb{R}^n$. Тогда равносильны следующие утверждения.*

1) *E компактно.*

2) *E замкнуто и ограничено.*

3) *Из любой последовательности элементов E можно выбрать сходящуюся подпоследовательность, предел которой также лежит в E.*

Замечание 1. Импликацию 2) $\Rightarrow$ 1) принято называть *теоремой Гейне – Бореля.*

Замечание 2. Множества, удовлетворяющие условию 3), часто называют *секвенциально компактными.*

Доказательство проведем по схеме 1) $\Rightarrow$ 3) $\Rightarrow$ 2) $\Rightarrow$ 1).

1) $\Rightarrow$ 3) Пусть $\left\{x^k\right\}_{k=1}^{\infty}$ — последовательность в E. По теореме 4 § 2 из нее можно выбрать подпоследовательность $\left\{x^{k_i}\right\}_{i=1}^{\infty}$, имеющую предел $a \in \overline{\mathbb{R}^n}$. Нам достаточно проверить включение $a \in E$. Предположим, что $a \notin E$. Рассмотрим семейство открытых множеств

$$\Omega = \left\{\mathbb{R}^n \setminus \overline{V_a(\varepsilon)}\right\}_{\varepsilon > 0}.$$

В силу закона де Моргана (см. § 1 главы 1)

$$\bigcup_{A\in\Omega} A = \mathbb{R}^n \setminus \bigcap_{\varepsilon>0} \overline{V_a(\varepsilon)} = \mathbb{R}^n \setminus \{a\} \supset E,$$

то есть Ω покрывает E. Пусть теперь $\widetilde{\Omega}$ — произвольное конечное подсемейство Ω. Тогда существуют $m \in \mathbb{N}$ и положительные числа $\varepsilon_1, \ldots, \varepsilon_m$, для которых

$$\widetilde{\Omega} = \left\{\mathbb{R}^n \setminus \overline{V_a(\varepsilon_k)}\right\}_{k=1}^{m}.$$

Множество $V_a = \bigcap\limits_{k=1}^{m} V_a(\varepsilon_k)$ является окрестностью точки a. Заметим, что

$$\bigcup_{A\in\widetilde{\Omega}} A = \mathbb{R}^n \setminus \bigcap_{k=1}^{m} \overline{V_a(\varepsilon_k)} \subset \mathbb{R}^n \setminus V_a.$$

Но по определению предела $x^{k_i} \in E \cap V_a$ при всех достаточно больших $i \in \mathbb{N}$. Таким образом, $\widetilde{\Omega}$ не покрывает E, что противоречит компактности E.

3) $\Rightarrow$ 2) Проверим вначале замкнутость E. Пусть $\left\{x^k\right\}$ — последовательность в E, сходящаяся к некоторой точке $a \in \mathbb{R}^n$. В силу теоремы 1 нам достаточно показать, что $a \in E$. Из условия 3) вытекает, что у $\left\{x^k\right\}$ есть подпоследовательность, предел которой лежит в E. Но любая подпоследовательность $\left\{x^k\right\}$ сходится к a, откуда $a \in E$.

Докажем теперь ограниченность E. Если E не ограничено, то по любому $k \in \mathbb{N}$ найдется $x^k \in E$, для которого $\left\|x^k\right\| \geqslant k$. Но тогда $x^k \to \infty$ при $k \to \infty$, и по теореме 3 § 2 все подпоследовательности $\left\{x^k\right\}$ также стремятся к бесконечности. Мы получили противоречие с условием 3).

2) $\Rightarrow$ 1) Так как множество E ограничено в $\mathbb{R}^n$, его проекции на все координатные оси ограничены в $\mathbb{R}$ (см. замечание 2 к определению 2 § 2). Поэтому существует такое $C > 0$, что $E \subset [-C, C]^n$. По лемме 2 куб $[-C, C]^n$ компактен. Тогда в силу замечания к определению 7 компактно и его замкнутое подмножество E. $\square$

Замечание. Из теоремы 5 вытекает, что рассмотренные ранее замкнутые множества (шары, сферы, параллелепипеды) являются также и компактными, поскольку их ограниченность очевидна. В частности, компактны отрезки вещественной оси.

§ 4. Предел отображения нескольких переменных

В этом параграфе будут рассмотрены отображения вида

$$f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m, \quad \text{где } m, n \in \mathbb{N}.$$

Мы будем называть их *отображениями нескольких переменных*. Если $m = 1$, то f называется *функцией нескольких переменных*, а при $n = 1$ — *вектор-функцией*. Под эту терминологию подпадают также *отображения комплексного аргумента* и *комплекснозначные отображения*, поскольку $\mathbb{C}$ как множество совпадает с $\mathbb{R}^2$. Наша задача — ввести понятие предела отображения нескольких переменных и доказать для таких отображений аналоги некоторых теорем главы 2.

Определение 1. Предел отображения по Коши. Пусть $E \subset \mathbb{R}^n$, $f\colon E \to \mathbb{R}^m$, точка $a \in \overline{\mathbb{R}^n}$ является предельной для E, $A \in \overline{\mathbb{R}^m}$. Будем говорить, что $\lim\limits_{x\to a} f(x) = A$ (или $f(x) \to A$ при $x \to a$) *в смысле Коши*, если

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in \dot{V}_a(\delta) \cap E \ \ f(x) \in V_A(\varepsilon).$$

Определение 1 называют еще *определением предела на языке окрестностей*. Оно выглядит так же, как для функции одной переменной, но окрестности в данном случае имеют другой смысл.

Замечание 1. Определение 1 можно сформулировать и на языке неравенств. Если $a \in \mathbb{R}^n$ и $A \in \mathbb{R}^m$, то соотношение $\lim\limits_{x\to a} f(x) = A$ означает, что

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in E\colon \ 0 < \|x - a\| < \delta \ \ \|f(x) - A\| < \varepsilon,$$

а при $a = \infty$ и $A = \infty$ оно эквивалентно

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in E\colon \ \|x\| > \delta \ \ \|f(x)\| > \varepsilon.$$

Рекомендуем читателю самостоятельно записать с помощью неравенств остальные случаи.

Замечание 2. Определение 1 содержит в себе понятия предела отображения комплексного аргумента и предела комплекснозначного отображения. Если, например, $f: E \subset \mathbb{C} \to \mathbb{C}$ и $a, A \in \mathbb{C}$, то равенство $\lim\limits_{x \to a} f(x) = A$ эквивалентно

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in E\colon \; 0 < |x - a| < \delta \;\; |f(x) - A| < \varepsilon,$$

поскольку модуль комплексного числа совпадает с его нормой в $\mathbb{R}^2$.

Определение 2. Предел отображения по Гейне. Пусть $E \subset \mathbb{R}^n$, $f: E \to \mathbb{R}^m$, точка $a \in \overline{\mathbb{R}^n}$ является предельной для E, $A \in \overline{\mathbb{R}^m}$. Будем говорить, что $\lim\limits_{x \to a} f(x) = A$ *в смысле Гейне*, если для любой последовательности $\left\{x^k\right\}_{k=1}^{\infty}$ в $E \setminus \{a\}$, стремящейся к a, выполняется условие $\lim\limits_{k \to \infty} f(x^k) = A$.

Замечание. Для доказательства существования предела f в смысле Гейне достаточно проверить, что любая последовательность $\left\{x^k\right\}_{k=1}^{\infty}$ в $E \setminus \{a\}$, стремящаяся к a, переводится отображением f в последовательность, имеющую предел. *Равенство* пределов последовательностей вида $\left\{f(x^k)\right\}$ доказывается так же, как в замечании 6 § 3 главы 2.

Теорема 1. Эквивалентность двух определений предела. *Пусть $E \subset \mathbb{R}^n$, $f: E \to \mathbb{R}^m$, точка $a \in \overline{\mathbb{R}^n}$ является предельной для E, $A \in \overline{\mathbb{R}^m}$. Тогда равносильны следующие утверждения.*

1) $\lim\limits_{x \to a} f(x) = A$ *в смысле Коши.*

2) $\lim\limits_{x \to a} f(x) = A$ *в смысле Гейне.*

Доказательство. 1) $\Rightarrow$ 2) Пусть A — предел f в смысле Коши. Возьмем последовательность $\left\{x^k\right\}_{k=1}^{\infty}$ в $E \setminus \{a\}$, стремящуюся к $\{a\}$. Нам нужно показать, что $f(x^k) \to A$ при $k \to \infty$. В силу 1) по любому $\varepsilon > 0$ можно подобрать $\delta > 0$, для которого

$$f(x) \in V_A(\varepsilon) \quad \text{для всех} \;\; x \in \dot{V}_a(\delta) \cap E.$$

Поскольку $x^k \to a$, существует такое $N \in \mathbb{N}$, что при всех $k > N$ справедливо включение $x^k \in V_a(\delta)$. Кроме того, $x^k \in E\setminus\{a\}$, откуда $x^k \in \dot{V}_a(\delta) \cap E$. Поэтому

$$f(x^k) \in V_A(\varepsilon) \quad \text{для любых} \;\; k > N.$$

Таким образом, $f(x^k) \to A$ при $k \to \infty$, то есть $\lim\limits_{x\to a} f(x) = A$ и в смысле Гейне.

2) $\Rightarrow$ 1) Пусть A — предел f по Гейне. Докажем, что предел f в смысле Коши также существует и равен A. Действительно, если это не так, то

$$\exists \varepsilon > 0 \;\; \forall \delta > 0 \;\; \exists x \in \dot{V}_a(\delta) \cap E \;\; f(x) \notin V_A(\varepsilon). \tag{2}$$

Положим

$$F = \{x \in E \setminus \{a\} \colon\, f(x) \notin V_A(\varepsilon)\}.$$

Условие (2) означает, что $\dot{V}_a(\delta) \cap F \neq \varnothing$ при любом $\delta > 0$, то есть a является предельной точкой F. В силу теоремы 6 § 2 найдется последовательность $\left\{x^k\right\}_{k=1}^{\infty}$ в F (и, значит, в $E\setminus\{a\}$), стремящаяся к a. Тогда по определению 2 $\lim\limits_{k\to\infty} f(x^k) = A$, что невозможно, так как $f(x^k) \notin V_A(\varepsilon)$ при всех $k \in \mathbb{N}$. $\square$

Следствие 1. Единственность предела. *Если $A, B \in \overline{\mathbb{R}^m}$, $f(x) \to A$ и $f(x) \to B$ при $x \to a$, то $A = B$.*

Доказательство. Пусть $\left\{x^k\right\}_{k=1}^{\infty}$ — последовательность в $E \setminus \{a\}$, стремящаяся к a. Тогда $\lim\limits_{k\to\infty} f(x^k) = A$ и $\lim\limits_{k\to\infty} f(x^k) = B$, откуда $A = B$ в силу единственности предела последовательности. $\square$

Следствие 2. Предел и покоординатный предел. *Пусть в условиях теоремы 1 $A \in \mathbb{R}^m$. Тогда равносильны следующие утверждения.*

1) $\lim\limits_{x\to a} f(x) = A$.

2) *Для любого $i \in \{1, \ldots, m\}$ верно равенство* $\lim\limits_{x\to a} f_i(x) = A_i$.

Доказательство. Пусть $\left\{x^k\right\}$ — последовательность в $E \setminus \{a\}$, стремящаяся к a. Утверждение 1) означает, что $\lim\limits_{k\to\infty} f(x^k) = A$, а утверждение 2) — что $\lim\limits_{k\to\infty} f_i(x^k) = A_i$ при любом $i \in \{1,\ldots,n\}$. Поэтому их равносильность вытекает из теоремы 1 § 2. $\square$

Пример 1. *Пусть $a \in \mathbb{R}^n$. Тогда $\lim\limits_{x\to a} \|x\| = \|a\|$.* Действительно, если $\{x^k\}$ — последовательность в $E\setminus\{a\}$, стремящаяся к a, то $\|x^k\| \to \|a\|$ при $k\to\infty$ (см. замечание 3 к определению 1 § 2). $\square$

Пример 2. *Пусть функция f задается формулой*

$$f(x,y) = \frac{xy}{x^2+y^2}, \quad (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}.$$

Тогда f не имеет предела в точке $(0,0)$. Действительно, положим

$$z^k = \left(\tfrac{1}{k}, 0\right), \quad w^k = \left(\tfrac{1}{k}, \tfrac{1}{k}\right) \quad (k\in\mathbb{N}).$$

По теореме 1 § 2 $z^k \to (0,0)$ и $w^k \to (0,0)$ при $k\to\infty$. Кроме того,

$$f(z^k) = 0 \to 0 \quad \text{и} \quad f(w^k) = \frac{1/k^2}{2/k^2} \to \frac{1}{2} \quad (k\to\infty).$$

Если бы предел f в точке $(0,0)$ существовал, то он был бы равен одновременно 0 и $\frac{1}{2}$, что невозможно. $\square$

Перейдем к обобщению теории пределов, построенной в главе 2. Рассмотрим вначале связь предела с арифметическими операциями в $\mathbb{R}^n$.

Теорема 2. Предел и арифметические операции. *Пусть $E \subset \mathbb{R}^n$, точка $a \in \overline{\mathbb{R}^n}$ является предельной для E, $f, g\colon E \to \mathbb{R}^m$. Предположим, что f, g имеют в точке a пределы $A, B \in \mathbb{R}^m$ соответственно. Тогда справедливы следующие утверждения.*

1) $\lim\limits_{x\to a}\bigl(f(x)+g(x)\bigr) = A+B$.
2) *Если* $\lambda\colon E\to\mathbb{R}$, $\alpha\in\mathbb{R}$, $\lim\limits_{x\to a}\lambda(x) = \alpha$, *то* $\lim\limits_{x\to a}\lambda(x)f(x) = \alpha A$.
3) $\lim\limits_{x\to a} f(x)\cdot g(x) = A\cdot B$.
4) *Если* $m=1$ *и* $B\neq 0$, *то* $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)} = \dfrac{A}{B}$.

Доказательство. Если $\{x^k\}$ — последовательность в $E\setminus\{a\}$, стремящаяся к a, то

$$f(x^k)\to A, \quad g(x^k)\to B, \quad \lambda(x^k)\to\alpha \quad \text{при} \quad k\to\infty.$$

Тогда по теореме 2 § 2

$$f(x^k) + g(x^k) \to A + B, \quad \lambda(x^k)\, f(x^k) \to \alpha A \quad (k \to \infty),$$

откуда вытекают утверждения 1) и 2). Для доказательства 3) заметим, что по теореме 1 § 2

$$f(x^k) \cdot g(x^k) = \sum_{i=1}^{m} f_i(x^k)\, g_i(x^k) \to \sum_{i=1}^{m} A_i B_i = A \cdot B \quad (k \to \infty).$$

Наконец, по теореме 5 § 1 главы 2

$$\lim_{k\to\infty} \frac{f(x^k)}{g(x^k)} = \frac{A}{B},$$

что и дает утверждение 4). □

Замечание 1. Из утверждения 2) теоремы 2 вытекает, что для любого $\lambda \in \mathbb{R}$ верно равенство

$$\lim_{x\to a} \lambda f(x) = \lambda A.$$

Комбинируя его с утверждением 1), мы получим при $\lambda, \mu \in \mathbb{R}$

$$\lim_{x\to a} \big(\lambda f(x) + \mu g(x)\big) = \lambda A + \mu B.$$

Это свойство называют *линейностью предела*.

Замечание 2. Утверждения 1), 2), 4) теоремы 2 верны и для *комплекснозначных функций* f, g и λ, если под арифметическими действиями понимать операции с комплексными числами. Докажем для примера комплексный вариант утверждения 4). Положим

$$f_1 = \operatorname{Re} f, \quad f_2 = \operatorname{Im} f, \quad g_1 = \operatorname{Re} g, \quad g_2 = \operatorname{Im} g.$$

Тогда

$$\operatorname{Re} \frac{f}{g} = \frac{f_1\, g_1 + f_2\, g_2}{g_1^2 + g_2^2}, \quad \operatorname{Im} \frac{f}{g} = \frac{f_2\, g_1 - f_1\, g_2}{g_1^2 + g_2^2}.$$

В силу следствия 2 теоремы 1 при $x \to a$

$$f_1(x) \to \operatorname{Re} A, \quad f_2(x) \to \operatorname{Im} A, \quad g_1(x) \to \operatorname{Re} B, \quad g_2(x) \to \operatorname{Im} B.$$

Тогда по теореме 2

$$\lim_{x \to a} \operatorname{Re} \frac{f(x)}{g(x)} = \frac{\operatorname{Re} A \cdot \operatorname{Re} B + \operatorname{Im} A \cdot \operatorname{Im} B}{(\operatorname{Re} B)^2 + (\operatorname{Im} B)^2} = \operatorname{Re} \frac{A}{B},$$

$$\lim_{x \to a} \operatorname{Im} \frac{f(x)}{g(x)} = \frac{\operatorname{Im} A \cdot \operatorname{Re} B - \operatorname{Re} A \cdot \operatorname{Im} B}{(\operatorname{Re} B)^2 + (\operatorname{Im} B)^2} = \operatorname{Im} \frac{A}{B},$$

и в силу следствия 2 теоремы 1

$$\frac{f(x)}{g(x)} \to \frac{A}{B} \quad \text{при} \ \ x \to a. \quad \square$$

Теорема 3. Предел композиции. *Пусть* $E \subset \mathbb{R}^n$, $F \subset \mathbb{R}^m$, $a \in \overline{\mathbb{R}^n}$ *и* $b \in F$ *— предельные точки* E *и* F *соответственно. Если* $f\colon E \to F$, $g\colon F \to \mathbb{R}^s$, $\lim\limits_{x \to a} f(x) = b$, $\lim\limits_{x \to b} g(x) = g(b)$, *то*

$$\lim_{x \to a} \bigl(g \circ f\bigr)(x) = g(b).$$

Доказательство. Пусть $\varepsilon > 0$. Существует такое $\sigma > 0$, что

$$g(y) \in V_{g(b)}(\varepsilon) \quad \text{при всех} \ \ y \in V_b(\sigma) \cap F. \tag{3}$$

Действительно, для $y \neq b$ это включение вытекает из определения предела g, а при $y = b$ оно очевидно. Подберем теперь такое $\delta > 0$, что

$$f(x) \in V_b(\sigma) \quad \text{для любого} \ \ x \in \dot{V}_a(\delta) \cap E.$$

Подставляя в (3) $y = f(x)$, мы получим

$$g\bigl(f(x)\bigr) \in V_{g(b)}(\varepsilon) \quad \text{при всех} \ \ x \in \dot{V}_a(\delta) \cap E.$$

Таким образом, $\lim\limits_{x \to a} \bigl(g \circ f\bigr)(x) = g(b)$. $\square$

Теорема 4. Критерий Больцано – Коши для отображений. *Пусть $E \subset \mathbb{R}^n$, точка $a \in \overline{\mathbb{R}^n}$ является предельной для E, $f\colon E \to \mathbb{R}^m$. Тогда равносильны следующие утверждения.*

1) *Отображение f имеет предел в $\mathbb{R}^m$.*

2) $\forall \varepsilon > 0 \;\; \exists V_a \;\; \forall x, y \in \dot{V}_a \cap E \quad \|f(x) - f(y)\| < \varepsilon$.

Замечание. Утверждение 2) называют *сходимостью f в себе в точке a.*

Доказательство. 1) $\Rightarrow$ 2) Положим $A = \lim\limits_{x\to a} f(x)$. По $\varepsilon > 0$ подберем окрестность V_a точки a, для которой

$$\|f(x) - A\| < \frac{\varepsilon}{2} \quad \text{при всех} \;\; x \in \dot{V}_a \cap E.$$

Тогда для любых $x, y \in \dot{V}_a \cap E$

$$\|f(x) - f(y)\| \leqslant \|f(x) - A\| + \|A - f(y)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

2) $\Rightarrow$ 1) По $\varepsilon > 0$ подберем окрестность V_a точки a в соответствии с условием 2). Пусть $\{x^k\}_{k=1}^{\infty}$ — последовательность в $E \setminus \{a\}$, стремящаяся к a. Тогда найдется такое $N \in \mathbb{N}$, что $x^k \in V_a$ при всех $k > N$. По выбору V_a

$$\left\|f(x^k) - f(x^m)\right\| < \varepsilon \quad \text{для любых} \;\; k, m > N.$$

Поэтому последовательность $\{f(x^k)\}$ сходится в себе и, согласно теореме 5 § 2, имеет предел в $\mathbb{R}^m$. В силу замечания к определению 2 отсюда вытекает, что отображение f имеет конечный предел в точке a. $\square$

§ 5. Непрерывные отображения нескольких переменных

В этом параграфе мы определим понятие непрерывности для отображений нескольких переменных и обобщим основные теоремы § 4 главы 2 на случай таких отображений.

Определение 1. Непрерывность отображения в точке. Пусть $E \subset \mathbb{R}^n$, $a \in E$, $f\colon E \to \mathbb{R}^m$. Отображение f называется *непрерывным в точке* a, если

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in V_a(\delta) \cap E \ \ f(x) \in V_{f(a)}(\varepsilon).$$

Это определение можно переписать и на языке неравенств:

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in E\colon \|x - a\| < \delta \ \ \|f(x) - f(a)\| < \varepsilon.$$

Сформулируем теперь несколько простых свойств непрерывности в точке. Предлагаем читателю либо доказать их самостоятельно, либо вспомнить рассуждения § 4 главы 2, которые без изменений переносятся на многомерный случай.

Замечание 1. Пусть точка a является предельной для E. Тогда непрерывность f в точке a эквивалентна равенству $\lim\limits_{x \to a} f(x) = f(a)$.

Замечание 2. Если a — изолированная точка множества E, то отображение f непрерывно в точке a.

Замечание 3. Непрерывность отображения f в точке a равносильна следующему утверждению:

$$\text{если } \left\{x^k\right\}_{k=1}^{\infty} \subset E, \ \ x^k \to a, \quad \text{то } \ f(x^k) \to f(a) \quad (k \to \infty).$$

Замечание 4. Непрерывность в точке a отображения f эквивалентна непрерывности всех его координатных функций $f_1, \dots, f_m$. В частности, непрерывность комплекснозначной функции f равносильна непрерывности $\operatorname{Re} f$ и $\operatorname{Im} f$. Эти утверждения вытекают из следствия 2 теоремы 1 § 4, если a является предельной точкой E, и из замечания 2 для изолированной точки a.

Перейдем к обобщению результатов § 4 главы 2 на многомерный случай. Изучим вначале локальные свойства непрерывных отображений.

Теорема 1. Непрерывность композиции. *Пусть $E \subset \mathbb{R}^n$, $F \subset \mathbb{R}^m$, $a \in E$, $f\colon E \to F$, $g\colon F \to \mathbb{R}^s$. Если f непрерывно в точке a и g непрерывно в точке $f(a)$, то $g \circ f$ непрерывно в точке a.*

Доказательство. Пусть $\{x^k\}_{k=1}^{\infty}$ — последовательность в E, стремящаяся к a. Применяя замечание 3 сначала к f, затем к g, мы получим

$$f(x^k) \to f(a), \quad g\big(f(x^k)\big) \to g\big(f(a)\big) \quad (k \to \infty).$$

Таким образом, $\big(g \circ f\big)(x^k) \to \big(g \circ f\big)(a)$ при $k \to \infty$. Отсюда в силу замечания 3 вытекает непрерывность $g \circ f$ в точке a. $\square$

Теорема 2. Непрерывность и арифметические операции. *Пусть $E \subset \mathbb{R}^n$, $a \in E$, отображения $f, g: E \to \mathbb{R}^m$ непрерывны в точке a. Тогда справедливы следующие утверждения.*

1) *Отображение $f + g$ непрерывно в точке a.*

2) *Если функция $\lambda: E \to \mathbb{R}$ непрерывна в точке a, то отображение λf непрерывно в точке a.*

3) *Отображение $f \cdot g$ непрерывно в точке a.*

4) *Если $m = 1$ и $g(a) \neq 0$, то функция $\dfrac{f}{g}$ непрерывна в точке a.*

Доказательство. Если a — предельная точка E, то все утверждения вытекают из теоремы 2 § 4, а для изолированной точки они верны в силу замечания 2. $\square$

Замечание 1. Утверждения 1), 2), 4) теоремы 2 справедливы и для *комплекснозначных функций* f, g и λ, если под арифметическими действиями понимать операции с комплексными числами (см. замечание 2 к теореме 2 § 4).

Замечание 2. Отметим два важных случая второго утверждения теоремы 2. Пусть отображения $f: E \to \mathbb{R}^m$ и $\lambda: E \to \mathbb{R}$ непрерывны в точке a. Тогда при любых $v \in \mathbb{R}^m$ и $\alpha \in \mathbb{R}$ непрерывными в точке a будут также отображения αf и λv.

Рассмотрим теперь глобальный вариант определения непрерывности.

Определение 2. Непрерывность отображения на множестве. Пусть $E \subset \mathbb{R}^n$, $f: E \to \mathbb{R}^m$. Отображение f называется *непрерывным на множестве* E, если оно непрерывно в любой точке E.

Для $F \subset \mathbb{R}^m$ множество непрерывных отображений, действующих из E в F, обозначают символами $C(E, F)$ или $C(E \to F)$. Как правило, в качестве F будет фигурировать все $\mathbb{R}^m$. Для класса *непрерывных функций* на E (то есть в случае $F = \mathbb{R}$) используется более краткое обозначение $C(E)$, что соответствует соглашениям § 4 главы 2. Множество комплекснозначных непрерывных функций на E мы также будем обозначать через $C(E)$. Это не вызовет путаницы, поскольку из контекста всегда ясно, о какой области значений идет речь.

При изучении глобальных свойств непрерывности в § 4 главы 2 мы рассматривали функции, заданные на промежутках. В многомерном случае промежутки необходимо заменить более общими классами множеств, описанными в § 3. Выясним вначале, как ведут себя эти множества под действием непрерывных отображений.

Теорема 3. *Пусть $E \subset \mathbb{R}^n$, $f\colon E \to \mathbb{R}^m$. Тогда равносильны следующие утверждения.*

1) *Отображение f непрерывно на E.*

2) *Если множество G открыто в $\mathbb{R}^m$, то $f^{-1}(G)$ открыто в E.*

Таким образом, непрерывность f на E означает, что *прообраз любого открытого в $\mathbb{R}^m$ множества открыт в E.* Подчеркнем, что в утверждении 2) речь идет об открытости $f^{-1}(G)$ именно в E, а не в $\mathbb{R}^n$ (см. определение 5 § 3).

Доказательство. 1) $\Rightarrow$ 2) Пусть множество G открыто в $\mathbb{R}^m$, $F = f^{-1}(G)$. Если $a \in F$, то $f(a) \in G$, и в силу открытости G найдется такое $\varepsilon > 0$, что $V_{f(a)}(\varepsilon) \subset G$. По непрерывности f существует $\delta > 0$, для которого $f\big(V_a(\delta) \cap E\big) \subset V_{f(a)}(\varepsilon)$. Тогда

$$V_a(\delta) \cap E \subset f^{-1}\big(V_{f(a)}(\varepsilon)\big) \subset f^{-1}(G) = F,$$

то есть любая точка $a \in F$ является внутренней для F в E. Поэтому множество F открыто в E.

2) $\Rightarrow$ 1) Пусть $a \in E$, $\varepsilon > 0$. В силу 2) множество $f^{-1}\big(V_{f(a)}(\varepsilon)\big)$ открыто в E. Поэтому существует такое $\delta > 0$, что

$$V_a(\delta) \cap E \subset f^{-1}\big(V_{f(a)}(\varepsilon)\big), \text{ то есть } f\big(V_a(\delta) \cap E\big) \subset V_{f(a)}(\varepsilon).$$

Отсюда и вытекает непрерывность f в точке a. $\square$

Замечание. Утверждение 2) не изменится, если предполагать множества G открытыми в $f(E)$, а не в $\mathbb{R}^m$.

Теорема 4. Непрерывный образ компакта. *Пусть $E \subset \mathbb{R}^n$, $f \in C(E, \mathbb{R}^m)$. Если множество E компактно, то $f(E)$ также компактно.*

Доказательство. Воспользуемся характеристикой компактности в терминах последовательностей. Пусть $\{y^k\}_{k=1}^{\infty}$ — последовательность в $f(E)$. Тогда существует такая последовательность $\{x^k\}_{k=1}^{\infty}$ в E, что $y^k = f(x^k)$ при всех $k \in \mathbb{N}$. Так как E секвенциально компактно (см. теорему 5 § 3), из $\{x^k\}$ можно выбрать подпоследовательность $\{x^{k_i}\}_{i=1}^{\infty}$, имеющую предел $a \in E$. По непрерывности f

$$y^{k_i} = f(x^{k_i}) \to f(a) \in f(E),$$

то есть подпоследовательность $\{y^{k_i}\}$ имеет предел в $f(E)$. Тогда по теореме 5 § 3 множество $f(E)$ компактно. $\square$

Следствие. Компактность отрезка в $\mathbb{R}^n$. *Пусть $a, b \in \mathbb{R}^n$. Тогда множество*

$$\Delta_{a,b} = \big\{a + t\,(b-a) : t \in [0,1]\big\} \tag{4}$$

компактно в $\mathbb{R}^n$.

Доказательство. В § 3 было показано, что отрезок $[0,1]$ компактен (см. замечание к теореме 5). Кроме того, отображение $\varphi(t) = a + t\,(b-a)$ непрерывно на $[0,1]$, так как его координатные функции суть многочлены. Отсюда по теореме 4 вытекает компактность множества $\varphi\big([0,1]\big)$, которое и есть $\Delta_{a,b}$. $\square$

Замечание. Если $n = 1$, то $\Delta_{a,b} = [a,b]$. Поэтому множество $\Delta_{a,b}$ называют *отрезком в $\mathbb{R}^n$ с концами a и b.*

Обобщим теперь на многомерный случай глобальные свойства непрерывных функций, доказанные в § 4 главы 2.

Теорема 5 (К. Вейерштрасс). *Пусть множество $E \subset \mathbb{R}^n$ компактно, $f \in C(E, \mathbb{R}^m)$. Тогда справедливы следующие утверждения.*

1) *Отображение f ограничено на E.*

2) *Если $m = 1$ (то есть f является функцией), то f достигает на E своего наибольшего и наименьшего значения.*

Доказательство. Положим $F = f(E)$. По теореме 4 множество F компактно и, значит, ограничено в $\mathbb{R}^m$. Тем самым доказана ограниченность f. Проверим теперь второе утверждение. Положим $M = \sup\limits_E f$. Из ограниченности F в $\mathbb{R}$ вытекает, что $M \in \mathbb{R}$. Тогда для любого $k \in \mathbb{N}$ найдется такое $y^k \in F$, что $M - \frac{1}{k} < y^k \leqslant M$. По теореме о сжатой последовательности $y^k \to M$ при $k \to \infty$. Так как F замкнуто, по теореме 1 § 3 $M \in F$, то есть $M \in f(E)$. Включение $\inf\limits_E f \in f(E)$ доказывается аналогично. $\square$

Замечание 1. Иногда утверждение 1) теоремы 5 называют первой теоремой Вейерштрасса, а утверждение 2) — второй теоремой Вейерштрасса о непрерывных отображениях.

Замечание 2. Так как отрезки в $\mathbb{R}$ являются компактными множествами, теорема 3 § 4 главы 2 вытекает из теоремы 5. Отметим, что даже в случае $n = 1$ теорема 5 сильнее, так как класс компактных подмножеств $\mathbb{R}$ содержит не только отрезки.

Рассмотрим теперь многомерный вариант понятия равномерной непрерывности. Рекомендуем читателю предварительно вспомнить материал § 4 главы 2, посвященный равномерной непрерывности функций.

Определение 3. Пусть $E \subset \mathbb{R}^n$, $f\colon E \to \mathbb{R}^m$. Отображение f называется *равномерно непрерывным на* E, если

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x, y \in E\colon \ \|x - y\| < \delta \quad \|f(x) - f(y)\| < \varepsilon.$$

Теорема 6 (Г. Кантор). *Пусть E — компактное подмножество $\mathbb{R}^n$, $f \in C(E, \mathbb{R}^m)$. Тогда f равномерно непрерывно на E.*

Доказательство. Пусть $\varepsilon > 0$. Необходимо подобрать по нему $\delta > 0$, удовлетворяющее определению 3. Для любого $a \in E$ существует такое $\delta_a > 0$, что

$$\text{если} \quad \|x - a\| < 2\delta_a, \quad \text{то} \quad \|f(x) - f(a)\| < \frac{\varepsilon}{2}. \tag{5}$$

Семейство $\left\{V_a(\delta_a)\right\}_{a \in E}$ образует открытое покрытие E. В силу компактности E из него можно выбрать конечное подпокрытие, то есть найти конечное подмножество $F \subset E$, для которого

$$E \subset \bigcup_{a \in F} V_a(\delta_a).$$

Положим $\delta = \min\limits_{a \in F} \delta_a$. Поскольку множество F конечно, число δ корректно определено и положительно. Покажем, что оно искомое. Действительно, пусть $x, y \in E$, $\|x - y\| < \delta$. Найдется такое $a \in F$, что $x \in V_a(\delta_a)$, то есть $\|x - a\| < \delta_a$. Тогда

$$\|y - a\| \leqslant \|y - x\| + \|x - a\| < \delta + \delta_a \leqslant 2\delta_a,$$

откуда в силу (5)

$$\|f(x) - f(y)\| \leqslant \|f(x) - f(a)\| + \|f(y) - f(a)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \square$$

Замечание 1. Теорема 6 сильнее теоремы 5 § 4 главы 2 даже при $n = m = 1$.

Замечание 2. Предлагаем читателю доказать теорему Кантора с использованием последовательностей, как это было сделано в одномерном случае.

§ 6. Линейные операторы

Для построения многомерного дифференциального исчисления нам потребуются отображения нескольких переменных, перестановочные со сложением и умножением на вещественные числа в $\mathbb{R}^n$. Эти отображения, называемые *линейными операторами*, будут играть ту же роль, что линейные функции в главе 3. В отличие от курса алгебры, мы будем преимущественно рассматривать аналитические свойства линейных операторов, связанные с их непрерывностью. Некоторые алгебраические теоремы, необходимые для изложения материала, будут формулироваться без доказательства.

Определение 1. Линейный оператор. Пусть отображение $T\colon \mathbb{R}^n \to \mathbb{R}^m$ удовлетворяет *условию линейности*:

$$T(\lambda x + \mu y) = \lambda T(x) + \mu T(y) \quad \text{при всех } x, y \in \mathbb{R}^n \text{ и } \lambda, \mu \in \mathbb{R}.$$

Тогда T называется *линейным оператором.*

Замечание. Следуя принятым в алгебре соглашениям, мы часто будем опускать скобки у аргумента линейного оператора, то есть писать Tx вместо $T(x)$.

Обозначим класс всех линейных операторов, действующих из $\mathbb{R}^n$ в $\mathbb{R}^m$, через $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ или $\mathcal{L}(\mathbb{R}^n \to \mathbb{R}^m)$. В случае $n = m$ вместо $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ мы будем писать просто $\mathcal{L}(\mathbb{R}^n)$.

Линейные операции над отображениями, определенные в § 1, без изменения переносятся на операторы. Отметим, что результатом этих операций также будут линейные операторы. Кроме того, введем *нулевой оператор*:

$$\mathbb{O}(x) = 0 \quad \text{при всех } x \in \mathbb{R}^n.$$

Для операторов $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ и $S \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^s)$ можно определить их *композицию* по обычному правилу:

$$(S \circ T)(x) = S\bigl(T(x)\bigr) \quad \text{при всех } x \in \mathbb{R}^n.$$

В алгебре принято называть композицию линейных операторов *произведением* и записывать ее в виде ST. Предлагаем читателю проверить, что $ST \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^s)$. Не следует путать произведение операторов с их скалярным произведением, введенным в § 1 (его мы обозначаем $S \cdot T$).

Каждому оператору $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ соответствует единственная матрица A размером $m \times n$, для которой

$$Tx = A \cdot x \quad \text{при всех } x \in \mathbb{R}^n$$

(здесь векторы в $\mathbb{R}^n$ и $\mathbb{R}^m$ записываются в виде столбцов). Если обозначить элементы матрицы A через a_{ij} для $i \in \{1, \dots, m\}$ и $j \in \{1, \dots, n\}$, то оператор T будет действовать по формуле

$$T_i\, x = \sum_{j=1}^{n} a_{ij}\, x_j, \quad \text{где } i \in \{1, \dots, m\} \text{ и } x \in \mathbb{R}^n. \tag{6}$$

Напомним, что через T_i мы обозначаем координатные функции T, которые, очевидно, также являются линейными операторами. Матрица A, называемая *матрицей оператора* T, строится следующим образом. Обозначим через $e^1, \dots, e^n$ *орты*, образующие *стандартный базис в* $\mathbb{R}^n$:

$$e^i_j = \begin{cases} 1, & j = i \\ 0, & j \neq i, \end{cases} \quad \text{где } i, j \in \{1, \dots, n\}. \tag{7}$$

Положим

$$a_{ij} = T_i(e^j) \quad \text{при всех } i \in \{1, \dots, m\} \text{ и } j \in \{1, \dots, n\}.$$

Иными словами, в качестве столбцов матрицы A нужно взять векторы $Te^1, \dots, Te^n$. Действительно, для любых $i \in \{1, \dots, m\}$ и $x \in \mathbb{R}^n$

$$T_i\, x = T_i \left(\sum_{j=1}^n x_j\, e^j \right) = \sum_{j=1}^n T_i\big(e^j\big)\, x_j = \sum_{j=1}^n a_{ij}\, x_j.$$

Замечание. Пусть $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, $S \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^s)$, A, B — матрицы T, S соответственно. Тогда матрицей ST будет $B \cdot A$, то есть при перемножении операторов их матрицы также перемножаются. Действительно, для любого $x \in \mathbb{R}^n$

$$STx = S(Tx) = B \cdot Tx = B \cdot (A \cdot x) = (B \cdot A) \cdot x.$$

Перейдем к изучению аналитических свойств линейных операторов. Докажем, что они обладают двумя важными свойствами: *непрерывностью* и *ограниченностью*.

Теорема 1. *Пусть $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Тогда справедливы следующие утверждения.*

1) *Оператор T непрерывен на $\mathbb{R}^n$.*

2) *Существует число $C \geqslant 0$, для которого*

$$\|Tx\| \leqslant C\, \|x\| \quad \textit{при всех } x \in \mathbb{R}^n. \tag{8}$$

3) *Если E — ограниченное подмножество $\mathbb{R}^n$, то $T(E)$ ограничено в $\mathbb{R}^m$.*

Доказательство. 1) Для любого $i \in \{1, \dots, n\}$ функция $x \mapsto x_i$ непрерывна как координатная функция тождественного отображения. Тогда в силу теоремы 2 § 5 и равенства (6) все координатные функции T непрерывны. Поэтому непрерывен и сам оператор T.

2) В § 3 показано, что сфера $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$ компактна в $\mathbb{R}^n$. Положим

$$C = \sup_{x \in S} \|Tx\|.$$

Так как оператор T непрерывен, по теореме Вейерштрасса $C < +\infty$ (см. § 5). Покажем, что C удовлетворяет условию (8). Действительно, при $x = 0$ неравенство (8) очевидно. Для $x \neq 0$ положим $z = \dfrac{x}{\|x\|}$. Тогда $z \in S$, и в силу линейности T

$$\|Tx\| = \|T(z\,\|x\|)\| = \|Tz\| \cdot \|x\| \leqslant C\,\|x\|.$$

3) Пусть множество E ограничено в $\mathbb{R}^n$. Тогда существует такое $\delta > 0$, что $E \subset V_0(\delta)$. В силу (8) для любого $x \in E$

$$\|Tx\| \leqslant C\,\|x\| \leqslant C\delta.$$

Поэтому $T(E) \subset V_0(C\delta)$, что и дает ограниченность $T(E)$. $\square$

Замечание. Свойство 3) называется *ограниченностью оператора* T. Не следует путать понятия *ограниченный оператор* и *ограниченное отображение*. Ограниченным отображением является только нулевой оператор. Действительно, если существует $x \in \mathbb{R}^n$, для которого $Tx \neq 0$, то

$$\|T(kx)\| = \|k \cdot Tx\| = k\,\|Tx\| \to +\infty \quad \text{при} \;\; k \to +\infty,$$

то есть множество $T(\mathbb{R}^n)$ не ограничено в $\mathbb{R}^m$.

Теорема 1 позволяет ввести важную характеристику линейного оператора — его *норму*.

Определение 2. Норма оператора. Пусть $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. *Нормой оператора* T называется величина

$$\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \in \mathbb{R}^n \setminus \{0\}\right\}. \tag{9}$$

Замечание 1. Из (8) вытекает, что $\|T\| < +\infty$. Кроме того, справедливо следующее утверждение: *если (8) выполняется при* *всех* $x \in \mathbb{R}^n$*, то* $\|T\| \leqslant C$*, а если для некоторого* $x \in \mathbb{R}^n \setminus \{0\}$ *верно неравенство* $\|Tx\| \geqslant C\,\|x\|$*, то* $\|T\| \geqslant C$.

Замечание 2. *Справедливы следующие равенства*:

$$\|T\| = \sup_{\|x\|\leqslant 1} \|Tx\| = \sup_{\|x\|=1} \|Tx\|\,. \tag{10}$$

Действительно, положим

$$A = \sup_{\|x\|\leqslant 1} \|Tx\|\,, \quad B = \sup_{\|x\|=1} \|Tx\|\,.$$

Тогда

$$\|T\| = \sup_{\|x\|\neq 0} \frac{\|Tx\|}{\|x\|} \geqslant \sup_{0<\|x\|\leqslant 1} \frac{\|Tx\|}{\|x\|} \geqslant \sup_{\|x\|\leqslant 1} \|Tx\| = A.$$

Неравенство $A \geqslant B$ очевидно. Кроме того,

$$\|T\| = \sup_{\|x\|\neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|\neq 0} \left\| T\left(\frac{x}{\|x\|}\right)\right\| \leqslant B,$$

поэтому $A = B = \|T\|$. $\square$

Замечание 3. Пусть $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, $S \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^s)$. Тогда

$$\|ST\| \leqslant \|S\| \cdot \|T\|\,. \tag{11}$$

Действительно, при любых $x \in \mathbb{R}^n$

$$\|ST(x)\| = \|S(Tx)\| \leqslant \|S\| \cdot \|Tx\| \leqslant \|S\| \cdot \|T\| \cdot \|x\|\,,$$

откуда в силу замечания 1 вытекает оценка (11).

В § 1 было доказано, что для евклидовой нормы в $\mathbb{R}^n$ выполнены три свойства: неотрицательность, положительная однородность и неравенство треугольника. Покажем, что этим условиям удовлетворяет и операторная норма, определенная на $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Теорема 2. *Справедливы следующие утверждения.*

1) *Если* $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, *то* $\|T\| \geqslant 0$, *причем равенство достигается только для* $T = \mathbb{O}$.

2) $\|\lambda T\| = |\lambda| \cdot \|T\|$ *при всех* $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ *и* $\lambda \in \mathbb{R}$.

3) $\|T + S\| \leqslant \|T\| + \|S\|$ *для любых* $T, S \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Доказательство. 1) Неравенство $\|T\| \geqslant 0$ очевидно. Если $\|T\| = 0$, то в силу (8)

$$\|Tx\| \leqslant \|T\| \cdot \|x\| = 0 \text{ при всех } x \in \mathbb{R}^n,$$

откуда $Tx = 0$ для любых $x \in \mathbb{R}^n$. Поэтому $T = \mathbb{O}$.

2) В силу (10)

$$\|\lambda T\| = \sup_{\|x\|=1} \|\lambda Tx\| = \sup_{\|x\|=1} |\lambda| \cdot \|Tx\| = |\lambda| \cdot \sup_{\|x\|=1} \|Tx\| = |\lambda| \cdot \|T\| .$$

3) Если $x \in \mathbb{R}^n$, $\|x\| = 1$, то

$$\|(T+S)(x)\| = \|Tx + Sx\| \leqslant \|Tx\| + \|Sx\| \leqslant \|T\| + \|S\| .$$

Поэтому

$$\|T + S\| = \sup_{\|x\|=1} \|(T+S)(x)\| \leqslant \|T\| + \|S\| . \quad \square$$

Следствие. *Для любых* $T, S \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ *справедливо неравенство*

$$\|T - S\| \geqslant \big| \|T\| - \|S\| \big| .$$

Доказывается это утверждение точно так же, как замечание 4 к определению 1 § 1.

С практической точки зрения формулы (10) неудобны. Было бы полезно получить выражение для нормы оператора в терминах его матрицы. Рассмотрим этот вопрос более подробно.

Теорема 3. Оценка нормы оператора. *Пусть* $T : \mathbb{R}^n \to \mathbb{R}^m$ – *линейный оператор с матрицей* $A = \{a_{ij}\}$. *Тогда справедливо неравенство*

$$\|T\| \leqslant \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} \,. \tag{12}$$

Доказательство. В силу замечания 1 к определению нормы достаточно проверить соотношение

$$\|Tx\| \leqslant \|x\| \cdot \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} \quad \text{при всех } x \in \mathbb{R}^n.$$

Пусть $x \in \mathbb{R}^n$. Используя (6) и неравенство Коши, мы получим

$$\|Tx\|^2 = \sum_{i=1}^{m}\big(T_i\,x\big)^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\,x_j\right)^2 \leqslant$$

$$\leqslant \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}^2\right)\cdot\left(\sum_{j=1}^{n} x_j^2\right) = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)\cdot\|x\|^2. \quad \square$$

Замечание 1. Если $n = 1$ или $m = 1$, то матрица A представляет собой соответственно столбец или строку, а правая часть (12) равна евклидовой норме A. Покажем, что в этих случаях в (12) реализуется равенство. Случай $n = 1$ очевиден. При $m = 1$

$$|T(A)| = \left|A \cdot A^T\right| = \|A\|^2 = \|A\|\cdot\|A\|,$$

откуда $\|T\| \geqslant \|A\|$ по замечанию 1 к определению операторной нормы. Напомним, что символом A^T в алгебре обозначают матрицу, транспонированную к A.

Замечание 2. В общем случае равенства в (12) нет. В следующей главе мы увидим, что на самом деле

$$\|T\| = \max\left\{\sqrt{\lambda} : \lambda \text{ — собственное число } A^TA\right\}.$$

Тем не менее для практических целей оценка (12) удобнее.

Рассмотрим теперь *обратимые* линейные операторы. Их свойства нам потребуются в главе 6 при решении задачи о локальной разрешимости систем уравнений.

Лемма 1. *Пусть $T \in \mathcal{L}(\mathbb{R}^n)$ — биективный линейный оператор. Тогда обратное к T отображение также является линейным оператором.*

Напомним, что обратное отображение принято обозначать символом T^{-1}.

Доказательство. Пусть $y, \tilde{y} \in \mathbb{R}^n$ и $\lambda, \mu \in \mathbb{R}$. В силу сюръективности T найдутся $x, \tilde{x} \in \mathbb{R}^n$, для которых $y = Tx$ и $\tilde{y} = T\tilde{x}$. Тогда

$$\lambda y + \mu\tilde{y} = \lambda Tx + \mu T\tilde{x} = T(\lambda x + \mu\tilde{x}),$$

откуда

$$T^{-1}(\lambda y + \mu\tilde{y}) = \lambda x + \mu\tilde{x} = \lambda T^{-1}(y) + \mu T^{-1}(\tilde{y}). \quad \square$$

Определение 3. Обратимый оператор. Пусть $T \in \mathcal{L}(\mathbb{R}^n)$ — биективный линейный оператор. Тогда T называется *обратимым* или *изоморфизмом*, а оператор T^{-1} — *обратным к* T. Множество всех изоморфизмов в $\mathcal{L}(\mathbb{R}^n)$ мы будем обозначать через $\mathcal{I}(\mathbb{R}^n)$.

Замечание 1. Из курса алгебры известно, что биективность оператора $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ возможна лишь при условии $m = n$. Поэтому только такая ситуация рассматривается в лемме 1 и в определении 3.

Замечание 2. Пусть $T \in \mathcal{L}(\mathbb{R}^n)$, A — матрица T. Тогда обратимость T равносильна обратимости A, причем матрицей T^{-1} является A^{-1}. Предлагаем читателю доказать это самостоятельно.

Замечание 3. Если $T, S \in \mathcal{I}(\mathbb{R}^n)$ то $ST \in \mathcal{I}(\mathbb{R}^n)$, причем

$$(ST)^{-1} = T^{-1}S^{-1}.$$

Проверка этого утверждения также предоставляется читателю.

Перейдем к описанию класса $\mathcal{I}(\mathbb{R}^n)$. Напомним вначале читателю условия разрешимости линейных систем, известные из курса алгебры. С матрицей A размером $n \times n$ мы свяжем две системы линейных уравнений относительно x:

$$A \cdot x = y, \quad \text{где } y \text{ — столбец } n \times 1; \tag{13}$$

$$A \cdot x = 0. \tag{14}$$

Системы (13) и (14) называют соответственно *неоднородной* и *однородной*.

Лемма 2. *Пусть A — матрица размером $n \times n$. Тогда равносильны следующие утверждения.*

1) $\det A \neq 0$.

2) *Неоднородная система (13) разрешима при любом $y \in \mathbb{R}^n$.*

3) *Однородная система (14) имеет только нулевое решение.*

Мы не будем приводить доказательство этой леммы.

Теорема 4. Критерий обратимости оператора в $\mathbb{R}^n$. *Пусть $T \in \mathcal{L}(\mathbb{R}^n)$, A — матрица T. Тогда равносильны следующие утверждения.*

1) *T обратим.*

2) *T сюръективен.*

3) *T инъективен.*

4) *$Tx \neq 0$ при всех $x \in \mathbb{R}^n \setminus \{0\}$.*

5) $\det A \neq 0$.

6) *Существует такое $C > 0$, что*

$$\|Tx\| \geqslant C\,\|x\| \quad \text{при всех } x \in \mathbb{R}^n. \tag{15}$$

Доказательство. Проверим, что 1) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 5) $\Rightarrow$ 1). Переход 1) $\Rightarrow$ 3) очевиден. Утверждение 4) можно записать в виде

$$T(x) \neq T(0) \quad \text{для любых } x \in \mathbb{R}^n \setminus \{0\},$$

поэтому оно вытекает из инъективности T. Если выполнено 4), то

$$A \cdot x \neq 0 \quad \text{при } x \in \mathbb{R}^n \setminus \{0\}.$$

Это означает, что однородная система (14) имеет только нулевое решение, и в силу леммы 2 $\det A \neq 0$. Наконец, из условия 5) следует обратимость матрицы A, а тогда и оператор T обратим. Таким образом, утверждения 1), 3), 4), 5) эквивалентны.

Покажем теперь, что 2) равносильно всем этим утверждениям. Действительно, переход 1) $\Rightarrow$ 2) очевиден. Если оператор T сюръективен, то неоднородная система (13) разрешима при любой правой части. Поэтому из леммы 2 следует импликация 2) $\Rightarrow$ 5).

Осталось проверить, что 6) эквивалентно утверждениям 1) – 5). Импликация 6) $\Rightarrow$ 4) очевидна, поскольку при $x \neq 0$

$$\|Tx\| \geqslant C\,\|x\| > 0.$$

Докажем переход 1) $\Rightarrow$ 6). Пусть оператор T обратим. Тогда для любого $x \in \mathbb{R}^n$

$$\|x\| = \left\|T^{-1}(Tx)\right\| \leqslant \left\|T^{-1}\right\| \cdot \|Tx\|. \tag{16}$$

Заметим, что $\left\|T^{-1}\right\| \neq 0$, иначе по теореме 2 оператор T^{-1} был бы нулевым и, значит, необратимым. Поэтому неравенство (15) выполняется с константой $C = \left\|T^{-1}\right\|^{-1}$. $\square$

Замечание 1. Если оператор T обратим, то $\left\|T^{-1}\right\|^{-1}$ — наибольшая из констант C, допустимых в (15). Действительно, оценка (15) верна для $C = \left\|T^{-1}\right\|^{-1}$ в силу (16). С другой стороны, из (15) вытекает, что для любого $x \in \mathbb{R}^n$

$$\left\|T^{-1}x\right\| = C^{-1}C\left\|T^{-1}x\right\| \leqslant C^{-1}\left\|T(T^{-1}x)\right\| = C^{-1}\|x\|.$$

Отсюда в силу замечания 1 к определению операторной нормы $\left\|T^{-1}\right\| \leqslant C^{-1}$, то есть $C \leqslant \left\|T^{-1}\right\|^{-1}$.

Замечание 2. *Если $T_0 \in \mathcal{I}(\mathbb{R}^n)$ и оператор $T \in \mathcal{L}(\mathbb{R}^n)$ удовлетворяет условию $\|T - T_0\| < \left\|T_0^{-1}\right\|^{-1}$, то он обратим.* Действительно, в силу замечания 1 для любого $x \in \mathbb{R}^n$

$$\|Tx\| \geqslant \|T_0x\| - \|(T - T_0)x\| \geqslant \left(\left\|T_0^{-1}\right\|^{-1} - \|T - T_0\|\right) \cdot \|x\|,$$

откуда по теореме 4 вытекает обратимость T.

Доказанное утверждение означает, что все операторы, близкие по норме к обратимому, тоже обратимы. Это свойство называют *открытостью* $\mathcal{I}(\mathbb{R}^n)$ *в* $\mathcal{L}(\mathbb{R}^n)$. Открытые подмножества $\mathcal{L}(\mathbb{R}^n)$ можно определить строго, задав на $\mathcal{L}(\mathbb{R}^n)$ окрестности так же, как мы это делали в $\mathbb{R}^n$.

§ 7. Многочлены нескольких переменных

Ключевым понятием в главе 3 был многочлен Тейлора. Поэтому для построения многомерного дифференциального исчисления нам потребуется изучить многочлены нескольких переменных. Этот параграф состоит из двух частей. Вначале мы определим многочлены в $\mathbb{R}^n$ и приведем различные способы их описания. Для этого будет введено понятие *мультииндекса*, которое окажется полезным и в следующей главе. Затем мы подробно обсудим *квадратичные формы*, играющие важную роль при исследовании функций нескольких переменных на экстремум. Излагаемый здесь материал относится преимущественно к курсу алгебры, тем не менее для удобства читателя мы его приводим.

Многочлены одной переменной выражаются через степенные функции с помощью операций сложения и умножения на числа. Поэтому для описания многочленов в $\mathbb{R}^n$ необходимо ввести многомерные аналоги степенных функций. Определим вначале *мультииндексы*, позволяющие записывать степени векторов в удобной форме.

Определение 1. Мультииндексы. Вектор $\alpha \in \mathbb{Z}_+^n$ называется *мультииндексом размерности* n. *Порядком мультииндекса* α называется величина

$$|\alpha| = \alpha_1 + \ldots + \alpha_n.$$

Множество всех мультииндексов порядка s мы будем обозначать символом M_s .

Замечание. Множество M_s зависит и от n, но в обозначении это не отражается, поскольку значение n ясно из контекста.

Теперь мы можем ввести степенные функции нескольких переменных.

Определение 2. Одночлены в $\mathbb{R}^n$. Пусть $s \in \mathbb{Z}_+$, $\alpha \in M_s$. Тогда функция

$$p(x) = x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n} \quad (x \in \mathbb{R}^n) \tag{17}$$

называется *одночленом в* $\mathbb{R}^n$ *степени* s, *порожденным мультииндексом* α, и обозначается символом x^α. Если при некотором $k \in \{1, \ldots, n\}$ индекс α_k равен нулю, то мы считаем $x_k^{\alpha_k} = 1$ даже в случае $x_k = 0$.

Замечание. Иногда одночлены записывают в другой форме. Положим

$$I_0 = \{\varnothing\}, \quad I_s = \{1, \ldots, n\}^s \ \ (s \in \mathbb{N}).$$

Пусть $s \in \mathbb{N}$, $i \in I_s$. Тогда функция

$$p(x) = x_{i_1} \cdot \ldots \cdot x_{i_s} \quad (x \in \mathbb{R}^n) \tag{18}$$

является одночленом степени s. Действительно, выражение (17) отличается от (18) только тем, что в нем одинаковые множители записаны в виде степеней. Напомним также, что по соглашению главы 1 при $s = 0$ правая часть (18) тождественно равна 1.

Рассмотрим взаимосвязь между (17) и (18) более подробно.

Определение 3. Пусть $s \in \mathbb{Z}_+$, $i \in I_s$. Предположим, что для любого $k \in \{1, \ldots, n\}$ число α_k указывает, сколько раз k встречается в семействе i. Тогда мультииндекс α называется *распределением* i и обозначается через $d(i)$.

Замечание 1. Отображение $i \mapsto d(i)$ действует из I_s в M_s (мы будем его также обозначать через d). Оно сюръективно, но не инъективно: например, при $s = 2$ семейства $(1, 2)$ и $(2, 1)$ имеют одинаковое распределение $(1, 1)$.

Замечание 2. Распределению можно дать следующую интерпретацию. Кассир, принимающий платежи, сортирует купюры по номиналу и раскладывает их по разным ящикам. Мультииндекс, указывающий количество купюр в каждом ящике, и есть распределение выручки.

Пусть $s \in \mathbb{Z}_+$. Тогда для любого $i \in I_s$ справедливо соотношение

$$x_{i_1} \cdot \ldots \cdot x_{i_s} = x^{d(i)} \quad (x \in \mathbb{R}^n).$$

Поэтому переход от (18) к (17) осуществляется однозначно, а от (17) к (18) — нет: любое семейство индексов с распределением α порождает одночлен x^α. Выясним, сколько в I_s существует различных семейств с заданным распределением. Положим

$$\alpha! = \alpha_1! \cdot \ldots \cdot \alpha_n! \quad (\alpha \in \mathbb{Z}_+^n). \tag{19}$$

Лемма 1. *Пусть $s \in \mathbb{Z}_+$, $\alpha \in M_s$. Тогда множество $d^{-1}\bigl(\{\alpha\}\bigr)$ содержит ровно $\dfrac{s!}{\alpha!}$ элементов.*

Иными словами, в I_s существует ровно $\dfrac{s!}{\alpha!}$ различных семейств с распределением α.

Доказательство. Обозначим через C_s^α число элементов множества $d^{-1}\big(\{\alpha\}\big)$. Нам нужно проверить равенство

$$C_s^\alpha = \frac{s!}{\alpha!} \quad \text{при всех } \alpha \in M_s. \tag{20}$$

Докажем его индукцией по s. База индукции очевидна: при $s = 0$

$$C_s^\alpha = 1 = \frac{s!}{\alpha!}\,.$$

Предположим, что для некоторого $s \in \mathbb{Z}_+$ равенство (20) доказано, и проверим его для $s+1$. Пусть $k \in \{1, \dots, n\}$. Если $i \in I_{s+1}$ и $i_{s+1} = k$, то распределением семейства $(i_1, \dots, i_s)$ будет мультииндекс $\beta \in M_s$, определяемый равенствами

$$\beta_k = \alpha_k - 1 \quad \text{и} \quad \beta_l = \alpha_l \ \text{ при } \ l \neq k.$$

По индукционному предположению количество семейств в I_s, имеющих распределение β, равно

$$C_s^\beta = \frac{s!}{\beta!} = \frac{s!}{\alpha!} \cdot \alpha_k.$$

Суммируя эти числа по $k \in \{1, \dots, n\}$, мы получим

$$C_{s+1}^\alpha = \frac{s!}{\alpha!} \sum_{k=1}^{n} \alpha_k = \frac{s!}{\alpha!} \cdot |\alpha| = \frac{s!}{\alpha!}\,(s+1) = \frac{(s+1)!}{\alpha!}. \quad \square$$

Замечание. Числа C_s^α называются *мультиномиальными коэффициентами*. Отметим, что при $n = 2$ мультииндекс $\alpha \in M_s$ имеет вид $(k, s-k)$, где $k \in \{0, \dots, s\}$. Тогда

$$C_s^\alpha = \frac{s!}{k!(s-k)!} = C_s^k,$$

то есть мультиноминальные коэффициенты в этом случае совпадают с биномиальными, введенными в § 2 главы 1.

Определение 4. Многочлены нескольких переменных. Пусть M — конечное множество мультииндексов размерности n, $\{c_\alpha\}_{\alpha\in M}$ — семейство элементов $\mathbb{R}^m$. Тогда отображение

$$p(x)=\sum_{\alpha\in M} c_\alpha\, x^\alpha \quad (x\in\mathbb{R}^n)$$

называется *многочленом n переменных со значениями в* $\mathbb{R}^m$. Если $p\not\equiv 0$, то *степенью* p называется число

$$\deg p=\max\bigl\{|\alpha| : \alpha\in M,\ c_\alpha\neq 0\bigr\}.$$

В противном случае мы полагаем степень p равной $-\infty$.

Замечание 1. Если степень p равна s, мы можем не умаляя общности считать, что M есть множество *всех* мультииндексов порядка не выше s, то есть

$$p(x)=\sum_{|\alpha|\leqslant s} c_\alpha\, x^\alpha. \tag{21}$$

Замечание 2. *Многочлены нескольких переменных непрерывны на* $\mathbb{R}^n$. Действительно, воспользуемся теоремой 2 § 5. Любой одночлен непрерывен на $\mathbb{R}^n$, поскольку он либо равен 1, либо является произведением линейных функций вида $x\mapsto x_k$, где $k\in\{1,\dots,n\}$. Осталось заметить, что многочлены получаются из одночленов с помощью операций сложения и умножения на векторы. $\square$

Замечание 3. *Если p — многочлен степени s, $a\in\mathbb{R}^n$, то и отображение $p(\cdot+a)$ будет многочленом степени s.* Действительно, положим $q=p(\cdot+a)$. Очевидно, что для любого мультииндекса α функция $x\mapsto(x+a)^\alpha$ есть многочлен степени не выше s (достаточно записать $(x+a)^\alpha$ по формуле (17) и раскрыть скобки). В силу (21) q также будет многочленом степени не выше s. Так как $p=q(\cdot-a)$, по доказанному степень p не больше степени q, то есть $\deg p=\deg q$. $\square$

Замечание 4. Если $s \in \mathbb{Z}_+$ и $a \in \mathbb{R}^n$, то любой многочлен p степени s, заданный на $\mathbb{R}^n$, можно записать в виде

$$p(x) = \sum_{|\alpha| \leqslant s} c_\alpha \, (x-a)^\alpha. \tag{22}$$

Действительно, по замечанию 3 отображение $z \mapsto p(z+a)$ есть многочлен степени s. Выражая его по формуле (21) и подставляя $z = x - a$, мы получим (22). Равенство (22) называется *разложением p по степеням $x - a$*.

Замечание 5. *Любой многочлен выражается через одночлены единственным образом.* Иными словами, коэффициенты c_α в (21) однозначно определяются по многочлену p. Это утверждение мы докажем в § 5 главы 6 с помощью частных производных.

В дифференциальном исчислении важное значение будут иметь многочлены специального вида, называемые *однородными*. Перейдем к их рассмотрению.

Определение 5. Однородные многочлены. Пусть $s \in \mathbb{Z}_+$, $\{c_\alpha\}_{\alpha \in M_s}$ — семейство элементов $\mathbb{R}^m$. Тогда многочлен p, действующий по формуле

$$p(x) = \sum_{\alpha \in M_s} c_\alpha x^\alpha \quad (x \in \mathbb{R}^n), \tag{23}$$

называется *однородным степени s* или *s-формой*.

Замечание 1. Не следует путать термины "степень однородности" и "степень многочлена": многочлен $p \equiv 0$ является однородным любой степени, а $\deg p = -\infty$. Во всех остальных случаях эти понятия совпадают.

Замечание 2. Однородный многочлен степени 2 называют *квадратичной формой*. Отметим также, что 0-форма есть постоянное отображение, а 1-форма — линейный оператор.

Замечание 3. Если $s \in \mathbb{Z}_+$ и p есть s-форма, то p удовлетворяет условию *однородности степени s*:

$$p(tx) = t^s \, p(x) \quad \text{для любых } x \in \mathbb{R}^n \text{ и } t \in \mathbb{R}. \tag{24}$$

Действительно, для $s \in \mathbb{N}$ в силу (17)

$$p(tx) = \sum_{\alpha \in M_s} c_\alpha \cdot (tx_1)^{\alpha_1} \cdot \ldots \cdot (tx_n)^{\alpha_n} = t^{\alpha_1 + \cdots + \alpha_n} p(x) = t^s\, p(x).$$

Несложно доказать, что верно и обратное: многочлен, удовлетворяющий условию (24), можно записать в виде (23). Эти рассуждения объясняют смысл термина “однородный многочлен”.

Замечание 4. Группируя в (21) слагаемые, соответствующие мультииндексам одного порядка, мы получим

$$p(x) = \sum_{k=0}^{s} p_k(x), \quad \text{где } \ p_k(x) = \sum_{\alpha \in M_k} c_\alpha\, x^\alpha.$$

Таким образом, любой многочлен p степени s представи́м в виде суммы k-форм p_k по $k \in \{0, \ldots, s\}$. Такое разложение единственно, поскольку коэффициенты c_α однозначно определяются многочленом p. Формы p_k называют *однородными компонентами* p.

Пусть $s \in \mathbb{Z}_+$, p — s-форма. Записывая одночлены по формуле (18), мы можем представить p в виде

$$p(x) = \sum_{i \in I_s} c_i \cdot x_{i_1} \cdot \ldots \cdot x_{i_s} \quad (x \in \mathbb{R}^n), \tag{25}$$

где $\{c_i\}_{i \in I_s}$ — семейство векторов из $\mathbb{R}^m$. Выражение, стоящее в правой части (25), содержит много подобных слагаемых. Тем не менее, оно обладает симметрией и бывает полезным для проведения преобразований. Коэффициенты при одночленах в такой записи уже не определяются многочленом p однозначно: например, выражения $2x_1x_2$ и $x_1x_2 + x_2x_1$ задают один и тот же многочлен в $\mathbb{R}^2$. Наложим на векторы c_i в (25) следующее *условие симметричности*:

$$\text{если } \ i, j \in I_s,\ d(i) = d(j), \ \text{ то } \ c_i = c_j. \tag{26}$$

Оно означает, что коэффициенты при одинаковых одночленах в правой части (25) совпадают. Иными словами, вектор c_i зависит не от самого семейства i, а только от его распределения. Поэтому будет корректным использовать для c_i также запись $c_{d(i)}$.

Теорема 1. *Пусть $s \in \mathbb{Z}_+$, p — s-форма. Тогда существует единственное семейство $\{c_i\}$, удовлетворяющее (25) и (26), и верно равенство*

$$p(x) = \sum_{\alpha \in M_s} \frac{s!}{\alpha!} c_\alpha x^\alpha \quad (x \in \mathbb{R}^n), \tag{27}$$

где c_α — общее значение коэффициентов c_i при $d(i) = \alpha$.

Доказательство. Многочлен p единственным образом записывается в виде

$$p(x) = \sum_{|\alpha|=s} C_\alpha x^\alpha \quad (x \in \mathbb{R}^n).$$

Пусть семейство $\{c_i\}$ удовлетворяет условию (25). Приведем в (25) подобные слагаемые:

$$\sum_{i \in I_s} c_i \cdot x_{i_1} \cdot \ldots \cdot x_{i_s} = \sum_{|\alpha|=s} \sum_{i\colon d(i)=\alpha} c_i x^\alpha = \sum_{|\alpha|=s} x^\alpha \sum_{i\colon d(i)=\alpha} c_i.$$

Приравнивая коэффициенты при x^α, мы получим

$$C_\alpha = \sum_{i\colon d(i)=\alpha} c_i \quad (\alpha \in M_s).$$

По лемме 1 сумма в правой части содержит ровно $\dfrac{s!}{\alpha!}$ слагаемых. Условие (26) означает, что все они равны, то есть

$$C_\alpha = \frac{s!}{\alpha!} c_i, \quad \text{если } d(i) = \alpha.$$

Из этих формул коэффициенты c_i однозначно находятся. Осталось заметить, что

$$p(x) = \sum_{|\alpha|=s} C_\alpha x^\alpha = \sum_{\alpha \in M_s} \frac{s!}{\alpha!} c_\alpha x^\alpha. \quad \square$$

Обобщим теперь бином Ньютона, доказанный в § 2 главы 1, на случай нескольких слагаемых.

Теорема 2. Полиномиальная формула Ньютона. *Пусть $s \in \mathbb{Z}_+$, $x \in \mathbb{R}^n$. Тогда справедливо равенство*

$$(x_1 + \ldots + x_n)^s = \sum_{\alpha \in M_s} \frac{s!}{\alpha!} \, x^\alpha.$$

Доказательство. Случай $s = 0$ тривиален. Для $s \in \mathbb{N}$ положим

$$p(x) = (x_1 + \ldots + x_n)^s \quad (x \in \mathbb{R}^n).$$

Раскрывая в правой части скобки, мы получим

$$p(x) = \sum_{i \in I_s} x_{i_1} \cdot \ldots \cdot x_{i_s}.$$

Таким образом, p является однородным многочленом степени s, у которого все коэффициенты в разложении (25) равны 1. Поэтому утверждение теоремы вытекает из равенства (27). $\square$

Замечание 1. Как уже отмечалось ранее, при $n = 2$ мультиномиальные коэффициенты превращаются в биномиальные. Предлагаем читателю вывести бином Ньютона из теоремы 2.

Замечание 2. Теорему 2 можно получить как следствие многомерной формулы Тейлора. Это будет показано в § 6 главы 6.

Следствие. Оценка однородного многочлена. *Пусть однородный многочлен p степени $s \in \mathbb{Z}_+$ задается формулой (27). Тогда*

$$\|p(x)\| \leqslant \max_{\alpha \in M_s} \|c_\alpha\| \cdot \left(\sqrt{n}\right)^s \cdot \|x\|^s \quad \text{для всех } x \in \mathbb{R}^n. \tag{28}$$

Доказательство. Достаточно рассмотреть случай $s \in \mathbb{N}$. Положим $C = \max\limits_{\alpha \in M_s} \|c_\alpha\|$. Тогда в силу теоремы 2 и неравенства Коши

$$\|p(x)\| \leqslant \sum_{\alpha \in M_s} \frac{s!}{\alpha!} \, C \cdot |x_1|^{\alpha_1} \cdot \ldots \cdot |x_n|^{\alpha_n} = C\big(|x_1| + \ldots + |x_n|\big)^s \leqslant$$

$$\leqslant C\left(\sqrt{n}\right)^s \cdot \left(\sqrt{\sum_{k=1}^n x_k^2}\right)^s = C\left(\sqrt{n}\right)^s \|x\|^s. \quad \square$$

В заключение изучим более подробно квадратичные формы со значениями в $\mathbb{R}$, поскольку они будут иметь важное значение в следующей главе при изучении экстремума функций нескольких переменных. Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$ — квадратичная форма. Записывая ее в виде (25), мы получим

$$f(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, x_i x_j \quad (x \in \mathbb{R}^n). \tag{29}$$

Условие (26) в этом случае эквивалентно

$$a_{ij} = a_{ji} \quad \text{для любых } \ i, j \in \{1, \ldots, n\}.$$

Оно означает, что *матрица* $A = \{a_{ij}\}_{i,j=1}^{n}$ *симметрична*. Назовем A *матрицей квадратичной формы*. Если записывать элементы $\mathbb{R}^n$ в виде столбцов, то можно получить выражение для f в матричном виде:

$$f(x) = x^T A x \quad (x \in \mathbb{R}^n).$$

Заметим, что слагаемые в правой части (29) не меняются при перестановке i и j. Поэтому для вычислительных целей удобнее другая формула:

$$f(x) = \sum_{i=1}^{n} a_{ii}\, x_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} a_{ij}\, x_i x_j \quad (x \in \mathbb{R}^n). \tag{30}$$

Она получается из (29), если разбить правую часть (29) на суммы по $i = j$, $i < j$, $i > j$ и заметить, что две последние суммы одинаковы.

Определение 6. Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$ — квадратичная форма. Она называется *положительно определенной*, если $f(x) > 0$ при всех $x \in \mathbb{R}^n \setminus \{0\}$, и *отрицательно определенной*, если $f(x) < 0$ для любых $x \in \mathbb{R}^n \setminus \{0\}$.

Замечание. Не следует путать *положительную определенность* и *неотрицательность* квадратичной формы. Рассмотрим, например, форму $f(x) = (x_1 + \ldots + x_n)^2$. Очевидно, $f(x) \geqslant 0$ для любых $x \in \mathbb{R}^n$. Тем не менее, f не будет положительно определенной, поскольку $f(1, -1, 0, \dots, 0) = 0$. Аналогичное замечание верно и для отрицательной определенности.

Покажем, что положительно и отрицательно определенные формы допускают оценку через квадрат нормы аргумента.

Теорема 3. *Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$ — квадратичная форма. Тогда справедливы следующие утверждения.*

1) *Если f положительно определена, то существует такое $C > 0$, что*

$$f(x) \geqslant C\ \|x\|^2 \quad \text{при всех } x \in \mathbb{R}^n.$$

2) *Если f отрицательно определена, то существует такое $C > 0$, что*

$$f(x) \leqslant -C\ \|x\|^2 \quad \text{при всех } x \in \mathbb{R}^n.$$

Замечание. Утверждения, обратные к 1) и 2), очевидно, тоже верны.

Доказательство. 1) В § 3 было показано, что единичная сфера в $\mathbb{R}^n$ с центром в нуле компактна. По теореме Вейерштрасса f достигает на сфере своего наименьшего значения, которое мы и обозначим через C. Покажем, что число C искомое. Действительно, $C > 0$ в силу положительной определенности f. Если $x \in \mathbb{R}^n \setminus \{0\}$, то по свойству однородности (24)

$$\frac{f(x)}{\|x\|^2} = f\left(\frac{x}{\|x\|}\right) \geqslant \min_{\|z\|=1} f(z) = C,$$

а для $x = 0$ утверждение очевидно.

2) Если форма f отрицательно определена, то форма $-f$ положительно определена. В силу утверждения 1) найдется такое $C > 0$, что $-f(x) \geqslant C\ \|x\|^2$ при всех $x \in \mathbb{R}^n$, откуда и вытекает требуемая оценка. $\square$

При $n = 1$ формула (29) примет вид $f(x) = ax^2$, где $a \in \mathbb{R}$, и положительная определенность f будет равносильна неравенству

$a > 0$. Каков аналог этого условия в многомерной ситуации? Сформулируем критерий положительной определенности формы f в терминах ее матрицы A. Положим

$$\Delta_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix}, \quad \text{где } k \in \{1, \dots, n\}. \tag{31}$$

Определители Δ_k называют *главными минорами матрицы* A.

Теорема 4. Критерий Сильвестра. *Пусть f — квадратичная форма с матрицей A. Тогда равносильны следующие утверждения.*

1) *Форма f положительно определена.*

2) *Все главные миноры* (31) *матрицы A положительны:*

$$\Delta_k > 0 \quad \text{при любом } k \in \{1, \dots, n\}.$$

Мы не будем приводить доказательства этой теоремы.

Следствие. Критерий отрицательной определенности формы. *Пусть f — квадратичная форма с матрицей A. Тогда равносильны следующие утверждения.*

1) *Форма f отрицательно определена.*

2) *При любом $k \in \{1, \dots, n\}$ главный минор Δ_k матрицы A имеет знак $(-1)^k$, то есть*

$$(-1)^k \Delta_k > 0 \quad \text{для всех } k \in \{1, \dots, n\}.$$

Доказательство. Заметим, что форма $-f$ положительно определена. По теореме 4 это равносильно положительности всех главных миноров матрицы $-A$, которые равны

$$\begin{vmatrix} -a_{11} & \dots & -a_{1k} \\ \dots & \dots & \dots \\ -a_{k1} & \dots & -a_{kk} \end{vmatrix} = (-1)^k \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} = (-1)^k \Delta_k,$$

где $k \in \{1, \dots, n\}$. $\square$

ГЛАВА 6. ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ОТОБРАЖЕНИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 1. Введение

Основная идея одномерного дифференциального исчисления, изложенного в главе 3, состояла в том, что вблизи некоторой точки функция с высокой точностью приближалась многочленом:

$$f(x) = T_{a,n}f(x) + o\big((x-a)^n\big) \quad (x \to a),$$

где $T_{a,n}f$ — многочлен степени не выше n, который назывался *многочленом Тейлора* f. Наиболее важный случай $n = 1$ приводил нас к понятию *дифференцируемости*. Многомерное дифференциальное исчисление основано на той же идее, но технически оно сложнее. Например, при изучении дифференцируемых отображений аналогом линейной функции будет линейный оператор, а дифференциалы старших порядков оказываются однородными многочленами. Поэтому мы будем использовать теорию, развитую в параграфах 6 и 7 главы 5.

Как и в предыдущей главе, объектом нашего рассмотрения будут *отображения нескольких переменных*, имеющие вид

$$f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m.$$

Здесь $n, m \in \mathbb{N}$, что будет всегда предполагаться и в дальнейшем. Напомним, что при $m = 1$ такие отображения называются *функциями нескольких переменных*, а при $n = 1$ — *вектор-функциями*.

Материал главы 6 можно разбить на три части. Вначале мы введем для отображений нескольких переменных базовые понятия дифференциального исчисления: *частные производные* и *дифференциалы* различных порядков, *многочлен Тейлора*, а также рассмотрим классы отображений, связанные с дифференцируемостью. Основными результатами этого раздела будут правила дифференцирования и *формула Тейлора*, которую мы выведем в двух вариантах — локальном и глобальном. Вторая часть главы посвящена вопросам разрешимости уравнений и систем. Здесь будут доказаны две ключевые теоремы — *о локальной обратимости гладкого отображения* и *о неявном отображении*. В качестве приложения этих результатов мы приведем различные способы описания гладких поверхностей в $\mathbb{R}^n$. В заключительной части главы 6 рассматривается понятие *экстремума*. Здесь будут выведены правила нахождения и исследования точек экстремума функций нескольких переменных. Кроме того, мы определим и изучим *условный экстремум*, то есть экстремум функции, аргументы которой удовлетворяют дополнительным *уравнениям связи*. Это понятие, имеющее важное значение в геометрии и механике, теряет смысл для функций одной переменной и потому не рассматривалось в главе 3. В заключение мы обсудим задачу о наибольшем и наименьшем значении функции на компактном подмножестве $\mathbb{R}^n$.

§ 2. Дифференцируемость отображения в точке

В этом параграфе мы определим для отображений нескольких переменных дифференцируемость в точке и обсудим связанные с ней понятия — *дифференциал*, *матрица Якоби*, *градиент*. Кроме того, мы введем для таких отображений *частные производные* и выразим через них дифференциал и матрицу Якоби отображения. В заключение будут разобраны примеры исследования функций нескольких переменных на дифференцируемость.

В дифференциальном исчислении локальную малость отображения принято оценивать в терминах степенных функций. Дадим соответствующее определение для отображений нескольких переменных.

Определение 1. Символы Ландау. Пусть $E \subset \mathbb{R}^n$, точка $a \in \mathbb{R}^n$ является предельной для E, $f\colon E \to \mathbb{R}^m$, $N \in \mathbb{N}$. Положим $\alpha(x) = f(x) \cdot \|x-a\|^{-N}$.

1) Если отображение $\alpha(x)$ ограничено в некоторой проколотой окрестности $\dot{V}_a$ точки a, то говорят, что

$$f(x) = O\big(\|x-a\|^N\big) \quad \text{при } x \to a.$$

2) Если $\lim\limits_{x\to a} \alpha(x) = 0$, то говорят, что

$$f(x) = o\big(\|x-a\|^N\big) \quad \text{при } x \to a.$$

Замечание 1. Пусть $a \in E$. Хотя в определении символа o значение $\alpha(a)$ несущественно, удобно полагать $\alpha(a) = 0$, то есть считать отображение α непрерывным в точке a. В дальнейшем мы всегда будем это делать.

Замечание 2. Если $N = 1$, то вместо $o(\|x-a\|)$ мы будем писать просто $o(x-a)$, как в случае функций одной переменной. Это замечание относится и к символу O.

Теперь мы можем ввести понятие дифференцируемости отображения в точке.

Определение 2. Дифференцируемость в точке. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$. Отображение f называется *дифференцируемым в точке* a, если существует линейный оператор $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, для которого

$$f(x) = f(a) + T(x-a) + o(x-a) \quad \text{при } x \to a. \tag{1}$$

Оператор T называется *дифференциалом* f *в точке* a и обозначается символом $d_a f$.

Замечание 1. Формулу (1) часто записывают с помощью приращения аргумента f:

$$f(a+h) = f(a) + d_a f(h) + o(h) \quad (h \to 0) \tag{2}$$

или, более подробно,

$$f(a+h) = f(a) + d_a f(h) + \alpha(h) \cdot \|h\|, \tag{3}$$

где отображение α непрерывно в нуле и $\alpha(0) = 0$.

Замечание 2. В случае $n = 1$ оператор $d_a f$ представляет собой *линейную вектор-функцию*, то есть отображение вида $h \mapsto k \cdot h$, где $k \in \mathbb{R}^m$. Коэффициент k называется *производной* f *в точке* a и обозначается символом $f'(a)$. При $m = 1$ это согласуется с определением 1 § 2 главы 3. Отметим также, что если E есть отрезок, то определение 2 можно распространить и на концевые точки E.

Замечание 3. *Из дифференцируемости отображения* f *в точке* a *вытекает его непрерывность в точке* a. Действительно, по формуле (3)

$$f(a+h) - f(a) = d_a f(h) + \alpha(h) \cdot \|h\| \to 0 \quad (h \to 0)$$

в силу непрерывности $d_a f$ и теоремы 2 § 4 главы 5. $\square$

Обратить замечание 3 нельзя даже для функций одной переменной, что было показано в главе 3.

Приведем два простых, но полезных примера дифференцируемых отображений.

Пример 1. *Если отображение $f\colon \mathbb{R}^n \to \mathbb{R}^m$ постоянно, то оно дифференцируемо в любой точке $a \in \mathbb{R}^n$, причем $d_a f = \mathbb{O}$.* Действительно, для любого $h \in \mathbb{R}^n$

$$f(a+h) = f(a) = f(a) + \mathbb{O}(h),$$

и мы получаем (3) с $d_a f = \mathbb{O}$ и $\alpha \equiv 0$. $\square$

Пример 2. *Если $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, то f дифференцируемо в любой точке $a \in \mathbb{R}^n$, причем $d_a f = f$.* Действительно, для любого $h \in \mathbb{R}^n$

$$f(a+h) - f(a) = f(a+h-a) = f(h),$$

и мы получаем (3) с $d_a f = f$ и $\alpha \equiv 0$. $\square$

Введем теперь понятия производной по вектору и производной по направлению.

Определение 3. Производная по вектору. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$, $e \in \mathbb{R}^n \setminus \{0\}$.

1) Предел $\lim\limits_{t\to 0} \dfrac{f(a+te) - f(a)}{t}$ называется *производной f по вектору e в точке a* и обозначается $\dfrac{\partial f}{\partial e}(a)$.

2) Если дополнительно предположить, что $\|e\| = 1$, то $\dfrac{\partial f}{\partial e}(a)$ называется *производной f по направлению e* или *производной f в направлении e* в точке a.

В дальнейшем, говоря о производной по вектору, мы будем всегда предполагать ее *конечной*.

Замечание 1. Производная f по направлению e имеет следующий смысл. Сужение f на прямую, проходящую через точку a вдоль вектора e, есть вектор-функция $g(t) = f(a+te)$. Производная g в нуле и есть $\dfrac{\partial f}{\partial e}(a)$. Отметим, что некоторая окрестность $V_a(\delta)$

точки a содержится в E, поэтому g определена по крайней мере на интервале $(-\delta, \delta)$.

Замечание 2. Для любого $e \in \mathbb{R}^n \setminus \{0\}$ справедливо равенство

$$\frac{\partial f}{\partial e}(a) = \left(\frac{\partial f_1}{\partial e}(a), \dots, \frac{\partial f_m}{\partial e}(a)\right),$$

которое следует понимать так: если имеет смысл одна из его частей, то определена и другая, и они равны. Доказательство сводится к утверждению о связи предела отображения с пределами его координатных функций (следствие 2 теоремы 1 § 4 главы 5).

Покажем, что дифференцируемое в точке отображение имеет в ней производную по любому ненулевому вектору.

Теорема 1. Вычисление производной по вектору. *Пусть* $E \subset \mathbb{R}^n$, a *— внутренняя точка* E, *отображение* $f\colon E \to \mathbb{R}^m$ *дифференцируемо в точке* a, $e \in \mathbb{R}^n \setminus \{0\}$. *Тогда*

$$\frac{\partial f}{\partial e}(a) = d_a f(e). \tag{4}$$

Доказательство. Обозначим $T = d_a f$. В силу (3)

$$f(a+h) - f(a) = Th + \alpha(h) \cdot \|h\|,$$

где отображение α непрерывно в нуле и $\alpha(0) = 0$. Полагая $h = te$ при $t \neq 0$, мы получим

$$\frac{f(a+te) - f(a)}{t} = \frac{T(te)}{t} + \frac{\alpha(te)\,\|te\|}{t} = T(e) + \frac{|t|}{t}\,\alpha(te) \cdot \|e\|.$$

Заметим, что

$$\left\|\frac{|t|}{t}\,\alpha(te)\right\| = \|\alpha(te)\| \to 0 \quad (t \to 0)$$

по теореме о пределе композиции (см. § 4 главы 5). Поэтому

$$\frac{\partial f}{\partial e}(a) = \lim_{t \to 0} \frac{f(a+te) - f(a)}{t} = T(e) = d_a f(e). \quad \square$$

Следствие. Единственность дифференциала. *Если отображение f дифференцируемо в точке a, то его дифференциал $d_a f$ определен однозначно.*

Доказательство. Пусть $e \in \mathbb{R}^n \setminus \{0\}$. В силу единственности предела (см. § 4 главы 5) производная $\dfrac{\partial f}{\partial e}(a)$ определена однозначно, и по формуле (4) значение $d_a f(e)$ также единственно. Осталось заметить, что $d_a f(0) = 0$. $\square$

Рассмотрим теперь производные f в направлении координатных осей, то есть по ортам $e^1, \ldots, e^n$, определенным формулой (7) главы 5. Это приведет нас к следующему важному понятию.

Определение 4. Частные производные. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$, $k \in \{1, \ldots, n\}$. Тогда $\dfrac{\partial f}{\partial e^k}(a)$ называется *частной производной f по k-й переменной в точке a* и обозначается $\dfrac{\partial f}{\partial x_k}(a)$ (а также $D_k f(a)$, $\partial_k f(a)$ или $f'_{x_k}(a)$).

Замечание 1. Для производной функции одной переменной использовалась запись $\dfrac{df}{dx}(a)$. Если f зависит от более чем одной переменной, то принято писать символ ∂ вместо d.

Замечание 2. Правило вычисления частных производных состоит в следующем. Для $k \in \{1, \ldots, n\}$ мы рассматриваем *вектор-функцию*

$$\varphi_k(t) = f(a_1, \ldots, a_{k-1}, t, a_{k+1}, \ldots, a_n).$$

Тогда $\dfrac{\partial f}{\partial x_k}(a) = \varphi'_k(a_k)$. Таким образом, мы фиксируем у f все аргументы, кроме k-го, а по k-й переменной вычисляем производную. Эти рассуждения объясняют смысл термина "частная производная".

Замечание 3. Для любого $k \in \{1, \ldots, n\}$ справедливо равенство

$$\frac{\partial f}{\partial x_k}(a) = \left(\frac{\partial f_1}{\partial x_k}(a), \ldots, \frac{\partial f_m}{\partial x_k}(a)\right),$$

если хотя бы одна из его частей имеет смысл. Это непосредственно вытекает из замечания 2 к определению 3.

Покажем, как дифференциал и производная по вектору дифференцируемого отображения выражаются через его частные производные.

Теорема 2. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, отображение $f \colon E \to \mathbb{R}^m$ дифференцируемо в точке a. Тогда*

$$d_a f(h) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) \cdot h_k \quad \textit{при всех } \ h \in \mathbb{R}^n; \tag{5}$$

$$\frac{\partial f}{\partial e}(a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) \cdot e_k \quad \textit{при всех } \ e \in \mathbb{R}^n \setminus \{0\}. \tag{6}$$

Доказательство. В силу теоремы 1 формулы (5) и (6) равносильны, поэтому достаточно проверить (5). Для любого $h \in \mathbb{R}^n$

$$d_a f(h) = d_a f\left(\sum_{k=1}^{n} h_k\, e^k\right) = \sum_{k=1}^{n} d_a f\big(e^k\big) \cdot h_k = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) \cdot h_k. \quad \square$$

Замечание 1. В литературе аргумент $d_a f$ часто обозначают символом dx. В этом случае формула (5) принимает вид

$$d_a f(dx) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) \cdot dx_k. \tag{7}$$

Запись (7) удобна в ситуации, когда отображение зависит от параметров и нужно явно указать, по каким переменным берется приращение.

Замечание 2. Правые части формул (5) и (6) имеют смысл всегда, когда у f есть частные производные по всем переменным в точке a. Тем не менее, ниже будет показано, что это условие не гарантирует дифференцируемости f в точке a и даже существования производной f в точке a по произвольному направлению.

Любой линейный оператор, действующий из $\mathbb{R}^n$ в $\mathbb{R}^m$, порождается единственной матрицей по формуле (6) главы 5. Это верно и для дифференциала f. Рассмотрим его матрицу более подробно.

Определение 5. Матрица Якоби, градиент. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, отображение $f\colon E \to \mathbb{R}^m$ дифференцируемо в точке a.

1) Матрица оператора $d_a f$ называется *матрицей Якоби f в точке a* и обозначается $f'(a)$.

2) Если $m = 1$, то матрица $f'(a)$ имеет размер $1 \times n$, то есть является вектором из $\mathbb{R}^n$. Этот вектор называют *градиентом f в точке a* и обозначают $\operatorname{grad} f(a)$ или $\nabla f(a)$ (вторая запись читается "набла эф от a"). Оператор ∇ называют *символом Гамильтона*.

Замечание 1. Если записывать векторы из $\mathbb{R}^n$ и $\mathbb{R}^m$ в виде столбцов, то равенство (2) примет вид

$$f(a+h) = f(a) + f'(a) \cdot h + o(h) \quad (h \to 0). \tag{8}$$

Замечание 2. В случае $n = 1$ символ $f'(a)$ использовался ранее для обозначения производной вектор-функции f в точке a. Если записывать эту производную как столбец, то она будет совпадать с матрицей Якоби f в точке a. Таким образом, новая трактовка символа $f'(a)$ согласуется со старой.

Замечание 3. В случае $m = 1$ выражение $f'(a) \cdot h$ можно понимать и как произведение матриц, и как скалярное умножение $\operatorname{grad} f(a)$ на h. Поскольку результаты этих операций одинаковы, такая запись корректна.

Покажем, что градиент и матрица Якоби дифференцируемого отображения выражаются через его частные производные. Для этого нам потребуется утверждение, устанавливающее связь между дифференцируемостью отображения и его координатных функций.

Лемма 1. Дифференцируемость координатных функций. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$, $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Тогда равносильны следующие утверждения.*

1) *Отображение f дифференцируемо в точке a и $d_a f = T$.*

2) *Все координатные функции f дифференцируемы в точке a, причем $d_a f_k = T_k$ при любом $k \in \{1, \dots, m\}$.*

Доказательство. Первое утверждение эквивалентно соотношению

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Th}{\|h\|} = 0,$$

а второе означает, что для любого $k \in \{1, \ldots, m\}$

$$\lim_{h \to 0} \frac{f_k(a+h) - f_k(a) - T_k h}{\|h\|} = 0.$$

Поэтому равносильность 1) и 2) вытекает из утверждения о связи предела отображения с пределами его координатных функций (следствие 2 теоремы 1 § 4 главы 5). $\square$

Теорема 3. Описание матрицы Якоби. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, отображение $f \colon E \to \mathbb{R}^m$ дифференцируемо в точке a. Тогда*

$$f'(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \ldots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \ldots & \ldots & \ldots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \ldots & \dfrac{\partial f_m}{\partial x_n}(a) \end{pmatrix}. \tag{9}$$

В частности, при $m = 1$ справедливо равенство

$$\operatorname{grad} f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a) \right). \tag{10}$$

Таким образом, строками матрицы Якоби f являются градиенты координатных функций f.

Доказательство. Положим $T = d_a f$. В § 6 главы 5 мы показали, что элемент $f'(a)$ с индексами (i, j) равен $T_i\big(e^j\big)$. По лемме 1 $T_i = d_a f_i$, откуда

$$T_i\big(e^j\big) = d_a f_i\big(e^j\big) = \frac{\partial f_i}{\partial x_j}(a). \quad \square$$

Обсудим теперь геометрические свойства градиента функции нескольких переменных. Пусть $f \colon E \subset \mathbb{R}^n \to \mathbb{R}$, a — внутренняя точка E, функция f дифференцируема в точке a. Рассмотрим множество

$$\big\{(x, y) : x \in \mathbb{R}^n,\ y = f(a) + \operatorname{grad} f(a) \cdot (x - a)\big\}.$$

В силу (10) это множество можно записать в виде

$$\left\{(x, y) : x \in \mathbb{R}^n,\ y = f(a) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) \cdot (x_k - a_k)\right\}. \tag{11}$$

Если $n = 1$, то (11) задает касательную прямую к графику f (см. § 2 главы 3). При $n = 2$ множество (11) определяет плоскость в $\mathbb{R}^3$, которая называется *касательной плоскостью к графику* f. В общем случае уравнение первой степени от $n + 1$ неизвестных описывает подмножество $\mathbb{R}^{n+1}$, называемое *гиперплоскостью*. Таким образом, множество (11) задает гиперплоскость, которая называется *касательной гиперплоскостью к графику* f.

Установим теперь одну важную геометрическую характеристику градиента функции.

Теорема 4. Экстремальное свойство градиента. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, функция $f\colon E \to \mathbb{R}$ дифференцируема в точке a, причем* $\operatorname{grad} f(a) \neq 0$. *Тогда для любого вектора $e \in \mathbb{R}^n$ единичной нормы*

$$\left|\frac{\partial f}{\partial e}(a)\right| \leqslant \|\operatorname{grad} f(a)\|,$$

причем равенство реализуется только для $e = \pm\dfrac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|}$.

Таким образом, вектор $\operatorname{grad} f(a)$ указывает *направление скорейшего изменения f в точке a*, а его длина равна модулю производной f по этому направлению в точке a.

Доказательство. Пусть $e \in \mathbb{R}^n$, $\|e\| = 1$. Тогда по неравенству Коши в $\mathbb{R}^n$

$$\left|\frac{\partial f}{\partial e}(a)\right| = \left|d_a f(e)\right| = \left|\operatorname{grad} f(a) \cdot e\right| \leqslant \|\operatorname{grad} f(a)\|.$$

Равенство реализуется лишь в случае коллинеарности $\operatorname{grad} f(a)$ и e, то есть для $e = \pm\dfrac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|}$. □

В заключение этого параграфа приведем несколько примеров исследования функций двух переменных на дифференцируемость.

Пример 1. *Пусть*

$$f(x,y) = \left(x^2+y^2\right)\,\sin\frac{1}{x^2+y^2} \quad \textit{при} \;\; x^2+y^2 \neq 0, \qquad f(0,0)=0.$$

Тогда функция f дифференцируема в точке $a=(0,0)$, причем производные $\dfrac{\partial f}{\partial x}$ и $\dfrac{\partial f}{\partial y}$ определены на $\mathbb{R}^2$ и разрывны в точке a.

Положим

$$\alpha(x,y) = \sqrt{x^2+y^2}\,\sin\frac{1}{x^2+y^2} \;\; \text{при} \;\; x^2+y^2 \neq 0.$$

Тогда

$$|\alpha(h)| = \|h\|\cdot\left|\sin\frac{1}{\|h\|^2}\right| \leqslant \|h\| \to 0 \quad \big(h\to(0,0)\big).$$

Таким образом,

$$f(a+h)-f(a) = \|h\|\cdot\alpha(h) = o(h) \quad \big(h\to(0,0)\big).$$

Поэтому функция f дифференцируема в точке $(0,0)$ и $d_{(0,0)}f = \mathbb{O}$. Из теоремы 1 вытекает, что $\dfrac{\partial f}{\partial x}(0,0) = \dfrac{\partial f}{\partial y}(0,0) = 0$. Кроме того, при $(x,y)\neq(0,0)$ верны равенства

$$\frac{\partial f}{\partial x}(x,y) = 2x\,\sin\frac{1}{x^2+y^2} - \frac{2x}{x^2+y^2}\,\cos\frac{1}{x^2+y^2},$$
$$\frac{\partial f}{\partial y}(x,y) = 2y\,\sin\frac{1}{x^2+y^2} - \frac{2y}{x^2+y^2}\,\cos\frac{1}{x^2+y^2},$$

которые проверяются формальным дифференцированием. Тогда

$$\frac{\partial f}{\partial x}\left(\frac{1}{\sqrt{2\pi k}},0\right) = -2\sqrt{2\pi k} \to -\infty \quad (k\to\infty),$$

откуда и получается разрывность $\dfrac{\partial f}{\partial x}$ в нуле. Для $\dfrac{\partial f}{\partial y}$ рассуждения аналогичны. $\square$

Пример 2. *Пусть*

$$f(x,y) = xy \exp \frac{1}{x^2+y^2} \quad \textit{при} \quad x^2+y^2 \neq 0, \quad f(0,0)=0.$$

Тогда справедливы следующие утверждения.

1) *В любой точке* $z \in \mathbb{R}^2$ *существуют частные производные* $\dfrac{\partial f}{\partial x}(z)$ *и* $\dfrac{\partial f}{\partial y}(z)$.

2) *Функция* f *имеет в точке* $(0,0)$ *производные только по векторам, параллельным* e^1 *или* e^2.

3) *Функция* f *разрывна (и, значит, не дифференцируема) в точке* $(0,0)$.

Действительно, $f(t,0) = f(0,t) = 0$ при всех $t \in \mathbb{R}$, откуда

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0.$$

При $(x,y) \neq (0,0)$ формальное дифференцирование дает

$$\frac{\partial f}{\partial x}(x,y) = y \exp \frac{1}{x^2+y^2} \cdot \left(1 - \frac{2x^2}{(x^2+y^2)^2}\right),$$
$$\frac{\partial f}{\partial y}(x,y) = x \exp \frac{1}{x^2+y^2} \cdot \left(1 - \frac{2y^2}{(x^2+y^2)^2}\right).$$

Тем самым первое утверждение доказано. Проверим 2). Пусть $e = (a,b)$, где $a, b \neq 0$. Тогда

$$\frac{f(te) - f(0,0)}{t} = t \cdot ab \cdot \exp \frac{1}{(a^2+b^2)t^2} \to \infty \quad (t \to 0)$$

(см. § 4 главы 3). Поэтому f не имеет в точке $(0,0)$ производной по вектору e. Наконец,

$$f\left(\frac{1}{k}, \frac{1}{k}\right) = \frac{\exp(k^2/2)}{k^2} \to +\infty \quad (k \to \infty),$$

откуда вытекает разрывность f в точке $(0,0)$. $\square$

Пример 3. *Пусть* $f(x,y)=1$, *если* $y=x^2$, $x>0$, *и* $f(x,y)=0$ *при остальных* $(x,y)\in\mathbb{R}^2$. *Тогда для любого* $e\in\mathbb{R}^2\setminus\{(0,0)\}$ *производная* $\dfrac{\partial f}{\partial e}(0,0)$ *существует и равна нулю, но функция* f *разрывна в точке* $(0,0)$.

Действительно, пусть $e=(a,b)\in\mathbb{R}^2\setminus\{(0,0)\}$. При достаточно малых $t\neq 0$ равенство $bt=a^2t^2$ не выполняется, поэтому $f(at,bt)=0$ и
$$\frac{\partial f}{\partial e}(0,0)=\lim_{t\to 0}\frac{f(at,bt)}{t}=0.$$
Кроме того, при $k\to\infty$
$$f\left(\frac{1}{k},\frac{1}{k^2}\right)=1\to 1\neq f(0,0),$$
откуда вытекает разрывность f в нуле. $\square$

Замечание. Из приведенных примеров можно сделать несколько выводов.

1) Пример 3 показывает, что теорему 1 нельзя обратить: существование в точке конечных производных по всем ненулевым векторам не гарантирует дифференцируемости (и, более того, непрерывности) отображения в этой точке.

2) Из примера 2 вытекает, что существование у отображения частных производных по всем переменным в некоторой точке является более слабым условием, чем наличие производной по любому направлению в этой точке. Таким образом, мы полностью обосновали замечание 2 к теореме 2.

3) В главе 3 было показано, что для функции одной переменной дифференцируемость в точке равносильна существованию конечной производной в этой точке. Из примера 2 вытекает, что уже в случае двух переменных наличие конечных частных производных необходимо, но не достаточно для дифференцируемости функции.

4) В § 4 мы докажем, что если все частные производные отображения f *непрерывны в точке* a, то f дифференцируемо в точке a. Тем не менее, пример 1 показывает, что это условие не является необходимым для дифференцируемости f в точке a.

§ 3. Правила дифференцирования

В этом параграфе мы выведем многомерные аналоги правил дифференцирования, доказанных в § 3 главы 3.

Теорема 1. Дифференцирование композиции. *Пусть $E \subset \mathbb{R}^n$, $F \subset \mathbb{R}^m$, $f\colon E \to F$, $g\colon F \to \mathbb{R}^s$. Предположим, что a — внутренняя точка E, $f(a)$ — внутренняя точка F, отображения f и g дифференцируемы соответственно в точках a и $f(a)$. Тогда отображение $g \circ f$ дифференцируемо в точке a и*

$$d_a(g \circ f) = d_{f(a)} g \circ d_a f. \tag{12}$$

Доказательство. Пусть $b = f(a)$, $T = d_a f$, $S = d_b g$. Запишем дифференцируемость f и g в соответствии с (3):

$$f(a+h) = f(a) + Th + \alpha(h) \cdot \|h\|\,, \tag{13}$$

$$g(b+H) = g(b) + SH + \beta(H) \cdot \|H\|\,, \tag{14}$$

где $\alpha(h) \to 0$ $(h \to 0)$ и $\beta(H) \to 0$ $(H \to 0)$. Положим теперь $H = f(a+h) - f(a)$. Отображение f непрерывно в точке a, поэтому $H(h) \to 0$ при $h \to 0$. Заметим, что в силу (13)

$$\begin{aligned} &g(b+H) = g\bigl(f(a+h)\bigr) = (g \circ f)(a+h), \\ &SH = S\bigl(Th + \alpha(h) \cdot \|h\|\bigr) = (ST)h + S\bigl(\alpha(h)\bigr) \cdot \|h\|\,. \end{aligned}$$

Поэтому (14) перепишется в виде

$$(g \circ f)(a+h) = (g \circ f)(a) + (ST)h + S\bigl(\alpha(h)\bigr) \cdot \|h\| + \beta(H) \cdot \|H\|\,.$$

Утверждение теоремы будет вытекать из соотношений

$$S\bigl(\alpha(h)\bigr) \to 0, \quad \beta(H) \cdot \frac{\|H\|}{\|h\|} \to 0 \quad \text{при } h \to 0. \tag{15}$$

Первое из них очевидно, поскольку

$$\bigl\|S\bigl(\alpha(h)\bigr)\bigr\| \leqslant \|S\| \cdot \|\alpha(h)\| \to 0 \quad (h \to 0).$$

Для доказательства второго заметим, что в силу (13)

$$\frac{\|H\|}{\|h\|} = \frac{\big\|Th + \alpha(h)\cdot\|h\|\big\|}{\|h\|} \leqslant \frac{\|Th\|}{\|h\|} + \|\alpha(h)\| \leqslant \|T\| + \|\alpha(h)\|\,,$$

По теореме о пределе композиции $\beta\big(H(h)\big) \to 0$ $(h \to 0)$, откуда

$$\|\beta(H)\|\cdot\frac{\|H\|}{\|h\|} \leqslant \|\beta(H)\|\cdot\big(\|T\| + \|\alpha(h)\|\big) \to 0 \quad (h \to 0),$$

и второе из соотношений (15) также доказано. □

Замечание 1. Правило цепочки. Формулу (12) можно записать в виде

$$(g\circ f)'(a) = g'\big(f(a)\big)\cdot f'(a), \tag{16}$$

поскольку матрица композиции линейных операторов равна произведению матриц этих операторов.

Замечание 2. Частные производные композиции. В условиях теоремы 1 для любых $i \in \{1,\ldots,s\}$ и $j \in \{1,\ldots,n\}$ справедлива формула

$$\frac{\partial(g\circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{m}\frac{\partial g_i}{\partial x_k}\big(f(a)\big)\cdot\frac{\partial f_k}{\partial x_j}(a). \tag{17}$$

Действительно, приравнивая элементы с индексами (i,j) у матриц, стоящих в левой и правой частях (16), мы получим (17).

Замечание 3. Если отображение g зависит от переменных $y_1,\ldots,y_m$, то его дифференциал можно записать в виде

$$d_y g(dy_1,\ldots,dy_m) = \sum_{k=1}^{m}\frac{\partial g}{\partial y_k}(y_1,\ldots,y_m)\,dy_k.$$

Предположим теперь, что $y_1,\ldots,y_m$ являются функциями от переменных $x_1,\ldots,x_n$. Тогда дифференциал композиции запишется в том же виде, только под $dy_1,\ldots,dy_m$ нужно понимать дифференциалы функций $y_1,\ldots,y_m$. Это свойство называют *инвариантностью формы первого дифференциала*.

Теорема 2. Дифференцирование и арифметические операции. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, отображения $f, g\colon E \to \mathbb{R}^m$ дифференцируемы в точке a. Тогда справедливы следующие утверждения.*

1) *Для любых $\lambda, \mu \in \mathbb{R}$ отображение $\lambda f + \mu g$ дифференцируемо в точке a и*

$$d_a(\lambda f + \mu g) = \lambda \cdot d_a f + \mu \cdot d_a g.$$

2) *Если функция $\lambda\colon E \to \mathbb{R}$ дифференцируема в точке a, то отображение λf дифференцируемо в точке a и*

$$d_a(\lambda \cdot f) = f(a) \cdot d_a\lambda + \lambda(a) \cdot d_a f.$$

3) *Функция $f \cdot g$ дифференцируема в точке a и*

$$d_a(f \cdot g) = g(a) \cdot d_a f + f(a) \cdot d_a g.$$

4) *Если $m = 1$ и $g(a) \neq 0$, то функция $\dfrac{f}{g}$ дифференцируема в точке a и*

$$d_a\left(\frac{f}{g}\right) = \frac{g(a) \cdot d_a f - f(a) \cdot d_a g}{g^2(a)}.$$

В формулировке теоремы 2 используются операции над отображениями, введенные в § 1 главы 5.

Доказательство. Положим $T = d_a f$, $S = d_a g$.

1) Запишем дифференцируемость f и g в соответствии с (3):

$$f(a+h) = f(a) + Th + \alpha(h) \cdot \|h\|, \quad \text{где } \alpha(h) \to 0 \ \ (h \to 0),$$
$$g(a+h) = g(a) + Sh + \beta(h) \cdot \|h\|, \quad \text{где } \beta(h) \to 0 \ \ (h \to 0).$$

Умножим первое из равенств на λ, второе на μ и сложим их. Тогда

$$(\lambda f + \mu g)(a+h) = (\lambda f + \mu g)(a) + (\lambda T + \mu S)h + \big(\lambda\alpha(h) + \mu\beta(h)\big) \cdot \|h\|.$$

Осталось заметить, что $\lim\limits_{h\to 0}\big(\lambda\alpha(h) + \mu\beta(h)\big) = 0$ по теореме 2 § 4 главы 5.

2) В силу леммы 1 § 2 достаточно проверить утверждение для координатных функций $\lambda f_1, \ldots, \lambda f_m$ отображения λf. Поэтому можно не умаляя общности считать $m = 1$. Пусть вначале $f = \lambda$. Положим $\varphi(t) = t^2$ при $t \in \mathbb{R}$. Тогда по теореме 1

$$d_a(\lambda f) = d_a f^2 = d_a(\varphi \circ f) = (d_{f(a)}\varphi)(d_a f) = 2f(a) \cdot d_a f, \qquad (18)$$

так как $d_{f(a)}\varphi(h) = 2f(a) \cdot h$ (см. главу 3). В общем случае заметим, что

$$\lambda f = \frac{1}{4}\big((\lambda + f)^2 - (\lambda - f)^2\big),$$

откуда в силу утверждения 1) и равенства (18)

$$d_a(\lambda f) = \frac{1}{4}\Big(2\big(\lambda(a) + f(a)\big) \cdot \big(d_a\lambda + d_a f\big) - \\ -2\big(\lambda(a) - f(a)\big) \cdot \big(d_a\lambda - d_a f\big)\Big) = f(a) \cdot d_a\lambda + \lambda(a) \cdot d_a f.$$

3) По лемме 1 § 2 координатные функции отображений f и g дифференцируемы в точке a, причем

$$d_a f_k = T_k \quad \text{и} \quad d_a g_k = S_k \quad \text{при всех } k \in \{1, \ldots, m\}.$$

В силу утверждений 1) и 2) мы получим

$$d_a(f \cdot g) = d_a\left(\sum_{k=1}^{m} f_k\, g_k\right) = \\ = \sum_{k=1}^{m} d_a\big(f_k\, g_k\big) = \sum_{k=1}^{m}\big(f_k(a)\, d_a g_k + g_k(a)\, d_a f_k\big) = \\ = \sum_{k=1}^{m} f_k(a)\, S_k + \sum_{k=1}^{m} g_k(a)\, T_k = f(a) \cdot S + g(a) \cdot T.$$

4) Пусть вначале $f \equiv 1$. Положим $\varphi(t) = \dfrac{1}{t}$ при $t \in \mathbb{R} \setminus \{0\}$. Тогда по теореме 1

$$d_a\left(\frac{1}{g}\right) = d_a(\varphi \circ g) = (d_{g(a)}\varphi)(d_a g) = -\frac{1}{g^2(a)}\, S, \qquad (19)$$

так как $d_{g(a)}\varphi(h) = -\dfrac{h}{g^2(a)}$ (см. главу 3). В общем случае в силу утверждения 2) и равенства (19)

$$d_a\left(\frac{f}{g}\right) = d_a\left(f\cdot\frac{1}{g}\right) = \frac{1}{g(a)}\cdot T + f(a)\cdot d_a\left(\frac{1}{g}\right) = \\ = \frac{1}{g(a)}\cdot T - f(a)\cdot\frac{S}{g^2(a)} = \frac{g(a)\,T - f(a)\,S}{g^2(a)}. \quad \square$$

Замечание 1. Утверждение 1) называют *линейностью операции дифференцирования.*

Замечание 2. Все утверждения теоремы 2 можно записать в терминах матриц Якоби. Например, правило дифференцирования произведения функции λ и отображения f примет вид

$$(\lambda f)'(a) = f(a)\cdot\lambda'(a) + \lambda(a)\cdot f'(a).$$

Действительно, для любого $h\in\mathbb{R}^n$

$$(\lambda f)'(a)h = d_a(\lambda f)(h) = \\ = f(a)\cdot d_a\lambda(h) + \lambda(a)\cdot d_a f(h) = \big(f(a)\lambda'(a) + \lambda(a)f'(a)\big)h.$$

При $m = n = 1$ мы получим формулу для производной произведения, доказанную в § 3 главы 3. В общем случае справедливо равенство

$$\frac{\partial(\lambda f)}{\partial x_j}(a) = f(a)\cdot\frac{\partial\lambda}{\partial x_j}(a) + \lambda(a)\cdot\frac{\partial f}{\partial x_j}(a),$$

где $j\in\{1,\dots,n\}$. Предлагаем читателю переписать остальные утверждения теоремы 2 в терминах матриц Якоби и частных производных.

Замечание 3. Утверждения 1), 2) и 4) теоремы 2 верны и для комплекснозначных функций f, g и λ, если под арифметическими действиями понимать операции над комплексными числами. Например, для доказательства утверждения 2) в такой редакции нужно в произведении λf отделить вещественную и мнимую части и применить к ним вещественный вариант этого утверждения. Формальную проверку мы предоставляем читателю.

Теорема 3. Дифференцирование обратного отображения. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, отображение $f: E \to \mathbb{R}^n$ дифференцируемо в точке a. Предположим, что выполнены следующие условия:*

1) *оператор $d_a f$ обратим;*
2) *отображение f инъективно;*
3) *точка $f(a)$ является внутренней для $f(E)$;*
4) *обратное отображение f^{-1} непрерывно в точке $f(a)$.*

Тогда отображение f^{-1} дифференцируемо в точке $f(a)$ и

$$d_{f(a)} f^{-1} = (d_a f)^{-1}. \tag{20}$$

Доказательство. Пусть $b = f(a)$, $T = d_a f$, $g = f^{-1}$. Запишем дифференцируемость f в соответствии с (3):

$$f(a + H) = f(a) + TH + \alpha(H) \cdot \|H\|, \tag{21}$$

где $\alpha(H) \to 0$ при $H \to 0$. Положим

$$H = g(b + h) - g(b), \quad \text{где } \; b + h \in f(E).$$

Если $h \to 0$, то в силу условия 4) $H \to 0$, и по теореме о пределе композиции $\alpha(H) \to 0$. По условию 3) найдется такая окрестность V_b точки b, что

$$V_b \subset f(E) \quad \text{и} \quad \|\alpha(H)\| < \frac{1}{2\,\|T^{-1}\|} \;\text{ при }\; h \in V_b. \tag{22}$$

Заметим, что для любого $h \in V_b$

$$f(a + H) - f(a) = f\big(g(b + h)\big) - b = b + h - b = h.$$

Поэтому из (21) мы получаем

$$TH = f(a + H) - f(a) - \alpha(H) \cdot \|H\| = h - \alpha(H) \cdot \|H\|,$$

откуда

$$g(b+h)-g(b)=H=T^{-1}h-T^{-1}\big(\alpha(H)\big)\cdot\|H\|. \tag{23}$$

Осталось проверить, что

$$\frac{T^{-1}\big(\alpha(H)\big)\cdot\|H\|}{\|h\|}\to 0 \quad \text{при } h\to 0. \tag{24}$$

В силу (22) для $h\in V_b$

$$\big\|T^{-1}\big(\alpha(H)\big)\big\|\cdot\|H\|\leqslant\big\|T^{-1}\big\|\cdot\|\alpha(H)\|\cdot\|H\|\leqslant\tfrac{1}{2}\,\|H\|.$$

Тогда из (23) вытекает, что

$$\frac{\|H\|}{\|h\|}\leqslant\frac{\big\|T^{-1}h\big\|}{\|h\|}+\frac{\big\|T^{-1}\big(\alpha(H)\big)\big\|\cdot\|H\|}{\|h\|}\leqslant\big\|T^{-1}\big\|+\frac{1}{2}\cdot\frac{\|H\|}{\|h\|},$$

откуда $\dfrac{\|H\|}{\|h\|}\leqslant 2\,\big\|T^{-1}\big\|$ при всех $h\in V_b$. Поэтому

$$\frac{\big\|T^{-1}\big(\alpha(H)\big)\big\|\cdot\|H\|}{\|h\|}\leqslant 2\,\big\|T^{-1}\big\|^2\cdot\|\alpha(H)\|\to 0 \quad (h\to 0),$$

что и доказывает (24). □

§ 4. Непрерывно дифференцируемые отображения

В главе 3 было показано, что для функции одной переменной дифференцируемость в точке эквивалентна существованию в ней производной. Это утверждение давало нам удобный способ исследования таких функций на дифференцируемость. Для отображений нескольких переменных нет столь простой связи между дифференцируемостью и существованием частных производных (см. примеры в § 2). Тем не менее мы увидим, что в терминах частных производных удобно описывается многомерный вариант *непрерывной дифференцируемости*. Поэтому с практической точки зрения непрерывная дифференцируемость оказывается более важным понятием, чем дифференцируемость. Изучению класса непрерывно дифференцируемых отображений и посвящен этот параграф.

Дадим вначале глобальный вариант определения дифференцируемости.

Определение 1. Дифференцируемость на множестве. Пусть множество E открыто в $\mathbb{R}^n$. Отображение $f: E \to \mathbb{R}^m$ называется *дифференцируемым на* E, если оно дифференцируемо в любой точке множества E.

Замечание. Множество E предполагается открытым, так как в соответствии с определением 2 § 2 любая точка E должна быть внутренней.

Теперь мы введем главное понятие этого параграфа — непрерывную дифференцируемость.

Определение 2. Непрерывно дифференцируемое отображение. Пусть $E \subset \mathbb{R}^n$, $f: E \to \mathbb{R}^m$.

1) Если $a \in E$, все частные производные f определены в некоторой окрестности a и непрерывны в точке a, то отображение f называется *непрерывно дифференцируемым в точке* a.

2) Если множество E открыто и отображение f непрерывно дифференцируемо в любой точке E, то оно называется *непрерывно дифференцируемым на* E.

Замечание. Пусть множество E открыто в $\mathbb{R}^n$, $F \subset \mathbb{R}^m$. Класс непрерывно дифференцируемых отображений, действующих из E в F, мы будем обозначать через $C^1(E, F)$ или $C^1(E \to F)$. Как правило, F будет совпадать со всем $\mathbb{R}^m$. Если $m = 1$, то мы будем использовать более краткое обозначение $C^1(E)$, что соответствует соглашениям главы 3. Вместо $C^1(E, \mathbb{C})$ мы также будем писать $C^1(E)$, так как из контекста обычно ясно, какова область значений отображения. Элементы $C^1(E)$ называют *непрерывно дифференцируемыми функциями*.

Непрерывная дифференцируемость функции одной переменной тривиальным образом влечет ее дифференцируемость. Для отображений нескольких переменных это уже требует обоснования, поскольку нет естественной связи между дифференцируемостью и поведением частных производных (см. § 2). Докажем вначале утверждение, позволяющее выразить приращение функции нескольких переменных через ее частные производные.

Лемма 1. *Пусть* $E \subset \mathbb{R}^n$, $a \in E$, $\delta > 0$, $V_a(\delta) \subset E$, $h \in V_0(\delta)$. *Предположим, что частные производные функции* $f\colon E \to \mathbb{R}$ *по всем переменным определены на* $V_a(\delta)$. *Тогда существуют такие* $c^1, \dots, c^n \in V_a(\delta)$, *что*

$$f(a+h) - f(a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\big(c^k\big) \cdot h_k. \tag{25}$$

Доказательство. Положим $x^1 = a$ и

$$x^k = \big(a_1 + h_1, \dots, a_{k-1} + h_{k-1}, a_k, \dots, a_n\big) \text{ при } k \in \{2, \dots, n\}.$$

Для $k \in \{1, \dots, n\}$ определим функцию F_k следующим образом:

$$F_k(t) = f\big(x^k + th_k e^k\big), \quad t \in [0,1].$$

Заметим, что при $t \in [0,1]$

$$\big\|x^k + te^k h_k - a\big\| = \sqrt{h_1^2 + \ldots + h_{k-1}^2 + t^2 h_k^2} \leqslant \|h\| < \delta, \tag{26}$$

то есть функции F_k корректно определены на $[0,1]$. По теореме 1 § 3 все F_k дифференцируемы на $[0,1]$, и в силу (17)

$$F_k'(t) = \frac{\partial f}{\partial x_k}\big(x^k + te^k h_k\big) \cdot h_k \quad \text{для любых } k \in \{1, \dots, n\} \text{ и } t \in [0,1].$$

Применим к функциям F_k теорему Лагранжа о среднем. Тогда найдутся числа $\theta_k \in (0,1)$, для которых

$$F_k(1) - F_k(0) = F_k'\big(\theta_k\big) = \frac{\partial f}{\partial x_k}\big(x^k + \theta_k h_k e^k\big) \cdot h_k.$$

Положим $c^k = x^k + \theta_k h_k e^k$ при любом $k \in \{1, \dots, n\}$. В силу (26) все векторы c^k попадают в $V_a(\delta)$. Осталось заметить, что

$$\begin{aligned} f(a+h) - f(a) &= \sum_{k=1}^{n} \Big(f\big(x^k + h_k e^k\big) - f\big(x^k\big)\Big) = \\ &= \sum_{k=1}^{n} \big(F_k(1) - F_k(0)\big) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\big(c^k\big) \cdot h_k. \quad \square \end{aligned}$$

Изучим теперь связь непрерывной дифференцируемости с двумя другими важными свойствами, рассмотренными ранее, — дифференцируемостью и непрерывностью.

Теорема 1. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, отображение $f\colon E \to \mathbb{R}^m$ непрерывно дифференцируемо в точке a. Тогда справедливы следующие утверждения.*

1) *f непрерывно в некоторой окрестности V_a точки a.*

2) *f дифференцируемо в точке a.*

Доказательство. Заметим вначале, что достаточно ограничиться случаем $m = 1$. Действительно, пусть для $m = 1$ теорема доказана. Применяя ее к координатным функциям f, мы получим, что все они непрерывны в некоторой окрестности a и дифференцируемы в точке a. Поэтому первое утверждение леммы вытекает из замечания 4 к определению непрерывности (см. § 5 главы 5), а второе — из леммы 1 § 2. Перейдем к рассмотрению случая $m = 1$.

1) Так как частные производные f непрерывны в точке a, они ограничены вблизи a, то есть существуют окрестность V_a точки a и $M > 0$, для которых

$$\left|\frac{\partial f}{\partial x_k}(c)\right| \leqslant M \quad \text{при всех } \ k \in \{1, \ldots, n\} \ \text{ и } \ c \in V_a.$$

Пусть $x \in V_a$ и $h \in \mathbb{R}^n$, причем $\overline{V_x\bigl(\|h\|\bigr)} \subset V_a$. По лемме 1 найдутся такие $c^1, \ldots, c^n \in V_a$, что

$$f(x+h) - f(x) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\bigl(c^k\bigr) \cdot h_k.$$

Тогда по неравенству Коши

$$\bigl|f(x+h) - f(x)\bigr| \leqslant \sqrt{\sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}\bigl(c^k\bigr)\right)^2} \cdot \sqrt{\sum_{k=1}^{n} h_k^2} \leqslant M\sqrt{n}\,\|h\|,$$

откуда $f(x+h) \to f(x)$ при $h \to 0$.

2) По любым $\varepsilon > 0$ и $k \in \{1, \dots, n\}$ подберем такое $\delta_k > 0$, что

$$\left| \frac{\partial f}{\partial x_k}(x) - \frac{\partial f}{\partial x_k}(a) \right| < \frac{\varepsilon}{\sqrt{n}} \quad \text{для всех } \ x \in V_a(\delta_k).$$

Положим $\delta = \min\{\delta_1, \dots, \delta_n\}$, и пусть $h \in V_0(\delta)$. По лемме 1 найдутся $c^1, \dots, c^n \in V_a(\delta)$, для которых выполнено равенство (25). Тогда

$$f(a+h) - f(a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a) \cdot h_k + R(h),$$

где

$$R(h) = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(c^k) - \frac{\partial f}{\partial x_k}(a) \right) h_k.$$

По неравенству Коши

$$\bigl|R(h)\bigr| \leqslant \sqrt{\sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(c^k) - \frac{\partial f}{\partial x_k}(a) \right)^2} \cdot \|h\| \leqslant \varepsilon \, \|h\| \, .$$

Поэтому $R(h) = o(h)$ при $h \to 0$, откуда и вытекает дифференцируемость функции f в точке a. $\square$

Следствие. *Если E открыто в $\mathbb{R}^n$ и $f \in C^1(E, \mathbb{R}^m)$, то отображение f дифференцируемо на E.*

Класс непрерывно дифференцируемых отображений можно также описать в терминах дифференциалов. Для этого нам потребуется следующее вспомогательное понятие.

Определение 3. Пусть множество E открыто в $\mathbb{R}^n$, отображение $f\colon E \to \mathbb{R}^m$ дифференцируемо на E. Если для любой точки $a \in E$

$$\|d_x f - d_a f\| \to 0 \ \text{ при } \ x \to a, \tag{27}$$

то говорят, что *дифференциал f непрерывно зависит от точки на множестве E.*

Замечание. Рассмотрим условие (27) более подробно. В случае $n = 1$ из формулы (10) главы 5 вытекает, что

$$\|d_x f - d_a f\| = \sup_{|h|=1} \bigl\| \bigl(f'(x) - f'(a)\bigr) \cdot h \bigr\| = \|f'(x) - f'(a)\| \, .$$

Таким образом, для вектор-функции условие (27) равносильно непрерывности в точке a отображения $x \mapsto f'(x)$. В частности, при $m = 1$ определение 3 эквивалентно определению непрерывно дифференцируемой функции (см. главу 3). Отметим также, что из (27) следует непрерывность в точке a функции $x \mapsto \|d_x f\|$. Действительно, по следствию теоремы 2 § 6 главы 5

$$\big|\|d_x f\| - \|d_a f\|\big| \leqslant \|d_x f - d_a f\| \to 0 \quad (x \to a).$$

Опишем теперь непрерывно дифференцируемые отображения в терминах дифференциалов.

Теорема 2. Характеристика непрерывно дифференцируемых отображений. *Пусть множество E открыто в $\mathbb{R}^n$, $f\colon E \to \mathbb{R}^m$. Тогда равносильны следующие утверждения.*

1) $f \in C^1(E, \mathbb{R}^m)$.

2) *Отображение f дифференцируемо на E и дифференциал f непрерывно зависит от точки на E.*

Доказательство. 1) $\Rightarrow$ 2) Дифференцируемость f на E вытекает из следствия теоремы 1. Пусть $a \in E$. Покажем, что f удовлетворяет в точке a условию (27). По теореме 3 § 2 элементами матрицы f' являются функции вида $\dfrac{\partial f_i}{\partial x_j}$, непрерывные в точке a. Тогда по неравенству (12) главы 5

$$\|d_x f - d_a f\| \leqslant \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)\right)^2} \to 0 \quad (x \to a).$$

2) $\Rightarrow$ 1) По теореме 1 § 2 у f существуют частные производные по всем переменным на E. Пусть $a \in E$. В силу условия (27) для любого $k \in \{1, \ldots, n\}$

$$\begin{aligned}\left\|\frac{\partial f}{\partial x_k}(x) - \frac{\partial f}{\partial x_k}(a)\right\| &= \big\|d_x f\big(e^k\big) - d_a f\big(e^k\big)\big\| = \\ &= \big\|\big(d_x f - d_a f\big)\big(e^k\big)\big\| \leqslant \|d_x f - d_a f\| \to 0 \quad (x \to a).\end{aligned}$$

Поэтому f непрерывно дифференцируемо в любой точке $a \in E$, то есть $f \in C^1(E, \mathbb{R}^m)$. $\square$

В § 3 мы показали, что результат композиций и арифметических операций над дифференцируемыми отображениями также будет дифференцируемым отображением. Оказывается, что аналогичные утверждения верны и для непрерывно дифференцируемых отображений. Сформулируем их.

Теорема 3. Композиция отображений класса C^1. *Пусть множества E и F открыты соответственно в $\mathbb{R}^n$ и $\mathbb{R}^m$. Если $f \in C^1(E, F)$ и $g \in C^1(F, \mathbb{R}^s)$, то $g \circ f \in C^1(E, \mathbb{R}^s)$.*

Мы не будем приводить доказательства этой теоремы, поскольку далее будет получен более общий результат (см. теорему $2'$ § 5).

Теорема 4. Арифметические операции в классе C^1. *Пусть множество E открыто в $\mathbb{R}^n$, $f, g \in C^1(E, \mathbb{R}^m)$. Тогда справедливы следующие утверждения.*

1) $\lambda f + \mu g \in C^1(E, \mathbb{R}^m)$ *при любых* $\lambda, \mu \in \mathbb{R}$.
2) *Если* $\lambda \in C^1(E)$, *то* $\lambda f \in C^1(E, \mathbb{R}^m)$.
3) $f \cdot g \in C^1(E)$.
4) *Если* $m = 1$ *и* $g \neq 0$ *на* E, *то* $\dfrac{f}{g} \in C^1(E)$.

Мы не будем приводить доказательства этой теоремы, поскольку далее будет получен более общий результат (см. теорему $1'$ § 5).

Замечание 1. Теорема 4 верна и для комплекснозначных функций, если под арифметическими действиями понимать операции над комплексными числами.

Замечание 2. Теорема 3 § 3 также допускает обобщение на случай непрерывно дифференцируемых отображений. Мы рассмотрим этот вопрос позже, в § 7, посвященном обратимым отображениям.

§ 5. Производные и дифференциалы высших порядков

В этом параграфе мы определим для отображений нескольких переменных частные производные и дифференциалы высших порядков, установим связь между ними. Кроме того, обобщая материал предыдущего параграфа, мы введем понятие многократной непрерывной дифференцируемости и изучим классы отображений, обладающих этим свойством.

Для описания частных производных высших порядков нам потребуется обозначение, введенное в § 7 главы 5:

$$I_s = \{1, \ldots, n\}^s \quad (s \in \mathbb{N}), \qquad I_0 = \{\varnothing\}.$$

Элементами I_s являются *семейства индексов*, лежащих в множестве $\{1, \ldots, n\}$. Очевидно, что I_s зависит и от n, но мы не отражаем этого в обозначении, поскольку число n всегда будет фиксированным.

Пусть $s \in \mathbb{N}$, $i \in I_s$. Введем *операцию дифференцирования порядка* s по переменным с индексами из семейства i, которая будет обозначаться символом ∂_i или, более подробно, $\partial_{i_1 \ldots i_s}$. Определим ее по индукции. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$. При $s = 1$ положим

$$\partial_i f(a) = \frac{\partial f}{\partial x_{i_1}}(a). \tag{28}$$

В этом случае наряду с ∂_i используется также запись $\dfrac{\partial}{\partial x_{i_1}}$. Предположим теперь, что $s > 1$ и дифференцирование порядка $s-1$ уже определено. Тогда положим

$$\partial_i f(a) = \frac{\partial}{\partial x_{i_s}}\bigl(\partial_{i_1 \ldots i_{s-1}} f\bigr)(a) \tag{29}$$

при условии, что правая часть (29) имеет смысл. Это означает, что отображение $x \mapsto \partial_{i_1 \ldots i_{s-1}} f(x)$ задано по крайней мере в некоторой окрестности a и имеет в точке a частную производную по переменной x_{i_s}. Таким образом, равенства (28) и (29) определяют операцию дифференцирования по произвольному набору переменных. Теперь мы можем ввести частные производные высших порядков.

Определение 1. Частные производные высших порядков. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$, $s \in \mathbb{N}$, $i \in I_s$. Тогда $\partial_i f(a)$ называют *частной производной порядка s по переменным $x_{i_1}, \ldots, x_{i_s}$ в точке a.* Для нее используются также обозначения $\dfrac{\partial^s f}{\partial x_{i_1} \ldots \partial x_{i_s}}(a)$, $D^s_{i_1 \ldots i_s} f(a)$, $f^{(s)}_{x_{i_1} \ldots x_{i_s}}(a)$.

Замечание 1. Если все координаты вектора i равны, то производная $\partial_i f(a)$ называется *чистой*. Ее принято обозначать символом $\dfrac{\partial^s f}{\partial x_k^s}(a)$, где k — общее значение индексов семейства i. В противном случае производная называется *смешанной*. При записи смешанных производных количество последовательных дифференцирований по одной переменной также указывается верхним индексом. Например, вместо $\dfrac{\partial^3 f}{\partial x_1 \partial x_1 \partial x_3}(a)$ принято писать $\dfrac{\partial^3 f}{\partial x_1^2 \partial x_3}(a)$.

Замечание 2. Удобно ввести также частную производную *нулевого порядка* в точке a, положив ее равной $f(a)$. Такое определение соответствует соглашениям главы 3. Отметим, что производная нулевого порядка порождается пустым семейством индексов и имеет смысл для любого отображения.

Замечание 3. Пусть $s, r \in \mathbb{N}$, $i \in I_s$, $j \in I_r$. Тогда

$$\partial_{i_1 \ldots i_s j_1 \ldots j_r} f(a) = \partial_{j_1 \ldots j_r}\big(\partial_{i_1 \ldots i_s} f\big)(a). \tag{30}$$

Это вытекает непосредственно из определения 1.

Замечание 4. *Пусть $s \in \mathbb{N}$, $i \in I_s$. Тогда равносильны следующие утверждения.*

1) *Существует производная $\partial_i f(a)$.*

2) *Для любого $k \in \{1, \ldots, m\}$ существует производная $\partial_i f_k(a)$.*

Если выполнены утверждения 1) и 2), то справедливо равенство

$$\big(\partial_i f\big)_k(a) = \partial_i f_k(a) \quad \text{при всех } \ k \in \{1, \ldots, m\}.$$

Доказательство проводится индукцией по s: достаточно воспользоваться замечанием 3 к определению 4 § 2 и равенством (29). Предлагаем читателю сделать это самостоятельно.

В § 2 мы определили дифференцируемые отображения. Введем теперь понятие дифференцируемости более высокого порядка.

Определение 2. Многократная дифференцируемость. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$, $s \in \mathbb{N}$. Отображение f называется *s раз дифференцируемым в точке a*, если выполняются два условия:

1) все частные производные f порядков $0, \ldots, s-2$ дифференцируемы в некоторой окрестности точки a;

2) все частные производные f порядка $s-1$ дифференцируемы в точке a.

Если множество E открыто и отображение f s раз дифференцируемо в любой точке E, то оно называется *s раз дифференцируемым на E*. Отображение, дифференцируемое на E любое число раз, называется *бесконечно дифференцируемым на E*.

Замечание 1. При $s = 1$ условие 1) тривиально, а 2) эквивалентно определению дифференцируемости в точке, данному в § 2. Таким образом, термины "однократная дифференцируемость" и "дифференцируемость" означают одно и то же.

Замечание 2. Из замечания 4 к определению 1 и леммы 1 § 2 вытекает, что s-кратная дифференцируемость f в точке a эквивалентна s-кратной дифференцируемости в точке a всех координатных функций отображения f.

Замечание 3. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f: E \to \mathbb{R}^m$, $k, s \in \mathbb{N}$, $k < s$. Тогда равносильны следующие утверждения.*

1) *Отображение f s раз дифференцируемо в точке a.*

2) *Отображение f k раз дифференцируемо в некоторой окрестности точки a и все его частные производные k-го порядка $s-k$ раз дифференцируемы в точке a.*

Действительно, k-кратная дифференцируемость f в окрестности V_a точки a означает, что все частные производные f порядка не выше $k-1$ дифференцируемы на V_a. Остальное вытекает из равенства (30). Это замечание будет полезным при проверке многократной дифференцируемости по индукции. Наиболее важными являются случаи $k = 1$ и $k = s-1$.

Условия дифференцируемости частных производных различных порядков не являются независимыми. Докажем утверждение, устанавливающее связь между ними.

Лемма 1. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $s \in \mathbb{Z}_+$, $f: E \to \mathbb{R}^m$ и все частные производные f порядка s дифференцируемы в окрестности V_a точки a. Тогда для любого $k \in \{0, \ldots, s\}$ частные производные f порядка k также дифференцируемы на V_a.*

Доказательство проведем индукцией по $s-k$. Для $k = s$ утверждение тривиально. Предположим, что $k \in \{0, \ldots, s-1\}$ и дифференцируемость на V_a производных f порядка $k+1$ уже доказана. Тогда частные производные f порядка $k+1$ непрерывны на V_a, откуда вытекает непрерывная дифференцируемость (и, тем более, дифференцируемость) на V_a частных производных f порядка k. $\square$

Замечание. Из леммы 1 вытекает, что при $s \geqslant 2$ условие 1) в определении 2 можно изменить следующим образом:

$1'$) *все частные производные f порядка $s-2$ дифференцируемы в некоторой окрестности точки a.*

Поэтому в дальнейшем, доказывая многократную дифференцируемость отображений, мы можем проверять условие $1'$) вместо 1), что обычно упрощает рассуждения.

В § 3 была изучена связь дифференцируемости с композициями и арифметическими операциями. Обобщим теперь эти результаты на случай многократной дифференцируемости.

Теорема 1. Многократная дифференцируемость и арифметические операции. *Пусть $s \in \mathbb{N}$, $E \subset \mathbb{R}^n$, a — внутренняя точка E, отображения $f, g\colon E \to \mathbb{R}^m$ s раз дифференцируемы в точке a. Тогда справедливы следующие утверждения.*

1) *При любых $\lambda, \mu \in \mathbb{R}$ отображение $\lambda f + \mu g$ s раз дифференцируемо в точке a.*

2) *Если функция $\lambda\colon E \to \mathbb{R}$ s раз дифференцируема в точке a, то этим свойством обладает и отображение λf.*

3) *Функция $f \cdot g$ s раз дифференцируема в точке a.*

4) *Если $m = 1$ и $g(a) \neq 0$, то функция $\dfrac{f}{g}$ s раз дифференцируема в точке a.*

Доказательство проводится индукцией по s. В случае $s = 1$ достаточно сослаться на теорему 2 § 3. Индукционный переход не представляет сложностей, и мы его подробно проведем только для второго утверждения. Предположим, что $s > 1$ и для $s-1$ теорема доказана. Пусть $k \in \{1, \ldots, n\}$. По замечанию 2 к теореме 2 § 3 в некоторой окрестности точки a справедливо равенство

$$\partial_k(\lambda f) = f \cdot \partial_k \lambda + \lambda \cdot \partial_k f.$$

В силу замечания 3 к определению 2 отображения λ, f, $\partial_k\lambda$ и $\partial_k f$ $s-1$ раз дифференцируемы в точке a. Тогда по индукционному предположению таковым будет и отображение $\partial_k(\lambda f)$. Еще раз применяя замечание 3 к определению 2, мы получим s-кратную дифференцируемость λf в точке a. $\square$

Теорема 2. Композиция многократно дифференцируемых отображений. *Пусть $s \in \mathbb{N}$, $E \subset \mathbb{R}^n$, $F \subset \mathbb{R}^m$, $f\colon E \to F$, $g\colon F \to \mathbb{R}^r$. Предположим, что a — внутренняя точка E, $f(a)$ — внутренняя точка F, отображения f и g s раз дифференцируемы соответственно в точках a и $f(a)$. Тогда отображение $g \circ f$ s раз дифференцируемо в точке a.*

Доказательство проводится индукцией по s. В случае $s=1$ достаточно сослаться на теорему 1 § 3. Предположим, что $s>1$ и для $s-1$ утверждение уже доказано. Из определения 2 вытекает, что отображения f и g дифференцируемы соответственно в некоторых окрестностях V_a точки a и $V_{f(a)}$ точки $f(a)$. В силу непрерывности f в точке a окрестность V_a можно выбрать так, что $f\bigl(V_a\bigr) \subset V_{f(a)}$. Из теоремы 1 § 3 следует дифференцируемость $g \circ f$ на V_a. Кроме того, по замечанию 2 к теореме 1 § 3 при всех $k \in \{1,\dots,n\}$ верна формула

$$\partial_k(g \circ f)(x) = \sum_{j=1}^{m} \partial_j g\bigl(f(x)\bigr) \cdot \partial_k f_j(x) \quad (x \in V_a).$$

В силу замечаний 2 и 3 к определению 2 отображения f и $\partial_k f_j$ будут $s-1$ раз дифференцируемыми в точке a, а $\partial_j g$ — в точке $f(a)$. Тогда по индукционному предположению $\partial_j g \circ f$ является $s-1$ раз непрерывно дифференцируемым в точке a, и по теореме 1 таковым же будет отображение $\partial_k(g \circ f)$. Еще раз применяя замечание 3 к определению 2, мы получим s-кратную дифференцируемость $g \circ f$ в точке a. $\square$

В § 4 мы рассмотрели понятие непрерывной дифференцируемости, которое, в отличие от дифференцируемости, легко формулируется в терминах частных производных. Определим теперь непрерывную дифференцируемость высших порядков.

Определение 3. Многократная непрерывная дифференцируемость. Пусть $E \subset \mathbb{R}^n$, $f\colon E \to \mathbb{R}^m$, $s \in \mathbb{N}$.

1) Предположим, что у точки $a \in E$ существует такая окрестность $V_a \subset E$, что при любом $i \in I_s$ производная $\partial_i f$ определена на V_a и непрерывна в точке a. Тогда отображение f называется *s раз непрерывно дифференцируемым в точке a.*

2) Если множество E открыто и отображение f s раз непрерывно дифференцируемо в любой точке E, то оно называется *s раз непрерывно дифференцируемым на E.*

Замечание 1. Согласно определению 2 § 4, случай $s = 1$ соответствует непрерывно дифференцируемому отображению.

Замечание 2. Назовем отображение f *0 раз непрерывно дифференцируемым* во внутренней точке a множества E, если оно непрерывно в точке a. Это соглашение оправдывается тем, что частная производная f нулевого порядка есть f.

Замечание 3. Из замечания 4 к определению 1 вытекает, что s-кратная непрерывная дифференцируемость f в точке a эквивалентна s-кратной непрерывной дифференцируемости в a всех координатных функций отображения f.

Замечание 4. Пусть множество E открыто в $\mathbb{R}^n$, $F \subset \mathbb{R}^m$, $s \in \mathbb{Z}_+$. Класс s раз непрерывно дифференцируемых отображений, действующих из E в F, мы будем обозначать через $C^s(E, F)$ или $C^s(E \to F)$. Положим также

$$C^\infty(E, F) = \bigcap_{s=0}^{\infty} C^s(E, F).$$

Как правило, F будет совпадать со всем $\mathbb{R}^m$. Если $m = 1$, то мы будем использовать более краткое обозначение $C^s(E)$, что соответствует соглашениям главы 3. Вместо $C^s(E, \mathbb{C})$ мы также будем писать $C^s(E)$, так как из контекста обычно ясно, какова область значений отображения. Элементы $C^s(E)$ называют *s раз непрерывно дифференцируемыми функциями.* В случае $s = 0$ вместо $C^0(E, F)$ пишут просто $C(E, F)$, что согласуется с обозначениями главы 5. Заметим также, что класс $C^\infty(E, F)$ состоит из отображений, бесконечно дифференцируемых на E.

Установим теперь некоторые важные свойства классов C^s.

Замечание 5. *Пусть множество E открыто в $\mathbb{R}^n$, $f: E \to \mathbb{R}^m$, $k, s \in \mathbb{N}$, $k < s$. Тогда равносильны следующие утверждения.*

1) $f \in C^s(E, \mathbb{R}^m)$.

2) *Все частные производные f порядка k принадлежат классу $C^{s-k}(E, \mathbb{R}^m)$.*

Для доказательства этого утверждения достаточно воспользоваться соотношением (30).

Лемма 2. *Пусть множество E открыто в $\mathbb{R}^n$. Тогда справедливы следующие утверждения.*

1) *Для любых $k, s \in \mathbb{Z}_+ \cup \{\infty\}$, $k \leqslant s$ верно включение*

$$C^s(E, \mathbb{R}^m) \subset C^k(E, \mathbb{R}^m).$$

2) *Если $s \in \mathbb{N}$ и $f \in C^s(E, \mathbb{R}^m)$, то отображение f s раз дифференцируемо на E.*

Доказательство. Мы ограничимся случаем $s \in \mathbb{Z}_+$, к которому общая ситуация легко сводится.

1) Пусть $f \in C^s(E, \mathbb{R}^m)$. Докажем индукцией по $s-k$ включение $f \in C^k(E, \mathbb{R}^m)$. Для $k = s$ оно очевидно. Предположим, что $k < s$ и $f \in C^{k+1}(E, \mathbb{R}^m)$. В силу замечаний 1 и 5 все частные производные f порядка k непрерывно дифференцируемы на E, и из теоремы 1 § 4 вытекает их непрерывность на E. Таким образом, $f \in C^k(E, \mathbb{R}^m)$.

2) Пусть $k \in \{0, \ldots, s-1\}$. Из утверждения 1) следует включение $f \in C^{k+1}(E, \mathbb{R}^m)$. В силу замечаний 1 и 5 все частные производные f порядка k непрерывно дифференцируемы на E и, по теореме 1 § 4, дифференцируемы на E. Тем самым доказана s-кратная дифференцируемость f на E. $\square$

Замечание. Справедлив также локальный вариант утверждения 2): *если $a \in E$ и отображение f s раз непрерывно дифференцируемо в точке a, то оно s раз дифференцируемо в точке a.* Предлагаем читателю проверить это самостоятельно.

В теоремах 1 и 2 утверждалось, что свойство многократной дифференцируемости сохраняется при композициях и арифметических операциях. Выведем теперь аналогичные результаты для отображений класса C^s.

Теорема 1$'$. Арифметические операции в классе C^s. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{Z}_+$, $f, g \in C^s(E, \mathbb{R}^m)$. Тогда справедливы следующие утверждения.*

1) $\lambda f + \mu g \in C^s(E, \mathbb{R}^m)$ *при любых* $\lambda, \mu \in \mathbb{R}$.
2) *Если* $\lambda \in C^s(E)$, *то* $\lambda f \in C^s(E, \mathbb{R}^m)$.
3) $f \cdot g \in C^s(E)$.
4) *Если* $m = 1$ *и* $g \neq 0$ *на* E, *то* $\dfrac{f}{g} \in C^s(E)$.

Доказательство проводится индукцией по s. В случае $s = 0$ достаточно сослаться на теорему 2 § 5 главы 5. Индукционный переход мы подробно проведем только для второго утверждения. Предположим, что $s \in \mathbb{N}$ и для $s-1$ теорема уже доказана. Пусть $k \in \{1, \ldots, n\}$. По замечанию 2 к теореме 2 § 3 на E выполняется равенство
$$\partial_k(\lambda f) = f \cdot \partial_k \lambda + \lambda \cdot \partial_k f.$$
В силу замечания 5 к определению 3 отображения λ, f, $\partial_k \lambda$ и $\partial_k f$ принадлежат классу C^{s-1}, откуда по индукционному предположению $\partial_k(\lambda f) \in C^{s-1}(E, \mathbb{R}^m)$. Еще раз применяя замечание 5 к определению 3, мы получим включение $\lambda f \in C^s(E, \mathbb{R}^m)$. $\square$

Теорема 2$'$. Композиция отображений класса C^s. *Пусть множества E и F открыты соответственно в $\mathbb{R}^n$ и $\mathbb{R}^m$, $s \in \mathbb{Z}_+$. Если $f \in C^s(E, F)$ и $g \in C^s(F, \mathbb{R}^r)$, то $g \circ f \in C^s(E, \mathbb{R}^r)$.*

Доказательство проводится индукцией по s. В случае $s = 0$ достаточно сослаться на теорему о непрерывности композиции. Предположим, что $s \in \mathbb{N}$ и для $s-1$ утверждение уже доказано. По лемме 2 отображения f и g s раз дифференцируемы (и, значит, дифференцируемы) соответственно на E и F. Пусть $k \in \{1, \ldots, n\}$. Воспользуемся формулами для частных производных композиции:
$$\partial_k(g \circ f)(x) = \sum_{j=1}^{m} \partial_j g\big(f(x)\big) \cdot \partial_k f_j(x) \quad (x \in E).$$
В силу замечания 5 к определению 3 и леммы 2 отображения f, $\partial_k f_j$ и $\partial_j g$ принадлежат классу C^{s-1}. Тогда по индукционному предположению $\partial_j g \circ f \in C^{s-1}(E, \mathbb{R}^r)$, и из теоремы 1$'$ вытекает

включение $\partial_k(g \circ f) \in C^{s-1}(E, \mathbb{R}^r)$. Еще раз применяя замечание 5 к определению 3, мы получим $g \circ f \in C^s(E, \mathbb{R}^r)$. $\square$

Перейдем к рассмотрению вопроса о симметрии смешанных производных. Поставим вначале задачу в простейшей ситуации. Пусть $E \subset \mathbb{R}^2$, a — внутренняя точка E, $f : E \to \mathbb{R}$. Договоримся переменные f обозначать через x и y. Предположим, что частные производные $\dfrac{\partial^2 f}{\partial x \partial y}(a)$ и $\dfrac{\partial^2 f}{\partial y \partial x}(a)$ имеют смысл. Будет ли они равны? Рассмотрим вначале два примера.

Пример 1. Положим

$$f(x, y) = x^y, \quad E = \{(x, y) \in \mathbb{R}^2 : x > 0\}.$$

Тогда

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= y\, x^{y-1}, \qquad & \frac{\partial^2 f}{\partial x \partial y}(x, y) &= x^{y-1} + y\, x^{y-1} \ln x, \\ \frac{\partial f}{\partial y}(x, y) &= x^y \ln x, \qquad & \frac{\partial^2 f}{\partial y \partial x}(x, y) &= \frac{x^y}{x} + y\, x^{y-1} \ln x \end{aligned}$$

для любых $(x, y) \in E$. Таким образом, смешанные производные f второго порядка совпадают на E.

Пример 2. Определим в $\mathbb{R}^2$ функцию f следующим образом:

$$f(x, y) = xy \cdot \frac{x^2 - y^2}{x^2 + y^2} \text{ при } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

Тогда при $(x, y) \neq (0, 0)$

$$\frac{\partial f}{\partial x}(x, y) = y \cdot \frac{x^4 - y^4 + 4x^2y^2}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y}(x, y) = x \cdot \frac{x^4 - y^4 - 4x^2y^2}{(x^2 + y^2)^2}$$

(предлагаем читателю проверить это). Кроме того,

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x} = 0 \quad \text{и} \quad \frac{\partial f}{\partial y}(0, 0) = 0.$$

Поэтому

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{y\to 0} \frac{\dfrac{\partial f}{\partial x}(0,y) - \dfrac{\partial f}{\partial x}(0,0)}{y} = \lim_{y\to 0} \frac{-y^5}{y^5} = -1.$$

Аналогичные вычисления показывают, что $\dfrac{\partial^2 f}{\partial y \partial x}(0,0) = 1$. Таким образом, производные $\dfrac{\partial^2 f}{\partial x \partial y}(0,0)$ и $\dfrac{\partial^2 f}{\partial y \partial x}(0,0)$ здесь различны.

Из примера 2 вытекает, что смешанные производные, вычисленные в различном порядке, могут не совпадать. Тем не менее оказывается, что у "хороших" функций они все-таки совпадают. Ниже мы увидим, что для этого достаточно либо непрерывности смешанных производных, либо двукратной дифференцируемости функции. Введем вначале одно вспомогательное понятие. Пусть $a = \big(x^0, y^0\big)$. Назовем *двойным приращением* f *в точке* a число

$$\Delta_a^2 f(u,v) = f\big(x^0+u, y^0+v\big) - f\big(x^0+u, y^0\big) - f\big(x^0, y^0+v\big) + f\big(x^0, y^0\big).$$

Смысл этого термина будет ясен из доказательства следующего утверждения.

Лемма 3. *Пусть $E \subset \mathbb{R}^2$, $a = \big(x^0, y^0\big)$ — внутренняя точка E, $f\colon E \to \mathbb{R}$, $\delta > 0$, частные производные f первого порядка определены на $V_a(\delta)$, $(u,v) \in V_{(0,0)}(\delta)$. Тогда справедливы следующие утверждения.*

1) *Существует число $\theta \in (0,1)$, для которого*

$$\Delta_a^2 f(u,v) = \left(\frac{\partial f}{\partial x}\big(x^0+\theta u, y^0+v\big) - \frac{\partial f}{\partial x}\big(x^0+\theta u, y^0\big)\right)\cdot u.$$

2) *Если производные $\dfrac{\partial^2 f}{\partial x \partial y}$ и $\dfrac{\partial^2 f}{\partial y \partial x}$ также определены на $V_a(\delta)$, то найдутся такие $\theta_x, \theta_y \in (0,1)$, что*

$$\Delta_a^2 f(u,v) = \frac{\partial^2 f}{\partial x \partial y}\big(x^0+\theta_x u, y^0+\theta_y v\big)\, uv.$$

Доказательство. Заметим, что

$$\Delta_a^2 f(u,v) = \varphi\bigl(x^0+u\bigr) - \varphi\bigl(x^0\bigr), \quad \text{где } \varphi(x) = f\bigl(x, y^0+v\bigr) - f\bigl(x, y^0\bigr)$$

(это объясняет смысл термина “двойное приращение”). Применяя к функции φ теорему Лагранжа, мы найдем $\theta \in (0,1)$, для которого

$$\Delta_a^2 f(u,v) = \varphi'\bigl(x^0+\theta u\bigr)\,u = \\ = \left(\frac{\partial f}{\partial x}\bigl(x^0+\theta u, y^0+v\bigr) - \frac{\partial f}{\partial x}\bigl(x^0+\theta u, y^0\bigr)\right)\cdot u,$$

что доказывает первое утверждение. Положим теперь $\theta_x = \theta$ и

$$\psi(y) = \frac{\partial f}{\partial x}\bigl(x^0+\theta_x u, y\bigr) \quad \text{при } \bigl|y-y^0\bigr| \leqslant |v|.$$

Функция ψ определена и дифференцируема на отрезке с концами y^0 и y^0+v. По теореме Лагранжа существует такое $\theta_y \in (0,1)$, что

$$\Delta_a^2 f(u,v) = \Bigl(\psi\bigl(y^0+v\bigr) - \psi\bigl(y^0\bigr)\Bigr)\,u = \\ = \psi'\bigl(y^0+\theta_y v\bigr)\,vu = \frac{\partial^2 f}{\partial x \partial y}\bigl(x^0+\theta_x u, y^0+\theta_y v\bigr)\,uv. \quad \square$$

Замечание. Меняя ролями x и y в доказательстве леммы 3, мы найдем такие числа $\tilde\theta$, $\tilde\theta_x$ и $\tilde\theta_y$ из $(0,1)$, что

$$\Delta_a^2 f(u,v) = \left(\frac{\partial f}{\partial y}\bigl(x^0+u, y^0+\tilde\theta v\bigr) - \frac{\partial f}{\partial y}\bigl(x^0, y^0+\tilde\theta v\bigr)\right)\cdot v,$$

$$\Delta_a^2 f(u,v) = \frac{\partial^2 f}{\partial y \partial x}\bigl(x^0+\tilde\theta_x u, y^0+\tilde\theta_y v\bigr)\,uv.$$

Перейдем теперь к задаче о симметрии смешанных производных функции двух переменных. Ключевым для ее решения будет следующее утверждение.

Лемма 4. *Пусть $E \subset \mathbb{R}^2$, a — внутренняя точка E, функция $f\colon E \to \mathbb{R}$ имеет в точке a смешанные производные второго порядка. Если при $(u,v) \to (0,0)$*

$$\Delta_a^2 f(u,v) = \frac{\partial^2 f}{\partial x \partial y}(a)\,uv + o\bigl(u^2+v^2\bigr),$$

$$\Delta_a^2 f(u,v) = \frac{\partial^2 f}{\partial y \partial x}(a)\,uv + o\bigl(u^2+v^2\bigr),$$

то справедливо равенство $\dfrac{\partial^2 f}{\partial x \partial y}(a) = \dfrac{\partial^2 f}{\partial y \partial x}(a)$.

Доказательство. Пусть $D = \dfrac{\partial^2 f}{\partial x \partial y}(a) - \dfrac{\partial^2 f}{\partial y \partial x}(a)$. Из условий леммы вытекает, что $D \cdot uv = o\big(u^2+v^2\big)$ при $(u,v) \to (0,0)$, то есть существует бесконечно малая в $(0,0)$ функция α, для которой

$$D\, uv = \alpha(u,v) \cdot \big(u^2+v^2\big).$$

Полагая $u = v = t$, мы получим

$$D = \frac{\alpha(t,t) \cdot 2t^2}{t^2} = 2\alpha(t,t) \to 0 \quad (t \to 0),$$

откуда $D = 0$. $\square$

Выведем теперь более удобные условия равенства смешанных производных функции двух переменных.

Теорема 3. *Пусть $E \subset \mathbb{R}^2$, a — внутренняя точка E, $f\colon E \to \mathbb{R}$ и выполнено одно из двух следующих условий.*

1) *Производные* $\dfrac{\partial^2 f}{\partial x \partial y}$ *и* $\dfrac{\partial^2 f}{\partial y \partial x}$ *определены в некоторой окрестности a и непрерывны в точке a.*

2) *Функция f дважды дифференцируема в точке a.*

Тогда $\dfrac{\partial^2 f}{\partial x \partial y}(a) = \dfrac{\partial^2 f}{\partial y \partial x}(a)$.

Доказательство будет вытекать из леммы 4. Проверку ее условий в каждом из двух случаев проведем по-своему.

1) Пусть $a = \big(x^0, y^0\big)$, а число $\delta > 0$ таково, что смешанные производные f определены на $V_a(\delta)$. В силу второго утверждения леммы 3 для $(u,v) \in V_a(\delta)$ найдутся такие $\theta_x, \theta_y \in (0,1)$, что

$$\Delta_a^2 f(u,v) = \frac{\partial^2 f}{\partial x \partial y}\big(x^0 + \theta_x u, y^0 + \theta_y v\big)\, uv = \frac{\partial^2 f}{\partial x \partial y}(a)\, uv + \alpha(u,v)\, uv,$$

где

$$\alpha(u,v) = \frac{\partial^2 f}{\partial x \partial y}\big(x^0 + \theta_x u, y^0 + \theta_y v\big) - \frac{\partial^2 f}{\partial x \partial y}\big(x^0, y^0\big).$$

Заметим, что $\alpha(u,v) \to 0$ при $(u,v) \to (0,0)$, поскольку

$$\left\|\left(x^0 + \theta_x u, y^0 + \theta_y v\right) - \left(x^0, y^0\right)\right\| = \sqrt{\theta_x^2 u^2 + \theta_y^2 v^2} \leqslant \|(u,v)\|,$$

а производная $\dfrac{\partial^2 f}{\partial x \partial y}$ непрерывна в точке a. Учитывая неравенство $|uv| \leqslant \frac{1}{2}\left(u^2 + v^2\right)$, мы получаем

$$\Delta_a^2 f(u,v) = \frac{\partial^2 f}{\partial x \partial y}(a)\, uv + o\left(u^2 + v^2\right) \quad \text{при} \ \ (u,v) \to (0,0).$$

Второе условие леммы 4 выводится аналогичным образом из замечания к лемме 3.

2) Будем не умаляя общности предполагать, что $a = (0,0)$, иначе заменим f на $f(\cdot + a)$. Кроме того, можно считать, что $d_a f = \mathbb{O}$. Действительно, функция $T = d_a f$ линейна, поэтому $\Delta_a^2 T \equiv 0$ и $\dfrac{\partial^2 T}{\partial x \partial y} \equiv 0$ на $\mathbb{R}^2$. Заменяя f на $f - T$, мы получим функцию с нулевым дифференциалом в точке a.

Проверим первое из условий леммы 4. Пусть число $\delta > 0$ таково, что частные производные f первого порядка определены на $V_a(\delta)$. В силу леммы 3 для $(u,v) \in V_a(\delta)$ найдется такое $\theta \in (0,1)$, что

$$\Delta_a^2 f(u,v) = \left(\frac{\partial f}{\partial x}(\theta u, v) - \frac{\partial f}{\partial x}(\theta u, 0)\right) \cdot u.$$

Так как функция $\dfrac{\partial f}{\partial x}$ дифференцируема в точке a и $\dfrac{\partial f}{\partial x}(a) = 0$,

$$\frac{\partial f}{\partial x}(\theta u, t) = \frac{\partial^2 f}{\partial x^2}(0,0)\, \theta u + \frac{\partial^2 f}{\partial x \partial y}(0,0)\, t + \alpha(\theta u, t) \cdot \|(\theta u, t)\|,$$

где α — бесконечно малая функция в точке $(0,0)$, $|t| \leqslant |v|$. Выразим отсюда разность $\dfrac{\partial f}{\partial x}(\theta u, v) - \dfrac{\partial f}{\partial x}(\theta u, 0)$ и подставим ее в предыдущую формулу. Мы получим

$$\Delta_a^2 f(u,v) = \frac{\partial^2 f}{\partial x \partial y}(a)\, uv + \beta(u,v) \cdot u,$$

где
$$\beta(u,v) = \alpha(\theta u, v)\sqrt{\theta^2u^2+v^2} - \alpha(\theta u, 0)\cdot|\theta u|.$$
Заметим, что
$$|\theta u|\cdot|u| \leqslant u^2 \leqslant u^2+v^2 \quad \text{и} \quad |u|\sqrt{\theta^2u^2+v^2} \leqslant u^2+v^2,$$
откуда
$$\beta(u,v)\cdot u = o\big(u^2+v^2\big) \quad \text{при } (u,v)\to(0,0).$$
Второе условие леммы 4 выводится аналогичным образом из замечания к лемме 3. $\square$

Следствие. *Если функция f дважды непрерывно дифференцируема в точке a, то* $\dfrac{\partial^2 f}{\partial x\partial y}(a) = \dfrac{\partial^2 f}{\partial y\partial x}(a)$.

Действительно, из двукратной непрерывной дифференцируемости f в точке a вытекает любое из двух условий теоремы 3.

Замечание 1. Ни одно из двух условий теоремы 3 не является необходимым для симметрии смешанных производных. Приведем пример, подтверждающий это. Положим $f(x,y) = \varphi(x)\cdot\varphi(y)$, где
$$\varphi(t) = t^2\sin\frac{1}{t^2} \text{ при } t\neq 0, \quad \varphi(0)=0.$$
Функция φ дифференцируема на $\mathbb{R}$, причем
$$\varphi'(t) = 2\left(t\sin\frac{1}{t^2} - \frac{1}{t}\cos\frac{1}{t^2}\right) \text{ при } t\neq 0, \quad \varphi'(0)=0.$$
Отсюда легко вытекает, что все частные производные f первого и второго порядка в точке $a=(0,0)$ равны нулю. Покажем, что ни одно из условий теоремы 3 не выполняется для f. Поскольку
$$\frac{\partial^2 f}{\partial x\partial y}\left(\frac{1}{\sqrt{2\pi k}}, \frac{1}{\sqrt{2\pi k}}\right) = \left(\varphi'\left(\frac{1}{\sqrt{2\pi k}}\right)\right)^2 = 8\pi k \to +\infty \quad (k\to\infty),$$
функция $\dfrac{\partial^2 f}{\partial x\partial y}$ разрывна в нуле. Проверим теперь, что $\dfrac{\partial f}{\partial x}$ не дифференцируема в точке a. Так как функция $\dfrac{\partial f}{\partial x}$ и ее частные

производные первого порядка равны нулю в точке a, дифференцируемость $\dfrac{\partial f}{\partial x}$ равносильна условию $\dfrac{\partial f}{\partial x}(x,y) = o\big(\|(x,y)\|\big)$ при $(x,y) \to (0,0)$. Это условие не выполняется, так как при $t \to 0+$

$$\frac{\dfrac{\partial f}{\partial x}(t,t)}{\|(t,t)\|} = \frac{\varphi(t)\,\varphi'(t)}{t\sqrt{2}} = \sqrt{2}\left(t^2 \sin^2 \frac{1}{t^2} - \sin\frac{1}{t^2}\cos\frac{1}{t^2}\right) \not\to 0.$$

Замечание 2. Условия 1) и 2) теоремы 3 независимы: ни одно из них не вытекает из другого. Действительно, пусть $a = (0,0)$ и

$$\varphi(t) = t^2 \sin\frac{1}{t} \ \text{ при } \ t \neq 0, \quad \varphi(0) = 0.$$

Предлагаем читателю проверить два утверждения.

1) Функция $f(x,y) = \varphi(x) + \varphi(y)$ не является дважды дифференцируемой в точке a, но имеет в ней непрерывные смешанные производные.

2) Функция $g(x,y) = \varphi(x)\,\varphi(y)$ дважды дифференцируема в точке a, но ее смешанные производные разрывны в a.

Рассмотрим теперь задачу о симметрии частных производных в полной общности. Для этого нам потребуются некоторые алгебраические понятия. Пусть $s \in \mathbb{N}$, $i \in I_s$. Если σ — перестановка множества $\{1,\ldots,s\}$, то семейство $\big(i_{\sigma(1)},\ldots,i_{\sigma(s)}\big)$ будет называться *перестановкой семейства* i*, порожденной* σ, и обозначаться символом i^σ. Перестановка σ называется *транспозицией*, если она удовлетворяет двум условиям:

1) Существует такое число $k \in \{2,\ldots,s\}$, что $\sigma(k-1) = k$ и $\sigma(k) = k-1$;

2) $\sigma(l) = l$ при всех l, отличных от $k-1$ и k.

Иными словами, транспозиция меняет местами два соседних элемента множества $\{1,\ldots,s\}$, а остальные отображает на себя. Перестановку семейства i, порожденную транспозицией, мы также будем называть транспозицией.

Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$, $s \in \mathbb{N}$, $i \in I_s$, σ — перестановка $\{1,\ldots,n\}$. В каком случае будет верно равенство $\partial_i f(a) = \partial_{i^\sigma} f(a)$? Иными словами, мы хотим получить достаточное условие, позволяющее менять порядок дифференцирований при вычислении частных производных высших порядков. Докажем утверждение, обобщающее теорему 3.

Теорема 4. Симметричность смешанных производных. *Пусть* $E \subset \mathbb{R}^n$, a *— внутренняя точка* E, $s \in \mathbb{N}$, *отображение* $f\colon E \to \mathbb{R}^m$ s *раз дифференцируемо в точке* a. *Тогда для любых* $i \in I_s$ *и перестановки* σ *множества* $\{1, \dots, s\}$ *справедливо равенство*

$$\partial_i f(a) = \partial_{i^\sigma} f(a). \tag{31}$$

Доказательство. Сделаем несколько замечаний, которые позволят свести соотношение (31) к теореме 3. Положим $j = i^\sigma$.

1) *Можно считать* $m = 1$. Предположим, что для $m = 1$ равенство (31) доказано. В силу замечания 4 к определению 1 все координатные функции f удовлетворяют условиям теоремы 4, поэтому $\partial_i f_k(a) = \partial_j f_k(a)$ для любых $k \in \{1, \dots, m\}$. Еще раз применяя замечание 4 к определению 1, мы получим (31).

2) *Можно считать перестановку* σ *транспозицией*. Действительно, любую перестановку множества $\{1, \dots, s\}$ можно реализовать как композицию конечного числа транспозиций.

3) *Достаточно рассмотреть случай* $s = 2$. Предположим, что для производных второго порядка равенство (31) доказано. Пусть транспозиция σ меняет местами $(k-1)$-й и k-й элементы i. Положим $F = \partial_{i_1 \dots i_{k-2}} f$. Напомним, что при $k = 2$ мы получим $F = f$ в силу замечания 2 к определению 1. Если $k < s$, то по замечанию 3 к определению 2 функция F дважды дифференцируема в некоторой окрестности V_a точки a. Тогда в силу соотношения (30) для любого $x \in V_a$

$$\begin{aligned} \partial_{i_1 \dots i_k} f(x) = \partial_{i_{k-1} i_k} F(x) = \partial_{i_k i_{k-1}} F(x) = \\ = \partial_{j_{k-1} j_k} \bigl(\partial_{j_1 \dots j_{k-2}} f\bigr)(x) = \partial_{j_1 \dots j_k} f(x) \end{aligned} \tag{32}$$

(во втором переходе мы использовали утверждение теоремы при $s = 2$). Дифференцируя равенство (32) по переменным $x_{i_{k+1}}, \dots, x_{i_s}$ и подставляя точку a, мы получим (31). В случае $k = s$ все преобразования в (32) остаются верными для $x = a$, что и дает (31).

4) *Достаточно рассмотреть случай* $n = 2$. Действительно, при вычислении производных $\partial_i f(a)$ и $\partial_j f(a)$ все переменные, кроме x_{i_1} и x_{i_2}, фиксированы, и можно считать, что f от них не зависит.

С учетом допущений 1) – 4) мы попадаем в условия теоремы 3, применение которой завершает доказательство. $\square$

Согласно определению 1, частные производные порядка $s \in \mathbb{N}$ задаются семействами индексов из I_s. Это не всегда удобно, так как в силу теоремы 4 разным семействам могут соответствовать одинаковые производные. Поэтому для записи частных производных используют также *мультииндексы*, введенные в § 7 главы 5. Делается это следующим образом. Пусть отображение f удовлетворяет условиям теоремы 4, $\alpha \in M_s$ (напомним, что символом M_s обозначается множество мультииндексов порядка s). Положим

$$\frac{\partial^\alpha f}{\partial x^\alpha}(a) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}(a). \tag{33}$$

Тогда $\dfrac{\partial^\alpha f}{\partial x^\alpha}(a)$ называется *частной производной f в точке a, порожденной мультииндексом α.* Заметим, что если $\alpha_k = 0$ для некоторого $k \in \{1, \ldots, n\}$, то по переменной x_k отображение f не дифференцируется ни разу. В частности, при $s = 0$ мы получим $\dfrac{\partial^\alpha f}{\partial x^\alpha}(a) = f(a)$, что соответствует определению частной производной нулевого порядка.

Покажем, что для любого $i \in I_s$ найдется такой мультииндекс $\alpha \in M_s$, что $\partial_i f(a) = \dfrac{\partial^\alpha f}{\partial x^\alpha}(a)$. Действительно, пусть в семействе i каждое число $k \in \{1, \ldots, n\}$ встречается ровно α_k раз. Переставим координаты i так, чтобы они шли по возрастанию. Мы получим семейство $j \in I_s$, для которого $\partial_j f(a) = \dfrac{\partial^\alpha f}{\partial x^\alpha}(a)$. С другой стороны, по теореме 4 $\partial_i f(a) = \partial_j f(a)$, откуда и вытекает искомое равенство. Напомним, что построенный таким образом мультииндекс α назывался в § 7 главы 5 *распределением* i и обозначался через $d(i)$.

Замечание. Проведенные рассуждения показывают, что в условиях теоремы 4

$$\text{если } i, j \in I_s,\ d(i) = d(j), \text{ то } \partial_i f(a) = \partial_j f(a). \tag{34}$$

Таким образом, производная $\partial_i f(a)$ зависит не от семейства $i \in I_s$, а только от его распределения.

Пример. Пусть $a \in \mathbb{R}^n$, α, β — мультииндексы. Тогда

$$\frac{\partial^\beta (x-a)^\alpha}{\partial x^\beta}\bigg|_{x=a} = \begin{cases} 0, & \beta \neq \alpha \\ \beta!, & \beta = \alpha. \end{cases} \tag{35}$$

Действительно, для $n = 1$ это доказано в главе 3. В общем случае

$$\frac{\partial^\beta (x-a)^\alpha}{\partial x^\beta} = \prod_{k=1}^{n} \frac{\partial^{\beta_k} (x_k - a_k)^{\alpha_k}}{\partial x_k^{\beta_k}}.$$

Если $\alpha = \beta$, то $\alpha_k = \beta_k$ при всех $k \in \{1, \dots, n\}$, и

$$\frac{\partial^\beta (x-a)^\alpha}{\partial x^\beta}\bigg|_{x=a} = \prod_{k=1}^{n} \beta_k! = \beta!\,.$$

Если же $\alpha_k \neq \beta_k$ для некоторого $k \in \{1, \dots, n\}$, то при $x = a$ k-й множитель в правой части равен нулю, поэтому нулевой будет и левая часть.

Частные производные, порожденные различными мультииндексами, никак не связаны друг с другом. Это будет вытекать из следующего утверждения.

Лемма 5. *Пусть $a \in \mathbb{R}^n$, $s \in \mathbb{Z}_+$, $\{c_\alpha\}$ — семейство векторов из $\mathbb{R}^m$, занумерованное мультииндексами порядка не выше s, и*

$$f(x) = \sum_{|\alpha| \leqslant s} \frac{c_\alpha}{\alpha!} (x-a)^\alpha \quad (x \in \mathbb{R}^n).$$

Тогда $\dfrac{\partial^\alpha f}{\partial x^\alpha}(a) = c_\alpha$ для всех мультииндексов α порядка не выше s.

Доказательство. Пусть β — мультииндекс порядка не выше s. Тогда в силу (35)

$$\frac{\partial^\beta f}{\partial x^\beta}(a) = \sum_{|\alpha| \leqslant s} \frac{c_\alpha}{\alpha!} \, \frac{\partial^\beta (x-a)^\alpha}{\partial x^\beta}\bigg|_{x=a} = \frac{c_\beta}{\beta!} \cdot \beta! = c_\beta. \quad \square$$

Замечание. Из леммы 5 вытекает единственность коэффициентов многочлена нескольких переменных, которая без доказательства утверждалась в § 7 главы 5. Действительно, записывая многочлен p порядка не выше s в виде

$$p(x) = \sum_{|\alpha| \leqslant s} c_\alpha \, x^\alpha,$$

по лемме 5 мы получим

$$c_\alpha = \alpha! \cdot \frac{\partial^\alpha p}{\partial x^\alpha}(0).$$

Поэтому единственность коэффициентов c_α вытекает из единственности частных производных.

Введем теперь *дифференциалы высших порядков*. Это будет сделано по индукции.

Определение 4. Дифференциалы высших порядков. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$.

1) Положим

$$d_a^0 f(h) = f(a) \quad \text{при всех} \;\; h \in \mathbb{R}^n. \tag{36}$$

Отображение $d_a^0 f$ называется *дифференциалом f нулевого порядка в точке a*. Для любого $h \in \mathbb{R}^n$ отображение $x \mapsto d_x^0 f(h)$, определенное по крайней мере в точке a, называется *дифференциалом f нулевого порядка на приращении h* и обозначается символом $d^0 f(h)$. Очевидно, что при всех $h \in \mathbb{R}^n$ отображения $d^0 f(h)$ и f совпадают во внутренних точках E.

2) Предположим, что $s \in \mathbb{N}$ и дифференциал f порядка $s-1$ уже определен. Пусть для любого $h \in \mathbb{R}^n$ отображение $d^{s-1} f(h)$ задано в некоторой окрестности a и дифференцируемо в точке a. Тогда положим

$$d_a^s f(h) = d_a\bigl(d^{s-1} f(h)\bigr)(h) \quad \text{при всех} \;\; h \in \mathbb{R}^n. \tag{37}$$

Отображение $d_a^s f$ называется *дифференциалом f порядка s в точке a*. Для любого $h \in \mathbb{R}^n$ отображение $x \mapsto d_x^s f(h)$, заданное по крайней мере в точке a, называется *дифференциалом f порядка s на приращении h* и обозначается символом $d^s f(h)$. Таким образом, формулы (36) и (37) индуктивно определяют дифференциал f произвольного порядка.

Замечание. При $s = 1$ мы получим

$$d_a^1 f(h) = d_a\big(d^0 f(h)\big)(h) = d_a f(h) \quad \text{для любых } h \in \mathbb{R}^n.$$

Таким образом, дифференциал f, введенный в § 2, совпадает с первым дифференциалом f. Поэтому вместо $d_a^1 f$ мы будем попрежнему использовать обозначение $d_a f$.

Перейдем к изучению дифференциалов высших порядков. Вначале мы обсудим вопрос об их существовании. Будет показано, что $d_a^s f$ определен для отображений f, s раз дифференцируемых в точке a. Кроме того, нужно вывести явную формулу для дифференциалов высших порядков, поскольку она нагляднее рекуррентного соотношения (37) и обычно удобнее для вычислительных целей. Мы получим эту формулу в двух вариантах, а также обсудим ее важнейшие частные случаи.

Теорема 5. Вычисление дифференциалов высших порядков. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $s \in \mathbb{Z}_+$, отображение $f\colon E \to \mathbb{R}^m$ s раз дифференцируемо в точке a. Тогда дифференциал f порядка s точке a существует и вычисляется по формулам*

$$d_a^s f(h) = \sum_{i \in I_s} \partial_i f(a)\, h_{i_1} \cdot \ldots \cdot h_{i_s} \quad \textit{для всех } h \in \mathbb{R}^n, \tag{38}$$

$$d_a^s f(h) = \sum_{\alpha \in M_s} \frac{s!}{\alpha!} \cdot \frac{\partial^\alpha f}{\partial x^\alpha}(a)\, h^\alpha \quad \textit{для всех } h \in \mathbb{R}^n. \tag{39}$$

Доказательство. Случай $s = 0$ тривиален, поэтому ограничимся натуральными s. Существование дифференциалов и равенство (38) мы докажем параллельно индукцией по s. Для $s = 1$ достаточно воспользоваться предыдущим замечанием и формулой (5). Предположим, что $s > 1$ и для дифференциалов порядка $s-1$ утверждения уже доказаны. По замечанию 3 к определению 2 найдется окрестность V_a точки a, в которой отображение f дифференцируемо $s-1$ раз. По индукционному предположению мы можем использовать равенство (38), заменив в нем s на $s-1$ и a на $x \in V_a$:

$$d_x^{s-1} f(h) = \sum_{i \in I_{s-1}} \partial_i f(x)\, h_{i_1} \cdot \ldots \cdot h_{i_{s-1}} \quad (h \in \mathbb{R}^n).$$

По определению 2 при любом $i \in I_{s-1}$ производная $\partial_i f$ дифференцируема в точке a. Поэтому для произвольного $h \in \mathbb{R}^n$ отображение $d^{s-1} f(h)$ дифференцируемо в точке a и

$$\begin{aligned} &d_a^s f(h) = d_a\big(d^{s-1} f(h)\big)(h) = \\ &= \sum_{k=1}^{n} \left(\sum_{i \in I_{s-1}} \partial_k\big(\partial_i f\big)(a)\, h_{i_1} \cdot \ldots \cdot h_{i_{s-1}} \cdot h_k \right) = \sum_{j \in I_s} \partial_j f(a)\, h_{j_1} \cdot \ldots \cdot h_{j_s}, \end{aligned}$$

где через j обозначен вектор $(i_1, \ldots, i_{s-1}, k)$. Тем самым индукционный переход завершен.

Осталось проверить формулу (39). В силу (38) отображение $d_a^s f$ представляет собой однородный многочлен степени s. Коэффициенты $d_a^s f$ удовлетворяют условию (34), которое для $d_a^s f$ совпадает с условием (26) главы 5. Напомним также, что если $i \in I_s$ и $\alpha = d(i)$, то $\partial_i f(a) = \dfrac{\partial^\alpha f}{\partial x^\alpha}(a)$. Поэтому формула (39) непосредственно вытекает из равенства (27) главы 5. $\square$

Замечание. Существование $d_a^s f$ на самом деле *эквивалентно* s-кратной дифференцируемости f в точке a. Мы не будем использовать это утверждение и потому оставляем его без доказательства.

Разберем подробнее два наиболее важных частных случая формул для дифференциалов высших порядков.

1) *Дифференциал второго порядка.* При $s = 2$ равенство (38) примет вид

$$d_a^2 f(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, h_i h_j \quad (h \in \mathbb{R}^n). \tag{40}$$

Матрица $\left\{\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)\right\}_{i,j=1}^{n}$, порождающая квадратичную форму $d_a^2 f$, называется *матрицей Гессе*. Применяя к $d_a^2 f$ формулу (30) главы 5, мы получим

$$d_a^2 f(h) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(a)\, h_i^2 + 2 \sum_{1 \leqslant i < j \leqslant n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, h_i h_j \quad (h \in \mathbb{R}^n). \quad (41)$$

Для вычислительных целей равенство (41) удобнее, чем (40), поскольку не содержит подобных слагаемых. Предлагаем читателю в качестве упражнения вывести (41) из (39).

2) *Случай двух переменных*. Если $n = 2$, то любой мультииндекс $\alpha \in M_s$ имеет вид $(k, s-k)$ при некотором $k \in \{0, \dots, s\}$. Тогда

$$\frac{s!}{\alpha!} = \frac{s!}{k!\,(s-k)!} = C_s^k.$$

Обозначая переменные f через x, y, а координаты h через dx, dy, можно записать (39) следующим образом:

$$d_a^s f(dx, dy) = \sum_{k=0}^{s} C_s^k \, \frac{\partial^s f}{\partial x^k \, \partial y^{s-k}}(a)\, dx^k \, dy^{s-k}. \quad (42)$$

§ 6. Формула Тейлора

В главе 3 был изучен вопрос о существовании и единственности у функции одной переменной многочлена Тейлора, хорошо приближающего ее вблизи некоторой точки. Теперь мы рассмотрим эту задачу для отображений нескольких переменных.

Пусть $f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m$, a — внутренняя точка E, $s \in \mathbb{Z}_+$. Мы хотим построить многочлен $p\colon \mathbb{R}^n \to \mathbb{R}^m$ степени не выше s, для которого

$$p(a) = f(a) \quad \text{и} \quad f(x) = p(x) + o\big(\|x-a\|^s\big) \ \ (x \to a). \quad (43)$$

Следуя главе 3, докажем сначала *единственность* такого многочлена.

Лемма 1. *Если многочлен p, удовлетворяющий условию (43), существует, то он единственный.*

Доказательство. Предположим, что многочлены p и q степени не выше s удовлетворяют условию (43). Положим $r = p - q$ и покажем, что $r \equiv 0$. Можно считать $a = 0$, иначе разложим p по степеням $x - a$ и заменим x на $x - a$ (см. равенство (22) главы 5). Так как $r(x) = o\big(\|x\|^s\big)$ при $x \to 0$, то

$$r(x) = \alpha(x) \cdot \|x\|^s, \quad \text{где } \alpha(x) \to 0 \ \ (x \to 0).$$

В соответствии с § 7 главы 5 разложим многочлен r на однородные компоненты:

$$r = \sum_{k=0}^{s} r_k, \quad \text{где } r_k \text{ — однородный многочлен степени } k.$$

Пусть $x \in \mathbb{R}^n$. Тогда в силу формулы (24) главы 5 для любого $t \in \mathbb{R}$

$$\sum_{k=0}^{s} r_k(x)\, t^k = \sum_{k=0}^{s} r_k(tx) = r(tx) = \alpha(tx)\, \|tx\|^s = |t|^s \cdot \alpha(tx)\, \|x\|^s .$$

В левой части этого равенства стоит векторнозначный многочлен относительно t степени не выше s, все координатные функции которого суть $o(t^s)$ при $t \to 0$ и, значит, тождественно равны нулю (см. главу 3). Отсюда следует, что $r_k(x) = 0$ при всех $k \in \{1, \dots, s\}$, то есть $r(x) = 0$. $\square$

Определение 1. Многочлен Тейлора. Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $f\colon E \to \mathbb{R}^m$, $s \in \mathbb{Z}_+$. Многочлен $p\colon \mathbb{R}^n \to \mathbb{R}^m$ степени не выше s, удовлетворяющий условию (43), называется *многочленом Тейлора f порядка s в точке a* и обозначается символом $T_{a,s}f$.

Замечание. Если многочлен Тейлора существует, то в силу леммы 1 он единственный.

Рассмотрим теперь вопрос о существовании многочлена Тейлора. Частичный ответ на этот вопрос мы уже получили: для $s = 0$

соотношение (43) равносильно непрерывности f в точке a, а при $s = 1$ — дифференцируемости f в точке a. Случай $s \geqslant 2$ исследуется сложнее. В главе 3 было показано, что для функций одной переменной существование $T_{a,s}f$ вытекает из s-кратной дифференцируемости f в точке a. Аналогичный результат будет справедлив и в многомерном случае. Для его доказательства нам потребуется вспомогательное утверждение.

Лемма 2. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $s \in \mathbb{N}$, отображение $f\colon E \to \mathbb{R}^m$ s раз дифференцируемо в точке a, причем f и все его частные производные порядка не выше s равны нулю в точке a. Тогда $f(a+h) = o\big(\|h\|^s\big)$ при $h \to 0$.*

Доказательство. Так как достаточно проверить утверждение для координатных функций f, можно считать $m = 1$. Доказательство проведем индукцией по s. Случай $s = 1$ тривиален, поскольку

$$f(a+h) = f(a) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a)\, h_k + o\big(\|h\|\big) = o\big(\|h\|\big) \quad (h \to 0).$$

Предположим, что для некоторого s утверждение уже доказано, и проверим его для $s+1$. Если функция f $s+1$ раз дифференцируема в точке a, то все частные производные f первого порядка s раз дифференцируемы в точке a. Применяя к ним индукционное предположение, мы получим

$$\frac{\partial f}{\partial x_k}(a+h) = o\big(\|h\|^s\big) \ \ (h \to 0) \quad \text{при всех} \ \ k \in \{1, \ldots, n\}.$$

По $\varepsilon > 0$ подберем такое $\delta > 0$, что для любого $k \in \{1, \ldots, n\}$

$$\left| \frac{\partial f}{\partial x_k}(a+h) \right| \leqslant \frac{\varepsilon}{n}\ \|h\|^s \quad \text{при всех} \ \ h \in V_0(\delta).$$

Пусть $h \in V_0(\delta)$. По лемме 1 § 4 найдутся такие $c^1, \ldots, c^n \in V_a(\delta)$, что

$$f(a+h) = f(a+h) - f(a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\big(c^k\big)\, h_k.$$

Тогда

$$|f(a+h)| \leqslant \sum_{k=1}^{n} \left| \frac{\partial f}{\partial x_k}\big(c^k\big) \right| \cdot \big|h_k\big| \leqslant \sum_{k=1}^{n} \frac{\varepsilon}{n}\ \|h\|^s \cdot \|h\| = \varepsilon \, \|h\|^{s+1} .$$

Поэтому $f(a+h) = o\big(\|h\|^{s+1}\big)$ при $h \to 0$. $\square$

Докажем теперь теорему о существовании многочлена Тейлора.

Теорема 1. Формула Тейлора – Пеано. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $s \in \mathbb{Z}_+$, отображение $f\colon E \to \mathbb{R}^m$ s раз дифференцируемо в точке a. Тогда справедливо равенство*

$$f(x) = \sum_{|\alpha| \leqslant s} \frac{1}{\alpha!} \cdot \frac{\partial^\alpha f}{\partial x^\alpha}(a)\,(x-a)^\alpha + o\big(\|x-a\|^s\big) \quad \textit{при}\ \ x \to a.$$

Иными словами, многочлен Тейлора f порядка s в точке a существует и вычисляется по формуле

$$T_{a,s} f(x) = \sum_{|\alpha| \leqslant s} \frac{1}{\alpha!} \cdot \frac{\partial^\alpha f}{\partial x^\alpha}(a)\,(x-a)^\alpha. \tag{44}$$

Суммирование здесь ведется по всем мультииндексам порядка не выше s.

Доказательство. Положим

$$p(x) = \sum_{|\alpha| \leqslant s} \frac{1}{\alpha!} \cdot \frac{\partial^\alpha f}{\partial x^\alpha}(a)\,(x-a)^\alpha, \quad F = f - p.$$

Если α — мультииндекс порядка не выше s, то по лемме 5 § 5

$$\frac{\partial^\alpha F}{\partial x^\alpha}(a) = \frac{\partial^\alpha f}{\partial x^\alpha}(a) - \frac{\partial^\alpha p}{\partial x^\alpha}(a) = 0.$$

Применяя к F лемму 2, мы получим

$$f(x) - p(x) = F(x) = F(a + x - a) = o\big(\|x-a\|^s\big) \quad (x \to a). \quad \square$$

Замечание 1. Отображение $R_{a,s}f = f - T_{a,s}f$ называют *остатком в формуле Тейлора*. Утверждение теоремы 1 записывают также следующим образом:

$$R_{a,s}f(x) = o\big(\|x-a\|^s\big) \quad \text{при} \ \ x \to a.$$

Замечание 2. Многочлен Тейлора можно выразить через дифференциалы высших порядков:

$$T_{a,s}f(x) = \sum_{k=0}^{s} \frac{d_a^k f(x-a)}{k!}. \tag{45}$$

Действительно, в силу формулы (39)

$$T_{a,s}f(x) = \sum_{k=0}^{s} \frac{1}{k!} \sum_{\alpha \in M_k} \frac{k!}{\alpha!} \cdot \frac{\partial^\alpha f}{\partial x^\alpha}(a)\,(x-a)^\alpha = \sum_{k=0}^{s} \frac{d_a^k f(x-a)}{k!}.$$

Замечание 3. Если отображение f есть многочлен степени не выше s, то для него $R_{a,s}f \equiv 0$, то есть

$$f(x) = \sum_{|\alpha| \leqslant s} \frac{1}{\alpha!} \cdot \frac{\partial^\alpha f}{\partial x^\alpha}(a)\,(x-a)^\alpha.$$

Это равенство, называемое *формулой Тейлора для многочленов*, является другой формой записи леммы 5 § 5.

Замечание 4. Из замечания 3 вытекает полиномиальная формула Ньютона (теорема 2 § 7 главы 5). Действительно, функция

$$f(x) = (x_1 + \ldots + x_n)^s \quad (x \in \mathbb{R}^n)$$

является однородным многочленом степени s, а все ее частные производные порядка s тождественно равны $s!$. Осталось сослаться на формулу Тейлора для многочленов.

Многочлен Тейлора $T_{a,s}f$ хорошо приближает отображение f вблизи точки a. Тем не менее, из теоремы 1 нельзя извлечь информации о малости остатка $R_{a,s}f(x)$ при *фиксированном* x. Чтобы формулу Тейлора можно было использовать в приближенных вычислениях, желательно получить какую-то оценку $R_{a,n}f(x)$. Для этого нам потребуется некоторая подготовка. Положим

$$\Delta_{a,b} = \{a + t\,(b-a) : t \in [0,1]\}, \quad \widetilde{\Delta}_{a,b} = \Delta_{a,b} \setminus \{a, b\}. \tag{46}$$

Напомним, что $\Delta_{a,b}$ мы называли в § 5 главы 5 отрезком с концами a и b. Поэтому множество $\widetilde{\Delta}_{a,b}$ называется *интервалом с концами a и b*.

Лемма 3. *Пусть множество E открыто в $\mathbb{R}^n$, $a \in E$, $h \in \mathbb{R}^n$, причем $\Delta_{a,a+h} \subset E$. Предположим, что $s \in \mathbb{N}$ и функция $f\colon E \to \mathbb{R}$ s раз дифференцируема на E. Тогда функция $F(t) = f(a+th)$ s раз дифференцируема на $[0,1]$ и*

$$F^{(k)}(t) = d^k_{a+th} f(h) \quad \textit{при всех} \ \ k \in \{0, \ldots, s\} \ \ \textit{и} \ \ t \in [0,1]. \tag{47}$$

Доказательство. В случае $k = 0$ утверждение очевидно, так как

$$F(t) = f(a+th) = d^0_{a+th} f(h).$$

Предположим, что $k \in \{1, \ldots, s\}$ и для $k-1$ искомое равенство уже доказано. По теореме 5 § 5 функция $d^{k-1}f(h)$ дифференцируема на E. Кроме того, вектор-функция φ, определяемая равенством $\varphi(t) = a + th$, дифференцируема на $[0,1]$, причем

$$d_t\varphi(dt) = h\,dt \quad \text{при всех} \ \ t \in [0,1] \ \ \text{и} \ \ dt \in \mathbb{R}.$$

По индукционному предположению

$$F^{(k-1)}(t) = d^{k-1}_{a+th} f(h) = d^{k-1} f(h) \circ \varphi.$$

Тогда по теореме 1 § 3 функция $F^{(k-1)}$ дифференцируема на $[0,1]$ и для любого $t \in [0,1]$

$$\begin{aligned} F^{(k)}(t) = d_t F^{(k-1)}(1) = d_{\varphi(t)}\big(d^{k-1}f(h)\big)\big(d_t\varphi(1)\big) = \\ = d_{a+th}\big(d^{k-1}f(h)\big)(h) = d^k_{a+th} f(h). \end{aligned} \tag{48}$$

Тем самым индукционный переход доказан. $\square$

Полезно также сформулировать локальный вариант леммы 3.

Следствие. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, $h \in \mathbb{R}^n$, $s \in \mathbb{N}$, функция $f: E \to \mathbb{R}$ s раз дифференцируема в точке a. Тогда функция $F(t) = f(a + th)$ s раз дифференцируема в нуле и*

$$F^{(k)}(0) = d_a^k f(h) \quad \textit{при всех } k \in \{0, \ldots, s\}.$$

Доказательство. В силу s-кратной дифференцируемости f в точке a существует такое $\delta > 0$, что функция f $s - 1$ раз дифференцируема на $V_a(\delta)$. Пусть вначале $\|h\| < \delta$. По лемме 3

$$F^{(k)}(t) = d_{a+th}^k f(h) \quad \text{при всех } k \in \{0, \ldots, s-1\} \text{ и } t \in [0,1].$$

Подставляя $t = 0$, мы получим требуемые равенства для $k \leqslant s - 1$. Так как функция $d^{s-1}f(h)$ дифференцируема в точке a, преобразования (48) будут верны для $t = 0$ и $k = s$, откуда $F^{(s)}(0) = d_a^s f(h)$.

Откажемся теперь от условия $\|h\| < \delta$. Положим

$$\lambda = \frac{\delta}{2\,\|h\|}, \qquad G(t) = f(a + t\lambda h) = F(t\lambda) \text{ при } t \in [0,1].$$

Тогда $\|\lambda h\| < \delta$, и по доказанному

$$\lambda^k F^{(k)}(0) = G^{(k)}(0) = d_a^k f(\lambda h) = \lambda^k d_a^k f(h).$$

Осталось сократить это равенство на λ^k. $\square$

Перейдем к глобальному варианту формулы Тейлора. Вначале мы докажем его для *функций* нескольких переменных. Случай отображений будет рассмотрен позже.

Теорема 2. Формула Тейлора – Лагранжа для функций. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{Z}_+$, функция $f: E \to \mathbb{R}$ $s+1$ раз дифференцируема на E, $a, x \in E$, причем $a \neq x$ и $\Delta_{a,x} \subset E$. Тогда существует точка $c \in \widetilde{\Delta}_{a,x}$, для которой*

$$f(x) = \sum_{k=0}^{s} \frac{d_a^k f(x-a)}{k!} + \frac{d_c^{s+1} f(x-a)}{(s+1)!}.$$

Напомним, что $\Delta_{a,x}$ и $\widetilde{\Delta}_{a,x}$ определяются формулами (46).

Замечание. Утверждение теоремы 2 можно записать в виде

$$R_{a,s}f(x) = \frac{d_c^{s+1} f(x-a)}{(s+1)!}.$$

Иными словами, найдется такое $\theta \in (0,1)$, что

$$R_{a,s}f(x) = \frac{d_{a+\theta(x-a)}^{s+1} f(x-a)}{(s+1)!}. \tag{49}$$

Таким образом, остаток в формуле Тейлора напоминает по виду $(s+1)$-е слагаемое в правой части (45), но дифференциал вычисляется не в точке a, а в некоторой средней точке, лежащей на интервале $\widetilde{\Delta}_{a,x}$.

Доказательство. Положим

$$h = x - a, \quad F(t) = f\bigl(a + th\bigr) \ \text{при} \ t \in [0,1].$$

По лемме 3 функция F $s+1$ раз дифференцируема на $[0,1]$, а ее производные задаются равенствами (47). Применяя к F формулу Тейлора – Лагранжа, доказанную в главе 3, мы найдем $\theta \in (0,1)$, для которого

$$F(1) = \sum_{k=0}^{s} \frac{F^{(k)}(0)}{k!}\,(1-0)^k + \frac{F^{(s+1)}(\theta)}{(s+1)!}\,(1-0)^{s+1},$$

то есть

$$f(x) = \sum_{k=0}^{s} \frac{d_a^k f(x-a)}{k!} + \frac{d_{a+\theta h}^{s+1} f(x-a)}{(s+1)!}.$$

Осталось положить $c = a + \theta h$. $\square$

Следствие 1. *В условиях теоремы 2 справедлива оценка*

$$|R_{a,s}f(x)| \leqslant \sup_{c \in \widetilde{\Delta}_{a,x}} \frac{|d_c^{s+1} f(x-a)|}{(s+1)!}.$$

Обобщим теперь на многомерный случай теорему Лагранжа о среднем.

Следствие 2. Теорема о среднем. *Пусть множество E открыто в $\mathbb{R}^n$, функция $f\colon E \to \mathbb{R}$ дифференцируема на E, $a, b \in E$, причем $a \neq b$ и $\Delta_{a,b} \subset E$. Тогда существует точка $c \in \widetilde{\Delta}_{a,b}$, для которой*

$$f(b) - f(a) = d_c f(b - a).$$

Это утверждение является частным случаем теоремы 2, соответствующим $s = 0$.

Перейдем к рассмотрению отображений нескольких переменных. Оказывается, что для них неверны и теорема 2, и даже следствие 2. Проиллюстрируем это на примере.

Пример. Пусть вектор-функция $f\colon \mathbb{R} \to \mathbb{R}^2$ действует по формуле

$$f(x) = (\cos x, \sin x) \quad (x \in \mathbb{R}).$$

Тогда отображение f дифференцируемо на $\mathbb{R}$ и

$$f'(x) = (-\sin x, \cos x) \quad \text{при всех} \ \ x \in \mathbb{R}.$$

Положим $a = 0$, $b = 2\pi$. Для них утверждение теоремы о среднем неверно. Действительно, $f(b) - f(a) = (0, 0)$, но при любом $c \in (a, b)$

$$\|d_c f(b - a)\| = \|f'(c)\| \cdot (b - a) = 2\pi\sqrt{\sin^2 c + \cos^2 c} = 2\pi \neq 0.$$

Хотя теорема 2 для отображений неверна, вытекающее из нее следствие 1 останется справедливым и для отображений, если заменить в его формулировке модули на нормы. В этом и заключается общий вариант формулы Тейлора – Лагранжа. Для его доказательства нам потребуется вспомогательное утверждение.

Лемма 4. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{Z}_+$, отображение $f\colon E \to \mathbb{R}^m$ s раз дифференцируемо на E, $H \in \mathbb{R}^m$. Тогда функция $H \cdot f$ s раз дифференцируема на E и*

$$d_a^k(H \cdot f) = H \cdot d_a^k f \quad \textit{при всех} \ \ a \in E \ \ \textit{и} \ \ k \in \{0, \ldots, s\}.$$

Доказательство проведем индукцией по k. Случай $k = 0$ тривиален. Предположим, что для некоторого $k < s$ утверждение доказано. Тогда для любого $h \in \mathbb{R}^n$

$$
\begin{aligned}
d_a^{k+1}(H \cdot f)(h) &= d_a\bigl(d^k(H \cdot f)(h)\bigr)(h) = \\
&= d_a\bigl(H \cdot d^k f(h)\bigr)(h) = H \cdot d_a\bigl(d^k f(h)\bigr)(h) = H \cdot d_a^{k+1} f(h).
\end{aligned}
$$

Мы воспользовались индукционным предположением и утверждением 3) теоремы 2 § 3. $\square$

Теорема 3. Формула Тейлора – Лагранжа для отображений. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{Z}_+$, отображение $f\colon E \to \mathbb{R}^m$ $s+1$ раз дифференцируемо на E, $a, x \in E$, причем $a \neq x$ и $\Delta_{a,x} \subset E$. Тогда*

$$
\left\| f(x) - \sum_{k=0}^{s} \frac{d_a^k f(x-a)}{k!} \right\| \leqslant \frac{1}{(s+1)!} \cdot \sup_{c \in \widetilde{\Delta}_{a,x}} \left\| d_c^{s+1} f(x-a) \right\|.
$$

Замечание. Утверждение теоремы можно записать так:

$$
\|R_{a,s} f(x)\| \leqslant \frac{1}{(s+1)!} \cdot \sup_{c \in \widetilde{\Delta}_{a,x}} \left\| d_c^{s+1} f(x-a) \right\|. \tag{50}
$$

Доказательство. Для $H \in \mathbb{R}^m$ положим

$$
F = H \cdot f, \quad h = x - a, \quad M = \sup_{c \in \widetilde{\Delta}_{a,x}} \left\| d_c^{s+1} f(h) \right\|.
$$

По лемме 4 функция F $s+1$ раз дифференцируема на E. Применяя к ней теорему 2, мы найдем такое $c \in \widetilde{\Delta}_{a,x}$ (зависящее от H), что

$$
F(x) = \sum_{k=0}^{s} \frac{d_a^k F(h)}{k!} + \frac{d_c^{s+1} F(h)}{(s+1)!},
$$

откуда в силу леммы 4

$$
H \cdot f(x) = \sum_{k=0}^{s} H \cdot \frac{d_a^k f(h)}{k!} + H \cdot \frac{d_c^{s+1} f(h)}{(s+1)!}.
$$

Тогда по неравенству Коши

$$|H \cdot R_{a,s}f(x)| =$$
$$= \left| H \cdot \left(f(x) - \sum_{k=0}^{s} \frac{d_a^k f(h)}{k!} \right) \right| = \left| H \cdot \frac{d_c^{s+1} f(h)}{(s+1)!} \right| \leqslant \frac{M}{(s+1)!} \, \|H\| \, .$$

Полагая здесь $H = R_{a,s}f(x)$, мы получим

$$\|R_{a,s}f(x)\|^2 \leqslant \frac{M}{(s+1)!} \, \|R_{a,s}f(x)\| \, .$$

Если $R_{a,s}f(x) \neq 0$, то достаточно сократить это неравенство на $\|R_{a,s}f(x)\|$, иначе утверждение теоремы очевидно. $\square$

Следствие. Оценка конечных приращений. *Пусть множество E открыто в $\mathbb{R}^n$, отображение $f\colon E \to \mathbb{R}^m$ дифференцируемо на E, $a, b \in E$, причем $a \neq b$ и $\Delta_{a,b} \subset E$. Тогда*

$$\|f(b) - f(a)\| \leqslant \sup_{c \in \widetilde{\Delta}_{a,b}} \|d_c f\| \cdot \|x - a\| \, .$$

Это утверждение непосредственно вытекает из теоремы 3 в случае $s = 0$.

В условиях теоремы 3 неясно, будет ли правая часть (50) конечна. Мы сейчас покажем, что ответ на этот вопрос положителен для отображений класса $C^{s+1}(E, \mathbb{R}^m)$. Кроме того, будет получена оценка остатка в терминах частных производных высших порядков.

Теорема 4. Оценка остатка в формуле Тейлора. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{Z}_+$, $f \in C^{s+1}(E, \mathbb{R}^m)$, $a, x \in E$, причем $a \neq x$ и $\Delta_{a,x} \subset E$. Положим*

$$M = \max_{\alpha \in M_{s+1}} \sup_{c \in \widetilde{\Delta}_{a,x}} \left\| \frac{\partial^\alpha f}{\partial x^\alpha}(c) \right\| .$$

Тогда $M < +\infty$ и справедливо неравенство

$$\|R_{a,s}f(x)\| \leqslant \frac{M}{(s+1)!} \left(\sqrt{n} \, \|x - a\| \right)^{s+1} . \tag{51}$$

Доказательство. В § 5 главы 5 было показано, что множество $\Delta_{a,x}$ компактно в $\mathbb{R}^n$. Все частные производные f порядка $s+1$ непрерывны на E и, в силу теоремы Вейерштрасса, ограничены на $\Delta_{a,x}$. Тем самым доказано, что $M < +\infty$. Пусть $h = x - a$. Используя формулу (39) и неравенство (28) главы 5, мы получим для $c \in \widetilde{\Delta}_{a,x}$

$$\left\| d_c^{s+1} f(h) \right\| \leqslant \max_{\alpha \in M_{s+1}} \left\| \frac{\partial^\alpha f}{\partial x^\alpha}(c) \right\| \left(\sqrt{n}\, \|h\| \right)^{s+1} \leqslant M \left(\sqrt{n}\, \|h\| \right)^{s+1},$$

откуда в силу (50) вытекает (51). $\square$

Замечание 1. Запишем явным образом случай $s = 0$:

$$\|f(x) - f(a)\| \leqslant \sup_{i \in \{1, \ldots, n\}} \sup_{c \in \widetilde{\Delta}_{a,x}} \left\| \frac{\partial f}{\partial x_i}(c) \right\| \cdot \sqrt{n}\, \|x - a\|.$$

Замечание 2. Если в условиях теоремы 1 дополнительно потребовать s-кратную непрерывную дифференцируемость f в точке a, то формулу Тейлора – Пеано можно вывести из теоремы 2. Покажем, как это делается. Будем снова считать, что $m = 1$ и $s > 0$. Пусть $\varepsilon > 0$. Так как все частные производные f порядка s непрерывны в точке a, существует такое $\delta > 0$, что

$$\left| \frac{\partial^\alpha f}{\partial x^\alpha}(x) - \frac{\partial^\alpha f}{\partial x^\alpha}(a) \right| < \frac{\varepsilon \cdot s!}{n^{s/2}} \quad \text{для всех } \alpha \in M_s \text{ и } x \in V_a(\delta).$$

Пусть $x \in V_a(\delta)$, $h = x - a$. По теореме 2 найдется $c \in \widetilde{\Delta}_{a,x}$, для которого

$$f(x) = \sum_{k=0}^{s-1} \frac{d_a^k f(h)}{k!} + \frac{d_c^s f(h)}{s!} = T_{a,s} f(x) + r(x),$$

где

$$r(x) = \frac{d_c^s f(h) - d_a^s f(h)}{s!} = \sum_{\alpha \in M_s} \frac{1}{\alpha!} \left(\frac{\partial^\alpha f}{\partial x^\alpha}(c) - \frac{\partial^\alpha f}{\partial x^\alpha}(a) \right) h^\alpha$$

(мы воспользовались формулой (39)). В силу оценки (28) главы 5

$$|r(x)| \leqslant \max_{\alpha \in M_{s+1}} \left| \frac{\partial^\alpha f}{\partial x^\alpha}(c) - \frac{\partial^\alpha f}{\partial x^\alpha}(a) \right| \frac{n^{s/2}}{s!} \|h\|^s < \varepsilon \|h\|^s .$$

Поэтому $r(x) = o\big(\|h\|^s\big)$ при $h \to 0$. $\square$

В условиях теорем 2 – 4 мы требовали, чтобы вместе с точками a и x в множестве E лежал соединяющий их отрезок. Формулировки этих теорем можно упростить, если отображение f задано на *выпуклом* множестве.

Определение 2. Выпуклое множество. Пусть $E \subset \mathbb{R}^n$. Множество E называется *выпуклым*, если для любых $a, b \in E$ отрезок $\Delta_{a,b}$, определенный формулой (46), содержится в E.

Приведем два примера выпуклых множеств.

Пример 1. *Если $c \in \mathbb{R}^n$ и $r > 0$, то $V_c(r)$ и $\overline{V_c(r)}$ — выпуклые множества.* Докажем это, например, для открытых шаров. Пусть $a, b \in V_c(r)$. Тогда при всех $t \in [0, 1]$

$$\|a + t(b-a) - c\| = \|(1-t)(a-c) + t(b-c)\| \leqslant$$
$$\leqslant (1-t)\|a-c\| + t\|b-c\| < (1-t)r + tr = r,$$

то есть $\Delta_{a,b} \subset V_c(r)$. $\square$

Пример 2. *Если $a, b \in \mathbb{R}^n$, то параллелепипеды вида $\langle a, b\rangle$, определенные в примере 5 § 3 главы 5, выпуклы.* Проверку этого утверждения мы оставляем читателю.

Замечание 1. Если множество E предполагать выпуклым, то теоремы 2 – 4 будут верны для любой пары различных точек $a, x \in E$.

Замечание 2. *Пусть E открыто в $\mathbb{R}^n$, $f \in C^1(E, \mathbb{R}^m)$, множество $F \subset E$ компактно и выпукло. Тогда существует такое $C > 0$, что*

$$\|f(b) - f(a)\| \leqslant C\|b-a\| \quad \text{при всех } a, b \in F. \tag{52}$$

Это вытекает из замечания 1 к теореме 4, поскольку по теореме Вейерштрасса все частные производные f первого порядка ограничены на F. Отображения, удовлетворяющие условию (52) при некотором $C > 0$, называются *липшицевыми на множестве F*.

§ 7. Локальная обратимость гладкого отображения

В этом параграфе мы изучим вопрос о разрешимости нелинейных уравнений и систем. Рассмотрим вначале случай уравнений, поскольку он технически проще и позволит читателю лучше понять постановку задачи.

Пусть $\langle A, B\rangle \subset \mathbb{R}$, $f\colon \langle A, B\rangle \to \mathbb{R}$, $y \in \mathbb{R}$. Рассмотрим уравнение относительно переменной x

$$f(x) = y, \quad x \in \langle A, B\rangle. \tag{53}$$

Оно разрешимо тогда и только тогда, когда $y \in f\big(\langle A, B\rangle\big)$. Разумеется, решение может не быть единственным. Достаточным условием *однозначной* разрешимости уравнения (53) является инъективность f, но необходимым оно не будет. Предлагаем читателю подтвердить сказанное примерами.

Во многих приложениях однозначная разрешимость (53) нужна не на всем $\langle A, B\rangle$, а только вблизи некоторой точки. Для изучения этой задачи можно использовать средства дифференциального исчисления. Предположим, что $f \in C^1\big(\langle A, B\rangle\big)$ и $a \in (A, B)$. Верно ли, что существует окрестность точки a, в которой уравнение (53) имеет не более одного решения? Приведем пример, показывающий, что это не так. Пусть

$$f(x) = x^3 \sin\frac{\pi}{x} \ \text{ при } \ x \neq 0 \quad \text{и} \quad f(0) = 0.$$

Предлагаем читателю проверить, что $f \in C^1(\mathbb{R})$ и $f'(0) = 0$. Заметим, однако, что $f\left(\frac{1}{k}\right) = 0$ при всех $k \in \mathbb{N}$, то есть уравнение $f(x) = 0$ имеет бесконечно много решений в любой окрестности нуля. Тем не менее мы сейчас докажем, что ответ на поставленный вопрос положителен при условии $f'(a) \neq 0$.

Теорема 1. Локальная обратимость гладкой функции. *Пусть $\langle A, B\rangle \subset \mathbb{R}$, $f \in C^1\big(\langle A, B\rangle\big)$, $a \in (A, B)$, $f'(a) \neq 0$. Тогда существует $\delta > 0$, для которого справедливы следующие утверждения.*

1) $f'(x) \neq 0$ *при всех* $x \in V_a(\delta)$.

2) *Функция* $f\big|_{V_a(\delta)}$ *взаимно однозначна.*

3) *Множество* $f\bigl(V_a(\delta)\bigr)$ *открыто в* $\mathbb{R}$.

4) $\Bigl(f\big|_{V_a(\delta)}\Bigr)^{-1} \in C^1\bigl(f\bigl(V_a(\delta)\bigr)\bigr)$.

Доказательство. Пусть для определенности $f'(a) > 0$. Так как f' непрерывна на (A, B), найдется такое $\delta > 0$, что

$$[a-\delta, a+\delta] \subset \langle A, B\rangle \quad \text{и} \quad f'(x) > 0 \ \text{ при всех } \ x \in V_a(\delta).$$

Покажем, что это δ — искомое. Действительно, утверждение 1) очевидно. По теореме 1 § 7 главы 3 функция f строго возрастает на $[a-\delta, a+\delta]$. Отсюда вытекают утверждения 2) и 3), поскольку $f\big|_{V_a(\delta)}$ взаимно однозначна и $f\bigl((a-\delta, a+\delta)\bigr) = \bigl(f(a-\delta), f(a+\delta)\bigr)$. Осталось проверить 4). Пусть $g = \Bigl(f\big|_{V_a(\delta)}\Bigr)^{-1}$. По теореме 9 § 4 главы 2 функция g непрерывна на $f\bigl(V_a(\delta)\bigr)$. По правилу дифференцирования обратной функции (см. § 3 главы 3)

$$g'(y) = \frac{1}{f'\bigl(g(y)\bigr)} \quad \text{для любых } \ y \in \bigl(f(a-\delta), f(a+\delta)\bigr).$$

Заметим, что $f' \circ g$ непрерывна на $f\bigl(V_a(\delta)\bigr)$ по теореме о непрерывности композиции и $f'\bigl(g(y)\bigr) \neq 0$ при всех $y \in f\bigl(V_a(\delta)\bigr)$. Поэтому функция g' непрерывна на $f\bigl(V_a(\delta)\bigr)$, то есть $g \in C^1\bigl(f\bigl(V_a(\delta)\bigr)\bigr)$. $\square$

Замечание 1. Без условия $f'(a) \neq 0$ утверждение теоремы неверно (рассмотрите функцию $f(x) = x^3$ и $a = 0$).

Замечание 2. Если в условиях теоремы 1 дополнительно предположить, что $s \in \mathbb{N}$ и $f \in C^s(\langle A, B\rangle)$, то мы получим включение $\Bigl(f\big|_{V_a(\delta)}\Bigr)^{-1} \in C^s\bigl(f\bigl(V_a(\delta)\bigr)\bigr)$. Таким образом, локально обратная к f функция имеет тот же класс гладкости, что и f. Это утверждение является частным случаем теоремы 5, которая будет доказана ниже.

Замечание 3. Утверждение теоремы 1 можно истолковать следующим образом:

1) *существуют такие* $\varepsilon, \delta > 0$, *что для любого* $y \in V_{f(a)}(\varepsilon)$ *уравнение (53) имеет на* $V_a(\delta)$ *единственное решение* $x(y)$;

2) *это решение гладко зависит от правой части уравнения, то есть отображение* $y \mapsto x(y)$ *принадлежит классу* $C^1\big(V_{f(a)}(\varepsilon)\big)$.

Действительно, выберем число δ так же, как в теореме 1. Поскольку $f\big(V_a(\delta)\big)$ открыто, найдется такое $\varepsilon > 0$, что $V_{f(a)}(\varepsilon) \subset f\big(V_a(\delta)\big)$. Если $y \in V_{f(a)}(\varepsilon)$, то $y \in f\big(V_a(\delta)\big)$, поэтому уравнение (53) имеет решение на $V_a(\delta)$. Единственность этого решения вытекает из инъективности $f\big|_{V_a(\delta)}$. Второе утверждение справедливо в силу гладкости $\Big(f\big|_{V_a(\delta)}\Big)^{-1}$. $\square$

Перейдем к изучению систем нелинейных уравнений. Пусть $E \subset \mathbb{R}^n$, $f_1, \ldots, f_n \colon E \to \mathbb{R}$, $y \in \mathbb{R}^n$. Рассмотрим систему уравнений относительно $x_1, \ldots, x_n$

$$\begin{cases} f_1(x_1, \ldots, x_n) = y_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_n(x_1, \ldots, x_n) = y_n. \end{cases} \tag{54}$$

Полагая $f = \big(f_1, \ldots, f_n\big)$, мы можем записать ее в виде векторного уравнения:

$$f(x) = y, \quad x \in E. \tag{55}$$

Верны ли утверждения, аналогичные теореме 1 и замечанию 3, для уравнений вида (55)? Чтобы ответить на этот вопрос, необходимо найти правильный многомерный аналог условия $f'(a) \neq 0$. Пусть a — внутренняя точка E, отображение f дифференцируемо в точке a. Запишем (55) в виде

$$f(a) + f'(a) \cdot (x - a) + o(x - a) = y.$$

Если отбросить в левой части слагаемое $o(x - a)$, мы получим *линейную* систему

$$f'(a) \cdot (x - a) = y - f(a).$$

Она однозначно разрешима тогда и только тогда, когда матрица $f'(a)$ обратима, то есть в случае $\det f'(a) \neq 0$. Приведенные рассуждения наводят на мысль, что при условии $\det f'(a) \neq 0$ нелинейная система (54) тоже будет однозначно разрешима вблизи точки a, если вектор y достаточно близок к $f(a)$. Ниже мы покажем, что теорема 1 верна и в многомерном случае, если заменить в ее формулировке условия вида $f'(x) \neq 0$ на *обратимость* $f'(x)$ или *обратимость* $d_x f$. Это утверждение является основным результатом параграфа.

Заметим, однако, что *доказательство* теоремы 1 обобщить на многомерный случай нельзя, поскольку понятие монотонности бессмысленно в $\mathbb{R}^n$ при $n > 1$. Нам потребуется другая, более сложная техника. Основной результат будет выведен из теорем 2 – 4, которые мы сформулируем ниже.

Докажем вначале некоторые оценки, дающие информацию о приращении отображения f вблизи точки a.

Лемма 1. *Пусть множество E открыто в $\mathbb{R}^n$, $f \in C^1(E, \mathbb{R}^n)$, $a \in E$, $d_a f$ обратим. Тогда для любого $\sigma > 0$ существует такое $r > 0$, что при всех $x, y \in \overline{V_a(r)}$ справедливы неравенство*

$$\|f(y) - f(x) - d_a f(y - x)\| \leqslant 3\sigma \|y - x\| \tag{56}$$

и двусторонняя оценка

$$\begin{aligned} &\|f(y) - f(x)\| \leqslant \big(\|d_a f\| + \sigma\big) \cdot \|y - x\|, \\ &\|f(y) - f(x)\| \geqslant \Big(\big\|(d_a f)^{-1}\big\|^{-1} - 3\sigma\Big) \cdot \|y - x\|. \end{aligned} \tag{57}$$

Доказательство. Положим

$$T = d_a f, \quad C = \big\|T^{-1}\big\|^{-1}.$$

Пусть $\sigma > 0$. В силу теоремы 2 § 4 дифференциал f непрерывно зависит от точки на E. Поэтому найдется $r > 0$, для которого

$$\|d_x f - d_a f\| < \sigma \ \text{ при всех } \ x \in \overline{V_a(r)}. \tag{58}$$

Покажем, что такое r искомое. Пусть $x, y \in \overline{V_a(r)}$. Проверим вначале (56). Положим

$$F(t) = f(t) - d_x f(t) \quad \text{при} \ \ t \in \overline{V_a(r)}.$$

Тогда для всех $c \in \overline{V_a(r)}$

$$\|d_c F\| = \|d_c f - d_x f\| \leqslant \|d_c f - d_a f\| + \|d_a f - d_x f\| < 2\sigma.$$

Используя оценку конечных приращений и выпуклость $\overline{V_a(r)}$, мы получим

$$\begin{aligned}\|f(y)-f(x)-T(y-x)\| &= \|F(y)-F(x)+(d_xf-d_af)(y-x)\| \leqslant \\ &\leqslant \|F(y)-F(x)\|+\|(d_xf-d_af)(y-x)\| \leqslant \\ &\leqslant \sup_{c\in\Delta_{x,y}} \|d_cF\|\cdot\|y-x\|+\sigma\,\|y-x\| \leqslant 3\sigma\,\|y-x\|\,.\end{aligned}$$

Докажем теперь (57). Заметим, что при всех $c \in \overline{V_a(r)}$

$$\|d_cf\| \leqslant \|T\|+\|d_cf-d_af\| < \|T\|+\sigma,$$

откуда

$$\|f(y)-f(x)\| \leqslant \sup_{c\in\Delta_{x,y}} \|d_cf\|\cdot\|y-x\| \leqslant \big(\|T\|+\sigma\big)\,\|y-x\|\,.$$

Кроме того, в силу (56) и замечания 1 к теореме 4 § 6 главы 5

$$\begin{aligned}\|f(y)-f(x)\| &\geqslant \|T(y-x)\|-\|f(y)-f(x)-T(y-x)\| \geqslant \\ &\geqslant C\,\|y-x\|-3\sigma\,\|y-x\| = (C-3\sigma)\,\|y-x\|\,,\end{aligned}$$

и второе из неравенств (57) тоже доказано. $\square$

Теорема 2. Существование и непрерывность локально обратного отображения. *Пусть множество E открыто в $\mathbb{R}^n$, $f \in C^1(E,\mathbb{R}^n)$, $a \in E$, d_af обратим. Тогда существует такая окрестность V_a точки a, что*

1) *для любой точки $x \in V_a$ оператор d_xf обратим;*
2) *отображение $f|_{V_a}$ инъективно;*
3) *обратное отображение $\big(f|_{V_a}\big)^{-1}$ непрерывно на $f\big(V_a\big)$.*

Доказательство. Положим

$$C = \left\|\big(d_af\big)^{-1}\right\|^{-1}, \quad \sigma = \tfrac{C}{4}, \quad V_a = V_a(r),$$

где число r выбрано по σ в соответствии с леммой 1. Покажем, что окрестность V_a искомая. Проверим первое утверждение. Если $x \in V_a$, то в силу (58) для любого $h \in \mathbb{R}^n$

$$\begin{aligned}\|d_xf(h)\| &\geqslant \|Th\|-\left\|\big(d_xf-T\big)(h)\right\| \geqslant \\ &\geqslant C\,\|h\|-\|d_xf-d_af\|\cdot\|h\| \geqslant (C-\sigma)\,\|h\|\,,\end{aligned}$$

что по теореме 4 § 6 главы 5 дает обратимость $d_x f$. Докажем теперь 2) и 3). В силу (57)

$$\|f(y) - f(x)\| \geqslant \sigma \|y - x\| \quad \text{при всех } x, y \in V_a.$$

Отсюда следует инъективность $f|_{V_a}$, так как равенство $f(x) = f(y)$ влечет $x = y$. Кроме того, пусть $u, v \in f(V_a)$. Полагая $x = f^{-1}(u)$, $y = f^{-1}(v)$, мы получим

$$\left\|f^{-1}(v) - f^{-1}(u)\right\| \leqslant \sigma^{-1} \|v - u\|.$$

Поэтому $f^{-1}(v) \to f^{-1}(u)$ при $v \to u$, то есть отображение $\left(f|_{V_a}\right)^{-1}$ непрерывно на $f(V_a)$. $\square$

Введем одно важное понятие, которое будет необходимо нам в дальнейшем.

Определение 1. Гомеоморфизм. Пусть $E \subset \mathbb{R}^n$, отображение $f\colon E \to \mathbb{R}^m$ инъективно. Если f непрерывно на E, а f^{-1} — на $f(E)$, то f называют *гомеоморфизмом множеств E и $f(E)$*.

Замечание. Утверждение теоремы 2 можно истолковать так: *если $d_a f$ обратим, то существует такая окрестность V_a точки a, что $f|_{V_a}$ — гомеоморфизм V_a на $f(V_a)$.*

Наша дальнейшая задача — получить многомерный аналог утверждения 3) теоремы 1. Выясним вначале, как устроен образ малой окрестности точки a, если $d_a f$ есть тождественный оператор. Положим

$$I(x) = x \quad \text{при всех } x \in \mathbb{R}^n.$$

Лемма 2. Образ малого шара при почти сдвиге. *Пусть множество E открыто в $\mathbb{R}^n$, $f \in C^1(E, \mathbb{R}^n)$, $a \in E$, $d_a f = I$, $\varepsilon \in (0, 1)$. Тогда существует такое $r > 0$, что $\overline{V_a(r)} \subset E$ и*

$$V_{f(a)}\big((1-\varepsilon)\, r\big) \subset f\big(V_a(r)\big) \subset V_{f(a)}\big((1+\varepsilon)\, r\big). \tag{59}$$

Замечание 1. В силу определения дифференцируемости

$$f(x) = f(a) + x - a + o(x - a) \approx f(a) + x - a \quad \text{при } x \approx a.$$

Отображение, стоящее в правой части этого приближенного равенства, действует как сдвиг аргумента на вектор $f(a) - a$. Поэтому мы называем f *почти сдвигом*.

Замечание 2. Очевидно, что сдвиг переводит $V_a(r)$ в шар радиуса r. Для почти сдвига это, вообще говоря, не так, но (59) показывает, что мы можем приблизить $f\bigl(V_a(r)\bigr)$ снаружи и изнутри концентрическими шарами, причем зазор между ними будет мал по сравнению с радиусами шаров.

Доказательство. Применяя лемму 1 для $\sigma = \dfrac{\varepsilon}{6}$, мы найдем такое $r > 0$, что для любых $x, y \in \overline{V_a(r)}$ выполнены неравенства

$$\|f(y) - f(x) - (y - x)\| \leqslant 3\sigma \|y - x\| \tag{60}$$

и

$$(1 - 3\sigma) \|y - x\| \leqslant \|f(y) - f(x)\| \leqslant (1 + \sigma) \|y - x\|. \tag{61}$$

Покажем, что это число r искомое. Правое включение в (59) очевидно. Действительно, если $x \in V_a(r)$, то в силу (61)

$$\|f(x) - f(a)\| \leqslant (1 + \sigma) \|x - a\| < (1 + \varepsilon)r,$$

то есть $f(x) \in V_{f(a)}\bigl(r(1 + \varepsilon)\bigr)$.

Перейдем к доказательству левого включения в утверждении леммы. Пусть $y \in V_{f(a)}\bigl(r(1 - \varepsilon)\bigr)$, то есть $\|y - f(a)\| < r(1 - \varepsilon)$. Нам нужно найти точку $x^* \in V_a(r)$, для которой $f\bigl(x^*\bigr) = y$. Положим $F(x) = \|f(x) - y\|$. Функция F непрерывна на компактном множестве $\overline{V_a(r)}$ и по теореме Вейерштрасса достигает на нем наименьшего значения в некоторой точке $x^* \in \overline{V_a(r)}$. Покажем, что она искомая, то есть $x^* \in V_a(r)$ и $F\bigl(x^*\bigr) = 0$. Проверим вначале, что x^* не лежит на границе шара $V_a(r)$. Положим $\widetilde{x} = a + y - f(a)$ (в точке $\widetilde{x}$ значение y принимает отображение $x \mapsto x + f(a) - a$, хорошо приближающее f вблизи a). Тогда

$$\|\widetilde{x} - a\| = \|y - f(a)\| < (1 - \varepsilon)r < r,$$

и в силу (60)

$$F\bigl(x^*\bigr) \leqslant F\bigl(\widetilde{x}\bigr) = \bigl\|f\bigl(\widetilde{x}\bigr) - y\bigr\| = \\ = \bigl\|f\bigl(\widetilde{x}\bigr) - f(a) - \bigl(\widetilde{x} - a\bigr)\bigr\| \leqslant 3\sigma \|\widetilde{x} - a\| < 3\sigma r.$$

Если же $\|x-a\|=r$, то по неравенству (61) и выбору y

$$\begin{aligned}\|f(x)-y\| &\geqslant \|f(x)-f(a)\|-\|f(a)-y\| \geqslant \\ &\geqslant (1-3\sigma)\,\|x-a\|-(1-\varepsilon)\,r=(\varepsilon-3\sigma)\,r=3\sigma r,\end{aligned}$$

откуда $F(x)>F(x^*)$.

Докажем теперь, что $F(x^*)=0$. Положим $e=y-f(x^*)$. Так как $x^*\in V_a(r)$, при всех достаточно малых $t\in(0,1)$ выполняется включение $x^*+te\in V_a(r)$. Проверим, что для таких t

$$F(x^*+te)\leqslant\big(1-t(1-3\sigma)\big)\,F(x^*). \tag{62}$$

Действительно,

$$f(x^*)-y+te=(t-1)\,e \quad \text{и} \quad \|e\|=F(x^*),$$

откуда в силу (60)

$$\begin{aligned}F(x^*+te)&=\big\|f(x^*+te)-y\big\|\leqslant \\ &\leqslant\big\|f(x^*+te)-f(x^*)-te\big\|+\big\|f(x^*)-y+te\big\|\leqslant \\ &\leqslant 3\sigma t\,\|e\|+(1-t)\,\|e\|=F(x^*)\big(1-t(1-3\sigma)\big).\end{aligned}$$

Так как $F(x^*)\leqslant F(x^*+te)$, из (62) мы получаем

$$F(x^*)\leqslant\big(1-t(1-3\sigma)\big)F(x^*).$$

Поскольку $1-t(1-2\sigma)<1$, такое неравенство возможно лишь в случае $F(x^*)=0$. $\square$

Теорема 3. Открытость образа гладкого отображения. *Пусть множество E открыто в $\mathbb{R}^n$, $f\in C^1(E,\mathbb{R}^n)$, причем в любой точке $a\in E$ дифференциал f обратим. Тогда образ отображения f открыт в $\mathbb{R}^n$.*

Доказательство. Пусть $a\in E$, $b=f(a)$. Мы должны проверить, что b является внутренней точкой $f(E)$.

Предположим вначале, что $d_af=I$. В силу леммы 2 по любому $\varepsilon\in(0,1)$ можно подобрать число $r>0$, для которого

$$f(E)\supset f\big(V_a(r)\big)\supset V_b\big((1-\varepsilon)\,r\big).$$

Таким образом, точка b является внутренней для $f(E)$.

Рассмотрим теперь общий случай. Положим

$$T = d_a f, \quad F = T^{-1} \circ f.$$

Напомним, что дифференциал линейного оператора T^{-1} в любой точке совпадает с ним самим (см. § 2). Тогда по теореме 1 § 3

$$d_a F(h) = d_b T^{-1}\big(d_a f(h)\big) = T^{-1}(Th) = h \quad \text{при всех } \ h \in \mathbb{R}^n,$$

то есть $d_a F = I$. По доказанному выше $F(E)$ содержит некоторую окрестность U точки $F(a)$. Тогда

$$f(E) = T\big(F(E)\big) \supset T(U).$$

Так как оператор T^{-1} непрерывен в $\mathbb{R}^n$ и $T(U) = \big(T^{-1}\big)^{-1}(U)$, по теореме 3 § 5 главы 5 множество $T(U)$ открыто в $\mathbb{R}^n$. Кроме того,

$$b = f(a) = T\big(F(a)\big) \in T(U),$$

поэтому точка b является внутренней для $f(E)$. $\square$

Изучим теперь вопрос о *гладкости обратного отображения*. Пусть $s \in \mathbb{N}$, $f\colon E \subset \mathbb{R}^n \to \mathbb{R}^n$ — биективное отображение класса C^s, дифференциал которого обратимым в любой точке множества E. Что можно сказать о гладкости обратного отображения? Оказывается, f^{-1} также будет принадлежать классу C^s. Сформулируем это утверждение.

Теорема 4. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{N}$, f — инъективное отображение класса $C^s(E, \mathbb{R}^n)$, причем его дифференциал в любой точке E обратим. Тогда $f^{-1} \in C^s\big(f(E), \mathbb{R}^n\big)$.*

Замечание. По теореме 3 множество $f(E)$ открыто в $\mathbb{R}^n$, поэтому класс $C^s\big(f(E), \mathbb{R}^n\big)$ корректно определен.

Доказательство. Положим $g = f^{-1}$, $G = f(E)$. Проверим по индукции, что для любого $k \in \{0, \ldots, s\}$ справедливо включение $g \in C^k\big(G, \mathbb{R}^n\big)$. Случай $k = 0$ непосредственно вытекает из теоремы 2. Предположим, что для некоторого $k < s$ утверждение уже доказано. Воспользуемся теоремой 3 § 3. Первые два условия этой теоремы очевидны, а два других вытекают из теорем 3 и 2 соответственно. Поэтому отображение g дифференцируемо на G и

$$g'(y) = \bigl(f'\bigl(g(y)\bigr)\bigr)^{-1} \quad \text{при всех } \ y \in G.$$

Для $x \in E$ положим $\Delta(x) = \det f'(x)$, а через $\Delta_{ij}(x)$ обозначим алгебраическое дополнение элемента $f'(x)$ с индексами i, j. Используя формулу для обратной матрицы, мы получим

$$\frac{\partial g_i}{\partial x_j}(y) = \frac{\Delta_{ji}\bigl(g(y)\bigr)}{\Delta\bigl(g(y)\bigr)} \quad \text{для любых } \ y \in G \ \text{ и } \ i, j \in \{1, \ldots, n\}.$$

В § 5 было показано, что все частные производные первого порядка от координатных функций f принадлежат классу C^{s-1} и, значит, классу C^k. Пусть $i, j \in \{1, \ldots, n\}$. В силу теоремы $1'$ § 5 $\Delta \in C^k(E)$, $\Delta_{ji} \in C^k(E)$, а также $\dfrac{\Delta_{ji}}{\Delta} \in C^k(E)$, поскольку $\Delta \neq 0$ на E. Из индукционного предположения и теоремы $2'$ § 5 вытекает включение $\dfrac{\Delta_{ji}}{\Delta} \circ g \in C^k(G)$. Тем самым доказано, что $\dfrac{\partial g_i}{\partial x_j} \in C^k(G)$ для любых $i, j \in \{1, \ldots, n\}$. Поэтому $g \in C^{k+1}\bigl(G, \mathbb{R}^n\bigr)$. $\square$

Теперь мы можем обобщить теорему 1 о локальной обратимости функций на многомерный случай.

Теорема 5. Локальная обратимость гладких отображений. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{N}$, $f \in C^s(E, \mathbb{R}^n)$, $a \in E$, $d_a f$ обратим. Тогда существует окрестность V_a точки a, для которой справедливы следующие утверждения.*

1) *При всех $x \in V_a$ оператор $d_x f$ обратим.*
2) *Отображение $f|_{V_a}$ инъективно.*
3) *Множество $f(V_a)$ открыто в $\mathbb{R}^n$.*
4) *Отображение $\bigl(f|_{V_a}\bigr)^{-1}$ принадлежит классу $C^s\bigl(f(V_a), \mathbb{R}^n\bigr)$.*

Доказательство. По теореме 2 существует окрестность V_a точки a, для которой справедливы утверждения 1) и 2). Применяя к отображению $f|_{V_a}$ теоремы 3 и 4, мы получим соответственно утверждения 3) и 4). $\square$

Замечание. В случае $s = 1$ полезно сравнить теорему 5 с теоремой 3 § 3. В теореме 5 мы предполагаем только обратимость $d_a f$, остальные условия теоремы 3 *доказываются.* При этом, однако, отображение f предполагается не просто дифференцируемым в точке a, а *непрерывно дифференцируемым в окрестности* a.

Введем еще одно важное понятие, характеризующее обратимые отображения.

Определение 2. Диффеоморфизм. Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{N} \cup \{\infty\}$, $f \in C^s(E, \mathbb{R}^n)$.

1) Отображение f называют *диффеоморфизмом множеств* E *и* $f(E)$ *класса* C^s, если f инъективно, $f(E)$ открыто в $\mathbb{R}^n$ и $f^{-1} \in C^s\big(f(E), E\big)$.

2) Отображение f называют *локальным диффеоморфизмом класса* C^s, если у любой точки $a \in E$ существует такая окрестность V_a, что $f|_{V_a}$ — диффеоморфизм класса C^s.

Замечание. Если отображение f удовлетворяет 1) при $s = 0$, то оно является гомеоморфизмом E и $f(E)$.

Переформулируем полученные выше результаты, используя понятие диффеоморфизма.

Теорема 6. *Пусть E открыто в $\mathbb{R}^n$, $s \in \mathbb{N}$, $f \in C^s(E, \mathbb{R}^n)$, причем дифференциал f в любой точке множества E обратим. Тогда справедливы следующие утверждения.*

1) *Отображение f — локальный диффеоморфизм класса C^s.*

2) *Если отображение f инъективно, то оно является диффеоморфизмом E и $f(E)$ класса C^s.*

Вернемся теперь к вопросу о локальной разрешимости системы уравнений (54). Хотя ответ на этот вопрос фактически содержится в теореме 5, мы сформулируем его явно.

Теорема 7. Локальная разрешимость систем уравнений. *Пусть множество E открыто в $\mathbb{R}^n$, $s \in \mathbb{N}$, $f_1, \ldots, f_n \in C^s(E)$, $a \in E$, $b = \big(f_1(a), \ldots, f_n(a)\big)$. Предположим, что*

$$\det \left\{ \frac{\partial f_i}{\partial x_j}(a) \right\}_{i,j=1}^{n} \neq 0. \tag{63}$$

Тогда справедливы следующие утверждения.

1) *Существуют такие окрестности V_a и V_b точек a и b, что для любого $y \in V_b$ система (54) имеет на V_a единственное решение $\big(x_1(y), \dots, x_n(y)\big)$.*

2) *При любом $k \in \{1, \dots, n\}$ функция $y \mapsto x_k(y)$ принадлежит классу $C^s(V_b)$.*

Доказательство. Перепишем систему (54) в векторной форме $f(x) = y$, где $f = (f_1, \dots, f_n)$. Тогда $f \in C^s(E, \mathbb{R}^n)$ (см. замечание 3 к определению 3 § 5). Условие (63) эквивалентно обратимости $d_a f$. Выберем V_a в соответствии с теоремой 5. Поскольку множество $f(V_a)$ открыто, найдется окрестность V_b, лежащая в $f(V_a)$. Если $y \in V_b$, то существование решения уравнения $f(x) = y$ очевидно, а единственность вытекает из инъективности $f|_{V_a}$. Второе утверждение справедливо в силу включения $\big(f|_{V_a}\big)^{-1} \in C^s\big(f(V_a), \mathbb{R}^n\big)$ и замечания 3 к определению 3 § 5. $\square$

Замечание. Условие (63) называют *независимостью уравнений системы (54) в точке a*. Если система (54) линейна, то матрица f' постоянна и совпадает с матрицей системы. Условие (63) в этом случае означает линейную независимость уравнений и гарантирует однозначную разрешимость (54) на всем $\mathbb{R}^n$ при любой правой части (см. лемму 2 § 6 главы 5). В общем случае теорема 7 утверждает только локальную разрешимость системы.

В заключение приведем несколько примеров диффеоморфизмов в $\mathbb{R}^2$ и в $\mathbb{R}^3$. Они будут играть важную роль при вычислении кратных интегралов.

Пример 1. *Полярные координаты.* Положим

$$f(r, \varphi) = (r\cos\varphi, r\sin\varphi) \quad (r, \varphi \in \mathbb{R}).$$

Очевидно, что $f \in C^\infty(\mathbb{R}^2)$. Кроме того,

$$\det f'(r, \varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r\cos^2\varphi + r\sin^2\varphi = r.$$

По теореме 6 f является локальным диффеоморфизмом на множестве $\big\{(r, \varphi) \in \mathbb{R}^2 : r \neq 0\big\}$. Заметим, однако, что отображение f

не будет диффеоморфизмом на этом множестве, поскольку оно 2π-периодично по φ и потому не инъективно. Несложно проверить, что сужение f на $(0,+\infty)\times(-\pi,\pi)$ является диффеоморфизмом класса C^∞, действующим на множество $\mathbb{R}^2 \setminus \{(x,0) : x \leqslant 0\}$.

Пример 2. *Цилиндрические координаты.* Рассмотрим отображение $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, сопоставляющее точке (r,φ,t) вектор (x,y,z) по формулам

$$x = r\cos\varphi, \quad y = r\sin\varphi, \quad z = t,$$

геометрический смысл которых ясен из рисунка 30.

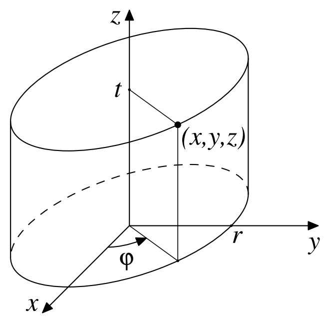

Рис. 30

Тогда $f \in C^\infty(\mathbb{R}^3)$ и

$$\det f'(r,\varphi,t) = \begin{vmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r\cos^2\varphi + r\sin^2\varphi = r.$$

Поэтому отображение f будет локальным диффеоморфизмом на множестве $\{(r,\varphi,t) \in \mathbb{R}^3 : r \neq 0\}$. Легко видеть, что оно является диффеоморфизмом класса C^∞ на множестве $(0,+\infty)\times(-\pi,\pi)\times\mathbb{R}$.

Пример 3. *Сферические координаты.* Рассмотрим отображение $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, сопоставляющее точке (r,φ,θ) вектор (x,y,z) по формулам

$$x = r\cos\theta\cos\varphi, \quad y = r\cos\theta\sin\varphi, \quad z = r\sin\theta.$$

При фиксированном r точка $f(r,\varphi,\theta)$ лежит на сфере радиуса r. Углы θ и φ, определяющие положение этой точки, принято называть соответственно *широтой* и *долготой* . Формулы, связывающие декартовы и сферические координаты, ясны из рисунка 31.

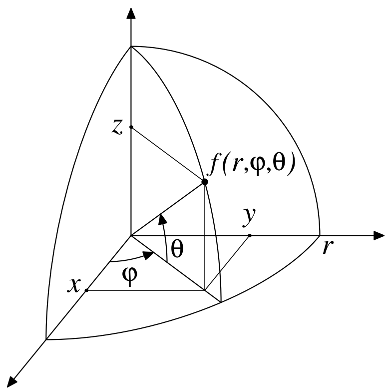

Рис. 31

Заметим, что $f \in C^\infty(\mathbb{R}^3)$ и

$$\det f'(r,\varphi,\theta) = \begin{vmatrix} \cos\theta\cos\varphi & -r\cos\theta\sin\varphi & -r\sin\theta\cos\varphi \\ \cos\theta\sin\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta & 0 & r\cos\theta \end{vmatrix} = r^2\cos\theta$$

(проверьте это). Поэтому отображение f будет локальным диффеоморфизмом на множестве

$$\bigl\{(r,\varphi,\theta) \in \mathbb{R}^3 : r \neq 0,\ \theta \neq \tfrac{\pi}{2} + \pi k\ \ (k \in \mathbb{Z})\bigr\}.$$

Легко видеть, что оно является диффеоморфизмом класса C^∞ на множестве $(0,+\infty) \times (-\pi,\pi) \times (-\frac{\pi}{2},\frac{\pi}{2})$.

§ 8. Неявные отображения

В этом параграфе мы изучим вопрос о разрешимости уравнений в более общей ситуации. Пусть вначале

$$f\colon E \subset \mathbb{R}^2 \to \mathbb{R}.$$

Рассмотрим уравнение относительно переменных x и y

$$f(x,y)=0, \quad (x,y)\in E. \tag{64}$$

Оно задает зависимость между x и y, которая называется *неявной*. Предположим, что из уравнения (64) можно однозначно выразить одну переменную через другую (например, y через x). Тогда множество решений (64) представляет собой график некоторой функции φ, то есть имеет вид

$$\bigl\{(x,y)\in E : y=\varphi(x)\bigr\}.$$

Такую функцию φ называют *неявной функцией, порожденной уравнением (64)*.

Вопрос о существовании неявной функции в такой постановке весьма сложен. Во многих приложениях достаточно рассматривать уравнение (64) не на всем E, а только в некоторой окрестности какой-либо точки множества E. Если вблизи любой точки E уравнение (64) порождает неявную функцию, то оно называется *локально разрешимым*, а сама неявная функция — *локальной*. Для изучения локальной разрешимости удобно использовать средства дифференциального исчисления. Пусть точка $a=(x^0,y^0)$ является внутренней для E и удовлетворяет (64), а функция f дифференцируема в точке a. Если (x,y) — решение уравнения (64), достаточно близкое к a, то мы можем записать

$$0=f(x,y)-f(x^0,y^0)\approx\frac{\partial f}{\partial x}(a)\bigl(x-x^0\bigr)+\frac{\partial f}{\partial y}(a)\bigl(y-y^0\bigr).$$

Рассмотрим линейное уравнение

$$\frac{\partial f}{\partial x}(a)\bigl(x-x^0\bigr)+\frac{\partial f}{\partial y}(a)\bigl(y-y^0\bigr)=0.$$

Из него можно однозначно выразить y через x, если $\dfrac{\partial f}{\partial y}(a)\neq 0$, и x через y в случае $\dfrac{\partial f}{\partial x}(a)\neq 0$. Это наводит на мысль, что достаточным условием однозначной разрешимости (64) вблизи точки a является соотношение

$$\operatorname{grad} f(a)\neq(0,0).$$

Далее мы увидим, что так оно и будет. Следующий пример показывает, что условие на $\operatorname{grad} f(a)$ существенно.

Пример. Пусть

$$f(x,y) = \left(x^2+y^2\right)^2 - \left(x^2-y^2\right), \quad a=(0,0).$$

Записывая уравнение $f(x,y)=0$ в полярных координатах, мы получим

$$r=\sqrt{\cos 2\varphi}, \quad \text{где } \varphi \in \left[-\tfrac{\pi}{4}+\pi k, \tfrac{\pi}{4}+\pi k\right] \ (k \in \mathbb{Z}).$$

Эта зависимость определяет кривую на плоскости, называемую *лемнискатой Бернулли* (см. рисунок 32).

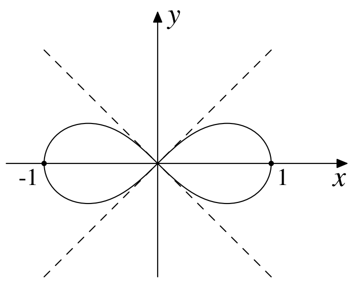

Рис. 32

Очевидно, что в окрестности нуля она не является графиком какой-либо функции. Предлагаем читателю самостоятельно проверить, что $\operatorname{grad} f(a) = (0,0)$.

Отметим, что вопрос о локальной обратимости функции, изученный в § 7, является частным случаем задачи о неявной функции. Действительно, пусть $\varphi\colon \langle A,B\rangle \to \mathbb{R}$, $c \in (A,B)$. Нахождение вблизи точки c обратной функции для φ сводится к разрешимости относительно x уравнения (64) при

$$f(x,y)=\varphi(x)-y$$

в окрестности точки $\bigl(c,\varphi(c)\bigr)$. Заметим, что условие $\varphi'(c) \neq 0$, наложенное в теореме 1 § 7, можно записать в виде

$$\frac{\partial f}{\partial x}\bigl(c,\varphi(c)\bigr) \neq 0.$$

Это согласуется с нашими предыдущими рассуждениями.

Обобщим теперь постановку задачи о неявной функции. Пусть $m, n \in \mathbb{N}$, $f\colon E \subset \mathbb{R}^{n+m} \to \mathbb{R}^m$. Рассмотрим систему уравнений

$$\begin{cases} f_1(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_m(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0. \end{cases} \tag{65}$$

Договоримся записывать элементы $\mathbb{R}^{n+m}$ в виде (x, y), где $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. Тогда система (65) эквивалентна векторному уравнению относительно двух векторных переменных:

$$f(x, y) = 0, \quad \text{где } (x, y) \in E. \tag{66}$$

В каком случае это уравнение локально разрешимо относительно переменной y? Пусть точка $a = (x^0, y^0)$ является внутренней для E и удовлетворяет (66), а отображение f дифференцируемо в точке a. Заметим, что для любого вектора $(dx, dy) \in \mathbb{R}^n \times \mathbb{R}^m$

$$d_a f(dx, dy) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, dx_i + \sum_{j=1}^{m} \frac{\partial f}{\partial y_j}(a)\, dy_j = \\ = f'_x(a) \cdot dx + f'_y(a) \cdot dy, \tag{67}$$

где

$$f'_x(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \ldots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \ldots & \ldots & \ldots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \ldots & \dfrac{\partial f_m}{\partial x_n}(a) \end{pmatrix},$$

$$f'_y(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1}(a) & \ldots & \dfrac{\partial f_1}{\partial y_m}(a) \\ \ldots & \ldots & \ldots \\ \dfrac{\partial f_m}{\partial y_1}(a) & \ldots & \dfrac{\partial f_m}{\partial y_m}(a) \end{pmatrix}. \tag{68}$$

Если (x, y) — решение уравнения (66), достаточно близкое к a, то

$$0 = f(x, y) - f(x^0, y^0) \approx d_a f\big(x - x^0, y - y^0\big).$$

Рассмотрим уравнение

$$d_a f\big(x - x^0, y - y^0\big) = 0.$$

Оно эквивалентно линейной системе

$$f'_x(a) \cdot \big(x - x^0\big) + f'_y(a) \cdot \big(y - y^0\big) = 0.$$

Эта система однозначно разрешима относительно y, если матрица $f'_y(a)$ обратима. Оказывается, что при этом условии уравнение (66) также однозначно разрешимо вблизи точки a относительно y. Докажем это утверждение.

Теорема 1. О неявном отображении. *Пусть множество E открыто в $\mathbb{R}^{n+m}$, $\big(x^0, y^0\big) \in E$, $s \in \mathbb{N}$, $f \in C^s(E, \mathbb{R}^m)$, причем*

$$f\big(x^0, y^0\big) = 0 \ \ \text{и} \ \ \det f'_y\big(x^0, y^0\big) \neq 0.$$

Тогда найдутся окрестность $V_{(x^0,y^0)}$ точки $\big(x^0, y^0\big)$, открытое в $\mathbb{R}^n$ множество U, содержащее x^0, и отображение $\psi \in C^s(U, \mathbb{R}^m)$, для которых

$$\big\{(x, y) \in V_{(x^0,y^0)} : f(x, y) = 0\big\} = \big\{\big(x, \psi(x)\big) : x \in U\big\}. \tag{69}$$

Таким образом, множество решений (66) вблизи $\big(x^0, y^0\big)$ представляет собой график отображения ψ, имеющего тот же класс гладкости, что и f.

Доказательство. Положим $a = \big(x^0, y^0\big)$. Рассмотрим отображение $\Phi\colon E \to \mathbb{R}^{n+m}$, действующее по формуле

$$\Phi(x, y) = \big(x, f(x, y)\big).$$

Матрицу $\Phi'(x, y)$ можно записать в виде

$$\Phi'(x, y) = \left(\begin{array}{c|c} I_n & \mathbb{O}_{nm} \\ \hline A & f'_y(x, y) \end{array}\right),$$

где I_n — единичная матрица $n \times n$, $\mathbb{O}_{nm}$ — нулевая матрица размера $n \times m$, а содержимое блока A не имеет значения. Из курса алгебры известно, что

$$\det \Phi' = \det I_n \cdot \det f'_y = \det f'_y, \tag{70}$$

откуда $\det \Phi'(a) \neq 0$. По теореме 5 § 7 существует такая окрестность V_a точки a, что $\Phi|_{V_a}$ — диффеоморфизм класса C^s. Положим

$$\Psi = \left(\Phi|_{V_a}\right)^{-1}, \quad \Psi_x = \left(\Psi_1, \ldots, \Psi_n\right), \quad \Psi_y = \left(\Psi_{n+1}, \ldots, \Psi_{n+m}\right)$$

и проверим справедливость равенств

$$\begin{aligned} y &= \Psi_y\big(x, f(x,y)\big) \quad \text{при всех } (x,y) \in V_a, \\ z &= f\big(x, \Psi_y(x,z)\big) \quad \text{при всех } (x,z) \in \Phi(V_a). \end{aligned} \tag{71}$$

Для любой точки $(x,y) \in V_a$

$$(x,y) = \Psi\big(\Phi(x,y)\big) = \Psi\big(x, f(x,y)\big) = \Big(\Psi_x\big(x, f(x,y)\big), \Psi_y\big(x, f(x,y)\big)\Big).$$

Приравнивая соответствующие координаты левой и правой частей, мы получим первое из равенств (71), а также тождество $\Psi_x(x,z) = x$, где $(x,z) \in \Phi(V_a)$. Поэтому при любых $(x,z) \in \Phi(V_a)$

$$(x,z) = \Phi\big(\Psi(x,z)\big) = \Phi\big(x, \Psi_y(x,z)\big) = \Big(x, f\big(x, \Psi_y(x,z)\big)\Big),$$

откуда вытекает второе из соотношений (71).

Положим

$$U = \left\{x \in \mathbb{R}^n : (x,0) \in \Phi(V_a)\right\}, \quad \psi(x) = \Psi_y(x,0) \ \ (x \in U).$$

Так как $\Phi(V_a)$ открыто в $\mathbb{R}^{n+m}$, множество $\Phi(V_a) \cap \left(\mathbb{R}^n \times \{0\}\right)$ открыто в $\mathbb{R}^n \times \{0\}$, то есть U открыто в $\mathbb{R}^n$. Кроме того, $x^0 \in U$, поскольку

$$\left(x^0, 0\right) = \Big(x^0, f\big(x^0, y^0\big)\Big) = \Phi\big(x^0, y^0\big) \in \Phi(V_a).$$

Проверим теперь, что $\psi \in C^s(U, \mathbb{R}^m)$. Заметим, что $\psi = g_2 \circ \Psi \circ g_1$, где отображения $g_1 \colon \mathbb{R}^n \to \mathbb{R}^{n+m}$ и $g_2 \colon \mathbb{R}^{n+m} \to \mathbb{R}^m$ действуют по формулам

$$g_1(x) = (x,0), \quad g_2(x,y) = y \quad (x \in \mathbb{R}^n, \ y \in \mathbb{R}^m).$$

Эти отображения, очевидно, бесконечно дифференцируемы. Тогда по теореме 2′ § 5 $\psi \in C^s(U, \mathbb{R}^m)$.

Осталось доказать равенство (69). Перепишем его в виде

$$\big\{(x,y) \in V_a : f(x,y) = 0\big\} = \big\{(x,y) \in \mathbb{R}^{n+m} : x \in U,\, y = \psi(x)\big\}.$$

Пусть $(x,y) \in V_a$, $f(x,y) = 0$. Тогда

$$(x,0) = \big(x, f(x,y)\big) = \Phi(x,y) \in \Phi(V_a),$$

то есть $x \in U$. Кроме того, первое из равенств (71) дает

$$y = \Psi_y(x,0) = \psi(x).$$

Проверим теперь обратное включение. Пусть $x \in U$, $y = \psi(x)$. Тогда $(x,0) \in \Phi(V_a)$, откуда

$$(x,y) = \big(x, \Psi_y(x,0)\big) = \Psi(x,0) \in V_a.$$

Применяя второе из соотношений (71) для $z = 0$, мы получим

$$f(x,y) = f\big(x,\psi(x)\big) = f\big(x, \Psi_y(x,0)\big) = 0. \quad \square$$

Определение 1. Неявное отображение. Отображение ψ, введенное в теореме 1, называется *неявным отображением, порожденным уравнением* $f(x,y) = 0$ *в окрестности точки* $\big(x^0, y^0\big)$.

Замечание 1. Из (70) вытекает, что $\det f'_y \neq 0$ на V_a. В частности, для любого $x \in U$ матрица $f'_y\big(x,\psi(x)\big)$ обратима.

Замечание 2. *Пусть в обозначениях теоремы 1 отображение* $F\colon E \to \mathbb{R}^m$ *дифференцируемо на* $V_{(x^0,y^0)}$ *и*

$$g(x) = F\big(x,\psi(x)\big) \quad (x \in U).$$

Тогда справедлива формула

$$g'(x) = F'_x\big(x,\psi(x)\big) + F'_y\big(x,\psi(x)\big)\cdot\psi'(x). \tag{72}$$

Действительно, рассмотрим отображение

$$\varphi(x) = \bigl(x, \psi(x)\bigr) \quad (x \in U).$$

Пусть $x \in U$, $z = \varphi(x)$. Несложно проверить, что

$$d_x\varphi(h) = \bigl(h, \psi'(x)\,h\bigr) \quad (h \in \mathbb{R}^n).$$

По теореме 1 § 3 и формуле (67) для любого $h \in \mathbb{R}^n$

$$\begin{aligned} d_x g(h) = d_x(F \circ \varphi)(h) = d_z F\bigl(d_x\varphi(h)\bigr) &= \\ = d_z F\bigl(h, \psi'(x)\,h\bigr) &= F'_x(z)\,h + F'_y(z)\,\psi'(x)\,h, \end{aligned}$$

откуда и вытекает (72). $\square$

Замечание 3. Доказательство теоремы 1 не является конструктивным, оно не дает ***формул*** для вычисления неявного отображения ψ. Тем не менее, выражения для матрицы Якоби ψ и частных производных функций $\psi_1, \ldots, \psi_m$ получить можно. Так как

$$f\bigl(x, \psi(x)\bigr) = 0 \quad \text{при всех } x \in U,$$

равенство (72), примененное к $F = f$ и $g \equiv 0$, дает

$$f'_y\bigl(x, \psi(x)\bigr) \cdot \psi'(x) = -f'_x\bigl(x, \psi(x)\bigr).$$

По замечанию 1 для любого $x \in U$ матрица $f'_y\bigl(x, \psi(x)\bigr)$ обратима, откуда

$$\psi'(x) = -\bigl(f'_y\bigl(x, \psi(x)\bigr)\bigr)^{-1} \cdot f'_x\bigl(x, \psi(x)\bigr).$$

Приравнивая соответствующие элементы матриц, стоящих в левой и правой частях этого равенства, мы получим формулы для производных $\dfrac{\partial \psi_i}{\partial x_j}(a)$, где $i \in \{1, \ldots, m\}$ и $j \in \{1, \ldots, n\}$. Дифференцируя тождество $f\bigl(x, \psi(x)\bigr) \equiv 0$ несколько раз и решая получившиеся линейные системы, можно вычислять частные производные отображения ψ высших порядков.

Переформулируем теперь теорему 1 для систем уравнений (65).

Теорема 2. Локальная разрешимость систем. *Пусть E открыто в $\mathbb{R}^{n+m}$, $s \in \mathbb{N}$, $f \in C^s\big(E, \mathbb{R}^m\big)$. Предположим, что точка $a = \big(x^0, y^0\big)$ множества E удовлетворяет системе (65), причем*

$$\begin{vmatrix} \dfrac{\partial f_1}{\partial y_1}(a) & \dots & \dfrac{\partial f_1}{\partial y_m}(a) \\ \dots & \dots & \dots \\ \dfrac{\partial f_m}{\partial y_1}(a) & \dots & \dfrac{\partial f_m}{\partial y_m}(a) \end{vmatrix} \neq 0.$$

Тогда существуют окрестность V_a точки a, открытое в $\mathbb{R}^n$ множество U, содержащее x^0, и отображение $\psi \in C^s\big(U, \mathbb{R}^m\big)$, для которых множество решений (65), лежащих в V_a, описывается системой

$$\begin{cases} y_1 = \psi_1(x_1, \dots, x_n) \\ \dots\dots\dots\dots\dots\dots \\ y_m = \psi_m(x_1, \dots, x_n), \end{cases} \quad \text{где } (x_1, \dots, x_n) \in U.$$

Перейдем к геометрическим приложениям теоремы 1. Мы определим многомерную гладкую поверхность в $\mathbb{R}^n$ и опишем различные способы ее задания. Приведем вначале простой пример. Пусть S — часть единичной сферы в $\mathbb{R}^3$ с центром в нуле, лежащая в открытом первом октанте. Как описать поверхность S аналитически? Для этого существует три способа.

1. Явное задание. Положим

$$f(x, y) = \sqrt{1 - x^2 - y^2}, \quad \text{где } x, y > 0, \ x^2 + y^2 < 1.$$

Тогда справедливо равенство

$$S = \big\{(x, y, z) : x, y > 0, \ x^2 + y^2 < 1, \ z = f(x, y)\big\}.$$

Таким образом, мы реализовали S как график гладкой функции, заданной на открытом подмножестве $\mathbb{R}^2$. Такое описание поверхности называется *явным*.

2. Неявное задание. Рассмотрим функцию

$$f(x, y, z) = x^2 + y^2 + z^2, \quad \text{где } x, y, z > 0.$$

Тогда
$$S = \{(x, y, z) : x, y, z > 0,\ f(x, y, z) = 1\}.$$
Иными словами, $S = f^{-1}(\{1\})$. Множество $f^{-1}(\{1\})$ называется *1-уровнем функции* f. Описание поверхности как множества уровня гладкой функции называют *неявным*.

3. Параметрическое задание. Используя сферические координаты, введенные в § 7, можно записать

$$S = \left\{(\cos\theta\cos\varphi, \cos\theta\sin\varphi, \sin\theta) : \varphi, \theta \in \left(0, \tfrac{\pi}{2}\right)\right\}.$$

Таким образом, S является образом гладкого биективного отображения, заданного на открытом подмножестве $\mathbb{R}^2$. Иными словами, все координаты точек S выражаются с помощью гладких функций от параметров φ и θ. Такой способ описания поверхности называется *параметрическим*.

Заметим, что указанные способы задания поверхностей не являются эквивалентными. Возьмем теперь в качестве S всю единичную сферу в $\mathbb{R}^3$ с центром в нуле. Она также допускает неявное задание
$$S = \{(x, y, z) : x^2 + y^2 + z^2 = 1\},$$
но реализовать ее как график гладкой функции нельзя, поскольку из уравнения $x^2 + y^2 + z^2 = 1$ ни одна из переменных не выражается однозначно через другие. Тем не менее, у любой точки $a \in S$ найдется такая окрестность V_a, что $S \cap V_a$ есть график гладкой функции. Говорят, что S является *локальным графиком* или *графиком вблизи любой точки* $a \in S$. Это же замечание верно и для параметрического задания. В дальнейшем нас будут интересовать именно локальные описания поверхностей.

Перейдем к формальному изложению теории. Введем вначале важное понятие, связанное с рангом матрицы Якоби. Для произвольной матрицы A символ $\operatorname{rank} A$ будет обозначать ранг A, то есть максимальное число k, для которого существует ненулевой минор A размера $k \times k$.

Определение 2. Регулярное отображение. Пусть множество E открыто в $\mathbb{R}^n$, $a \in E$, $f\colon E \to \mathbb{R}^m$. Отображение f называют *регулярным в точке* a, если оно дифференцируемо в точке a

и $\operatorname{rank} f'(a) = \min\{m, n\}$. При этом a называют *точкой регулярности* f. Если отображение f регулярно в любой точке E, то его называют *регулярным на* E.

Таким образом, матрица Якоби отображения f в точке регулярности имеет максимально возможный ранг. Геометрический смысл этого условия мы изучим в нашем курсе позже.

Лемма 1. *Пусть множество* E *открыто в* $\mathbb{R}^n$, $f \in C^1(E, \mathbb{R}^m)$, $a \in E$. *Если отображение* f *регулярно в точке* a, *то оно регулярно и в некоторой окрестности* a.

Доказательство. Поскольку $f \in C^1(E, \mathbb{R}^m)$, производные $\dfrac{\partial f_i}{\partial x_j}$ непрерывны на E при всех $i \in \{1, \ldots, m\}$ и $j \in \{1, \ldots, n\}$. Положим $N = \min\{m, n\}$ и обозначим через Δ такой минор матрицы f' размера $N \times N$, что $\Delta(a) \neq 0$. Заметим, что Δ является многочленом от частных производных координатных функций f, поэтому $\Delta \in C(E)$. Тогда существует окрестность V_a точки a, в которой $\Delta \neq 0$. Таким образом, $\operatorname{rank} f'(x) \geqslant N$ при любом $x \in V_a$, а обратное неравенство верно всегда. $\square$

Определение 3. Явное описание поверхности. Пусть $m, n, s \in \mathbb{N}$, $m < n$, $M \subset \mathbb{R}^n$, $a \in M$. Предположим, что существуют окрестность V_a точки a в $\mathbb{R}^n$, открытое множество U в $\mathbb{R}^m$, содержащее $(a_1, \ldots, a_m)$, и отображение $\psi \in C^s(U, \mathbb{R}^{n-m})$, для которых $M \cap V_a$ с точностью до перестановки координат совпадает с графиком ψ. Тогда M называют *m-мерным графиком класса C^s вблизи точки a*.

Замечание. Точный смысл определения 3 состоит в следующем: семейство $(1, \ldots, n)$ можно разбить на непересекающиеся подсемейства возрастающих индексов $(i_1, \ldots, i_m)$ и $(j_1, \ldots, j_{n-m})$ так, что $M \cap V_a$ представимо в виде

$$\bigl\{x \in \mathbb{R}^n : \bigl(x_{i_1}, \ldots, x_{i_m}\bigr) \in U,\\ x_{j_k} = \psi_k\bigl(x_{i_1}, \ldots, x_{i_m}\bigr) \text{ при всех } k = 1, \ldots, n-m\bigr\}.$$

В дальнейшем мы будем не умаляя общности считать $i_k = k$ $(k = 1, \ldots, m)$ и $j_k = k + m$ $(k = 1, \ldots, n - m)$.

Определение 4. Неявное описание поверхности. Пусть $m, n, s \in \mathbb{N}$, $m < n$, $M \subset \mathbb{R}^n$, $a \in M$. Предположим, что существуют окрестность V_a точки a, открытое множество U, содержащее a, и отображение $f \in C^s(U, \mathbb{R}^{n-m})$, регулярное в точке a, для которых

$$M \cap V_a = \bigl\{x \in U : f(x) = f(a)\bigr\}.$$

Тогда M называется *m-мерным уровнем класса C^s вблизи точки a.*

Замечание. В определении 4 можно считать, что отображение f регулярно на всем множестве U, а не только в точке a. Действительно, по лемме 1 найдется окрестность W точки a, на которой $\operatorname{rank} f' \equiv n - m$. Заменяя V_a и U на $V_a \cap W$ и $U \cap W$, мы получим требуемое.

Определение 5. Параметрическое описание поверхности. Пусть $m, n, s \in \mathbb{N}$, $m < n$, $M \subset \mathbb{R}^n$, $a \in M$. Предположим, что существуют окрестность V_a точки a, открытое в $\mathbb{R}^m$ множество U и гомеоморфизм φ класса C^s множеств U и $M \cap V_a$, регулярный в точке $\varphi^{-1}(a)$. Тогда говорят, что множество M *допускает m-мерную параметризацию класса C^s вблизи точки a*, а φ называют *m-мерной параметризацией M вблизи точки a.*

Напомним читателю смысл условия гомеоморфности: φ является биекцией U на $M \cap V_a$ и отображение φ^{-1} непрерывно на $M \cap V_a$.

Замечание 1. Правильнее было бы использовать термин “регулярная параметризация”, но мы для краткости будем говорить просто “параметризация”. Это не приведет к путанице, поскольку в дальнейшем рассматриваются только регулярные параметризации.

Замечание 2. В определении 5 можно считать, что параметризация φ регулярна на всем U, а не только в точке $\varphi^{-1}(a)$. Действительно, пусть $c = \varphi^{-1}(a)$. По лемме 1 найдется такое $\varepsilon > 0$, что $\operatorname{rank} \varphi' \equiv m$ на $V_c(\varepsilon)$. В силу непрерывности φ^{-1} существует $\delta > 0$, для которого $\varphi^{-1}\bigl(V_a(\delta) \cap M\bigr) \subset V_c(\varepsilon)$. По замечанию к теореме 3 § 5 главы 5 множество $\varphi^{-1}\bigl(V_a(\delta) \cap M\bigr)$ открыто в $\mathbb{R}^m$. Заменяя U на $U \cap \varphi^{-1}\bigl(V_a(\delta) \cap M\bigr)$, мы получим требуемое.

Покажем теперь, что определения 3 – 5 дают эквивалентные способы локального описания поверхностей. Для этого нам потребуются следующие обозначения. Пусть $m, n \in \mathbb{N}$, $m < n$. Договоримся записывать элементы $\mathbb{R}^n$ в виде (x, y), где $x \in \mathbb{R}^m$, $y \in \mathbb{R}^{n-m}$. Определим отображения-проекции

$$P_x(x, y) = x, \quad P_y(x, y) = y, \quad \text{где } (x, y) \in \mathbb{R}^n.$$

Заметим, что отображения P_x и P_y бесконечно дифференцируемы в $\mathbb{R}^n$, поскольку их координатные функции линейны.

Теорема 3. *Пусть $m, n, s \in \mathbb{N}$, $m < n$, $M \subset \mathbb{R}^n$, $a \in M$. Тогда равносильны следующие утверждения.*

1) *M является m-мерным графиком класса C^s вблизи a.*

2) *M является m-мерным уровнем класса C^s вблизи a.*

3) *M допускает m-мерную параметризацию класса C^s вблизи a.*

Доказательство. Запишем a в виде $\left(x^0, y^0\right)$.

1) $\Rightarrow$ 2) Пусть окрестность V_a точки a, открытое в $\mathbb{R}^m$ множество U, содержащее x^0, и отображение $\psi \in C^s(U, \mathbb{R}^{n-m})$ таковы, что

$$M \cap V_a = \left\{(x, y) \in \mathbb{R}^n : x \in U,\ y = \psi(x)\right\}.$$

Положим

$$f(x, y) = y - \psi(x), \quad \text{где } x \in U, \ \ y \in \mathbb{R}^{n-m}.$$

Тогда $f(a) = 0$ и

$$M \cap V_a = \left\{(x, y) \in U \times \mathbb{R}^{n-m} : f(x, y) = 0\right\}.$$

Кроме того, матрицу $f'(a)$ размера $(n-m) \times n$ можно записать в виде

$$f'(a) = \left(-\psi'\left(x^0\right) \,\middle|\, I_{n-m}\right),$$

где I_{n-m} — единичная матрица размера $(n-m) \times (n-m)$. Так как $\det I_{n-m} = 1$, матрица $f'(a)$ имеет максимальный ранг, равный $n - m$. Поэтому M является m-мерным уровнем класса C^s вблизи точки a.

2) $\Rightarrow$ 1) Пусть существуют окрестность $\widetilde{V}_a$ точки a, открытое в $\mathbb{R}^n$ множество E, содержащее a, и регулярное в точке a отображение $\Phi \in C^s(E, \mathbb{R}^{n-m})$, для которых

$$M \cap \widetilde{V}_a = \bigl\{(x, y) \in E : \Phi(x, y) = \Phi(a)\bigr\}.$$

Можно считать ненулевым минор $\det \Phi'_y(a)$, иначе переставим переменные. Положим

$$f(x, y) = \Phi(x, y) - \Phi(a), \quad \text{где } (x, y) \in E.$$

Тогда $\det f'_y(a) \neq 0$, и по теореме 1 существуют окрестность V_a точки a, открытое в $\mathbb{R}^m$ множество U, содержащее x^0, и отображение $\psi \in C^s(U, \mathbb{R}^{n-m})$, для которых

$$\bigl\{(x, y) \in E : f(x, y) = 0\bigr\} \cap V_a = \bigl\{\bigl(x, \psi(x)\bigr) : x \in U\bigr\}.$$

Тогда

$$M \cap \widetilde{V}_a \cap V_a = \bigl\{\bigl(x, \psi(x)\bigr) : x \in U\bigr\},$$

то есть M является m-мерным графиком класса C^s вблизи точки a.

1) $\Rightarrow$ 3) Пусть окрестность V_a точки a, открытое в $\mathbb{R}^m$ множество U, содержащее x^0, и отображение $\psi \in C^s(U, \mathbb{R}^{n-m})$ таковы, что

$$M \cap V_a = \bigl\{\bigl(x, \psi(x)\bigr) : x \in U\bigr\}.$$

Положим

$$\varphi(x) = \bigl(x, \psi(x)\bigr) \quad (x \in U)$$

и проверим, что φ удовлетворяет условиям определения 5. Покажем вначале, что φ — гомеоморфизм. Если $\varphi(v) = \varphi(u)$, то

$$v = P_x\bigl(\varphi(v)\bigr) = P_x\bigl(\varphi(u)\bigr) = u,$$

откуда следует инъективность φ. Непрерывность φ^{-1} вытекает из соотношения

$$\varphi^{-1}(z) = P_x(z) \quad \text{для любого } z \in \varphi(U).$$

Проверим теперь регулярность φ в точке x^0. Матрицу $\varphi'(x^0)$ размера $n \times m$ можно записать в виде

$$\varphi'(x^0) = \left(\frac{I_m}{\psi'(x^0)} \right),$$

где I_m — единичная матрица размера $m \times m$. Поэтому $\varphi'(x^0)$ имеет максимальный ранг, равный m. Таким образом, отображение φ является m-мерной параметризацией M класса C^s вблизи точки a.

3) $\Rightarrow$ 1) Пусть отображение φ является параметризацией M вблизи a. Обозначим $c = \varphi^{-1}(a)$. Поскольку ранг $\varphi'(c)$ равен m, мы можем занумеровать переменные так, что

$$\det \left\{ \frac{\partial \varphi_i}{\partial x_j}(c) \right\}_{i,j=1}^{m} \neq 0, \quad \text{то есть} \quad \det(P_x \circ \varphi)'(c) \neq 0.$$

В силу теоремы 5 § 7 существует окрестность V_c точки c, на которой $P_x \circ \varphi$ будет диффеоморфизмом класса C^s. По непрерывности φ^{-1} и φ найдется такая окрестность V_a точки a, что множество $W = \varphi^{-1}(M \cap V_a)$ открыто в V_c. Положим

$$\gamma = (P_x \circ \varphi)\big|_W, \quad \tau = \gamma^{-1}, \quad U = \gamma(W).$$

Так как $\tau \in C^s(U, W)$, по теореме $2'$ § 5 классу C^s принадлежит и отображение

$$\psi(x) = (P_y \circ \varphi \circ \tau)(x) \quad \big(x \in U\big).$$

Заметим, что для любого $x \in U$

$$(\varphi \circ \tau)(x) = \big((P_x \circ \varphi \circ \tau)(x), (P_y \circ \varphi \circ \tau)(x)\big) = \big(x, \psi(x)\big).$$

Тогда

$$M \cap V_a = \varphi(W) = (\varphi \circ \tau)(U) = \big\{ \big(x, \psi(x)\big) : x \in U \big\}. \quad \square$$

Теперь мы можем определить гладкую поверхность в $\mathbb{R}^n$.

Определение 6. Поверхность в $\mathbb{R}^n$. Пусть $m, n, s \in \mathbb{N}$, $m < n$, $M \subset \mathbb{R}^n$. Предположим, что для любой точки $a \in M$ множество M

удовлетворяет одному из определений 3 – 5. Тогда M называется *гладкой поверхностью* или *гладким многообразием* размерности m класса C^s в $\mathbb{R}^n$.

Определение корректно в силу теоремы 3.

Замечание 1. Поверхности размерности 1 называют также *регулярными гладкими кривыми*. Они допускают вблизи каждой точки параметризацию, имеющую ненулевую производную. Таким образом, регулярная кривая локально удовлетворяет определению кривой как множества точек (см. § 6 главы 4).

Замечание 2. В дальнейшем мы обобщим определение 6 и введем *поверхности с краем*. Таковыми окажутся, например, отрезок в $\mathbb{R}$ или полусфера в $\mathbb{R}^3$, содержащая диаметральную окружность.

В заключение докажем теорему о гладкой согласованности различных параметризаций.

Теорема 4. *Пусть $m, n, s \in \mathbb{N}$, $m < n$, M — m-мерная поверхность класса C^s в $\mathbb{R}^n$, $a \in M$. Предположим, что $\varphi \in C^s(U, M)$ и $\widetilde{\varphi} \in C^s(\widetilde{U}, M)$ — m-мерные параметризации M вблизи a, регулярные соответственно на U и $\widetilde{U}$, причем $\varphi(U) = \widetilde{\varphi}(\widetilde{U})$. Тогда отображение $\widetilde{\varphi}^{-1} \circ \varphi$ является диффеоморфизмом класса C^s множеств U и $\widetilde{U}$.*

Отметим, что силу замечания 2 к определению 5 требование регулярности параметризаций на U и $\widetilde{U}$ не является дополнительным ограничением.

Доказательство. Положим $\theta = \widetilde{\varphi}^{-1} \circ \varphi$. Пусть $x^0 \in U$, $\widetilde{x}^0 = \theta(x^0)$. Будем не умаляя общности считать, что

$$\det \left\{ \frac{\partial \widetilde{\varphi}_i}{\partial x_j}(\widetilde{x}^0) \right\}_{i,j=1}^{m} \neq 0.$$

Рассмотрим отображения $\Phi \colon U \to \mathbb{R}^m$ и $\widetilde{\Phi} \colon \widetilde{U} \to \mathbb{R}^m$, действующие по формулам

$$\Phi(x) = \big(\varphi_1(x), \dots, \varphi_m(x)\big) \quad (x \in U),$$
$$\widetilde{\Phi}(x) = \big(\widetilde{\varphi}_1(x), \dots, \widetilde{\varphi}_m(x)\big) \quad (x \in \widetilde{U}).$$

Заметим, что из равенства $\widetilde{\varphi}\circ\theta=\varphi$ вытекает, что $\widetilde{\Phi}\circ\theta=\Phi$ на U. Так как $\det\widetilde{\Phi}'\bigl(\widetilde{x}^0\bigr)\ne 0$, по теореме 5 § 7 найдется такая окрестность $V\subset\widetilde{U}$ точки $\widetilde{x}^0$, что $\widetilde{\Phi}|_V$ — диффеоморфизм класса C^s. Тогда множество $W=\Phi^{-1}\bigl(\widetilde{\Phi}(V)\bigr)$ открыто по теореме 5 § 7 и теореме 3 § 5 главы 5. Кроме того, $\theta=\widetilde{\Phi}^{-1}\circ\Phi$ на W. Заметим, что $x^0\in W$, так как

$$\widetilde{\Phi}^{-1}\bigl(\Phi(x^0)\bigr)=\theta\bigl(x^0\bigr)=\widetilde{x}^0\in V.$$

По теореме 2′ § 5 $\theta|_W\in C^s(W,\mathbb{R}^m)$. Отсюда в силу произвольности точки $x^0\in U$ вытекает включение $\theta\in C^s\bigl(U,\widetilde{U}\bigr)$. Меняя в этом рассуждении местами φ и $\widetilde{\varphi}$, мы докажем, что $\theta^{-1}\in C^s\bigl(\widetilde{U},U\bigr)$. □

Замечание. Пары (U,φ) и $(\widetilde{U},\widetilde{\varphi})$ в геометрии называют *картами на M*, а отображение $\widetilde{\varphi}^{-1}\circ\varphi$ — *сличением карт* или *переходом от параметризации φ к $\widetilde{\varphi}$*. Таким образом, сличение двух карт, изображающих одну и ту же часть поверхности M, имеет тот же класс гладкости, что и сама поверхность.

§ 9. Экстремум функций нескольких переменных

Основной целью параграфа является обобщение результатов § 7 главы 3 на многомерный случай. Исследование функции одной переменной на экстремум происходило в два этапа. Вначале с помощью *необходимого условия* отбрасывались точки, в которых экстремума заведомо не может быть. Затем для каждой из оставшихся точек мы с помощью *достаточного условия* проверяли, есть ли в ней экстремум и какого он типа. Аналогичная схема будет реализована и для функций нескольких переменных.

Другая важная задача — нахождение наибольшего и наименьшего значений функции, заданной на компактном подмножестве $\mathbb{R}^n$. Напомним, что для нахождения наибольшего значения функции на отрезке мы сравнивали ее значения в критических точках и на концах промежутка. Если функция задана на компактном множестве $E\subset\mathbb{R}^n$, то вместо концов отрезка нужно рассматривать множество всех точек E, не являющихся внутренними, то есть *границу* E. Это множество, определенное в § 3 главы 5, может иметь весьма сложный вид. Исследование функции на границе E приводит нас к понятию *условного экстремума*, которое также будет подробно изучено. В одномерной ситуации это понятие бессодержательно, поэтому оно не вводилось в главе 3.

Рассмотрим вначале первую из поставленных задач — исследование функции нескольких переменных на экстремум.

Определение 1. Экстремум функции. Предположим, что $E \subset \mathbb{R}^n$, $a \in E$, $f: E \to \mathbb{R}$.

1) Пусть существует такая окрестность V_a точки a, что

$$f(x) \geqslant f(a) \quad \text{при всех} \;\; x \in V_a \cap E.$$

Тогда a называется *точкой минимума* f. Если же

$$f(x) \leqslant f(a) \quad \text{при всех} \;\; x \in V_a \cap E,$$

то a называется *точкой максимума* f.

2) Пусть существует такая окрестность V_a точки a, что

$$f(x) > f(a) \quad \text{при всех} \;\; x \in \dot{V}_a \cap E.$$

Тогда a называется *точкой строгого минимума* f. Если же

$$f(x) < f(a) \quad \text{при всех} \;\; x \in \dot{V}_a \cap E,$$

то a называется *точкой строгого максимума* f.

3) Если a является точкой минимума или максимума f, то ее называют *точкой экстремума* f.

Замечание 1. Значение f в точке минимума является *локально наименьшим*, то есть наименьшим в некоторой окрестности этой точки. На всем множестве E оно может не быть наименьшим (см. § 7 главы 3). Аналогичное замечание верно и для точек максимума.

Замечание 2. Если точка a является внутренней для E, то в определении 1 вместо $V_a \cap E$ можно писать просто V_a, что мы и будем делать в дальнейшем.

Теорема 1. Необходимое условие экстремума функции нескольких переменных. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, функция $f: E \to \mathbb{R}$ дифференцируема в точке a. Если a является точкой экстремума f, то $d_a f = \mathbb{O}$.*

Замечание. Утверждение теоремы означает, что $\operatorname{grad} f(a) = 0$ или, более подробно,

$$\frac{\partial f}{\partial x_k}(a) = 0 \quad \text{при всех} \ \ k \in \{1, \dots, n\}.$$

Такую точку a называют *стационарной для функции* f. При $n = 1$ это определение стационарной точки согласуется с одномерным, которое было дано в § 7 главы 3.

Доказательство. Будем для определенности считать, что a — точка минимума f. По определению 1 найдется такое $\delta > 0$, что $f(x) \geqslant f(a)$ при всех $x \in V_a(\delta)$. Пусть $h \in \mathbb{R}^n \setminus \{0\}$. Положим

$$F(t) = f(a + th), \quad \text{где} \ \ t \in \left(-\frac{\delta}{\|h\|}, \frac{\delta}{\|h\|}\right).$$

По следствию леммы 3 § 6 функция F дифференцируема в нуле, причем $F'(0) = d_a f(h)$. Заметим, что при $|t| < \dfrac{\delta}{\|h\|}$ справедливо неравенство $\|a + th - a\| < \delta$, откуда

$$F(t) = f(a + th) \geqslant f(a) = F(0).$$

Поэтому функция F имеет минимум в нуле. В силу теоремы 2 § 7 главы 3 $F'(0) = 0$, то есть

$$d_a f(h) = 0 \quad \text{при любых} \ \ h \in \mathbb{R}^n \setminus \{0\}.$$

Таким образом, $d_a f = \mathbb{O}$. $\square$

Стационарность точки не гарантирует наличия в ней экстремума даже для функции одной переменной. Поэтому нужно вывести какое-то достаточное условие экстремума, которое позволяло бы исследовать стационарные точки. В одномерной ситуации мы разобрали два таких условия. Первое из них было основано на изучении монотонности f слева и справа от точки a. Однако уже для функций двух переменных понятие монотонности теряет смысл. Второе условие заключалось в исследовании знака второй производной или *второго дифференциала* f в точке a. Это условие мы и обобщим на многомерный случай.

Напомним, что с алгебраической точки зрения второй дифференциал является *квадратичной формой*. Поэтому мы рекомендуем читателю вспомнить сведения о квадратичных формах, изложенные в § 7 главы 5.

Теорема 2. Достаточное условие экстремума функции нескольких переменных. *Пусть $E \subset \mathbb{R}^n$, a — внутренняя точка E, функция $f\colon E \to \mathbb{R}$ дважды дифференцируема в точке a, причем $d_a f = \mathbb{O}$. Тогда справедливы следующие утверждения.*

1) *Если $d_a^2 f$ положительно определен, то a — точка строгого минимума f.*

2) *Если $d_a^2 f$ отрицательно определен, то a — точка строгого максимума f.*

3) *Если $d_a^2 f$ принимает как положительные, так и отрицательные значения, то f не имеет экстремума в точке a.*

Замечание. Теорема 2 не охватывает ситуации, когда второй дифференциал сохраняет знак, но не является положительно или отрицательно определенным (см. замечание к определению 6 § 7 главы 5). В частности, теорема ничего не утверждает в случае $d_a^2 f \equiv 0$.

Доказательство. По формуле Тейлора – Пеано (теорема 1 § 6)

$$f(a+h) = f(a) + d_a f(h) + \frac{d_a^2 f(h)}{2} + o\big(\|h\|^2\big) \quad (h \to 0).$$

С учетом условия $d_a f = \mathbb{O}$ мы получим

$$f(a+h) - f(a) = \frac{d_a^2 f(h)}{2} + \alpha(h)\,\|h\|^2, \tag{73}$$

где $\alpha(h) \to 0$ при $h \to 0$. Перейдем к доказательству утверждений теоремы.

1) По теореме 3 § 7 главы 5 существует $C > 0$, для которого

$$d_a^2 f(h) \geqslant C\,\|h\|^2 \quad \text{при всех } h \in \mathbb{R}^n.$$

Кроме того, найдется такое $\delta > 0$, что $V_a(\delta) \subset E$ и

$$|\alpha(h)| < \frac{C}{4} \quad \text{для любого } h \in \dot{V}_0(\delta).$$

Пусть $x \in \dot{V}_a(\delta)$. Полагая $h = x - a$, в силу (73) мы получим

$$f(x)-f(a) \geqslant \frac{d_a^2 f(h)}{2} - |\alpha(h)| \cdot \|h\|^2 \geqslant \frac{C\,\|h\|^2}{2} - \frac{C\,\|h\|^2}{4} = \frac{C\,\|h\|^2}{4} > 0.$$

Таким образом, $f(x) > f(a)$ при всех $x \in \dot{V}_a(\delta)$, то есть a является точкой строгого минимума f.

2) Заметим, что квадратичная форма $d_a^2(-f)$ положительно определена. Тогда по первому утверждению функция $(-f)$ имеет в точке a строгий минимум, а функция f, соответственно, строгий максимум.

3) По условию существуют векторы $h_-, h_+ \in \mathbb{R}^n$, для которых

$$d_a^2 f(h_-) < 0, \quad d_a^2 f(h_+) > 0.$$

Положим $C_\pm = d_a^2 f\bigl(h_\pm\bigr)$. Достаточно проверить, что в любой окрестности V_a точки a функция $f - f(a)$ принимает как положительные, так и отрицательные значения. Пусть $V_a \subset E$. По теореме о пределе композиции $\lim\limits_{t\to 0} \alpha(th_+) = 0$. Поэтому при достаточно малых $t \in \mathbb{R} \setminus \{0\}$ выполняются условия

$$a + th_+ \in \dot{V}_a \quad \text{и} \quad \bigl|\alpha(th_+)\bigr| \cdot \|h_+\|^2 \leqslant \frac{C_+}{4}.$$

Тогда в силу (73)

$$\begin{aligned} f(a+th_+) - f(a) &= \frac{t^2\, d_a^2 f(h_+)}{2} + \alpha(th_+)\, t^2\, \|h_+\|^2 \geqslant \\ &\geqslant t^2 \left(\frac{C_+}{2} - \bigl|\alpha(th_+)\bigr| \cdot \|h_+\|^2 \right) \geqslant \frac{t^2 C_+}{4} > 0. \end{aligned}$$

Аналогичные рассуждения показывают, что при достаточно малых $t \neq 0$ выполняется неравенство $f(a + th_-) - f(a) < 0$. $\square$

Разберем более подробно случай функций двух переменных. Мы будем записывать элементы $\mathbb{R}^2$ в виде (x, y).

Следствие. *Пусть $E \subset \mathbb{R}^2$, a — внутренняя точка E, функция $f\colon E \to \mathbb{R}$ дважды дифференцируема в точке a, $d_a f = \mathbb{O}$. Положим*

$$A = \frac{\partial^2 f}{\partial x^2}(a), \quad B = \frac{\partial^2 f}{\partial x \partial y}(a), \quad C = \frac{\partial^2 f}{\partial y^2}(a).$$

Тогда справедливы следующие утверждения.

1) *Если* $A > 0$ *и* $AC > B^2$, *то* a *– точка строгого минимума* f.
2) *Если* $A < 0$ *и* $AC > B^2$, *то* a *– точка строгого максимума* f.
3) *Если* $AC < B^2$, *то* f *не имеет экстремума в точке* a.

Доказательство. Проверим вначале утверждения 1) и 2). В силу формулы (40) второй дифференциал f в точке a — квадратичная форма с матрицей Гессе

$$G = \begin{pmatrix} A & B \\ B & C \end{pmatrix}.$$

Вычислим главные миноры матрицы G:

$$\Delta_1 = A, \quad \Delta_2 = \det G = AC - B^2.$$

В силу критерия Сильвестра (теорема 4 § 7 главы 5) и его следствия условия утверждения 1) гарантируют положительную определенность $d_a^2 f$, а условия утверждения 2) — отрицательную определенность $d_a^2 f$. Осталось сослаться на теорему 2.

Докажем теперь утверждение 3). По формуле (41)

$$d_a^2 f(dx, dy) = A dx^2 + 2B dx dy + C dy^2.$$

При фиксированном $dy \neq 0$ выражение, стоящее в правой части, является квадратным трехчленом относительно dx с положительным дискриминантом, равным $4\,dy^2(B^2 - AC)$. Поэтому $d_a^2 f$ принимает как положительные, так и отрицательные значения. Тогда по теореме 2 у f нет экстремума в точке a. $\square$

Замечание. Приведенная в следствии классификация не охватывает случай $AC = B^2$. Он соответствует ситуации, когда второй дифференциал f является полным квадратом, то есть имеет вид

$$d_a^2 f(dx, dy) = \pm(\alpha dx + \beta dy)^2, \quad \text{где } \alpha, \beta \in \mathbb{R}.$$

В частности, при *нулевом* втором дифференциале мы не можем сделать выводов о наличии у f экстремума в точке a.

Как же исследовать стационарную точку функции, если теорема 2 не позволяет этого сделать? Мы не будем формулировать каких-либо общих утверждений, а приведем несколько примеров, иллюстрирующих различные подходы к задаче.

Пример 1. Пусть $f(x,y) = x^4+3xy^2+y^4$, $a = (0,0)$. Предлагаем читателю проверить, что $d_a f \equiv 0$ и $d_a^2 f \equiv 0$. В этом случае можно рассмотреть *третий дифференциал* f. Покажем, что $d_a^3 f \not\equiv 0$. Действительно, по формуле (42)

$$d_a^3 f(dx, dy) = C_3^1 \, \frac{\partial^3 f}{\partial x \partial y^2}(a) \, dx \, dy^2 = 18 dx \, dy^2,$$

поскольку $\dfrac{\partial^3 f}{\partial x \partial y^2}(a) = 6$, а все остальные частные производные f третьего порядка равны нулю в точке a. Пусть $h \in \mathbb{R}^2$ — произвольный вектор, для которого $d_a^3 f(h) \neq 0$. Положим $\varphi(t) = f(th)$ $(t \in \mathbb{R})$. В силу следствия леммы 3 § 6

$$\varphi'(0) = \varphi''(0) = 0, \quad \varphi'''(0) = d_a^3 f(h) \neq 0.$$

Тогда по теореме 5 § 7 главы 3 φ не имеет экстремума в нуле и, тем более, у f нет экстремума в точке $(0,0)$.

Пример 2. Пусть $f(x,y) = x^2 e^y - 2xy + y^2 e^x$, $a = (0,0)$. Предлагаем читателю проверить, что $d_a f \equiv 0$ и

$$\frac{\partial^2 f}{\partial x^2} = 2e^y + y^2 e^x, \quad \frac{\partial^2 f}{\partial x \partial y} = 2xe^y - 2 + 2ye^x, \quad \frac{\partial^2 f}{\partial y^2} = 2e^x + x^2 e^y.$$

Используя формулу (41), мы получим

$$d_a^2 f(dx, dy) = 2 \, dx^2 - 4 \, dx \, dy + 2 \, dy^2 = 2 \, (dx - dy)^2.$$

Дифференциал $d_a^2 f$ неотрицателен, но он обращается в ноль на прямой ℓ, задаваемой уравнением $y = x$. Это наводит на мысль, что необходимо исследовать поведение функции f на ℓ. Положим $\varphi(t) = f(t,t)$ $(t \in \mathbb{R})$. Тогда $\varphi(t) = 2t^2 \, (e^t - 1)$. Очевидно, что $\varphi(t) > 0$ при $t > 0$ и $\varphi(t) < 0$ при $t < 0$. Поэтому φ не имеет экстремума в нуле и, тем более, у f нет экстремума в точке $(0,0)$.

Заметим, что сужение f на любую прямую, проходящую через $(0,0)$ и отличную от ℓ, имеет в нуле минимум. Действительно, пусть $h \in \mathbb{R}^2 \setminus \ell$. Положим $\psi(t) = f(th)$. Тогда по формуле (47)

$$\psi'(0) = d_a f(h) = 0, \quad \psi''(0) = d_a^2 f(h) > 0.$$

Осталось сослаться на теорему 4 § 7 главы 3.

Пример 3. Пусть $f(x,y) = y^4 - y^2 + 2x^2 y$, $a = (0,0)$. Предлагаем читателю проверить, что $d_a f \equiv 0$ и

$$\frac{\partial^2 f}{\partial x^2} = 4y, \quad \frac{\partial^2 f}{\partial x \partial y} = 4x, \quad \frac{\partial^2 f}{\partial y^2} = 12y^2 - 2.$$

По формуле (41) $d_a^2 f(dx, dy) = -2\,dy^2$. Повторяя рассуждения предыдущего примера, можно доказать, что сужение f на любую прямую, проходящую через $(0,0)$ и отличную от OX, имеет максимум в нуле. Так как $f(x,0) \equiv 0$, это верно и для OX.

Покажем, однако, что f не имеет максимума в точке $(0,0)$. Заметим, что $f(x,y) = y^4 + x^4 - \left(y - x^2\right)^2$. Поэтому

$$f(t, t^2) = t^8 + t^4 > 0 = f(0,0) \quad \text{при всех } \ t \neq 0.$$

Пример 4. Пусть $f(x,y) = 4 \sin xy - (x+y)^2$, $a = (0,0)$. Тогда $d_a^2 f(dx, dy) = -2\,(dx - dy)^2$. Покажем, что f имеет максимум в точке $(0,0)$. Предположим, что $(x,y) \in V_{(0,0)}(1)$. Если $xy \leqslant 0$, то $\sin xy \leqslant 0$ и

$$f(x,y) \leqslant -(x+y)^2 \leqslant 0 = f(0,0).$$

Если же $xy \geqslant 0$, то $\sin xy \leqslant xy$ (см. § 5 главы 2), откуда

$$f(x,y) \leqslant 4xy - (x+y)^2 = -(x-y)^2 \leqslant 0 = f(0,0).$$

Таким образом, $f(x,y) \leqslant f(0,0)$ при всех $(x,y) \in V_{(0,0)}(1)$.

Рассмотрим теперь задачу об *условном экстремуме*. Пусть $f \colon E \subset \mathbb{R}^n \to \mathbb{R}$, $a \in E$ — точка экстремума f. Это означает, что при любом x из некоторой окрестности V_a точки a в случае минимума выполняется неравенство $f(x) \geqslant f(a)$, а в случае максимума — $f(x) \leqslant f(a)$. Предположим теперь, что нас интересуют не все точки $x \in V_a$, а только те, которые удовлетворяют некоторым дополнительным равенствам, называемым *уравнениями связи*. Тогда говорят об *условном* или *относительном экстремуме*, подчиненном этим уравнениям связи. Определим понятие условного экстремума строго.

Определение 2. Условный (относительный) экстремум. Пусть $E \subset \mathbb{R}^n$, $a \in E$, $m \in \mathbb{N}$, $m < n$, $f, \Phi_1, \dots, \Phi_m \colon E \to \mathbb{R}$. Предположим, что a является точкой экстремума (минимума или максимума) сужения функции f на множество

$$\big\{x \in E : \Phi_k(x) = \Phi_k(a) \ \text{ при всех } \ k = 1, \dots, m\big\}. \tag{74}$$

Тогда a называется точкой *условного* или *относительного экстремума* f соответствующего типа, подчиненного *уравнениям связи* $\Phi_k(x) = \Phi_k(a)$, где $k \in \{1, \dots, m\}$, $x \in E$.

Замечание 1. Если ввести функции $\widetilde{\Phi}_k = \Phi_k - \Phi_k(a)$, то уравнения связи примут вид $\widetilde{\Phi}_k(x) = 0$. В такой форме их и принято записывать. Поэтому в дальнейшем мы будем часто предполагать, что $\Phi_k(a) = 0$ при всех $k \in \{1, \dots, m\}$.

Замечание 2. Предположим, что a является внутренней точкой E и функции $\Phi_1, \dots, \Phi_m$ дифференцируемы в точке a. Обозначим через Φ отображение с координатными функциями $\Phi_1, \dots, \Phi_m$. Мы будем считать, что Φ регулярно в точке a, то есть

$$\operatorname{rank} \Phi'(a) = m. \tag{75}$$

Если $\Phi_1, \dots, \Phi_m$ линейны, то условие (75) означает линейную независимость уравнений системы $\Phi(x) = \Phi(a)$. В общем случае мы будем говорить, что *уравнения связи независимы в точке* a.

Замечание 3. Предположим, что множество E открыто в $\mathbb{R}^n$, отображение Φ класса $C^1\big(E, \mathbb{R}^m\big)$ регулярно на E. В соответствии с определениями 4 и 6 § 8 множество (74) вблизи точки a является $(n-m)$-мерной поверхностью класса C^1. Таким образом, задачу об условном экстремуме можно понимать как *исследование на экстремум функции, заданной на гладкой поверхности*. Мы будем использовать эту трактовку при нахождении наибольшего и наименьшего значений f на компакте.

Изучим теперь вопрос о поиске точек условного экстремума. Эта задача решается в два этапа: вначале с помощью необходимого условия мы отбираем точки, подозрительные на экстремум, а затем исследуем их, используя достаточное условие. Рассмотрим эту схему более подробно.

Теорема 3. Необходимое условие относительного экстремума. *Пусть $m, n \in \mathbb{N}$, $m < n$, множество E открыто в $\mathbb{R}^n$, $a \in E$, $f \in C^1(E)$, $\Phi \in C^1(E, \mathbb{R}^m)$, $\operatorname{rank} \Phi'(a) = m$. Предположим, что a является точкой условного экстремума f, подчиненного уравнениям связи $\Phi = 0$. Тогда существуют $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$, для которых*

$$\operatorname{grad} f(a) = \sum_{k=1}^{m} \lambda_k \cdot \operatorname{grad} \Phi_k(a). \tag{76}$$

Замечание. Числа $\lambda_1, \ldots, \lambda_m$, удовлетворяющие (76), называют *множителями Лагранжа*.

Доказательство. Договоримся записывать элементы $\mathbb{R}^n$ в виде (x, y), где $x \in \mathbb{R}^{n-m}$, $y \in \mathbb{R}^m$, и пусть $a = \left(x^0, y^0\right)$. Обозначим через M множество (74). Будем не умаляя общности считать, что $\det \Phi'_y(a) \neq 0$ (матрицы Φ'_x и Φ'_y определяются в соответствии с формулами (68)). Тогда по теореме о неявном отображении существуют окрестность V_a точки a, открытое в $\mathbb{R}^{n-m}$ множество U, содержащее x^0, и отображение $\psi \in C^1(U, \mathbb{R}^m)$, для которых

$$M \cap V_a = \left\{ \left(x, \psi(x)\right) : x \in U \right\}.$$

Рассмотрим функцию F, действующую по формуле

$$F(x) = f\left(x, \psi(x)\right) \quad (x \in U).$$

Так как $f|_M$ имеет экстремум в точке a, из определения ψ вытекает, что x^0 является точкой экстремума для F. Тогда теореме 1 $F'\left(x^0\right) = 0$, и в силу (72)

$$f'_x(a) + f'_y(a) \cdot \psi'\left(x^0\right) = 0. \tag{77}$$

Кроме того,

$$\Phi\left(x, \psi(x)\right) = 0 \quad \text{при всех } x \in U.$$

Отсюда в силу (72) вытекает матричное равенство

$$\Phi'_x(a) + \Phi'_y(a) \cdot \psi'\bigl(x^0\bigr) = 0.$$

Умножим его слева на строку $\lambda = (\lambda_1, \ldots, \lambda_m)$ и вычтем результат из (77). Тогда

$$\bigl(f'_x(a) - \lambda \cdot \Phi'_x(a)\bigr) + \bigl(f'_y(a) - \lambda \cdot \Phi'_y(a)\bigr) \cdot \psi'\bigl(x^0\bigr) = 0. \tag{78}$$

Положим теперь $\lambda = f'_y(a) \cdot \bigl(\Phi'_y(a)\bigr)^{-1}$ (напомним, что матрица $\Phi'_y(a)$ обратима ввиду условия $\det \Phi'_y(a) \neq 0$). При таком выборе λ второе слагаемое в (78) обращается в ноль, а тогда нулевым будет и первое слагаемое. Таким образом, для любых $i \in \{1, \ldots, n-m\}$ и $j \in \{1, \ldots, m\}$

$$\frac{\partial f}{\partial x_i}(a) = \sum_{k=1}^{m} \lambda_k \cdot \frac{\partial \Phi_k}{\partial x_i}(a)\,, \qquad \frac{\partial f}{\partial y_j}(a) = \sum_{k=1}^{m} \lambda_k \cdot \frac{\partial \Phi_k}{\partial y_j}(a).$$

Из этих двух равенств и получается (76). $\square$

Замечание. Условия теоремы 3 можно ослабить: достаточно предполагать непрерывную дифференцируемость отображений f и Φ лишь в некоторой окрестности точки a.

На практике необходимое условие относительного экстремума часто записывают как условие стационарности для некоторой вспомогательной функции. Сформулируем теорему 3 в такой редакции.

Теорема 4. *Пусть $m, n \in \mathbb{N}$, $m < n$, множество E открыто в $\mathbb{R}^n$, $a \in E$, $f \in C^1(E)$, $\Phi \in C^1(E, \mathbb{R}^m)$, $\operatorname{rank} \Phi'(a) = m$. Предположим, что a является точкой условного экстремума f, подчиненного уравнениям связи $\Phi = 0$. Обозначим*

$$L(x, \lambda) = f(x) - \sum_{k=1}^{m} \lambda_k \, \Phi_k(x) \quad (x \in E, \ \lambda \in \mathbb{R}^m). \tag{79}$$

Тогда существует вектор $\lambda \in \mathbb{R}^m$, для которого

$$\operatorname{grad} L(a, \lambda) = 0. \tag{80}$$

Замечание. Функция L, определенная равенством (79), называется *функцией Лагранжа*. Точка a, удовлетворяющая условию (80), называется *критической для f относительно уравнения связи* $\Phi = 0$.

Доказательство. Выберем для функции f множители Лагранжа $\lambda_1, \ldots, \lambda_m$ в соответствии с теоремой 3 и покажем, что вектор $\lambda = (\lambda_1, \ldots, \lambda_m)$ искомый. Действительно, для всех $i \in \{1, \ldots, n\}$ в силу (76)

$$\frac{\partial L}{\partial x_i}(a, \lambda) = \frac{\partial f}{\partial x_i}(a) - \sum_{k=1}^{m} \lambda_k \frac{\partial \Phi_k}{\partial x_i}(a) = 0.$$

Кроме того, для любого $i \in \{1, \ldots, m\}$

$$\frac{\partial L}{\partial \lambda_i}(a, \lambda) = -\Phi_i(a) = 0.$$

Из этих двух равенств и вытекает утверждение теоремы. $\square$

С помощью теорем 3 и 4 удобно искать *наибольшее и наименьшее значения функции на гладкой поверхности*. Предположим, что M — поверхность класса C^1 в $\mathbb{R}^n$ и $f\colon M \to \mathbb{R}$. Пусть нам известно, что функция f должна достигать на M своего наибольшего значения в некоторой точке $a \in M$. Тогда a является для f точкой условного экстремума (см. замечание 3 к определению 2). Используя теоремы 3 или 4, мы можем найти множество C всех точек M, подозрительных на условный экстремум, а затем вычислить $\max\limits_{x \in C} f(x)$, который равен $\max\limits_{x \in M} f(x)$. Аналогичным образом ищется наименьшее значение f на M. В качестве приложения этой схемы рассмотрим задачу о максимуме и минимуме квадратичной формы на единичной сфере в $\mathbb{R}^n$.

Теорема 5. *Пусть A — симметричная матрица размера $n \times n$, Λ — множество собственных чисел A, f — квадратичная форма с матрицей A, $M = \{x \in \mathbb{R}^n : \|x\| = 1\}$. Тогда*

$$\min_{x \in M} f(x) = \min \Lambda \quad \text{и} \quad \max_{x \in M} f(x) = \max \Lambda.$$

Замечание. Из курса алгебры известно, что у симметричной матрицы все собственные числа вещественны.

Доказательство. В главе 5 было показано, что сфера M компактна в $\mathbb{R}^n$, а функция f непрерывна на $\mathbb{R}^n$. По теореме Вейерштрасса f достигает наибольшего и наименьшего значения в некоторых точках M. В соответствии с описанной выше схемой нам нужно найти точки, удовлетворяющие необходимому условию относительного экстремума, подчиненного уравнению связи $\Phi(x) = 0$, где

$$\Phi(x_1, \dots, x_n) = \sum_{i=1}^n x_i^2 - 1.$$

Заметим, что $\operatorname{grad} \Phi \neq 0$ на M. Функция Лагранжа (79) для этой задачи имеет вид

$$L(x, \lambda) = f(x) - \lambda\, \Phi(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}\, x_i\, x_j - \lambda \left(\sum_{i=1}^n x_i^2 - 1 \right),$$

где через a_{ij} обозначаются элементы матрицы A. В силу симметричности матрицы A для любого $k \in \{1, \dots, n\}$ мы получим

$$\frac{\partial L}{\partial x_k} = \sum_{i=1}^n a_{ik}\, x_i + \sum_{j=1}^n a_{kj}\, x_j - 2\lambda\, x_k = 2 \left(\sum_{j=1}^n a_{kj}\, x_j - \lambda\, x_k \right),$$

а равенство $\dfrac{\partial L}{\partial \lambda} = 0$ есть уравнение связи. Поэтому условие (80) примет вид

$$\sum_{j=1}^n a_{kj}\, x_j = \lambda\, x_k \ \text{ при всех } \ k \in \{1, \dots, n\} \quad \text{и} \quad \sum_{i=1}^n x_i^2 = 1,$$

или, в матричной форме,

$$A \cdot x = \lambda\, x \quad \text{и} \quad \|x\|^2 = 1.$$

Таким образом, подозрительными на условный экстремум являются собственные векторы x матрицы A единичной нормы. Для них

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}\, x_j\, x_i = \sum_{i=1}^n \lambda\, x_i \cdot x_i = \lambda \sum_{i=1}^n x_i^2 = \lambda.$$

Отсюда и вытекают равенства $\min\limits_M f = \min \Lambda$ и $\max\limits_M f = \max \Lambda$. $\square$

Замечание. В процессе доказательства установлено, что критическими точками f на единичной сфере являются собственные векторы A.

Теорема 5 дает возможность вычислять норму линейного оператора, действующего из $\mathbb{R}^n$ в $\mathbb{R}^m$, в терминах его матрицы.

Следствие. *Пусть $m, n \in \mathbb{N}$, $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, A — матрица T. Тогда*

$$\|T\| = \max\bigl\{\sqrt{\lambda} : \lambda \text{ — } \textit{собственное число матрицы } A^T A\bigr\}.$$

Отметим, что матрица $A^T A$ симметрична.

Доказательство. Пусть f — квадратичная форма с матрицей $A^T A$. Если записать вектор $x \in \mathbb{R}^n$ в виде столбца, то

$$f(x) = x^T \left(A^T A\right) x = (Ax)^T \cdot Ax = \sum_{i=1}^n (Ax)_i^2 = \|Ax\|^2 = \|Tx\|^2 .$$

Переходя в этом равенстве к супремуму по всем x единичной нормы, мы получим справа $\|T\|^2$, а слева — максимальное собственное число матрицы $A^T A$. $\square$

Замечание. Все собственные числа матрицы $A^T A$ неотрицательны, так как они являются значениями f на соответствующих собственных векторах.

Рассмотрим теперь *достаточное* условие относительного экстремума. Для этого нам потребуется ввести еще одно алгебраическое понятие.

Определение 3. Пусть N — линейное подпространство в $\mathbb{R}^n$, f — квадратичная форма в $\mathbb{R}^n$. Она называется *положительно определенной на* N, если $f(x) > 0$ при всех $x \in N \setminus \{0\}$, и *отрицательно определенной на* N, если $f(x) < 0$ для любых $x \in N \setminus \{0\}$.

Замечание. Если f положительно определена в $\mathbb{R}^n$, то она положительно определена и на любом линейном подпространстве $\mathbb{R}^n$.

Обратное неверно: форма $f(x, y) = x^2 - y^2$ положительно определена на оси OX, но не в $\mathbb{R}^2$.

Теорема 6. Достаточное условие относительного экстремума. *Пусть $m, n \in \mathbb{N}$, $m < n$, множество E открыто в $\mathbb{R}^n$, $a \in E$, $f \in C^2(E)$, $\Phi \in C^2(E, \mathbb{R}^m)$, $\Phi(a) = 0$, $\operatorname{rank} \Phi'(a) = m$. Положим*

$$N = \{h \in \mathbb{R}^n : d_a\Phi(h) = 0\},$$
$$L(x) = f(x) - \sum_{k=1}^{m} \lambda_k \, \Phi_k(x) \quad (x \in E),$$

где $\lambda_1, \ldots, \lambda_m$ удовлетворяют условию (76) в точке a. Тогда справедливы следующие утверждения.

1) *Если $d_a^2 L$ положительно определен на N, то f имеет в точке a условный минимум, подчиненный уравнениям связи $\Phi = 0$.*

2) *Если $d_a^2 L$ отрицательно определен на N, то f имеет в точке a условный максимум, подчиненный уравнениям связи $\Phi = 0$.*

3) *Если $d_a^2 L$ принимает на N как положительные, так и отрицательные значения, то f не имеет в точке a условного экстремума, подчиненного уравнениям связи $\Phi = 0$.*

Доказательство мы проведем только для случая $n = 2$, $m = 1$. Это позволит нам без сложных вычислений объяснить читателю смысл условий теоремы и принцип рассуждений.

Условие (76) при $n = 2$ примет вид

$$\frac{\partial f}{\partial x}(a) = \lambda \frac{\partial \Phi}{\partial x}(a), \quad \frac{\partial f}{\partial y}(a) = \lambda \frac{\partial \Phi}{\partial y}(a). \tag{81}$$

Так как $\operatorname{rank} \Phi'(a) = 1$, мы можем считать $\dfrac{\partial \Phi}{\partial y}(a) \neq 0$. Пусть $a = (x^0, y^0)$. По теореме 1 § 8 существуют окрестность V_a точки a, открытое в $\mathbb{R}$ множество U, содержащее x^0, и функция $\psi \colon U \to \mathbb{R}$, для которых

$$\{(x, y) \in V_a : \Phi(x, y) = 0\} = \{(x, \psi(x)) : x \in U\}.$$

Положим $F(x) = f(x, \psi(x))$. Наличие у f условного экстремума в точке a означает, что F имеет экстремум того же типа в точке x^0.

Дифференцируя на U функцию F и тождество $\Phi\big(x, \psi(x)\big) \equiv 0$, мы получим

$$\begin{aligned} F'(x) &= \frac{\partial f}{\partial x}\big(x, \psi(x)\big) + \frac{\partial f}{\partial y}\big(x, \psi(x)\big)\, \psi'(x), \\ 0 &= \frac{\partial \Phi}{\partial x}\big(x, \psi(x)\big) + \frac{\partial \Phi}{\partial y}\big(x, \psi(x)\big)\, \psi'(x). \end{aligned} \tag{82}$$

Подставим в первое из равенств (82) $x = x^0$. Тогда в силу (81)

$$F'\big(x^0\big) = \frac{\partial f}{\partial x}(a) + \frac{\partial f}{\partial y}(a)\, \psi'\big(x^0\big) = \lambda \left(\frac{\partial \Phi}{\partial x}(a) + \frac{\partial \Phi}{\partial y}(a)\, \psi'\big(x^0\big) \right) = 0.$$

Дифференцируя тождества (82) и подставляя $x = x^0$, мы получим

$$F''\big(x^0\big) = \frac{\partial^2 f}{\partial x^2}(a) + 2\, \frac{\partial^2 f}{\partial x \partial y}(a)\, \psi'\big(x^0\big) + \\ + \frac{\partial^2 f}{\partial y^2}(a)\, \big(\psi'\big(x^0\big)\big)^2 + \frac{\partial f}{\partial y}(a)\, \psi''\big(x^0\big), \tag{83}$$

$$0 = \frac{\partial^2 \Phi}{\partial x^2}(a) + 2\, \frac{\partial^2 \Phi}{\partial x \partial y}(a)\, \psi'\big(x^0\big) + \\ + \frac{\partial^2 \Phi}{\partial y^2}(a)\, \big(\psi'\big(x^0\big)\big)^2 + \frac{\partial \Phi}{\partial y}(a)\, \psi''\big(x^0\big). \tag{84}$$

Умножим (84) на λ и вычтем его из (83). Тогда в силу (81)

$$F''\big(x^0\big) = \frac{\partial^2 L}{\partial x^2}(a) + 2\, \frac{\partial^2 L}{\partial x \partial y}(a)\, \psi'\big(x^0\big) + \frac{\partial^2 L}{\partial y^2}(a)\, \big(\psi'\big(x^0\big)\big)^2. \tag{85}$$

Опишем теперь подпространство N. Включение $(dx, dy) \in N$ эквивалентно равенству $\dfrac{\partial \Phi}{\partial x}(a)\, dx + \dfrac{\partial \Phi}{\partial y}(a)\, dy = 0$, откуда в силу (82)

$$dy = -\left(\frac{\partial \Phi}{\partial y}(a) \right)^{-1} \frac{\partial \Phi}{\partial x}(a)\, dx = \psi'\big(x^0\big)\, dx.$$

Таким образом, подпространство N имеет вид

$$N = \big\{ (x, y) \in \mathbb{R}^2 : y = \psi'\big(x^0\big)\, x \big\}.$$

Домножая (85) на h^2, мы получим

$$d^2_{x^0} F(h) = F''\bigl(x^0\bigr)h^2 = d^2_a L\bigl(h, \psi'\bigl(x^0\bigr)h\bigr) \quad (h \in \mathbb{R}).$$

Поэтому положительная определенность $d^2_a L$ на N равносильна положительной определенности $d^2_{x^0} F$, а отрицательная определенность $d^2_a L$ на N — отрицательной определенности $d^2_{x^0} F$. Отсюда по теореме 2 вытекают утверждения 1) и 2), поскольку $F'\bigl(x^0\bigr) = 0$. Очевидно также, что $d^2_a L$ не может принимать на N значения разных знаков, то есть третий случай в теореме при $n = 2$ и $m = 1$ не реализуется. $\square$

Рассмотрим теперь задачу о наибольшем и наименьшем значениях функции нескольких переменных. Пусть E — компактное подмножество $\mathbb{R}^n$. Следуя § 3 главы 5, обозначим через ∂E границу E. Множество ∂E может иметь весьма сложный вид. Мы ограничимся изучением ситуации, когда

$$\partial E = M_1 \cup \ldots \cup M_k,$$

где $M_1, \ldots, M_k$ — поверхности в $\mathbb{R}^n$ класса C^1 различных размерностей, причем $M_i \cap M_j = \varnothing$ при $i \neq j$.

Пусть функция f непрерывно дифференцируема на открытом множестве $\widetilde{E}$, содержащем E. По теореме Вейерштрасса f достигает на E наибольшего или наименьшего значения в некоторой точке $a \in E$. Если a является внутренней точкой E, то по теореме 1 она должна быть стационарной для f. Предположим теперь, что $a \in \partial E$ и $i \in \{1, \ldots, k\}$ таково, что $a \in M_i$. Будем не умаляя общности считать, что поверхность M_i вблизи точки a описывается системой уравнений $\Phi = 0$, где $\Phi \in C^1(\widetilde{E}, \mathbb{R}^m)$ $(m \in \mathbb{N},\ m < n)$, $\operatorname{rank} \Phi'(a) = m$. Тогда по теореме 4 точка a является критической для f на поверхности M_i. Из этих рассуждений вытекает следующий алгоритм поиска наибольшего и наименьшего значений f на E.

1) Из уравнения $\operatorname{grad} f(x) = 0$ ищутся стационарные точки f, внутренние для E (обозначим множество таких точек через C_0).

2) Для любого $i \in \{1, \ldots, k\}$ мы находим критические точки f на поверхности M_i, решая систему (80). Пусть C_i — множество всех таких точек.

3) Наибольшее и наименьшее значения f на E вычисляются по формулам

$$\max_E f = \max_{i\in\{0,\dots,k\}} \max_{C_i} f, \quad \min_E f = \min_{i\in\{0,\dots,k\}} \min_{C_i} f.$$

На практике множества C_i часто оказываются конечными, поэтому $\max\limits_{C_i} f$ и $\min\limits_{C_i} f$ находятся простым перебором. Если какое-то из множеств C_i пусто, то мы считаем $\max\limits_{C_i} f = -\infty$ и $\min\limits_{C_i} f = +\infty$.

Приведем пример, иллюстрирующий описанную схему.

Пример. Пусть множество E в $\mathbb{R}^3$ определяется неравенствами

$$x^2 + y^2 + z^2 \leqslant 1, \quad x^2 + y^2 \leqslant x.$$

Геометрически оно представляет собой пересечение цилиндра и шара (см. рисунок 33).

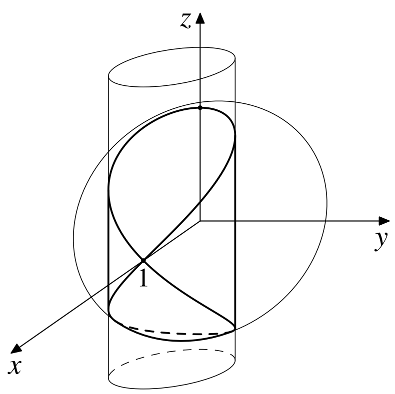

Рис. 33

Необходимо вычислить наибольшее значение на E функции

$$f(x, y, z) = 3x^2 + 2y^2 + 2z^2 + 2yz.$$

Решение поставленной задачи мы опишем кратко, опуская некоторые стандартные вычисления. Тем не менее, все ключевые шаги будут подробно разобраны.

1) *Поиск стационарных точек.* Внутренними для E являются все точки, координаты которых удовлетворяют неравенствам

$$x^2 + y^2 + z^2 < 1, \quad x^2 + y^2 < x.$$

Приравнивая к нулю градиент f, мы получим уравнения

$$6x = 0, \quad 4y + 2z = 0, \quad 4z + 2y = 0.$$

Решением этой системы будет только точка $(0,0,0)$, которая не является внутренней для E. Поэтому $C_0 = \varnothing$, то есть наибольшее значение f реализуется на границе E.

Множество ∂E состоит из трех непересекающихся подмножеств:

$$M_1\colon \quad x^2 + y^2 + z^2 = 1, \quad x^2 + y^2 < x; \tag{86}$$
$$M_2\colon \quad x^2 + y^2 + z^2 < 1, \quad x^2 + y^2 = x; \tag{87}$$
$$M_3\colon \quad x^2 + y^2 + z^2 = 1, \quad x^2 + y^2 = x. \tag{88}$$

Разберем их по отдельности.

2) *Сферическая часть границы.* Заметим, что f является квадратичной формой, порожденной матрицей

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

В силу замечания к теореме 5 критическими точками f на сфере будут собственные векторы матрицы A. Выпишем пары собственных чисел и собственных векторов матрицы A:

$$\begin{aligned} &\lambda = 1, \quad v = (0, \alpha, -\alpha) \quad (\alpha \in \mathbb{R}); \\ &\lambda = 3, \quad v = (\alpha, \beta, \beta) \qquad (\alpha, \beta \in \mathbb{R}). \end{aligned}$$

Первый случай не подходит, так как в силу (86) $\alpha^2 < 0$, что невозможно. Если $\lambda = 3$, то условия (86) примут вид

$$\alpha^2 + 2\beta^2 = 1, \quad \alpha^2 + \beta^2 < \alpha.$$

Исключая β^2, мы получим $(\alpha - 1)^2 < 0$, чего также не может быть. Поэтому наибольшее значение f достигается не на M_1 (то есть $C_1 = \varnothing$).

3) *Цилиндрическая часть границы.* Положим

$$\Phi(x,y,z) = x^2 + y^2 - x, \quad \text{где } x^2 + y^2 + z^2 < 1.$$

Заметим, что равенство $\Phi'(x,y,z) = (0,0,0)$ выполняется только в точках $\left(\frac{1}{2}, 0, z\right)$, которые не удовлетворяют условиям (87). Поэтому M_2 является двумерной поверхностью класса C^1. Выпишем для f и Φ условие (80). Функция Лагранжа (79) в данном случае равна

$$L(x,y,z,\lambda) = 3x^2 + 2y^2 + 2z^2 + 2yz - \lambda\,(x^2 + y^2 - x),$$

поэтому условие (80) запишется в виде

$$\begin{cases} 6x - 2\lambda x + \lambda = 0 \\ 4y + 2z - 2\lambda y = 0 \\ 4z + 2y = 0 \\ x^2 + y^2 = x, \end{cases}$$

где $x^2 + y^2 + z^2 < 1$. Эта система имеет два решения:

$$x = y = z = \lambda = 0 \quad \text{и} \quad x = 1,\ y = z = 0,\ \lambda = 6.$$

Второе решение нам не подходит, поскольку оно не удовлетворяет условиям (87). Таким образом, $C_2 = \{(0,0,0)\}$, $\max\limits_{C_2} f = 0$.

4) *Пересечение цилиндра и сферы.* Рассмотрим отображение

$$\Phi(x,y,z) = (x^2 + y^2 + z^2 - 1, x^2 + y^2 - x), \quad \text{где } (x,y,z) \in \mathbb{R}^3.$$

Тогда

$$\Phi'(x,y,z) = \begin{pmatrix} 2x & 2y & 2z \\ 2x - 1 & 2y & 0 \end{pmatrix}.$$

Несложно проверить, что $\operatorname{rank}\Phi' = 2$ во всех точках M_3, кроме $(1,0,0)$, а $\operatorname{rank}\Phi'(1,0,0) = 1$. Таким образом, $M_3 \setminus \{(1,0,0)\}$ является одномерной поверхностью в $\mathbb{R}^3$ класса C^1 (мы договорились называть ее *регулярной кривой класса* C^1). Значение в точке $(1,0,0)$ необходимо посчитать отдельно. Можно трактовать эту точку как "поверхность нулевой размерности".

Функция Лагранжа для f с уравнениями связи (88) равна

$$L(x,y,z,\lambda,\mu) = 3x^2+2y^2+2z^2+2yz-\lambda\,(x^2+y^2+z^2-1)-\mu\,(x^2+y^2-x).$$

Условие (80) для нее после приведения подобных слагаемых запишется в виде

$$\begin{cases} 2x(\lambda+\mu-3) = \mu \\ y(\lambda+\mu-2) - z = 0 \\ y - z(\lambda-2) = 0 \\ x^2+y^2+z^2 = 1 \\ x^2+y^2 = x. \end{cases} \tag{89}$$

Заметим, что второе и третье уравнения (89) представляют собой линейную однородную систему относительно y и z, причем ее нулевое решение нам не подходит, так как оно дает точку $(1,0,0)$. Поэтому

$$0 = \begin{vmatrix} \lambda+\mu-2 & -1 \\ 1 & -(\lambda-2) \end{vmatrix} = (2-\lambda)(\lambda+\mu-2)+1,$$

откуда $\lambda \neq 2$ и

$$\mu = -\frac{(\lambda-1)(\lambda-3)}{\lambda-2}, \quad \lambda+\mu-3 = -\frac{\lambda-3}{\lambda-2}.$$

Подставляя эти соотношения в первое уравнение системы (89), мы получим

$$2x = \lambda - 1 \quad \text{или} \quad \lambda = 3.$$

Второй случай не подходит, так как он приводит к точке $(1,0,0)$. Если $\lambda = 2x+1$, то из двух последних уравнений (89) вытекает, что

$$1-x = z^2 = \frac{y^2}{(\lambda-2)^2} = \frac{x(1-x)}{(\lambda-2)^2} = \frac{x(1-x)}{(2x-1)^2},$$

откуда в силу условия $x \neq 1$ получается решение

$$x = \frac{1}{4}, \quad y = \pm\frac{\sqrt{3}}{4}, \quad z = \mp\frac{\sqrt{3}}{2}, \quad \lambda = \frac{3}{2}, \quad \mu = -\frac{3}{2}.$$

Таким образом,

$$C_3 = \left\{ \left(\frac{1}{4}, \frac{\sqrt{3}}{4}, -\frac{\sqrt{3}}{2} \right), \left(\frac{1}{4}, -\frac{\sqrt{3}}{4}, \frac{\sqrt{3}}{2} \right) \right\}, \quad \max_{C_3} f = \frac{21}{16}.$$

5) *Вычисление наибольшего значения* f. До сих пор мы не исследовали точку $(1,0,0)$. Так как $f(1,0,0) = 3$, то

$$\max_E f = \max\bigl\{0, \tfrac{21}{16}, 3\bigr\} = 3.$$

Таким образом, наибольшее значение f реализуется в точке $(1,0,0)$ и равно 3.

Замечание. При решении задачи на наибольшее значение нам было неважно, есть ли в точках множеств C_2 и C_3 условный экстремум и какого он типа. Тем не менее, для лучшего понимания теоремы 6 исследуем эти точки с помощью достаточного условия относительного экстремума.

1) Рассмотрим точку $a = (0,0,0)$ на M_2, соответствующую $\lambda = 0$. Заметим, что

$$d_a\Phi(dx, dy, dz) = (2a_1 - 1)\, dx + 2a_2\, dy = -dx.$$

Поэтому в обозначениях теоремы 6

$$N = \bigl\{(x,y,z) \in \mathbb{R}^3 : x = 0\bigr\}.$$

Подставляя в функцию Лагранжа $\lambda = 0$, мы получим

$$L(x,y,z) = 3x^2 + 2y^2 + 2z^2 + 2yz.$$

Тогда в соответствии с формулой (41)

$$d_a^2 L(dx, dy, dz) = 6dx^2 + 4dy^2 + 4dz^2 + 4dydz.$$

Исследование $d_a^2 L$ на N сводится к изучению в $\mathbb{R}^2$ квадратичной формы

$$g(dy, dz) = 4\,(dy^2 + dz^2 + dydz).$$

Она порождается матрицей $\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}$, главные миноры которой равны 4 и 12. В силу критерия Сильвестра форма g положительно определена в $\mathbb{R}^2$, то есть $d_a^2 L$ положительно определен на N. Тогда по теореме 6 функция f имеет в точке $(0,0,0)$ относительный минимум, подчиненный уравнению связи $x^2 + y^2 = x$. Несложно проверить, что $d_a^2 L$ положительно определен не только на N, но даже в $\mathbb{R}^3$.

2) Рассмотрим точку $a = \left(\frac{1}{4}, \frac{\sqrt{3}}{4}, -\frac{\sqrt{3}}{2}\right)$ на M_3, соответствующую $\lambda = \frac{3}{2}$ и $\mu = -\frac{3}{2}$. Заметим, что

$$d_a\Phi(dx, dy, dz) = \bigl(2a_1\,dx + 2a_2\,dy + 2a_3\,dz, (2a_1 - 1)\,dx + 2a_2\,dy\bigr) = \\ = \bigl(\tfrac{1}{2}\,dx + \tfrac{\sqrt{3}}{2}\,dy - \sqrt{3}\,dz, -\tfrac{1}{2}\,dx + \tfrac{\sqrt{3}}{2}\,dy\bigr).$$

Равенство нулю $d_a\Phi$ эквивалентно соотношениям $dz = dy = \frac{dx}{\sqrt{3}}$, поэтому

$$N = \bigl\{(x, y, z) : z = y = \tfrac{x}{\sqrt{3}}\bigr\}.$$

Подставляя в функцию Лагранжа $\lambda = \frac{3}{2}$ и $\mu = -\frac{3}{2}$, мы получим

$$L(x, y, z) = 3x^2 + 2y^2 + \tfrac{1}{2}\,z^2 + 2yz - \tfrac{3}{2}\,x + \tfrac{3}{2}.$$

Тогда в соответствии с формулой (41)

$$d_a^2 L(dx, dy, dz) = 6dx^2 + 4dy^2 + dz^2 + 4dydz.$$

Чтобы исследовать $d_a^2 L$ на N, нужно рассмотреть квадратичную форму

$$d_a^2 L\bigl(dx, \tfrac{dx}{\sqrt{3}}, \tfrac{dx}{\sqrt{3}}\bigr) = 9dx^2.$$

Она, очевидно, положительно определена на $\mathbb{R}$. Тогда $d_a^2 L$ положительно определен на N, и по теореме 6 функция f имеет в точке $\left(\frac{1}{4}, \frac{\sqrt{3}}{4}, -\frac{\sqrt{3}}{2}\right)$ условный минимум, подчиненный уравнениям связи $x^2 + y^2 + z^2 = 1$ и $x^2 + y^2 = x$.

Точка $\left(\frac{1}{4}, -\frac{\sqrt{3}}{4}, \frac{\sqrt{3}}{2}\right)$ исследуется аналогично. Предлагаем читателю сделать это самостоятельно.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ