Оглавление

Предисловие

Порядок изложения материала в этой книге вполне традиционный, т.е. строго противоположный историческому. Исторически первыми были разработаны методы интегрирования. Это сделал ещё Архимед. Затем в работах Ньютона и Лейбница (1665-1675) были разработаны методы дифференцирования. Затем появились понятия предела и непрерывной функции. В основном это было сделано Коши (1821) и Вейерштрассом. Наконец, в работах Дедекинда (1872) появилось понятие вещественного числа как дедекиндова сечения, а в работах Кантора (1875) появились понятия множества и отображения.

Важную часть этой книги составляют задачи. Все они снабжены решениями, но сначала желательно попробовать их решить самостоятельно. Наиболее трудные задачи помечены звёздочкой.

Предварительный план этой книги я обсуждал с Сергеем Ландо.

Часть I

Функции одной переменной

Глава 1.

Вещественные числа. Счётные и несчётные множества

К определению вещественных чисел есть разные подходы. Одно из таких определений основано на дедекиндовых сечениях, которые ввёл Юлиус Вильгельм Рихард Дедекинд (1831-1916) в 1872 году. Мы приведём все необходимые определения и наметим некоторые доказательства, но в подробности доказательств вдаваться не будем.

Свойства рациональных чисел при этом предполагаются известными. Мы будем пользоваться следующими стандартными обозначениями:

$\mathbb{N}$ — множество натуральных (т.е. целых положительных) чисел;
$\mathbb{Z}$ — множество целых чисел;
$\mathbb{Q}$ — множество рациональных чисел;
$\varnothing$ — пустое множество;
$A \cup B$ — объединение множеств A и B;
$A \cap B$ — пересечение множеств A и B;
$A \subset B$ — множество A содержится в множестве B;
$a \in A$ — элемент a принадлежит множеству A;
$a \notin A$ — элемент a не принадлежит множеству A.

Нам понадобится также *разность множеств* $A \setminus B = \{x \in A \mid x \notin B\}$. Если $B \subset A$, то разность множеств $A \setminus B$ называют *дополнением* множества B (в множестве A). Дополнение множества B будем обозначать CB.

1.1. Дедекиндовы сечения

Дедекиндово сечение — это упорядоченная пара непустых непересекающихся подмножеств A и B в множестве рациональных чисел $\mathbb{Q}$, обладающих следующими свойствами:

- $A \cup B = \mathbb{Q}$, $A \cap B = \varnothing$;
- если $a \in A$ и $b \in B$, то $a < b$;
- в A нет наибольшего элемента.

Эту упорядоченную пару множеств будем обозначать (A, B), т.е. множество A первое, а множество B второе.

Вещественное число — это дедекиндово сечение. Рациональному число r соответствует дедекиндово сечение, в котором A — это множество всех рациональных чисел, меньших r (а B — это все остальные рациональные числа).

Пусть $x = (A, B)$ и $y = (C, D)$ — дедекиндовы сечения (вещественные числа). Мы говорим, что $x \leq y$, если $A \subset C$. Это позволяет определить, что такое положительные и отрицательные числа.

Вещественное число M называется *верхней гранью* множества S вещественных чисел, если $M \geq s$ для любого $s \in S$. Верхняя грань множества S, которая меньше всех остальных точных граней, называется *точной верхней гранью* множества S. Нижняя и точная нижняя грань определяются аналогично.

Множество вещественных чисел, имеющее верхнюю грань, называется *ограниченным сверху*.

Теорема 1.1. *Любое непустое множество S вещественных чисел, ограниченное сверху, имеет точную верхнюю грань.*

Доказательство. Точной верхней гранью множества S является сечение, для которого множество A — объединение всех первых множеств всех сечений, входящих в множество S. □

Сумма двух вещественных чисел $x = (A, B)$ и $y = (C, D)$ определяется как дедекиндово сечение (E, F), где E содержит все суммы $a + c$, $a \in A$, $c \in C$.

Для вещественного числа $x = (A, B)$ число $-x = (C, D)$ определяется следующим образом: чтобы получить C, нужно рассмотреть все рациональные числа $b \in B$, кроме наименьшего (если оно есть), и для каждого из них взять $c = -b$.

Если вещественные числа $x = (A, B)$ и $y = (C, D)$ неотрицательны, то их произведение xy определяется как сечение (E, F), где E состоит из всех отрицательных чисел и из всех произведений ac, где a и c — неотрицательные числа, входящие в A и C соответственно. Если, например, число x отрицательное, то xy можно определить как $(-x)y$.

Можно доказать, что если $x \neq 0$, то для любого y существует единственное вещественное число z, для которого $xz = y$. Это число z обозначается y/x.

Множество всех вещественных чисел обозначается $\mathbb{R}$.

Задача 1.1. а) Пусть α — иррациональное число. Докажите, что для любых чисел $a < b$ можно выбрать целые числа m и n, для которых $a < m\alpha - n < b$.

б) Докажите, что числа m и n можно выбрать натуральными.

Задача 1.2. Каждой точке x отрезка $[a, b]$ сопоставлено некоторое положительное число $\delta(x)$. Докажите, что можно выбрать конечное число точек $x_i \in [a, b]$ так, что отрезок $[a, b]$ содержится в объединении отрезков $I(x_i) = [x_i - \delta(x_i), x_i + \delta(x_i)]$.

1.2. Десятичная запись вещественного числа

Десятичная запись — это выражение вида $n, a_1a_2a_3\dots$, где n — целое число, a_i — одна из цифр 0, 1, …, 9. Если все a_i, начиная с некоторого, равны 0, то такую десятичную запись будем называть конечной. Каждому дедекиндову сечению (A, B) можно сопоставить десятичную запись следующим образом. Сначала находим целое число n, исходя из следующего условия: $n \in A$ и $n + 1 \notin A$. Затем находим цифра a_1: $n + \frac{a_1}{10} \in A$ и $n + \frac{a_1+1}{10} \notin A$. Затем находим цифра a_2: $n + \frac{a_1}{10} + \frac{a_2}{10^2} \in A$ и $n + \frac{a_1}{10} + \frac{a_2+1}{10^2} \notin A$, и т.д. Каждому сечению сопоставляется одна десятичная запись, но никакому сечению не сопоставляется конечная десятичная запись.

Наоборот, десятичной записи $n, a_1a_2a_3\dots$ можно сопоставить сечение, которое состоит из рациональных чисел r, удовлетворяющих хотя бы одному из неравенств: $r < n$, $r < n + \frac{a_1}{10}$, $r < n + \frac{a_1}{10} + \frac{a_2}{10^2}$, …При этом двум разным десятичным записям может соответствовать одно и то же дедекиндово сечение. Например, десятичным записям 1,000…и 0,999…соответствует одно и то же дедекиндово сечение.

Задача 1.3. Докажите, что десятичным записям, которые после некоторого места периодически повторяются, сопоставляются рациональные числа.

Задача 1.4. Докажите, что дедекиндову сечению, соответствующему рациональному числу, сопоставляется десятичная запись, которая периодически повторяется после некоторого места.

Задача 1.5. Докажите, что никакому дедекиндову сечению не сопоставляется конечная десятичная запись.

Задача 1.6. Докажите, что если двум разным десятичным записям соответствует одно и то же дедекиндово сечение, то одна из этих десятичных записей конечная, а другая оканчивается девятками.

1.3. Некоторые неравенства

В дальнейшем нам часто будут нужны различные неравенства. Приведём здесь некоторые основные неравенства в качестве задач. Их решения можно найти на с. 17–18

Задача 1.7. а) Докажите, что для любого $\alpha > 0$ и любого натурального $n > 1$ выполняется неравенство $(1+\alpha)^n > 1 + n\alpha$.

б) Докажите, что если $0 < \alpha \leq 1/n$ и n — натуральное число, то выполняется неравенство $(1+\alpha)^n < 1 + n\alpha + n^2\alpha^2$.

Задача 1.8. Докажите *неравенство Коши*

$$\left(\sum_{i=1}^{k} a_i b_i\right)^2 \leq \left(\sum_{i=1}^{k} a_i^2\right)\left(\sum_{i=1}^{k} b_i^2\right).$$

Задача 1.9. Пусть $a_1, \ldots, a_n$ — положительные числа. Докажите *неравенство между средним арифметическим и средним геометрическим*:

$$\frac{a_1 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \cdots a_n}.$$

(Если не все данные числа равны, то неравенство строгое.)

Задача 1.10. а) Пусть a — рациональное число, причём $a > 1$ или $a < 0$. Докажите, что для любого $x > 0,\ x \neq 1$, выполняется неравенство $x^a - ax + a - 1 > 0$.

б) Пусть a — рациональное число, причём $0 < a < 1$. Докажите, что для любого $x > 0,\ x \neq 1$, выполняется неравенство $x^a - ax + a - 1 < 0$.

Задача 1.11. Пусть A и B — положительные числа, а p и q — рациональные числа, связанные соотношением $\frac{1}{p} + \frac{1}{q} = 1$. Докажите, что если $p > 1$, то $A^{1/p}B^{1/q} \leq \frac{A}{p} + \frac{B}{q}$, а если $p < 1$, то $A^{1/p}B^{1/q} \geq \frac{A}{p} + \frac{B}{q}$. (Если $A \neq B$, то оба неравенства строгие.)

Задача 1.12. Пусть x_i и y_i — положительные числа, а p и q — рациональные числа, связанные соотношением $\frac{1}{p} + \frac{1}{q} = 1$. Докажите *неравенство Гёльдера*: если $p > 1$, то

$$x_1y_1 + \cdots + x_ny_n \leq (x_1^p + \cdots + x_n^p)^{1/p}(y_1^q + \cdots + y_n^q)^{1/q},$$

а если $p < 1$, то

$$x_1y_1 + \cdots + x_ny_n \geq (x_1^p + \cdots + x_n^p)^{1/p}(y_1^q + \cdots + y_n^q)^{1/q}.$$

Задача 1.13. Пусть x_i и y_i — положительные числа, а $p > 1$ — рациональное число. Докажите *неравенство Минковского*:

$$\left(\sum_{i=1}^{n}(x_i + y_i)^p\right)^{1/p} \leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p} + \left(\sum_{i=1}^{n} y_i^p\right)^{1/p}.$$

Если $p < 1$, то неравенство заменяется на противоположное.

1.4. Отображения множеств

Отображение f множества X в множество Y сопоставляет каждому элементу $x \in X$ некоторый элемент $f(x) \in Y$ — *образ* элемента x. Отображение $f\colon X \to Y$ *инъективно*, если образы любых двух различных элементов не совпадают. Отображение *сюръективно* (или является *отображением на* Y) если каждый элемент множества Y является образом некоторого элемента множества X. Отображение *биективно*, если оно одновременно инъективно и сюръективно. Биективное отображение называют также *взаимно однозначным соответствием множеств* X и Y.

Образ подмножества $A \subset X$ при отображение f — это множество всех $y \in Y$, которые являются образами элементов A. *Прообраз* подмножества $B \subset Y$ — это множество всех элементов X, отображающихся в элементы множества B.

Композиция отображений $f\colon X \to Y$ и $g\colon Y \to Z$ — это отображение $g \circ f\colon X \to Z$, заданное формулой $g \circ f(x) = g(f(x))$.

Задача 1.14. Докажите, что композиция отображений ассоциативна, т.е. $h \circ (g \circ f) = (h \circ g) \circ f$.

1.5. Счётные и несчётные множества

Бесконечное множество, между элементами которого и натуральными числами можно установить взаимно однозначное соответствие, называется *счётным*. Бесконечное множество, которое не является счётным, называется *несчётным*.

Теорема 1.2. *Бесконечное подмножество B счётного множества A счётно.*

Доказательство. Элементы множества A занумерованы, т.е. их можно записать так: a_1, a_2, a_3, … Запишем элементы множества B в порядке возрастания их номеров: a_{i_1}, a_{i_2}, a_{i_3}, … В результате элементы множества B окажутся занумерованными. □

Теорема 1.3. *Пусть A_1, A_2, A_3, … — счётные множества. Тогда их объединение счётно.*

Доказательство. Занумеруем элементы множества A_i: a_{i1}, a_{i2}, a_{i3}, … Рассмотрим теперь следующую последовательность: a_{11}; a_{12}, a_{21}; a_{13}, a_{22}, a_{31}; … (каждый раз выбираются элементы a_{ij} с одной и той же суммой $i+j$ и располагаются в порядке возрастания i). Если какие-то множества A_p и A_q пересекаются, то в полученной последовательности есть повторяющиеся элементы. Исключив повторяющиеся элементы, получим объединение множеств A_1, A_2, A_3, … Таким образом, это объединение множеств — подмножество счётного множества, а потому счётно. □

Если в качестве A_n взять множество несократимых дробей вида m/n, то в качестве следствия теоремы 1.3 получим, что множество рациональных чисел счётно.

Задача 1.15 даёт ещё один способ доказать счётность множества рациональных чисел.

Задача 1.15. Пусть p/q — несократимая дробь, p и q — натуральные числа. Положим $f\left(\frac{p}{q}\right) = \frac{p^2q^2}{q_1\cdots q_n}$, где $q_1, \ldots, q_n$ — различные простые делители числа q. Докажите, что f — взаимно однозначное отображение множества положительных рациональных чисел на множество всех натуральных чисел.

Теорема 1.4. *Множество A всех последовательностей, элементы которых — числа 0 и 1, несчётно.*

Доказательство. Рассмотрим в A произвольное счётное подмножество E. Пусть оно состоит из последовательностей s_1, s_2, s_3, … Построим последовательность s следующим образом: её n-й член не совпадает с n-м членом последовательности s_n, т.е. вместо 0 мы берём 1, а вместо 1 мы берём 0. Последовательность s отличается от любой из последовательностей s_i хотя бы одним членом, поэтому $s \notin E$.

Мы доказали, что никакое счётное подмножество множества A не может совпадать со всем множеством A, поэтому множество A несчётно. □

Способ, которым доказана теорема 1.4, называется *диагональный процесс Кантора*: если последовательности s_1, s_2, s_3, … записать в виде строк матрицы, то при построении последовательности s мы учитываем только диагональные элементы этой матрицы.

Диагональный процесс Кантора позволяет доказать, что множество вещественных чисел несчётно. Но при этом нужно действовать более аккуратно, чтобы учесть, что разные десятичные записи могут соответствовать одному и тому же числу. Например, 0,1000… =0,09999…

Теорема 1.5. *Множество вещественных чисел несчётно.*

Доказательство. Докажем, что даже подмножество вещественных чисел x, удовлетворяющих неравенствам $0 \le x < 1$, несчётно. Предположим, элементы этого множества можно записать в виде последовательности $x_1, x_2, x_3, \ldots$ Рассмотрим вещественное число ξ, десятичными знаками которого являются цифры от 1 до 8, причём n-й десятичный знак числа ξ отличен от n-го десятичного знака числа x_n.

Ясно, что $0 < \xi < 1$ и $\xi \neq x_n$ для всех n. Действительно, в десятичной записи числа ξ нет ни нулей, ни девяток. Поэтому $\xi = x_n$ тогда и только тогда, когда числа ξ и x_n имеют одинаковые десятичные записи. Но n-й десятичный знак числа ξ отличен от n-го десятичного знака числа x_n. □

Множества A и B называют *равномощными*, если между ними можно установить взаимно однозначное соответствие. Мощность множества A *не больше* мощности множества B, если можно установить взаимно однозначное соответствие между множеством A и подмножеством множества B. Мощность множества A *меньше* мощности множества B, если можно установить взаимно однозначное соответствие между множеством A и подмножеством множества B, но множества A и B не равномощны.

Мощность множества вещественных чисел называют мощностью *континуума*.

Конструкция, аналогичная диагональному процессу Кантора, позволяет доказать и следующее утверждение.

Теорема 1.6. *Никакое множество X не равномощно множеству всех своих подмножеств.*

Доказательство. Пусть $P(X)$ — множество всех подмножеств множества X. Предположим, что существует взаимно однозначное отображение $\varphi\colon X \to P(x)$. Рассмотрим множество Z, состоящее из тех элементов $x \in X$, для которых $x \notin \varphi(x)$. Покажем, что подмножество $Z \subset X$ не соответствует никакому элементу $x \in X$. Действительно, пусть $Z = \varphi(z)$ для некоторого $z \in X$. Если $z \in Z$, то $z \notin \varphi(z) = Z$ по построению множества Z, а если $z \notin Z$, то $z \notin \varphi(z)$, поскольку $Z = \varphi(z)$ по предположению, поэтому $z \in Z$ по построению множества Z. Полученное противоречие показывает, что не существует взаимно однозначного соответствия между X и $P(X)$. □

Теорема 1.7. *Если множество A бесконечно, а множество B конечно или счётно, то множество $A \cup B$ равномощно A.*

Доказательство. Можно считать, что множества A и B не пересекаются (иначе из B можно выбросить пересечение; после этого по-прежнему останется конечное или счётное множество).

Выделим в A счётное подмножество P, оставшееся множество обозначим Q. Нужно доказать, что множества $P \cup Q \cup B$ и $P \cup Q$ равномощны (множества B, P и Q попарно не пересекаются). Оба множества $P \cup B$ и P счётные, поэтому между ними можно установить взаимно однозначное соответствие. После этого можно установить взаимно однозначное соответствие между множествами $P \cup B \cup Q$ и $P \cup Q$ (каждый элемент множества Q соответствует самому себе). □

Следствие. *Множество $\mathbb{R}$, из которого удалено конечное или счётное подмножество, равномощно $\mathbb{R}$.*

Доказательство. Применим теорему в случае, когда B — удалённое конечное или счётное множество, $A \cup B$ — это множество $\mathbb{R}$. □

Задача 1.16. Докажите, что множество всех последовательностей вещественных чисел равномощно множеству $\mathbb{R}$.

Теорема 1.8 (Кантор, 1877). *Квадрат (с внутренностью) и отрезок равномощны.*

Доказательство. Квадрат равномощен множеству пар чисел, лежащих на отрезке $[0, 1]$. Сопоставим паре чисел $0,x_1x_2\dots$ и $0,y_1y_2\dots$ число $0,x_1y_1x_2y_2\dots$. Если исключить из рассмотрения числа, оканчивающиеся бесконечной последовательностью девяток, то это соответствие будет взаимно однозначным. А исключение счётного множества из более чем счётного множества не меняет мощность множества. □

Историческое замечание. Множество алгебраических чисел счётно, поэтому существуют вещественные числа, не являющиеся алгебраическими (их называют *трансцендентными*). Такое доказательство существования трансцендентных чисел предложил Георг Кантор (1845-1918) в 1874 году. Но первым доказал существования трансцендентных чисел Жозеф Лиувилль (1809-1882) в 1844 году; он указал конкретные примеры трансцендентных чисел.

Задача 1.17. Докажите, что на плоскости нельзя расположить более чем счётное множество попарно непересекающихся кругов.

Задача 1.18. Докажите, что на плоскости нельзя расположить более чем счётное множество попарно непересекающихся восьмёрок.

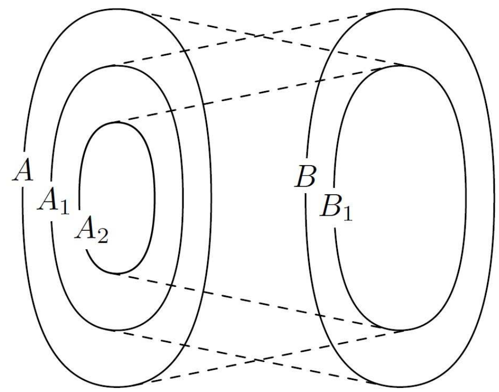

Рис. 1.1.

Задача 1.19. *а) Пусть α — корень неприводимого многочлена $f(x) = a_n x^n + a_{n-1}x^{n-1} + \dots + a_0$, где $n \geq 2$. Докажите, что существует такое число $c > 0$ (зависящее только от α), что $\left|\alpha - \frac{p}{q}\right| > \frac{c}{q^n}$ для любого целого p и натурального q (*теорема Лиувилля*).

б) Докажите, что число $\alpha = \sum\limits_{k=0}^{\infty} 2^{-k!}$ трансцендентное (*Лиувилль*).

1.6. Теорема Кантора–Бернштейна

Теорема 1.9 (Кантор–Бернштейн). *Если множество A равномощно некоторому подмножеству множества B, а B равномощно некоторому подмножеству множества A, то множества A и B равномощны.*

Доказательство. Пусть множество A равномощно множеству $B_1 \subset B$, а множество B равномощно множеству $A_1 \subset A$. При взаимно однозначном соответствии между B и A_1 множество B_1 соответствует некоторому множеству $A_2 \subset A$ (рис. 1.1). Из условия теоремы следует, что все три множества A, B_1 и A_2 равномощны, и требуется доказать, что они равномощны множеству B, которое, в свою очередь, равномощно множеству A_1.

Таким образом, мы получаем следующую эквивалентную формулировку теоремы: *если $A_2 \subset A_1 \subset A_0$ и A_2 равномощно A_0, то все три множества равномощны* (для единообразия множество A мы обозначили A_0).

Пусть $f\colon A_0 \to A_2$ — взаимно однозначное соответствие. Введём множества $A_{i+2} = f(A_i)$, полагая последовательно $i = 0, 1, 2, \dots$. Из включения $A_2 \subset A_1 \subset A_0$ следует включение $A_{i+2} \subset A_{i+1} \subset A_i$, поэтому мы получаем последовательность включений $\dots A_4 \subset A_3 \subset A_2 \subset A_1 \subset A_0$. Множество A разбивается на попарно непересекающиеся множества $C = \bigcap_i A_i$ и $C_i = A_i \setminus A_{i+1}$, где $i = 0, 1, 2, \dots$ (рис. 1.2). При этом отображение f осуществляет взаимно однозначное соответствие между C_i и C_{i+2}, поэтому все множества с чётными номерами равномощны и все множества с нечётными номерами также равномощны.

Взаимно однозначное соответствие $g\colon A_0 \to A_1$ строится следующим образом. Пусть $x \in A_0$. Тогда $g(x) = f(x)$ при $x \in C_{2k}$ и $g(x) = x$ при $x \in C_{2k+1}$ и при $x \in C$ (рис. 1.3).

□

Историческое замечание. Георг Кантор (1845-1918) сформулировал эту теорему в 1887 году. Первые доказательства получили в 1897 году Феликс Бернштейн (1878-1956) и Эрнст Шрёдер (1841-1902). Поэтому иногда её называют теоремой Шрёдера–Бернштейна.

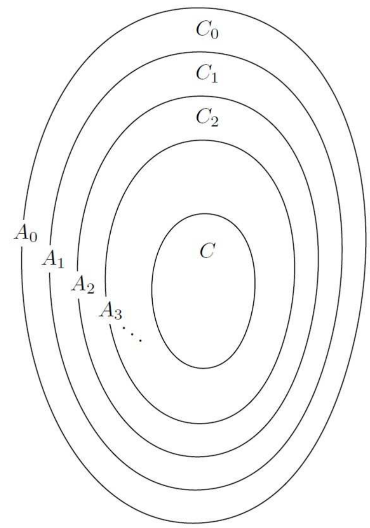

Рис. 1.2.

$$
\begin{tikzcd}
A_0 & = & C_0 \arrow[drrrr] & + & C_1 \arrow[d] & + & C_2 \arrow[drrrr] & + & C_3 \arrow[d] & + & C_4 & + & \cdots & + & C \arrow[d] \\
A_1 & = & & & C_1 & + & C_2 & + & C_3 & + & C_4 & + & \cdots & + & C
\end{tikzcd}
$$

Рис. 1.3.

1.7. Решения задач

1.1. а) Пусть $\Delta = b - a$. Для каждого целого числа m_1 можно выбрать целое число n_1 так, что $0 \leq m_1\alpha - n_1 \leq 1$. Разделим отрезок $[0, 1]$ на равные отрезки, длина каждого из которых меньше Δ. Пусть количество этих отрезков равно k. Тогда среди чисел $m_1\alpha - n_1, \ldots, m_{k+1}\alpha - n_{k+1}$ есть два числа, попадающих в один и тот же отрезок. Вычтем из большего числа меньшее: $m_p\alpha - n_p - (m_q\alpha - n_q) = t$. Ясно, что $0 \leq t < \Delta$. Более того, $t \neq 0$, поскольку иначе $\alpha = \frac{n_p - n_q}{m_p - m_q}$ — рациональное число.

Рассмотрим числа вида Nt, где N — целое число. Каждое из этих чисел имеет вид $m\alpha - n$. А из того, что $0 < t < \Delta$, следует, что хотя бы одно из этих чисел расположено строго между a и b.

б) Нужно доказать, что число t можно выбрать как вида $M\alpha - N$, так и вида $-M\alpha + N$, где M и N — натуральные числа. Пусть $t = M\alpha - N$. Между 0 и t есть бесконечно много чисел вида $m\alpha - n$, m и n — целые. Предположим, что у всех этих чисел $m > 0$. Тогда для одного из них $m > M$, а в таком случае число $t = M\alpha - N - (m\alpha - n)$ искомое. Если же $t = -M\alpha + N$, то рассуждения аналогичны.

1.2. Пусть E — множество точек $x \in [a, b]$, для которых на отрезке $[a, x]$ можно выбрать конечное множество точек так, чтобы отрезок $[a, x]$ содержался в объединении рассматриваемых отрезков. Множество E содержит все точки $x \in [a, a + \delta(a)]$, удовлетворяющие неравенству $x \leq b$. Действительно, в этом случае достаточно выбрать одну точку — точку a. Ясно также, что множество E ограниченное (оно содержится в отрезке $[a, b]$). Пусть u — точная верхняя грань множества E. Тогда $a < u \leq b$; нужно доказать, что $u = b$.

Докажем сначала, что $u \in E$. Из неравенства $u - \delta(u) < u = \sup E$ следует, что в E можно выбрать точку v, удовлетворяющую неравенствам $u - \delta(u) < v < u$. Выберем конечное множество точек так, чтобы соответствующие им отрезки покрывали отрезок $[a, v]$, и добавим к этим точкам точку u. Тогда соответствующие им отрезки покроют отрезок $[a, u]$.

Предположим, что $u < b$. Тогда на отрезке $[a, b]$ можно выбрать точку w так, что $u < w < u + \delta(u)$. Воспользовавшись тем, что $u \in E$, выберем конечное множество точек так, чтобы соответствующие им отрезки покрывали отрезок $[a, u]$, и добавим к ним точку w. Тогда соответствующие им отрезки покроют отрезок $[a, w]$, поэтому $w \in E$. Это противоречит тому, что $\sup E = u < w$.

1.3. Конечная десятичная дробь соответствует рациональному числу, поэтому достаточно рассмотреть периодически повторяющуюся часть. После умножения на подходящую степень числа 10, можно считать, что период начинается с первого знака после запятой, т.е. десятичная дробь соответствует числу $r(1 + 10^{-k} + 10^{-2k} + 10^{-3k} + \ldots) = r\frac{1}{1 - 10^{-k}}$. Это число рациональное.

1.4. Достаточно рассмотреть рациональные числа вида p/q, где $1 \leq p < q$. В таком случае цифры десятичной записи $p/q = 0, a_1a_2a_3\ldots$ находятся следующим образом: $a_k = \left[\dfrac{10^k p}{q}\right] - 10\left[\dfrac{10^{k-1}p}{q}\right]$. Пусть $10^{k-1}p = Mq + r_k$, где r_k — остаток от деления $10^{k-1}p$ на q. Тогда $a_k = 10M + \left[\dfrac{10r_k}{q}\right] - 10M = \left[\dfrac{10r_k}{q}\right]$. Среди чисел $r_1, \ldots, r_{q+1}$ есть два одинаковых: $r_i = r_{i+n}$. Но r_{k+1} — остаток от деления $10r_k$ на q. Поэтому $r_{i+1} = r_{i+n+1}$ и т.д. Значит, последовательность цифр a_k тоже периодически повторяется (после номера i).

1.5. Предположим, что дедекиндову сечению (A, B) сопоставлена конечная десятичная запись $n, a_1a_2a_3\ldots a_r$. Эта конечная десятичная запись соответствует рациональному числу p, причём по определению $p \in A$, но $p + 10^{-r} \notin A$, $p + 10^{-r-1} \notin A$, …Следовательно, A состоит из всех рациональных чисел, не превосходящих p. Но этого не может быть, потому что в таком A существует наибольший элемент — число p.

1.6. Предположим, что десятичным записям $a_0, a_1a_2a_3\ldots$ и $b_0, b_1b_2b_3\ldots$ сопоставлено одно и то же вещественное число x. Пусть первые знаки этих записей совпадают, а различаются только a_r и b_r и, возможно, последующие знаки (случай $r = 0$ не исключается). Можно считать, что $a_r > b_r$. Тогда $x \geq a_0, a_1a_2a_3\ldots a_r$ и $x \leq b_0, b_1b_2b_3\ldots b_r + 10^{-r} \leq a_0, a_1a_2a_3\ldots a_r$. Следовательно, все встречающиеся здесь нестрогие неравенства должны обращаться в равенства, а это возможно лишь при следующих условиях: $a_r = b_r + 1$, $a_{r+1} = a_{r+2} = \cdots = 0$, $b_{r+1} = b_{r+2} = \cdots = 9$.

1.7. а) Покажем, что если $\alpha > 0$ и $(1+\alpha)^n > 1+n\alpha$, то $(1+\alpha)^{n+1} > 1+(n+1)\alpha$. Действительно, $(1+\alpha)^{n+1} = (1+\alpha)(1+\alpha)^n > (1+\alpha)(1+n\alpha) = 1+(n+1)\alpha+n\alpha^2 > 1+(n+1)\alpha$.

б) Для $n = 1$ требуемое неравенство выполняется. Предположим, что если $0 < \alpha \leq 1/n$, то выполняется неравенство $(1+\alpha)^n < 1+n\alpha+n^2\alpha^2$. Пусть число β таково, что $0 < \beta \leq 1/(n+1)$. Тогда $\beta \leq 1/n$, поэтому $(1+\beta)^{n+1} < (1+n\beta+n^2\beta^2)(1+\beta) = 1+(n+1)\beta+(n^2+n+n^2\beta)\beta^2$. Остаётся проверить, что $n^2+n+n^2\beta \leq (n+1)^2$, т.е. $n^2\beta \leq n+1$. Это неравенство следует из того, что $\beta \leq 1/(n+1)$ и $n^2 < (n+1)^2$.

1.8. Пусть $A = \sqrt{\sum_{i=1}^k a_i^2}$ и $B = \sqrt{\sum_{i=1}^k b_i^2}$. Воспользовавшись неравенством $xy \leq (x^2+y^2)/2$, получим

$$\sum_{i=1}^k \frac{a_i}{A}\frac{b_i}{B} \leq \sum_{i=1}^k \frac{1}{2}\left(\frac{a_i^2}{A^2}+\frac{b_i^2}{B^2}\right) = \frac{1}{2A^2}\sum_{i=1}^k a_i^2 + \frac{1}{2B^2}\sum_{i=1}^k b_i^2 = 1.$$

Значит, $\left(\sum_{i=1}^k a_ib_i\right)^2 \leq A^2B^2 = \left(\sum_{i=1}^k a_i^2\right)\left(\sum_{i=1}^k b_i^2\right)$.

1.9. Первое решение. Применим индукцию по n. При $n = 1$ требуемое утверждение верно. Предположим, что требуемое неравенство доказано для любых n положительных чисел. Рассмотрим положительные числа $b_1 = \sqrt[n+1]{a_1}, \ldots, b_{n+1} = \sqrt[n+1]{a_{n+1}}$. Ясно, что $(b_i^n - b_j^n)(b_i - b_j) \geq 0$. Просуммируем такие неравенства для всех пар i, j, для которых $i > j$. В результате получим

$$n\sum_{i=1}^{n+1} b_i^{n+1} \geq \sum_{i=1}^{n+1} b_i \sum_{j\neq i} b_j^n.$$

(Если не все данные числа равны, то неравенство строгое.) Воспользовавшись неравенством между средним арифметическим и средним геометрическим для n чисел, получим

$$\sum_{i=1}^{n+1} b_i \sum_{j\neq i} b_j^n \geq n\sum_{i=1}^{n+1} b_i \prod_{j\neq i} b_j = n(n+1)\prod_{j=1}^{n+1} b_j.$$

Следовательно,

$$\sum_{i=1}^{n+1} b_i^{n+1} \geq (n+1)\prod_{j=1}^{n+1} b_j,$$

что и требовалось.

Второе решение. Расположим данные числа в порядке возрастания: $a_1 \leq a_2 \leq \cdots \leq a_n$. Если $a_1 = a_n$, то $a_1 = a_2 = \cdots = a_n$. В таком случае $A = G$. Поэтому будем предполагать, что $a_1 < a_n$. Тогда $a_1 < A < a_n$, а значит,

$$A(a_1+a_n-A) - a_1a_n = (a_1-A)(A-a_n) > 0,$$

т.е. $A(a_1+a_n-A) > a_1a_n$. В частности, $a_1+a_n-A > 0$.

Применим индукцию по n. Предположим, что неравенство между средним арифметическим и средним геометрическим верно для любых $n-1$ положительных чисел. Применим его к числам a_2, $a_3, \ldots, a_{n-1}$, a_1+a_n-A. В результате получим

$$\left(\frac{a_2+a_3+\cdots+a_{n-1}+a_1+a_n-A}{n-1}\right)^{n-1} \geq a_2a_3\cdots a_{n-1}(a_1+a_n-A).$$

Здесь

$$\frac{a_2+a_3+\cdots+a_{n-1}+a_1+a_n-A}{n-1} = \frac{nA-A}{n-1} = A.$$

Умножим теперь обе части полученного неравенства на A:

$$A^n \geq a_2a_3\cdots a_{n-1}A(a_1+a_n-A) > a_1a_2\cdots a_n;$$

мы воспользовались доказанным выше неравенством $A(a_1+a_n-A) > a_1a_n$.

1.10. Пусть n — натуральное число. Тождество

$$\frac{y^{n+1}-1}{n+1}-\frac{y^n-1}{n}=\frac{y-1}{n(n+1)}(ny^n-y^{n-1}-\dots-y-1)$$

показывает, что $\frac{y^{n+1}-1}{n+1}-\frac{y^n-1}{n}>0$ при $y>0,\ \ y\neq 1$. Действительно, если $y>1$, то $y-1>0$ и $ny^n>y^{n-1}+\dots+y+1$, а если $0<y<1$, то $y-1<0$ и $ny^n<y^{n-1}+\dots+y+1$.

Пусть m — натуральное число, причём $m>n$. Тогда $(y^m-1)/m>(y^n-1)/n$ при $y>0,\ \ y\neq 1$. Положим $y=x^{1/n}$, где $x>0,\ \ x\neq 1$. Тогда $(x^{m/n}-1)/m>(x-1)/n$, т.е. $x^{m/n}-1-\frac{m}{n}(x-1)>0$. Мы получили требуемое неравенство для $a=m/n>1$.

Чтобы получить остальные неравенства, поступим следующим образом. Для $a>1$ положим $x^a=x^{1-b}=y^{b-1}$, где $b=1-a<0$ и $y=x^{-1}$. Тогда неравенство $x^a-ax+a-1>0$ запишется в виде $y^{b-1}-(1-b)y^{-1}+1-b-1>0$, т.е. $y^{-1}(y^b-by+b-1)>0,\ \ b<0$.

Положим теперь $b=1/a$ и $y=x^a$ (здесь снова $a>1$ и $x>0$). Тогда неравенство $x^a-ax+a-1>0$ запишется в виде $y-\frac{1}{b}y^b+\frac{1}{b}-1>0$, т.е. $y^b-by+b-1<0,\ \ 0<b<1$.

1.11. Положим $x=A/B,\ \ a=1/p$ и $1-a=1/q$. Тогда $B(x^a-ax+a-1)=A^{1/p}B^{1/q}-\frac{A}{p}-\frac{B}{q}$. Остаётся воспользоваться результатом задачи 1.10.

1.12. Положим $X=x_1^p+\dots+x_n^p,\ \ Y=y_1^q+\dots+y_n^q,\ \ A=x_i^p/X$ и $B=y_i^q/Y$. Если $p>1$, то согласно задаче 1.11 $\dfrac{x_i}{X^{1/p}}\dfrac{y_i}{Y^{1/q}}\le\dfrac{x_i^p}{pX}+\dfrac{y_i^q}{qY}$. Сложив такие неравенства для $i=1,\ \dots,\ n$, получим

$$\frac{x_1y_1+\dots+x_ny_n}{X^{1/p}Y^{1/q}}\le\frac{X}{pX}+\frac{Y}{qY}=\frac{1}{p}+\frac{1}{q}=1.$$

В случае, когда $p<1$, доказательство аналогично.

1.13. Пусть $p>1$. Ясно, что

$$\sum_{i=1}^n(x_i+y_i)^p=\sum_{i=1}^n x_i(x_i+y_i)^{p-1}+\sum_{i=1}^n y_i(x_i+y_i)^{p-1}.$$

Согласно неравенству Гёльдера (задача 1.12)

$$\sum_{i=1}^n x_i(x_i+y_i)^{p-1}\le\Big(\sum_{i=1}^n x_i^p\Big)^{1/p}\Big(\sum_{i=1}^n(x_i+y_i)^p\Big)^{1/q},$$

где $q=p/(p-1)$, т.е. $\frac{1}{p}+\frac{1}{q}=1$. Запишем такое же неравенство для второй суммы и сложим оба этих неравенства. В результате получим

$$\sum_{i=1}^n(x_i+y_i)^p\le\bigg(\Big(\sum_{i=1}^n x_i^p\Big)^{1/p}+\Big(\sum_{i=1}^n y_i^q\Big)^{1/q}\bigg)\Big(\sum_{i=1}^n(x_i+y_i)^p\Big)^{1/q}.$$

Это неравенство эквивалентно требуемому.

В случае, когда $p<1$, доказательство аналогично.

1.14. Ясно, что $h\circ(g\circ f)(x)=h((g\circ f)(x))=h(g(f(x)))=(h\circ g)(f(x))=((h\circ g)\circ f)(x)$.

1.15. Пусть $p=p_1^{a_1}\cdots p_m^{a_m}$, $q=q_1^{b_1}\cdots q_n^{b_n}$. Тогда $f(p/q)=p_1^{2a_1}\cdots p_m^{2a_m}q_1^{2b_1-1}\cdots q_n^{2b_n-1}$. Ясно, что каждое натуральное число представимо в таком виде ровно одним способом.

1.16. Рассмотрим десятичную запись n-го числа в последовательности: $\sum a_{nm}10^m$. Множество пар (n,m) счётно, поэтому их можно занумеровать и, тем самым, сопоставить последовательности чисел число. Такое сопоставление не является взаимно однозначным отображением, потому что может встретиться десятичная дробь, оканчивающаяся одними девятками. Но таких дробей счётное множество, а исключение счётного множества не влияет на мощность рассматриваемых множеств.

1.17. Сопоставим кругу, расположенному на плоскости, какую-нибудь точку с рациональными координатами, лежащую внутри этого круга. Для непересекающихся кругов такие точки не могут совпадать.

1.18. Сопоставим восьмёрке, расположенной на плоскости, пару точек с рациональными координатами, выбрав по одной точке внутри каждой из петель. Для непересекающихся восьмёрок такие пары не могут совпадать.

1.19. а) Если $\left|\alpha - \frac{p}{q}\right| \geq 1$, то требуемое неравенство выполняется при $c = 1$. Поэтому будем считать, что $\left|\alpha - \frac{p}{q}\right| < 1$, а значит, $\left|\frac{p}{q}\right| < |\alpha| + 1$. Запишем многочлен $f(x)$ в виде $f(x) = a_n \prod_{i=1}^{n}(x - \alpha_i)$, где $\alpha_1 = \alpha$. Тогда

$$\begin{aligned}\left|f\left(\frac{p}{q}\right)\right| &= |a_n| \left|\alpha - \frac{p}{q}\right| \prod_{i=2}^{n} \left|\frac{p}{q} - \alpha_i\right| < \\ &< |a_n| \left|\alpha - \frac{p}{q}\right| \prod_{i=2}^{n}(|\alpha| + 1 + |\alpha_i|) = c_1 \left|\alpha - \frac{p}{q}\right|,\end{aligned}$$

где c_1 —положительное число, зависящее только от $|a_n|$ и α.

Будем считать, что $a_0, \ldots, a_n$ — целые числа, взаимно простые в совокупности. Тогда число $|a_n|$ полностью определяется числом α. Кроме того, число

$$q^n f(p/q) = a_n p^n + a_{n-1} p^{n-1} q + \cdots + a_0 q^n$$

целое, поэтому $|q^n f(p/q)| \geq 1$. Следовательно,

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{c_1} \left|f\left(\frac{p}{q}\right)\right| \geq \frac{1}{c_1 q^n} = \frac{c}{q^n},$$

где $c = c_1^{-1}$.

б) Для каждого натурального n рассмотрим число $\alpha = \sum_{k=0}^{n} 2^{-k!} = p/q$, где p — целое число и $q = 2^{n!}$. При этом

$$\left|\alpha - \frac{p}{q}\right| = \frac{1}{2^{(n+1)!}} \left(1 + \frac{1}{2^{n+2}} + \frac{1}{2^{(n+2)(n+3)}} + \cdots\right) < \frac{2}{2^{(n+1)!}} = \frac{2}{q^{n+1}}.$$

Предположим, что α — алгебраическое число степени N. Тогда согласно задаче а)

$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c}{q^N},$$

а значит, $2q^{-n-1} > cq^{-N}$, т.е. $c < 2q^{N-n-1} = 2 \cdot 2^{n!(N-n-1)}$. Но $\lim\limits_{n\to\infty} 2^{n!(N-n-1)} = 0$. Приходим к противоречию.

Глава 2.

Предел последовательности

Для бесконечной последовательности чисел $a_1,\ a_2,\ a_3,\ \ldots$ обычно используют обозначение $\{a_n\}$. Число a называют *пределом* последовательности $\{a_n\}$, если для любого $\varepsilon > 0$ можно выбрать номер N так, что $|a_n - a| < \varepsilon$ при $n > N$. Предел последовательности $\{a_n\}$ обозначают $\lim_{n\to\infty} a_n$. Предел последовательности не всегда существует. Например, у последовательности $a_n = (-1)^n$ предела нет.

Вместо обозначения $\lim_{n\to\infty} a_n = a$ иногда используют обозначение $a_n \to a$ при $n \to \infty$ или даже просто $a_n \to a$.

2.1. Свойства пределов

Теорема 2.1. *Если предел последовательности существует, то он единствен.*

Доказательство. Предположим, что a и b — пределы последовательности $\{a_n\}$, причём $a \neq b$. Пусть $\varepsilon = \frac{1}{3}|a-b|$. Согласно определению предела можно выбрать номера N_1 и N_2 так, что $|a_n - a| < \varepsilon$ при $n > N_1$ и $|a_n - b| < \varepsilon$ при $n > N_2$. Пусть N — наибольшее из чисел N_1 и N_2. Тогда $|a-b| \leq |a - a_N| + |a_N - b| < 2\varepsilon = \frac{2}{3}|a-b|$. Приходим к противоречию. □

Теорема 2.2. *Пусть $\{a_n\}$ и $\{b_n\}$ — две последовательности, причём $\lim_{n\to\infty} a_n = a$ и $\lim_{n\to\infty} b_n = b$. Тогда:*

а) $\lim_{n\to\infty}(a_n + c) = a + c$ и $\lim_{n\to\infty}(ca_n) = ca$ для любого числа c;

б) $\lim_{n\to\infty}(a_n + b_n) = a + b$;

в) $\lim_{n\to\infty}(a_n b_n) = ab$;

г) если $a \neq 0$ и $a_n \neq 0$ для всех n, то $\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{a}$.

Доказательство. а) Утверждение о том, что $\lim_{n\to\infty}(a_n + c) = a + c$ очевидно: в качестве N для последовательности $\{a_n + c\}$ можно выбрать то же самое число, что и для последовательности $\{a_n\}$.

При $c = 0$ утверждение о том, что $\lim_{n\to\infty}(ca_n) = 0$ очевидно. Если же $c \neq 0$, то в качестве числа N, соответствующего ε для последовательности $\{ca_n\}$, можно взять N, соответствующее $\varepsilon/|c|$ для последовательности $\{a_n\}$. А именно, если $|a_n - a| < \varepsilon/|c|$, то $|ca_n - ca| < \varepsilon$.

б) Для данного $\varepsilon > 0$ выберем N_1 и N_2 так, что $|a_n - a| < \varepsilon/2$ при $n > N_1$ и $|b_n - b| < \varepsilon/2$ при $n > N_2$. Пусть N — наибольшее из чисел N_1 и N_2. Тогда $|(a_n + b_n) - (a+b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ при $n > N$.

в) Воспользуемся тождеством

$$a_n b_n - ab = (a_n - a)(b_n - b) + a(b_n - b) + b(a_n - a). \tag{2.1}$$

Для данного $\varepsilon > 0$ выберем N_1 и N_2 так, что $|a_n - a| < \sqrt{\varepsilon}$ при $n > N_1$ и $|b_n - b| < \sqrt{\varepsilon}$ при $n > N_2$. Тогда $|(a_n - a)(b_n - b)| < \varepsilon$ при $n > \max(N_1, N_2)$, поэтому $\lim_{n\to\infty}(a_n - a)(b_n - b) = 0$. Из а) и б) следует, что $\lim_{n\to\infty} a(b_n - b) = 0$ и $\lim_{n\to\infty} b(a_n - a) = 0$. Поэтому тождество (2.1) показывает, что $\lim_{n\to\infty}(a_n b_n - ab) = 0$.

г) Выберем сначала N_1 так, что $|a_n - a| < \frac{1}{2}|a|$ при $n > N_1$. Отметим, что при этом $|a_n| > \frac{1}{2}|a|$. Выберем затем для данного $\varepsilon > 0$ номер N_2 так, что $|a_n - a| < \frac{1}{2}|a|^2\varepsilon$ при $n > N_2$. Тогда если

$n > \max(N_1, N_2)$, то

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \left|\frac{a_n - a}{a_n a}\right| < \frac{1}{2}|a|^2\varepsilon \cdot \frac{2}{|a|^2} = \varepsilon.$$

□

Теорема 2.3. *Если $a_n < b_n < c_n$ для всех n и $\lim_{n\to\infty} a_n = a = \lim_{n\to\infty} c_n$, то $\lim_{n\to\infty} b_n = a$.*

Доказательство. Для данного $\varepsilon > 0$ выберем числа N_1 и N_2 так, что $|a - a_n| < \varepsilon$ при $n > N_1$ и $|a - c_n| < \varepsilon$ при $n > N_2$. Тогда если $n > \max(N_1, N_2)$, то $a_n > a - \varepsilon$ и $c_n < a + \varepsilon$. Поэтому $a - \varepsilon < a_n \leq b_n \leq c_n < a + \varepsilon$, а значит, $|b_n - a| < \varepsilon$. □

Теорема 2.4. *Если $\lim_{n\to\infty} a_n = a$ и $\lim_{n\to\infty} b_n = b$, причём $a_n \leq b_n$ для всех n, то $a \leq b$.*

Доказательство. Выберем N так, что если $n > N$, то $|a - a_n| < \varepsilon$ и $|b - b_n| < \varepsilon$. Тогда $a - \varepsilon < a_n \leq b_n < b + \varepsilon$. Таким образом, для любого $\varepsilon > 0$ выполняется неравенство $b > a - 2\varepsilon$. Поэтому $b \geq a$. □

Задача 2.1. Дано положительное число a и последовательность положительных чисел x_n. Докажите, что если $\lim_{n\to\infty} \frac{x_n - a}{x_n + a} = 0$, то $\lim_{n\to\infty} x_n = a$.

Задача 2.2. Докажите, что если $\lim_{n\to\infty} x_n = a$, то $\lim_{n\to\infty} \frac{x_1 + \cdots + x_n}{n} = a$ (Коши, 1821).

Задача 2.3. Докажите, что если $\lim_{n\to\infty}(x_{n+1} - x_n) = a$, то $\lim_{n\to\infty} \frac{x_n}{n} = a$.

Задача 2.4. *Пусть $P(x)$ — многочлен с целыми коэффициентами. Докажите, что если $\lim_{n\to\infty}\{P(n)\alpha\} = 0$, где $\{\ \}$ обозначает дробную часть, то число α рационально.

2.2. Возрастающие последовательности. Теорема Вейерштрасса

Последовательность $\{a_n\}$ называют *ограниченной*, если можно выбрать числа c_1 и c_2 так, что $c_1 \leq a_n \leq c_2$ для всех n.

Точку c называют *предельной точкой* последовательности $\{a_n\}$, если для любого $\varepsilon > 0$ неравенство $|c - a_n| < \varepsilon$ выполняется для бесконечного множества членов последовательности.

Теорема 2.5 (Больцано–Вейерштрасс). *Любая ограниченная последовательность имеет по крайней мере одну предельную точку.*

Доказательство. Можно считать, что числа c_1 и c_2 из определения ограниченной последовательности целые. Отрезок $[c_1, c_2]$ разбивается на конечное число отрезков длины 1. Хотя бы в одном из этих отрезков содержится бесконечно много членов рассматриваемой последовательности. Выберем такой отрезок и разделим его на 10 отрезков длины 1/10. Среди этих отрезков выберем тот, который содержит бесконечно много членов последовательности. Этот отрезок снова разделим на 10 отрезков равной длины и т.д. Такая последовательность операций определяет некоторое число $c = m_0 + m_1 \cdot 10^{-1} + m_2 \cdot 10^{-2} + \ldots$ Действительно, на первом шаге мы определяем m_0, на втором m_1 и т.д. Число c является предельной точкой данной последовательности. □

Историческое замечание. Теорему 2.5 доказал Бернард Больцано (1781-1848) в 1817 году, но значение этой теоремы впервые в полной мере было выявлено в 1874 году в лекциях Карла Вейерштрасса (1815-1897).

Пусть $\{a_n\}$ — некоторая последовательность. Предположим, что можно выбрать строго возрастающую последовательность натуральных чисел n_1, n_2, n_3, … так, что последовательность $\{b_k\}$, где $b_k = a_{n_k}$, сходится. Тогда говорят, что последовательность $\{a_n\}$ содержит сходящуюся *подпоследовательность* $\{a_{n_k}\}$.

Теорема 2.6. *Любая ограниченная последовательность $\{a_n\}$ содержит сходящуюся подпоследовательность $\{a_{n_k}\}$.*

Доказательство. Пусть c — предельная точка последовательности $\{a_n\}$. Тогда неравенство $|a_n-c|<1/k$ выполняется для бесконечного множества членов a_n. Поэтому можно выбрать a_{n_k} так, что $|a_{n_k}-c|<1/k$. Более того, это можно сделать так, что $n_1<n_2<n_3<\dots$ Тогда $c=\lim_{k\to\infty}a_{n_k}$. □

Последовательность $\{a_n\}$ называют *возрастающей* (*строго возрастающей*), если $a_{n+1}\geq a_n$ ($a_{n+1}>a_n$) для всех n.

Последовательность $\{a_n\}$ называют *ограниченной сверху*, если можно выбрать число c так, что $a_n\leq c$ для всех n.

Теорема 2.7 (Вейерштрасс). *Любая возрастающая ограниченная сверху последовательность $\{a_n\}$ имеет предел.*

Доказательство. Возрастающая ограниченная сверху последовательность ограничена, поскольку $a_n\geq a_1$ для всех n. Поэтому согласно теореме Больцано–Вейерштрасса (теорема 2.5) последовательность $\{a_n\}$ имеет предельную точку c. Прежде всего докажем, что предельная точка единственна. Предположим, что есть две предельные точки: $c_1<c_2$. Пусть $c_2-c_1=\varepsilon$. Выберем N так, что $|x_N-c_2|<\varepsilon/2$. Тогда $x_N>c_2-\frac{\varepsilon}{2}$. Значит, $x_n>c_2-\frac{\varepsilon}{2}$ при $n\geq N$. Но тогда $x_n-c_1>\frac{\varepsilon}{2}$, поэтому c_1 не может быть предельной точкой.

Пусть c — предельная точка. Тогда $x_n\leq c$ для всех n (иначе все точки x_m, где $m\geq n$, лежат вне некоторой окрестности точки c). Поэтому для каждого $\varepsilon>0$ можно выбрать N так, что $c-\varepsilon<x_N\leq c$. Тогда $c-\varepsilon<x_n\leq c$ при $n\geq N$. □

Задача 2.5. Ограничена ли последовательность $a_n=1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{n}$?

Задача 2.6. Пусть a_n и b_n — две последовательности, причём b_n — строго возрастающая неограниченная последовательность положительных чисел. Докажите, что если существует предел $\lim_{n\to\infty}\frac{a_{n+1}-a_n}{b_{n+1}-b_n}$, то существует и предел $\lim_{n\to\infty}\frac{a_n}{b_n}$, и оба эти предела равны (Штольц, 1885).

2.3. Последовательности Коши

Теорема 2.8 (критерий Коши). *Последовательность $\{x_n\}$ сходится тогда и только тогда, когда для любого $\varepsilon>0$ можно выбрать номер N так, что $|x_n-x_m|<\varepsilon$ для любых $m,n\geq N$.*

Доказательство. Если последовательность $\{x_n\}$ имеет предел a, то для любого $\varepsilon>0$ можно выбрать N так, что $|x_n-a|<\varepsilon/2$ для любого $n\geq N$. Тогда если $m,n\geq N$, то $|x_n-x_m|\leq|x_n-a|+|a-x_m|<\varepsilon$.

Предположим теперь, что для любого $\varepsilon>0$ можно выбрать номер N так, что $|x_n-x_m|<\varepsilon$ для любых $m,n\geq N$. Положим $\varepsilon=1$ и выберем соответствующий номер N. Тогда $|x_n-x_N|<1$ при $n\geq N$. Поэтому последовательность $\{x_n\}$ ограничена. По теореме Больцано–Вейерштрасса (теорема 2.5) такая последовательность имеет предельную точку a. Для положительного числа $\varepsilon/2$ выберем номер N так, что $|x_n-x_m|<\varepsilon/2$ при $m,n\geq N$. Точка a предельная, поэтому $|x_n-a|<\varepsilon/2$ для бесконечно многих n; в частности, это неравенство выполняется для некоторого $n_0\geq N$. Поэтому для любого $m\geq N$ получаем $|x_m-a|\leq|x_m-x_{n_0}|+|x_{n_0}-a|<\varepsilon$. Следовательно, $\lim_{m\to\infty}x_m=a$. □

Историческое замечание. Последовательности Коши и признак Коши сходимости последовательности впервые появились в работах чешского философа и богослова Бернарда Больцано (1781-1848) в 1817 году. В работах Огюстена Луи Коши (1789-1857) они появились лишь 4 года спустя, независимо от Больцано.

Последовательности Коши можно определить и использовать в более общей ситуации — для метрических пространств. Множество X, элементы которого будем называть точками, называют *метрическим пространством*, если для любых двух точек x и y задано число $d(x,y)$ (*расстояние* между точками x и y) и выполняются следующие свойства:

- $d(x,y)>0$, если $x\neq y$; $d(x,x)=0$;
- $d(x,y)=d(y,x)$;

- $d(x, y) \le d(x, z) + d(z, y)$ (*неравенство треугольника*).

Последовательность $\{x_n\}$ точек метрического пространства называют *последовательностью Коши* (или *фундаментальной последовательностью*), если для любого $\varepsilon > 0$ существует натуральное число N, для которого $d(x_n, x_m) < \varepsilon$ при $n > N$ и $m > N$.

Точку a метрического пространства называют *пределом* последовательности $\{x_n\}$, если для любого $\varepsilon > 0$ можно выбрать номер N так, что $|x_n - a| < \varepsilon$ при $n > N$.

Задача 2.7. Докажите, что если последовательность точек метрического пространства имеет предел, то эта последовательность фундаментальная.

В метрическом пространстве фундаментальная последовательность не всегда имеет предел. Например, рациональные числа можно рассматривать как метрическое пространство, и в нём можно выбрать последовательность рациональных чисел, сходящуюся к какому-нибудь иррациональному числу. Эта последовательность фундаментальная, но у неё нет предела.

Метрическое пространство называют *полным*, если в нём любая фундаментальная последовательность имеет предел.

Любому метрическому пространству можно сопоставить полное метрическое пространство, добавив все пределы фундаментальных последовательностей. Эту операцию называют *пополнением*. Вещественные числа получаются из рациональных с помощью пополнения. Это даёт ещё один подход к определению вещественных чисел.

2.4. Вычисление некоторых пределов

Задача 2.8. Вычислите $\lim_{n\to\infty}(\sqrt{n+1} - \sqrt{n})$.

Задача 2.9. Вычислите $\lim_{n\to\infty}(\sqrt{n^2+n} - n)$.

Задача 2.10. а) Докажите, что $\lim_{n\to\infty} \frac{n}{10^n} = 0$.
б) Докажите, что $\lim_{n\to\infty} \frac{\lg n}{n} = 0$.

Задача 2.11. а) Докажите, что $\lim_{n\to\infty} \frac{n^k}{a^n} = 0$ для любого натурального k и любого $a > 1$.
б) Докажите, что $\lim_{n\to\infty} \frac{(\log_a n)^k}{n} = 0$ для любого натурального k и любого $a > 1$.

Задача 2.12. Докажите, что $\lim_{n\to\infty} \frac{x^n}{n!} = 0$.

Задача 2.13. Докажите, что $\lim_{n\to\infty} \sqrt[n]{x} = 1$ для любого положительного числа x.

Задача 2.14. Докажите, что $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Задача 2.15. Докажите, что $\lim_{n\to\infty} \sqrt[n]{n!} = \infty$.

Задача 2.16. Пусть $a_0 = a$, $b_0 = b$, где $0 < a < b$. Положим $a_{n+1} = \frac{2a_nb_n}{a_n+b_n}$ и $b_{n+1} = \frac{a_n+b_n}{2}$ при $n \ge 0$. Докажите, что $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \sqrt{ab}$.

Задача 2.17. Пусть среди чисел $2^1, 2^2, \ldots, 2^n$ ровно a_n чисел начинается с единицы. Вычислите предел $\lim_{n\to\infty} \frac{a_n}{n}$.

Задача 2.18. Пусть $x_1 = \sqrt{a}$, где a — некоторое положительное число, и $x_{n+1} = \sqrt{a + x_n}$ при $n \ge 1$. Докажите, что существует предел $\lim_{n\to\infty} x_n$ и вычислите его.

Задача 2.19. Пусть x_1 и a — положительные числа, $x_{n+1} = \frac{a}{1+x_n}$ при $n \ge 1$. Докажите, что существует предел $\lim_{n\to\infty} x_n$ и вычислите его.

Задача 2.20. а) Пусть x_1 и a — положительные числа, $x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$ при $n \ge 1$. Вычислите предел $\lim_{n\to\infty} x_n$.
б) Пусть x_1 и a — положительные числа, m — натуральное число,

$$x_{n+1} = \frac{1}{m}\left((m-1)x_n + \frac{a}{x_n^{m-1}}\right)$$

при $n \ge 1$. Вычислите предел $\lim_{n\to\infty} x_n$.

Задача 2.21. Пусть $x_1 = a$, $x_2 = b$ и $x_{n+1} = \frac{1}{2}(x_n + x_{n-1})$ при $n \geq 1$. Вычислите $\lim_{n\to\infty} x_n$.

Задача 2.22. Пусть $p_1, \ldots, p_k$ — простые числа, a_n — количество натуральных чисел, не превосходящих n и делящихся только на данные простые числа. Докажите, что $\lim_{n\to\infty} \frac{a_n}{n} = 0$.

Задача 2.23. Пусть x_0 и y_0 — некоторые положительные числа, причём $x_0 > y_0$, $x_{n+1} = \frac{x_n+y_n}{2}$ и $y_{n+1} = \frac{2x_ny_n}{x_n+y_n}$ при $n \geq 0$. Докажите, что $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \sqrt{x_0y_0}$.

Задача 2.24. Пусть x_0 и y_0 — некоторые положительные числа, причём $x_0 > y_0$, $x_{n+1} = \frac{x_n+y_n}{2}$ и $y_{n+1} = \sqrt{x_ny_n}$ при $n \geq 0$. Докажите, что обе последовательности $\{x_n\}$ и $\{y_n\}$ сходятся к одному и тому же пределу, называемому *средним арифметико-геометрическим* чисел x_0 и y_0 (*Гаусс*).

Задача 2.25. *Пусть $a_0 = 2\sqrt{3}$, $b_0 = 3$ и $a_{n+1} = \frac{2a_nb_n}{a_n+b_n}$, $b_{n+1} = \sqrt{a_{n+1}b_n}$ при $n \geq 0$. Докажите, что $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \pi$.

Задача 2.26. *Пусть x_0 и y_0 — некоторые неотрицательные числа, $x_{n+1} = \frac{x_n+y_n}{2}$ и $y_{n+1} = \sqrt{x_{n+1}y_n}$ при $n \geq 0$. Докажите, что если $0 \leq x_0 < y_0$, то

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \frac{\sqrt{y_0^2 - x_0^2}}{\arccos(x_0/y_0)} \quad \text{(Борхардт).}$$

Задача 2.27. Найдите предел дробной части числа $(2+\sqrt{3})^n$ при $n \to \infty$.

Задача 2.28. Пусть $(1+\sqrt{2}+\sqrt{3})^n = q_n + r_n\sqrt{2} + s_n\sqrt{3} + t_n\sqrt{6}$, где q_n, r_n, s_n, t_n — натуральные числа. Вычислите пределы $\lim_{n\to\infty} \frac{r_n}{q_n}$, $\lim_{n\to\infty} \frac{s_n}{q_n}$ и $\lim_{n\to\infty} \frac{t_n}{q_n}$.

2.5. Число e

Важную роль в анализе играет число e, которое можно определить многими разными способами. Здесь мы обсудим некоторые определения и свойства этого числа, которые можно получить, опираясь только на понятие предела.

Сначала докажем, что последовательность $\left(1+\frac{1}{n}\right)^n$ ограничена (число e определяется как предел этой последовательности).

Лемма 2.1. *Для любого натурального n выполняются неравенства* $2 \leq \left(1+\frac{1}{n}\right)^n < 3$.

Первое доказательство. Согласно задаче 1.7, если $0 < \alpha \leq 1/n$, то выполняется неравенство $1 + n\alpha \leq \left(1+\frac{1}{n}\right)^n < 1 + n\alpha + n^2\alpha^2$. При $\alpha = 1/n$ получаем требуемое. □

Второе доказательство. Ясно, что

$$\left(1+\frac{1}{n}\right)^n = 1 + n\cdot\frac{1}{n} + \frac{n(n-1)}{2}\cdot\frac{1}{n^2} + \frac{n(n-1)(n-2)}{2\cdot 3}\cdot\frac{1}{n^3} + \cdots < \\ < 1 + 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots,$$

поскольку $\frac{n-1}{2n} < \frac{1}{2}$, $\frac{n-2}{3n} < \frac{1}{2}$, $\frac{n-3}{4n} < \frac{1}{2}$ и т.д. □

Задача 2.29. Найдите первую цифру числа 2^{400}.

Теперь докажем, что последовательность $\left(1+\frac{1}{n}\right)^n$ монотонно возрастает, а последовательность $\left(1+\frac{1}{n}\right)^{n+1}$ монотонно убывает.

Лемма 2.2. *Докажите, что если m и n — натуральные числа, причём $m > n$, то*

$$\left(1+\frac{1}{m}\right)^m > \left(1+\frac{1}{n}\right)^n \ \text{ и } \ \left(1+\frac{1}{m}\right)^{m+1} < \left(1+\frac{1}{n}\right)^{n+1}.$$

Первое доказательство. Достаточно доказать требуемые неравенства в случае, когда $m = n+1$.

Пусть $0 < a < b$. Тогда

$$(n+1)a^n < \frac{b^{n+1}-a^{n+1}}{b-a} < (n+1)b^n, \tag{2.2}$$

поскольку $\frac{b^{n+1}-a^{n+1}}{b-a} = b^n + b^{n-1}a + \dots + a^n$. Неравенства (2.2) можно записать в виде

$$a^n\big((n+1)b - na\big) < b^{n+1}, \quad b^n\big((n+1)a - nb\big) < a^{n+1}.$$

Подставим в эти неравенства $a = 1 + \frac{1}{n+1}$ и $b = 1 + \frac{1}{n}$. В результате получим

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

$$\left(1+\frac{1}{n}\right)^n < \left(1+\frac{1}{n+1}\right)^{n+1}.$$

Остаётся заметить, что

$$\left(\frac{n+1}{n+2}\right)^2 \left(\frac{(n+1)^2}{n} - \frac{n(n+2)}{n+1}\right) = \frac{n^3+4n^2+4n+1}{n(n+2)^2} > 1.$$

$\square$

Второе доказательство. Неравенство $\left(1+\frac{1}{n+1}\right)^{n+1} > \left(1+\frac{1}{n}\right)^n$ эквивалентно неравенству $\left(1+\frac{1}{n+1}\right) > \left(\frac{(n+1)^2}{n(n+2)}\right)^n$. Согласно задаче 1.7 б)

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

Неравенство $\frac{1}{n+2} + \frac{1}{(n+2)^2} < \frac{1}{n+1}$ легко проверяется.

Неравенство $\left(1+\frac{1}{n+1}\right)^{n+2} < \left(1+\frac{1}{n}\right)^{n+1}$ эквивалентно неравенству $1+\frac{1}{n+1} < \left(1+\frac{1}{n^2+2n}\right)^{n+1}$. Согласно задаче 1.7 а)

$$\left(1+\frac{1}{n^2+2n}\right)^{n+1} > 1 + \frac{n+1}{n^2+2n}.$$

Неравенство $\frac{n+1}{n^2+2n} > \frac{1}{n+1}$ легко проверяется. $\square$

Итак, согласно лемме 2.2 последовательность $a_n = \left(1+\frac{1}{n}\right)^n$ монотонно возрастает, а согласно лемме 2.1 эта последовательность ограничена сверху. Поэтому существует предел $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$.

Задача 2.30. а) Докажите, что

$$\left(1+\frac{1}{n}\right)^n < 2 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}.$$

б) Докажите, что

$$e \geq 2 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{k!}$$

для любого k.

в) Докажите, что $e = \lim_{n\to\infty}\left(2 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}\right)$.

Задача 2.31. Докажите, что

$$0 < e - \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}\right) < \frac{1}{n!n}.$$

Задача 2.32. Сходится ли последовательность $a_n = \sin(2\pi n! e)$?

Задача 2.33. Докажите, что число e иррационально.

2.6. Верхний и нижний пределы

Расширим понятия точной верхней и точной нижней грани: будем считать, что точная верхняя грань неограниченного сверху множества равна $+\infty$, а точная нижняя грань неограниченного снизу множества равна $-\infty$.

Для последовательности $\{a_n\}$ можно рассмотреть две последовательности: $b_n = \sup_{m\geq n} a_m$ и $c_n = \inf_{m\geq n} a_m$. Если из какого-то множества чисел убрать одно число, то точная верхняя грань полученного множества может только уменьшиться, а точная нижняя грань — только увеличиться. Поэтому $b_{n+1} \leq b_n$ и $c_{n+1} \geq c_n$, т.е. последовательность $\{b_n\}$ убывающая, а последовательность $\{c_n\}$ возрастающая. Если последовательность $\{a_n\}$ ограниченная, то пределы этих последовательностей называют соответственно *верхним* и *нижним пределом* последовательности $\{a_n\}$ и обозначают $\overline{\lim}_{n\to\infty} a_n$ и $\underline{\lim}_{n\to\infty} a_n$.

Историческое замечание. Понятия верхнего и нижнего предела последовательности ввёл Огюстен Луи Коши (1789-1857) в 1821 году.

Если последовательность $\{a_n\}$ неограниченная сверху, то $b_n = +\infty$ для всех n. В этом случае $\overline{\lim}_{n\to\infty} a_n = +\infty$. Если последовательность $\{a_n\}$ неограниченная снизу, то $c_n = -\infty$ для всех n. В этом случае $\underline{\lim}_{n\to\infty} a_n = -\infty$.

Если последовательность $\{a_n\}$ ограниченная сверху, но неограниченная снизу, то последовательность $\{b_n\}$ либо имеет конечный предел b (и тогда $\overline{\lim}_{n\to\infty} a_n = b$), либо неограниченно убывает (и тогда $\overline{\lim}_{n\to\infty} a_n = -\infty$). Если последовательность $\{a_n\}$ ограниченная снизу, но неограниченная сверху, то последовательность $\{c_n\}$ либо имеет конечный предел c (и тогда $\underline{\lim}_{n\to\infty} a_n = c$), либо неограниченно возрастает (и тогда $\underline{\lim}_{n\to\infty} a_n = +\infty$).

Расширим понятие предельной точки последовательности. Будем считать предельными точками последовательности не только пределы сходящихся подпоследовательностей, но и точки $+\infty$ (для неограниченной сверху последовательности) и $-\infty$ (для неограниченной снизу последовательности).

Теорема 2.9. *Верхний (нижний) предел последовательности — это точная верхняя (нижняя) грань её предельных точек.*

Доказательство. Докажем требуемое утверждение для верхнего предела последовательности $\{a_n\}$ (для нижнего предела доказательство аналогично). Пусть $\overline{\lim}_{n\to\infty} a_n = b$. Тогда $b = \inf_n b_n$, где $b_n = \sup_{m\geq n} a_m$. Рассмотрим три случая.

1) $b = -\infty$. Тогда для любого числа M можно выбрать n так, что $b_n \leq M$, поэтому $a_m \leq b_n \leq M$ при $m \leq n$. Это означает, что у последовательности $\{a_n\}$ есть только одна предельная точка, а именно, точка $-\infty$.

2) $b \in \mathbb{R}$. Если $x > b$, то $b_n < x$ для некоторого n. Тогда $a_m \leq b_n < x$ при $m \geq n$. Поэтому у последовательности $\{a_n\}$ нет предельных точек которые, больше b.

Докажем теперь, что b — предельная точка последовательности $\{a_n\}$. Выберем $\varepsilon > 0$. Из того, что $\sup_{m\geq n} a_m = b_n \geq b$, следует, что для каждого n можно выбрать m_n так, что $a_{m_n} > b - \varepsilon$ и $m_n \geq n$. Тогда $b - \varepsilon < a_{m_n} \leq b_n \leq b_1$. Таким образом, на отрезке $[b-\varepsilon, b_1]$ расположено бесконечно много точек последовательности $\{a_n\}$, причём у этой последовательности нет предельных точек, которые больше b. Сдедовательно, для любого $\varepsilon > 0$ найдётся бесконечно много точек a_n, для которых $b - \varepsilon < a_n < b + \varepsilon$.

3) $b = +\infty$. Тогда $\inf_n b_n = +\infty$, поэтому $b_n = +\infty$ для всех n. Таким образом, $\sup_{m\geq n} a_m = b_n = +\infty$, т.е. для любого числа M можно выбрать $m \geq n$ так, что $a_m > M$. Следовательно, $+\infty$ — предельная точка последовательности $\{a_n\}$. Ясно, что эта предельная точка — точная верхняя грань всех предельных точек. $\square$

Задача 2.34. Докажите, что $\lim_{n\to\infty} a_n = a$ тогда и только тогда, когда $\overline{\lim}_{n\to\infty} a_n = \underline{\lim}_{n\to\infty} a_n = a$.

Задача 2.35. Докажите, что

$$\underline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n \leq \overline{\lim}_{n\to\infty}(a_n + b_n) \leq \overline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n,$$

$$\underline{\lim}_{n\to\infty} a_n + \overline{\lim}_{n\to\infty} b_n \geq \underline{\lim}_{n\to\infty}(a_n + b_n) \geq \underline{\lim}_{n\to\infty} a_n + \underline{\lim}_{n\to\infty} b_n.$$

Задача 2.36. Докажите, что если последовательность a_n сходится, то

$$\overline{\lim}_{n\to\infty}(a_n+b_n)=\lim_{n\to\infty}a_n+\overline{\lim}_{n\to\infty}b_n.$$

Задача 2.37. Докажите, что если для некоторой последовательности a_n и для любой последовательности b_n

$$\overline{\lim}_{n\to\infty}(a_n+b_n)=\overline{\lim}_{n\to\infty}a_n+\overline{\lim}_{n\to\infty}b_n,$$

то последовательность a_n сходится.

Задача 2.38. Докажите, что $\overline{\lim}_{n\to\infty}\frac{a_1+a_2+\dots+a_n}{n}\le\overline{\lim}_{n\to\infty}a_n$.

Задача 2.39. *Докажите, что для любой последовательности $\{x_n\}$ положительных чисел выполняется неравенство $\overline{\lim}_{n\to\infty}\left(\frac{x_1+x_{n+1}}{x_n}\right)^n\ge e$.

2.7. Теорема Тёплица

Пусть для всех пар натуральных i,j, где $i\ge j$, заданы некоторые числа t_{ij}. Рассмотрим следующее преобразование последовательностей: сопоставим последовательности $\{a_n\}$ последовательность $\{a'_n\}$, где $a'_n=t_{n1}a_1+t_{n2}a_2+\dots+t_{nn}a_n$. Нас интересуют преобразования, сопоставляющие любой сходящейся последовательности последовательность, которая сходится к тому же пределу.

С двумя примерами таких преобразований мы уже встречались. Первый пример — теорема Коши (задача 2.2). В этом случае $t_{nm}=\frac{1}{n}$. Второй пример — теорема Штольца (задача 2.6). Теорема Штольца является обобщением теоремы Коши: чтобы получить теорему Коши, достаточно в теореме Штольца положить $b_n=n$. Чтобы записать теорему Штольца в указанном выше виде, нужно положить $t_{nm}=\frac{b_m-b_{m-1}}{b_n}$ и применить преобразование к последовательности $\left\{\frac{a_n-a_{n-1}}{b_n-b_{n-1}}\right\}$; в результате получим последовательность $\left\{\frac{a_n}{b_n}\right\}$ (в формулах нужно положить $a_0=b_0=0$).

Несложно проверить, что должны выполняться следующие свойства:

1) $t_{nm}\to 0$ при $n\to\infty$ и фиксированном m;

2) $|t_{n1}|+|t_{n2}|+\dots+|t_{nn}|\le K$ для некоторой константы K (не зависящей от n);

3) $T_n=t_{n1}+t_{n2}+\dots+t_{nn}\to 1$ при $n\to\infty$.

Теорема 2.10 (Тёплиц). *а) Если выполняются свойства 1) и 2), то при преобразовании последовательность $\{a_n\}$, сходящаяся к нулю, переходит в последовательность $\{a'_n\}$, сходящуюся к нулю.*

б) Если выполняются свойства 1), 2) и 3), то при преобразовании последовательность $\{a_n\}$, сходящаяся к конечному пределу a, переходит в последовательность $\{a'_n\}$, сходящуюся к a.

Доказательство. а) Для заданного $\varepsilon>0$ выберем m так, что $|a_n|<\frac{\varepsilon}{2K}$ при $n>m$. Тогда из свойства 1) при $n>m$ получаем

$$|a'_n|<|t_{n1}a_1+t_{n2}a_2+\dots+t_{nm}a_m|+\frac{\varepsilon}{2}.$$

Из свойства 2) следует, что можно выбрать $N>m$ так, что при $n>M$ выполняется неравенство

$$|t_{n1}a_1+t_{n2}a_2+\dots+t_{nm}a_m|<\frac{\varepsilon}{2},$$

поэтому $|a'_n|<\varepsilon$.

б) Выражение для a'_n можно записать в виде

$$a'_n=t_{n1}(a_1-a)+t_{n2}(a_2-a)+\dots+t_{nn}(a_n-a)+T_na.$$

Применив утверждение а) к последовательности $\{a_n-a\}$, сходящейся к нулю, и воспользовавшись тем, что $T_n\to 1$ при $n\to\infty$, получаем требуемое. □

Историческое замечание. Отто Тёплиц (1881-1940) доказал теорему о сходящихся последовательностях в 1911 году.

2.8. Решения задач

2.1. Пусть $\left|\frac{x_n-a}{x_n+a}\right| < \varepsilon < 1$. Тогда

$$\frac{-2\varepsilon a}{1+\varepsilon} < x_n - a < \frac{2\varepsilon a}{1-\varepsilon}.$$

2.2. Вместо последовательности $\{x_n\}$ можно рассмотреть последовательность $\{x_n - a\}$, поэтому можно считать, что $a = 0$. Для любого $\varepsilon > 0$ можно выбрать N так, что если $n \geq N$, то $|x_n| < \varepsilon$. Тогда при $n > N$

$$\left|\frac{1}{n}(x_1 + \dots + x_n)\right| \leq \frac{C}{n} + \frac{n-N}{n}\varepsilon,$$

где $C = x_1 + \dots + x_N$. Ясно, что $\lim_{n\to\infty}\left(\frac{C}{n} + \frac{n-N}{n}\varepsilon\right) = \varepsilon$. Поэтому можно выбрать $M > N$ так, что если $n > M$, то $\left|\frac{1}{n}(x_1 + \dots + x_n)\right| < 2\varepsilon$.

2.3. Пусть $y_1 = x_1$ и $y_n = x_n - x_{n-1}$ при $n \geq 2$. Тогда $x_n = y_1 + y_2 + \dots + y_n$. По условию $\lim_{n\to\infty} y_n = a$. Поэтому согласно задаче 2.2 $\lim_{n\to\infty} \frac{x_n}{n} = \lim_{n\to\infty} \frac{y_1+y_2+\dots+y_n}{n} = a$.

2.4. Удобнее доказывать более общее утверждение: если $\{P(n)\alpha\} = a_n + \varepsilon_n$, где $\lim_{n\to\infty}\varepsilon_n = 0$ и a_n принимает лишь конечное число значений (когда n пробегает все натуральные значения), то α рационально.

Применим индукцию по m — степени многочлена $P(x)$. При $m = 1$ получаем $\{An\alpha + B\alpha\} = a_n + \varepsilon_n$. Поэтому $An\alpha + B\alpha = a_n + \varepsilon_n + k_n$ и $A(n+1)\alpha + B\alpha = a_{n+1} + \varepsilon_{n+1} + k_{n+1}$, где k_n и k_{n+1} — целые числа. Следовательно,

$$A\alpha = (a_{n+1} - a_n) + (\varepsilon_{n+1} - \varepsilon_n) + (k_{n+1} - k_n).$$

Число $k_{n+1} - k_n$ целое, а разность $a_{n+1} - a_n$ может принимать лишь конечное число значений. Поэтому из того, что $\lim_{n\to\infty}\varepsilon_n = 0$, следует, что $\varepsilon_{n+1} - \varepsilon_n = 0$ при $n \geq n_0$. Но тогда

$$a_n = a_{n_0} + l_n + (n - n_0)A\alpha,$$

где l_n — целое число. По условию a_n принимает конечное число значений. Следовательно, $l_n + nA\alpha$ принимает конечное число значений, поэтому $\{nA\alpha\}$ принимает конечное число значений. Тогда число $A\alpha$ рационально (задача 1.1), а значит, число α тоже рационально.

Шаг индукции доказывается совсем просто. Многочлен $Q(x) = P(x+1) - P(x)$ имеет степень $m-1$. Кроме того, если a_n принимает конечное число значений, то $a_{n+1} - a_n$ тоже принимает конечное число значений. Поэтому, применив к многочлену $Q(x)$ предположение индукции, получаем, что α рационально.

2.5. Ответ: Нет. Ясно, что $a_{2^{n+1}} - a_{2^n} \geq 2^n \cdot \frac{1}{2^{n+1}}$, поэтому $a_{2^{n+1}} \geq \frac{n}{2} + 1$.

2.6. Для заданного $\varepsilon > 0$ можно выбрать N так, что при $n \geq N$ выполняются неравенства

$$l - \varepsilon < \frac{a_{n+1} - a_n}{b_{n+1} - b_n} < l + \varepsilon.$$

По условию $b_{n+1} - b_n > 0$, поэтому

$$(l-\varepsilon)(b_{n+1} - b_n) < a_{n+1} - a_n < (l+\varepsilon)(b_{n+1} - b_n).$$

Запишем такие неравенства для $n = N, N+1, \dots, k$ и сложим их. В результате получим

$$(l-\varepsilon)(b_{k+1} - b_N) < a_{k+1} - a_N < (l+\varepsilon)(b_{k+1} - b_N),$$

т.е.

$$(l-\varepsilon)\left(1 - \frac{b_N}{b_{k+1}}\right) + \frac{a_N}{b_{k+1}} < \frac{a_{k+1}}{b_{k+1}} < (l+\varepsilon)\left(1 - \frac{b_N}{b_{k+1}}\right) + \frac{a_N}{b_{k+1}}.$$

Последовательность b_n возрастает неограниченно, поэтому для достаточно больших k получаем

$$l - \varepsilon \leq \frac{a_{k+1}}{b_{k+1}} \leq l + \varepsilon.$$

2.7. Пусть последовательность $\{x_n\}$ сходится к точке a. Выберем номер N так, что $|x_n - a| < \varepsilon/2$ при $n > N$. Тогда $|x_n - x_m| \le |x_n - a| + |x_m - a| < \varepsilon/2 + \varepsilon/2 = \varepsilon$ при $n, m > N$.

2.8. Ответ: 0. Ясно, что $0 < \sqrt{n+1} - \sqrt{n} = \frac{(\sqrt{n+1}-\sqrt{n})(\sqrt{n+1}+\sqrt{n})}{\sqrt{n+1}+\sqrt{n}} = \frac{1}{\sqrt{n+1}+\sqrt{n}} < \frac{1}{\sqrt{n}}$. Поэтому $\lim_{n\to\infty}(\sqrt{n+1} - \sqrt{n}) = 0$.

2.9. Ответ: 1/2. Домножив и поделив $\sqrt{n^2+n} - n$ на $\sqrt{n^2+n} + n$, получим

$$\sqrt{n^2+n} - n = \frac{n}{\sqrt{n^2+n}+n} = \frac{1}{\sqrt{1+\frac{1}{n}}+1} \to \frac{1}{2}.$$

2.10. а) Ясно, что

$$\frac{n}{10^n} = \frac{n}{n-1} \cdot \frac{n-1}{n-2} \cdots \frac{2}{1} \cdot \frac{1}{10^n} < 2^{n-1}\frac{1}{10^n} = \frac{1}{2} \cdot \frac{1}{5^n} \to 0.$$

б) Пусть $10^k \le n \le 10^{k+1}$. Тогда $\frac{\lg n}{n} \le \frac{k+1}{10^k} \to 0$ согласно задаче а).

2.11. а) Если n достаточно велико, то $\sqrt[k]{a} > \frac{n+1}{n} = 1 + \frac{1}{n}$. Выберем число $q > 1$ и натуральное число N так, что $q < \left(\frac{n}{n+1}\right)^k a$ при $n \ge N$. Тогда $q < \left(\frac{N}{N+1}\right)^k a$, $q < \left(\frac{N+1}{N+2}\right)^k a$, $q < \left(\frac{N+m-1}{N+m}\right)^k a$, поэтому $q^m < \left(\frac{N}{N+m}\right)^k a^m$, а значит, $\frac{(N+m)^k}{a^{N+m}} < Cq^{-m}$, где $C = \frac{N^k}{a^N}$. Остаётся заметить, что если $q > 1$, то $\lim_{m\to\infty} q^{-m} = 0$.

б) Пусть $a^{m-1} \le n \le a^m$. Тогда $\frac{(\log_a n)^k}{n} \le \frac{m^k}{a^{m-1}} \to 0$ согласно задаче а).

2.12. Выберем натуральное число N так, что $\frac{|x|}{N+1} \le \frac{1}{2}$. Если $n > N$, то $\frac{|x|^n}{n!} \le \frac{|x|^N}{N!}\left(\frac{1}{2}\right)^{n-N} = \frac{C}{2^n}$, где $C = \frac{2^N|x|^N}{N!}$ — постоянное число.

2.13. Пусть $x > 1$. Рассмотрим вспомогательную последовательность $a_n = x^{1/n} - 1$. Ясно, что $a_n > 0$, Поэтому согласно задаче 1.7 а) $1 + na_n \le (1 + a_n)^n = x$. Таким образом, $0 < a_n < \frac{x-1}{n}$. Поэтому $\lim_{n\to\infty} a_n = 0$.

Если $x = 1$, то утверждение очевидно, а если $0 < x < 1$, то рассмотрим $y = 1/x > 1$. Тогда $\lim_{n\to\infty} y^{1/n} = 1$, поэтому $\lim_{n\to\infty} x^{1/n} = 1$.

2.14. Первое решение. Пусть $a_n = \sqrt[n]{n} - 1$. Тогда $a_n \ge 0$, поэтому $n = (1+a_n)^n \ge 1 + na_n + \frac{n(n-1)}{2}a_n^2 \ge \frac{n(n-1)}{2}a_n^2$. Следовательно, $0 \le a_n \le \sqrt{\frac{2}{n-1}}$.

Второе решение. Воспользовавшись неравенством между средним арифметическим и средним геометрическим, получим:

$$\sqrt[n]{n} - 1 = \frac{n-1}{1 + n^{\frac{1}{n}} + \cdots + n^{\frac{n-1}{n}}} \le \frac{n-1}{n\sqrt[n]{n^{\frac{1}{n}+\frac{2}{n}+\cdots+\frac{n-1}{n}}}} = \frac{n-1}{n^{(3n-1)/2n}}.$$

Если $n \ge 3$, то $\frac{n-1}{n^{(3n-1)/2n}} \le n^{\frac{-n+1}{2n}} \le n^{-1/3}$.

2.15. Из неравенства $(2n)! \ge n(n+1)\cdots(2n) > n^{n+1}$ следует, что $\sqrt[2n]{(2n)!} > \sqrt[2n]{n^{n+1}} > \sqrt{n}$ и $\sqrt[2n+1]{(2n+1)!} > \sqrt[2n+1]{n^{n+1}} > \sqrt{n}$.

2.16. Если $0 < a_n < b_n$, то $a_{n+1} > a_n$ и $b_{n+1} < b_n$. Кроме того, $4a_nb_n < (a_n+b_n)^2$, т.е. $a_{n+1} < b_{n+1}$. Поэтому $a_0 < a_1 < a_2 < \cdots < b_2 < b_1 < b_0$. Значит, последовательности $\{a_n\}$ и $\{b_n\}$ сходятся. Пусть $\lim_{n\to\infty} a_n = \alpha$ и $\lim_{n\to\infty} b_n = \beta$. Из того, что $b_{n+1} = \frac{a_n+b_n}{2}$, следует, что $\beta = \frac{\alpha+\beta}{2}$, т.е. $\alpha = \beta$.

Перемножив равенства $a_{n+1} = \frac{2a_nb_n}{a_n+b_n}$ и $b_{n+1} = \frac{a_n+b_n}{2}$, получим $a_{n+1}b_{n+1} = a_nb_n$. Значит, $a_nb_n = ab$ при всех n. Поэтому $\alpha\beta = ab$, т.е. $\alpha^2 = ab$.

2.17. Рассмотрим все натуральные числа a, для которых $10^{k-1} \le a < 10^k$. Среди них есть хотя бы одна степень двойки, поскольку не может оказаться, что $2^m < 10^{k-1}$, а $2^{m+1} > 10^k$. Наименьшая из степеней двойки, заключённых в этих пределах, начинается с единицы, поскольку иначе мы могли бы поделить число $a = 2^m$ на 2 и получить число в тех же пределах. Степень двойки, следующая за наименьшей, начинается с цифры 2 или 3. Поэтому среди рассматриваемых чисел есть ровно одна степень двойки, начинающаяся с единицы. Значит, если $10^{k-1} \le 2^n < 10^k$, то $a_n = k - 1$. Таким образом, $a_n \le n\lg 2 < a_n + 1$, т.е. $\lg 2 - \frac{1}{n} < \frac{a_n}{n} \le \lg 2$. Поэтому $\lim_{n\to\infty} \frac{a_n}{n} = \lg 2$.

2.18. Если бы мы знали, что существует предел $\lim_{n\to\infty} x_n = c$, то найти c было бы легко. Действительно, $c = \sqrt{a+c}$, поэтому $c^2 - c - a = 0$. Следовательно, $c = \frac{1}{2} \pm \sqrt{a + \frac{1}{4}}$. Ясно также, что $c \geq 0$, поэтому $c = \frac{1}{2} + \sqrt{a + \frac{1}{4}}$.

Докажем теперь, что рассматриваемый предел существует. Прежде всего заметим, что последовательность $\{x_n\}$ монотонно возрастает. Действительно, учитывая, что все числа x_n положительны, получаем: $x_{n+1} > x_n \iff x_{n+1}^2 > x_n^2 \iff a + x_n > x_n^2 \iff a + x_n > a + x_{n-1} \iff x_n > x_{n-1}$. Остаётся заметить, что $x_2 = \sqrt{a + \sqrt{a}} > \sqrt{a} = x_1$.

Докажем, что последовательность $\{x_n\}$ ограничена сверху, а именно, $x_n < c$. Действительно, $x_1 < c$. Кроме того, $x_n < c \Rightarrow x_{n+1} < \sqrt{a+c} = c$.

2.19. Если бы мы знали, что существует предел $\lim_{n\to\infty} x_n = c$, то найти его было бы легко. Действительно, $c = \frac{a}{1+c}$, поэтому $c^2 + c - a = 0$. Ясно также, что $c \geq 0$. Поэтому c — положительный корень уравнения $x^2 + x - a = 0$. Такой корень единствен.

Докажем теперь существование предела. Пусть c — единственный положительный корень уравнения $x^2 + x - a = 0$. Если $x_n \geq c$, то $x_n^2 + x_n \geq a$, т.е. $x_n \geq \frac{a}{1+x_n} = x_{n+1}$. Поэтому $x_{n+1} = \frac{a}{1+x_n} \leq \frac{a}{1+x_{n+1}}$, т.е. $x_{n+1} \leq c$. Аналогично доказывается, что если $x_n \leq c$, то $x_{n+1} \geq c$.

Далее, $x_{n+2} - x_n = \frac{a}{1+\frac{a}{1+x_n}} - x_n = \frac{a - x_n - x_n^2}{1 + a + x_n}$. Поэтому если $x_n \leq c$, то $x_{n+2} - x_n \geq 0$, а если $x_n \geq c$, то $x_{n+2} - x_n \leq 0$. Таким образом, одна из последовательностей $x_1, x_3, x_5, \ldots$ и $x_2, x_4, x_6, \ldots$ монотонно возрастает, а другая монотонно убывает. Предел каждой из этих последовательностей является положительным корнем уравнения $x = \frac{a}{1+\frac{a}{1+x}}$, которое эквивалентно уравнению $x^2 + x - a = 0$.

2.20. а) Ответ: $\lim_{n\to\infty} x_n = \sqrt{a}$. Ясно, что

$$\frac{x_{n+1} - \sqrt{a}}{x_{n+1} + \sqrt{a}} = \frac{x_n^2 - 2\sqrt{a}x_n + a}{x_n^2 + 2\sqrt{a}x_n + a} = \left(\frac{x_n - \sqrt{a}}{x_n + \sqrt{a}}\right)^2.$$

Следовательно, $\frac{x_{n+1}-\sqrt{a}}{x_{n+1}+\sqrt{a}} = \left(\frac{x_1-\sqrt{a}}{x_1+\sqrt{a}}\right)^{2^n}$. Пусть $q = \frac{x_1-\sqrt{a}}{x_1+\sqrt{a}}$. Ясно, что $|q| < 1$. Поэтому $\lim_{n\to\infty} q^{2^n} = 0$. Значит, $\lim_{n\to\infty} \frac{x_n-\sqrt{a}}{x_n+\sqrt{a}} = 0$. Остаётся воспользоваться результатом задачи 2.1.

б) Ответ: $\lim_{n\to\infty} x_n = \sqrt[m]{a}$.

2.21. Ответ: $\lim_{n\to\infty} x_n = \frac{a+2b}{3}$. Докажем, что $x_n = \frac{a+2b}{3} + \left(\frac{-1}{2}\right)^{n-2} \frac{b-a}{3}$. При $n = 1$ и 2 требуемое равенство легко проверяется. Соотношение $x_{n+1} = \frac{1}{2}(x_n + x_{n-1})$ тоже легко проверяется.

2.22. Пусть $p_1^{\alpha_1} \cdots p_k^{\alpha_k} \leq n$. Тогда $p_i^{\alpha_i} \leq n$, поэтому $\alpha_i \leq \frac{\lg n}{\lg p_i}$. Значит, $a_n \leq \left(\frac{\lg n}{\lg p_1} + 1\right) \cdots \left(\frac{\lg n}{\lg p_k} + 1\right)$. Если $n \geq p_1, \ldots, p_k$, то $a_n \leq \frac{2^k}{\lg p_1 \cdots \lg p_k} (\lg n)^k$. Остаётся воспользоваться результатом задачи 2.11 б).

2.23. Легко проверить, что $x_0 > \frac{x_0+y_0}{2} > \frac{2x_0y_0}{x_0+y_0} > y_0$, т.е. $x_0 > x_1 > y_1 > y_0$. Аналогично $x_n > x_{n+1} > y_{n+1} > y_n$ для любого n. Таким образом, последовательности $\{x_n\}$ и $\{y_n\}$ монотонные и ограниченные, поэтому они сходятся к некоторым числам x и y. Если в равенстве $x_{n+1} = \frac{x_n+y_n}{2}$ перейти к пределу, то получим $x = \frac{x+y}{2}$, т.е. $x = y$. Ясно также, что $x_{n+1}y_{n+1} = x_ny_n$. Поэтому $xy = x_0y_0$. А так как $x = y$, получаем требуемое.

2.24. Легко проверить, что $x_0 > \frac{x_0+y_0}{2} > \sqrt{x_0y_0} > y_0$, т.е. $x_0 > x_1 > y_1 > y_0$. Аналогично $x_n > x_{n+1} > y_{n+1} > y_n$ для любого n. Таким образом, последовательности $\{x_n\}$ и $\{y_n\}$ монотонные и ограниченные, поэтому они сходятся к некоторым числам x и y. Если в равенстве $x_{n+1} = \frac{x_n+y_n}{2}$ перейти к пределу, то получим $x = \frac{x+y}{2}$, т.е. $x = y$.

2.25. Пусть $A_n = 6 \cdot 2^n \operatorname{tg} \frac{\pi}{6\cdot 2^n}$ и $B_n = 6 \cdot 2^n \sin \frac{\pi}{6\cdot 2^n}$, т.е. A_n — полупериметр правильного $6 \cdot 2^n$-угольника, описанного вокруг окружности радиуса 1, B_n — полупериметр правильного $6 \cdot 2^n$-угольника, вписанного в окружность радиуса 1. Покажем, что $a_n = A_n$ и $b_n = B_n$. При $n = 0$ это очевидно. Пусть $\alpha = \frac{\pi}{6\cdot 2^{n+1}}$. Тогда $\frac{A_{n+1}}{A_n} = \frac{2 \operatorname{tg} \alpha}{\operatorname{tg} 2\alpha} = 1 - \operatorname{tg}^2 \alpha$ и $\frac{B_{n+1}}{B_n} = \frac{2 \sin \alpha}{\sin 2\alpha} = \frac{1}{\cos \alpha}$. Кроме того, $\frac{A_n}{B_n} = \frac{1}{\cos 2\alpha}$ и $\frac{A_{n+1}}{B_{n+1}} = \frac{1}{\cos \alpha}$. Значит, $\frac{B_{n+1}}{B_n} = \frac{1}{\cos \alpha} = \frac{A_{n+1}}{B_{n+1}}$, т.е. $B_{n+1} = \sqrt{A_{n+1}B_n}$. Далее, $\frac{2A_nB_n}{A_n+B_n} = \frac{2A_n \cos 2\alpha}{1+\cos 2\alpha}$. Легко проверить, что $\frac{2\cos 2\alpha}{1+\cos 2\alpha} = 1 - \operatorname{tg}^2 \alpha$. Поэтому $\frac{2A_nB_n}{A_n+B_n} = A_{n+1}$.

2.26. Положим $x_0 = y_0 \cos\alpha$, т.е. $\alpha = \arccos(x_0/y_0)$. Тогда $x_1 = y_0 \frac{1+\cos\alpha}{2} = y_0 \cos^2\frac{\alpha}{2}$ и $y_1 = \sqrt{y_0\cos^2\frac{\alpha}{2} y_0} = y_0 \cos\frac{\alpha}{2}$. Поэтому $x_1 = y_1 \cos\frac{\alpha}{2}$. Продолжая такие рассуждения дальше, получаем $y_2 = y_0 \cos\frac{\alpha}{2}\cos\frac{\alpha}{2^2}$, …, $y_n = y_0 \cos\frac{\alpha}{2}\cos\frac{\alpha}{2^2}\dots\cos\frac{\alpha}{2^n} = \frac{y_0\sin\alpha}{2^n \sin\frac{\alpha}{2^n}} \to \frac{y_0\sin\alpha}{\alpha}$ и $x_n = y_n\cos\frac{\alpha}{2^n}$. Остаётся заметить, что $\cos\frac{\alpha}{2^n} \to 1$ при $n\to\infty$.

2.27. Ответ: 1. Легко проверить, что $0 < 2-\sqrt{3} < 1$, поэтому $(2-\sqrt{3})^n \to 0$ при $n\to\infty$. Число $(2-\sqrt{3})^n + (2+\sqrt{3})^n$ при всех целых n целое, поэтому при больших n дробная часть числа $(2+\sqrt{3})^n$ имеет вид, 0,9999…. Следовательно, дробная часть этого числа стремится к 1 при $n\to\infty$.

2.28. Пусть $\lambda_1 = 1+\sqrt{2}+\sqrt{3}$, $\lambda_2 = 1-\sqrt{2}+\sqrt{3}$, $\lambda_3 = 1+\sqrt{2}-\sqrt{3}$, $\lambda_4 = 1-\sqrt{2}-\sqrt{3}$. Тогда

$$\begin{aligned}
\lambda_1^n &= q_n + r_n\sqrt{2} + s_n\sqrt{3} + t_n\sqrt{6},\\
\lambda_2^n &= q_n - r_n\sqrt{2} + s_n\sqrt{3} - t_n\sqrt{6},\\
\lambda_3^n &= q_n + r_n\sqrt{2} - s_n\sqrt{3} - t_n\sqrt{6},\\
\lambda_4^n &= q_n - r_n\sqrt{2} - s_n\sqrt{3} + t_n\sqrt{6}.
\end{aligned}$$

Поэтому

$$\begin{aligned}
\lambda_1^n + \lambda_2^n + \lambda_3^n + \lambda_4^n &= 4q_n,\\
\lambda_1^n + \lambda_2^n - \lambda_3^n - \lambda_4^n &= 4s_n\sqrt{3},\\
\lambda_1^n - \lambda_2^n + \lambda_3^n - \lambda_4^n &= 4r_n\sqrt{2},\\
\lambda_1^n - \lambda_2^n - \lambda_3^n + \lambda_4^n &= 4t_n\sqrt{6}.
\end{aligned}$$

Ясно также, что $\lim_{n\to\infty}\frac{\lambda_2^n}{\lambda_1^n} = \lim_{n\to\infty}\frac{\lambda_3^n}{\lambda_1^n} = \lim_{n\to\infty}\frac{\lambda_4^n}{\lambda_1^n} = 0$. Поэтому $\lim_{n\to\infty}\frac{\lambda_1^n}{q_n} = 4$, $\lim_{n\to\infty}\frac{r_n}{\lambda_1^n} = \frac{1}{4\sqrt{2}}$, $\lim_{n\to\infty}\frac{s_n}{\lambda_1^n} = \frac{1}{4\sqrt{3}}$, $\lim_{n\to\infty}\frac{t_n}{\lambda_1^n} = \frac{1}{4\sqrt{6}}$, а значит, $\lim_{n\to\infty}\frac{r_n}{q_n} = \frac{1}{\sqrt{2}}$, $\lim_{n\to\infty}\frac{s_n}{q_n} = \frac{1}{\sqrt{3}}$ и $\lim_{n\to\infty}\frac{t_n}{q_n} = \frac{1}{\sqrt{6}}$.

2.29. Ответ: 2. Первая цифра числа $2^{400} = (2^{10})^{40} = (1024)^{40}$ совпадает с первой цифрой числа $1,024^{40}$. С одной стороны, согласно лемме 2.1 $1,024^{40} < 1,025^{40} = \left(1+\frac{1}{40}\right)^{40} < 3$. С другой стороны, $1,024^{40} > 1 + 40\cdot 0,024 + \frac{40\cdot 39}{2}0,024^2 = 2,40928 > 2$.

2.30. а) По формуле бинома

$$\left(1+\frac{1}{n}\right)^n = 1 + n\frac{1}{n} + \binom{n}{2}\frac{1}{n^2} + \dots + \binom{n}{k}\frac{1}{n^k} + \dots + \binom{n}{n}\frac{1}{n^n}.$$

Кроме того, $\binom{n}{k} = \frac{n(n-1)\cdots(n-k-1)}{k!} < \frac{n^k}{k!}$.

б) Если $k \le n$, то

$$\begin{aligned}
\left(1+\frac{1}{n}\right)^n &= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \cdots\\
&\cdots + \frac{1}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right) + \dots + \frac{1}{n!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{n-1}{n}\right) \ge\\
&\ge 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \dots + \frac{1}{k!}\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right) = b_n.
\end{aligned}$$

Ясно, что $\lim_{n\to\infty} b_n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{k!}$. Поэтому

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \ge \lim_{n\to\infty} b_n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{k!}.$$

в) Очевидным образом следует из задач а) и б).

2.31. Согласно задаче 2.30

$$e-\left(1+1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}\right)=\frac{1}{(n+1)!}+\frac{1}{(n+2)!}+\frac{1}{(n+3)!}+\cdots=$$
$$=\frac{1}{(n+1)!}\left(1+\frac{1}{n+2}+\frac{1}{(n+2)(n+3)}+\cdots\right)<$$
$$<\frac{1}{(n+1)!}\left(1+\frac{1}{n+2}+\frac{1}{(n+2)^2}+\cdots\right)=\frac{1}{(n+1)!}\frac{1}{1-\frac{1}{n+2}}=$$
$$=\frac{1}{(n+1)!}\frac{n+2}{n+1}=\frac{1}{n!}\frac{n+2}{(n+1)^2}.$$

Остаётся заметить, что $\frac{n+2}{(n+1)^2}<\frac{1}{n}$, поскольку $n^2+2n<n^2+2n+1$.

2.32. Ответ: да, сходится к 0. Согласно задаче 2.31 дробная часть числа $n!e$ заключена между 0 и $1/n$. Поэтому $0<\sin(2\pi n!e)<\sin(2\pi/n)$ при $n\geq 4$.

2.33. Предположим, что $e=m/n$, где m и n — натуральные числа. Тогда согласно задаче 2.31

$$0<\frac{m}{n}-\left(2+\frac{1}{2!}+\cdots+\frac{1}{n!}\right)<\frac{1}{n!n}.$$

После умножения на $n!$ получим, что $0<a<\frac{1}{n}$, где a — целое число. Этого не может быть.

2.34. Предположим, сначала что $\lim_{n\to\infty}a_n=a$. Тогда последовательность $\{a_n\}$ имеет единственную предельную точку (точку a). Поэтому согласно теореме 2.9 $\overline{\lim}_{n\to\infty}a_n=\underline{\lim}_{n\to\infty}a_n=a$.

Предположим теперь, что $\overline{\lim}_{n\to\infty}a_n=\underline{\lim}_{n\to\infty}a_n=a$. Тогда согласно теореме 2.9 последовательность $\{a_n\}$ имеет единственную предельную точку (точку a). Для любого $\varepsilon>0$ вне ε-окрестности точки a лежит конечное число точек последовательности $\{a_n\}$, поэтому $\lim_{n\to\infty}a_n=a$.

2.35. Докажем только первые два неравенства; два других неравенства доказываются аналогично.

Ясно, что $\sup_{m\geq n}(a_m+b_m)\leq\sup_{m\geq n}a_m+\sup_{m\geq n}b_m$, поэтому $\inf_n\sup_{m\geq n}(a_m+b_m)\leq\inf_n\sup_{m\geq n}a_m+\inf_n\sup_{m\geq n}b_m$, т.е.

$$\overline{\lim}_{n\to\infty}(a_n+b_n)\leq\overline{\lim}_{n\to\infty}a_n+\overline{\lim}_{n\to\infty}b_n.$$

Последовательность $c_n=\inf_{m\geq n}a_n$ сходится к $\underline{\lim}_{n\to\infty}a_n$, поэтому для любого $\varepsilon>0$ можно выбрать N так, что $a_k>\underline{\lim}_{n\to\infty}a_n-\varepsilon$ при $k>N$. Следовательно, $a_k+b_k>\underline{\lim}_{n\to\infty}a_n-\varepsilon+b_k$ при $k>N$. При добавлении константы к каждому члену последовательности к верхнему пределу добавляется та же константа, поэтому

$$\overline{\lim}_{n\to\infty}(a_n+b_n)\geq\underline{\lim}_{n\to\infty}a_n-\varepsilon+\overline{\lim}_{n\to\infty}b_n.$$

Это неравенство выполняется для всех $\varepsilon>0$, что и доказывает требуемое неравенство.

2.36. Согласно задаче 2.35

$$\underline{\lim}_{n\to\infty}a_n+\overline{\lim}_{n\to\infty}b_n\leq\overline{\lim}_{n\to\infty}(a_n+b_n)\leq\overline{\lim}_{n\to\infty}a_n+\overline{\lim}_{n\to\infty}b_n,$$

Кроме того, согласно задаче 2.34 $\overline{\lim}_{n\to\infty}a_n=\underline{\lim}_{n\to\infty}a_n=\lim_{n\to\infty}a_n$. Следовательно, $\overline{\lim}_{n\to\infty}(a_n+b_n)=\lim_{n\to\infty}a_n+\overline{\lim}_{n\to\infty}b_n$.

2.37. Предположим, что $\underline{\lim}_{n\to\infty}a_n<\overline{\lim}_{n\to\infty}a_n$. Выберем подпоследовательность a_{n_k}, сходящуюся к $\underline{\lim}_{n\to\infty}a_n$. Рассмотрим последовательность b_n, для которой $b_{n_k}=0$ и

$$b_n=\underline{\lim}_{n\to\infty}a_n-\overline{\lim}_{n\to\infty}a_n<0$$

в остальных случаях. Тогда $a_{n_k}+b_{n_k}=a_{n_k}$ и

$$a_n+b_n=a_n+\underline{\lim}_{n\to\infty}a_n-\overline{\lim}_{n\to\infty}a_n$$

в остальных случаях. Последовательность $a_n + b_n$ состоит из двух подпоследовательностей, верхний предел каждой из которых равен $\underline{\lim}_{n\to\infty} a_n$, поэтому

$$\overline{\lim}_{n\to\infty}(a_n + b_n) = \underline{\lim}_{n\to\infty} a_n.$$

Ясно также, что $\overline{\lim}_{n\to\infty} b_n = 0$. Поэтому из условия задачи следует, что $\underline{\lim}_{n\to\infty} a_n = \overline{\lim}_{n\to\infty} a_n$. Приходим к противоречию.

2.38. Пусть $\overline{\lim}_{n\to\infty} a_n = a < \infty$ и $A > a$. Тогда можно выбрать m так, что $a_n < A$ при $n > m$. При $n > m$ получаем:

$$\frac{a_1 + a_2 + \dots + a_n}{n} < \frac{a_1 + a_2 + \dots + a_m}{n} + \frac{(n-m)A}{n} \to A \quad \text{при} \quad n \to \infty.$$

Поэтому $\overline{\lim}_{n\to\infty} \frac{a_1+a_2+\dots+a_n}{n} \leq A$.

2.39. Предположим, что указанное неравенство не выполняется. Тогда можно выбрать m так, что при $n > m$ выполняется неравенство

$$\left(\frac{x_1 + x_{n+1}}{x_n}\right)^n < \left(\frac{1+n}{n}\right)^n,$$

т.е. $\frac{x_1}{n+1} < \frac{x_n}{n} - \frac{x_{n+1}}{n+1}$. Сложив такие неравенства для $n = m, m+1, \dots, m+k$, получим:

$$x_1\left(\frac{1}{m+1} + \frac{1}{m+2} + \dots + \frac{1}{m+k+1}\right) < \frac{x_m}{m} - \frac{x_{m+k+1}}{m+k+1} < \frac{x_m}{m}.$$

Устремляя k к бесконечности, приходим к противоречию, поскольку согласно задаче 2.5 $\frac{1}{m+1} + \frac{1}{m+2} + \dots + \frac{1}{m+k+1} \to \infty$ при фиксированном m и $k \to \infty$.

Глава 3.

Непрерывные функции

Мы будем использовать следующие обозначения:

$[a, b]$ — *отрезок*; он состоит из точек x, для которых $a \le x \le b$;

(a, b) — *интервал*; он состоит из точек x, для которых $a < x < b$;

$[a, b)$ или $(a, b]$ — *полуинтервал*; он состоит из точек x, для которых $a \le x < b$ или $a < x \le b$.

Функция f на отрезке (или интервале) сопоставляет каждой точке x этого отрезка некоторое число $f(x)$.

Функция f на отрезке (или интервале) *ограничена сверху*, если можно выбрать число M так, что $f(x) \le M$ для любой точки x этого отрезка. Функция f *ограничена снизу*, если можно выбрать число m так, что $f(x) \ge m$ для любой точки x этого отрезка. Функция f *ограниченная*, если она ограничена сверху и снизу.

Историческое замечание. Понятие функции формировалось очень долго. На начальном этапе развития анализа под функцией часто подразумевалось какое-то аналитическое выражение, задание функции определённой формулой. Термин *функция* ввёл Готфрид Вильгельм Лейбниц (1646-1716) в 1673 году. К 1697 году у него сформировалось понятие о функции как величины, зависящей от переменной. Современное определение функции как отображения предложил Петер Густав Лежён Дирихле (1805-1859) в 1837 году.

3.1. Предел функции

Число a называют *пределом* функции $f(x)$ в точке x_0, если для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что для любого $x \ne x_0$ из неравенства $|x - x_0| < \delta$ следует неравенство $|f(x) - a| < \varepsilon$. Если предел функции $f(x)$ в точке x_0 существует, то его обозначают $\lim_{x \to x_0} f(x)$.

Можно также определить пределы $\lim_{x\to\infty} f(x) = a$ и $\lim_{x\to-\infty} f(x) = a$. В первом случае для любого $\varepsilon > 0$ можно выбрать $N > 0$ так, что из неравенства $x > N$ следует неравенство $|f(x) - a| < \varepsilon$. Во втором случае для любого $\varepsilon > 0$ можно выбрать $N < 0$ так, что из неравенства $x < N$ следует неравенство $|f(x) - a| < \varepsilon$.

Иногда возникает необходимость рассматривать *односторонние пределы* функции $\lim_{x\to x_0+0} f(x) = f(x_0 + 0)$ и $\lim_{x\to x_0-0} f(x) = f(x_0 - 0)$. Они определяются почти так же, как и обычный предел, но в первом из них рассматриваются только $x > x_0$, а во втором — только $x < x_0$. Например, если функция $f(x)$ определена только на отрезке $[p, q]$, то в концах отрезка имеют смысл только односторонние пределы $\lim_{x\to p+0} f(x) = f(p + 0)$ и $\lim_{x\to q-0} f(x) = f(q - 0)$. Для краткости $0 + 0$ и $0 - 0$ часто обозначают $+0$ и -0.

Теорема 3.1. *Если предел функции $f(x)$ в точке x_0 существует, то он единствен.*

Доказательство. Предположим, что у функции $f(x)$ в точке x_0 есть два предела: a и b, причём $a \ne b$. Для $\varepsilon = \frac{1}{2}|a - b|$ выберем δ_1 и δ_2 так, что $|f(x) - a| < \varepsilon$ при $|x - x_0| < \delta_1$ и $|f(x) - b| < \varepsilon$ при $|x - x_0| < \delta_2$. Тогда если $x \ne x_0$ и $|x - x_0| < \min(\delta_1, \delta_2)$, то $|a - b| \le |f(x) - a| + |f(x) - b| < 2\varepsilon = |a - b|$. Приходим к противоречию. □

Теорема 3.2. *Пусть все точки a_n принадлежат области определения функции f. Тогда следующие следующие свойства эквивалентны:*

- $\lim_{x\to x_0} f(x) = a$;
- $\lim_{n\to\infty} f(a_n) = a$ *для любой последовательности* $\{a_n\}$, *для которой* $\lim_{n\to\infty} a_n = x_0$ *и* $a_n \neq x_0$ *при всех* n.

Доказательство. Предположим сначала, что $\lim_{x\to x_0} f(x) = a$. Для заданного $\varepsilon > 0$ выберем $\delta > 0$ так, что из неравенства $|x - x_0| < \delta$ следует неравенство $|f(x) - a| < \varepsilon$. Затем для данной последовательности $\{a_n\}$, сходящейся к x_0, выберем N так, что $|a_n - x_0| < \delta$ при $n > N$. Тогда $|f(a_n) - a| < \varepsilon$, поэтому $\lim_{n\to\infty} f(a_n) = a$.

Предположим теперь, что равенство $\lim_{x\to x_0} f(x) = a$ не имеет места. Тогда существует $\varepsilon > 0$, обладающее следующим свойством: для любого $\delta > 0$ существует $x \neq x_0$, для которого $|x - x_0| < \delta$, но $|f(x) - a| > \varepsilon$. Для $\delta = 1/n$ соответствующее x обозначим a_n. В результате получим последовательность $\{a_n\}$, для которой $\lim_{n\to\infty} a_n = x_0$, но равенство $\lim_{n\to\infty} f(a_n) = a$ не имеет места. □

Теорема 3.3. *Пусть $f(x)$ и $g(x)$ — две функции, причём* $\lim_{x\to x_0} f(x) = a$ *и* $\lim_{x\to x_0} g(x) = b$. *Тогда:*

а) $\lim_{x\to x_0}\big(f(x) + g(x)\big) = a + b$;
б) $\lim_{x\to x_0}\big(f(x)g(x)\big) = ab$;
в) $\lim_{x\to x_0}\big(f(x)/g(x)\big) = a/b$, *если* $b \neq 0$.

Доказательство. Согласно теореме 3.2 эти свойства пределов функций следуют из соответствующих свойств пределов последовательностей (теорема 2.2). □

Функцию $f(x)$ называют *возрастающей*, если $f(x) \geq f(y)$ при $x > y$, и *строго возрастающей*, если $f(x) > f(y)$ при $x > y$. Функцию $f(x)$ называют *убывающей*, если $f(x) \leq f(y)$ при $x > y$, и *строго убывающей*, если $f(x) < f(y)$ при $x > y$.

Функцию называют *монотонной*, если она возрастающая или убывающая. Функцию называют *строго монотонной*, если она строго возрастающая или строго убывающая.

Задача 3.1. Докажите, что для монотонной ограниченной функции существуют оба односторонних предела.

Функциональное уравнение вида $f(x + y) = f(x) + f(y)$ называют *уравнением Коши*.

Задача 3.2. Докажите, что если функция f, удовлетворяющая уравнению Коши, ограничена на некотором отрезке $[a, b]$, то $f(x) = cx$, где $c = f(1)$.

3.2. Непрерывность

Функцию $f(x)$ называют *непрерывной в точке* x_0, если $\lim_{x\to x_0} f(x) = f(x_0)$. Если функция $f(x)$ непрерывна в каждой точке (из области определения), то её называют просто *непрерывной*.

Функцию $f(x)$, определённую на отрезке $[a, b]$, мы будем называть непрерывной, если она непрерывна в каждой внутренней точке отрезка, а в концах отрезка существуют односторонние пределы, причём $f(a + 0) = f(a)$ и $f(b - 0) = f(b)$.

Задача 3.3. а) Пусть $P(x)$ — многочлен. Докажите, что функция $P(x)$ непрерывна.

б) Пусть $P(x)$ и $Q(x)$ — многочлены, причём $Q(x_0) \neq 0$. Докажите, что функция $P(x)/Q(x)$ непрерывна в точке x_0.

Теорема 3.4. *Пусть функция $f(x)$ непрерывна в точке x_0, а функция $g(y)$ непрерывна в точке $y_0 = f(x_0)$. Тогда функция $g\big(f(x)\big)$ непрерывна в точке x_0.*

Доказательство. Для данного $\varepsilon > 0$ выберем $\delta > 0$ так, что если $|y - y_0| < \delta$, то $|g(y) - g(y_0)| < \varepsilon$. Затем для $\varepsilon_1 = \delta$ выберем δ_1 так, что если $|x - x_0| < \delta_1$, то $|f(x) - f(x_1)| < \varepsilon_1$. Для данного $\varepsilon > 0$ мы выбрали $\delta_1 > 0$ так, что если $|x - x_0| < \delta_1$, то $\big|g\big(f(x)\big) - g\big(f(x_0)\big)\big| < \varepsilon$. □

Задача 3.4. Докажите, что $\sin\alpha < \alpha < \operatorname{tg}\alpha$ для $0 < \alpha < \pi/2$.

Задача 3.5. Докажите, что функция $f(x) = \sin x$ непрерывна.

Задача 3.6. Докажите, что $\lim_{\alpha\to 0} \frac{\sin\alpha}{\alpha} = 1$.

Задача 3.7. Пусть $f(x) = x \sin\frac{1}{x}$ при $x \neq 0$ и $f(0) = 0$. Докажите, что функция $f(x)$ непрерывна (во всех точках x).

Пусть f — ограниченная функция на отрезке $[a, b]$ и $x \in [a, b]$. Для каждого $\delta > 0$ рассмотрим множество точек отрезка $[a, b]$, расстояние от которых до точки x меньше δ, и пусть $\omega_f(x, \delta)$ — это разность между точной верхней и точной нижней гранями функции f на это множестве. *Колебание* функции f в точке x — это точная нижняя грань чисел $\omega_f(x, \delta)$ для всех $\delta > 0$; обозначим эту точную нижнюю грань $\omega_f(x)$. При фиксированных f и x функция $\omega(\delta) = \omega_f(x, \delta)$ возрастающая, поэтому $\omega_f(x) = \lim_{\delta\to+0} \omega_f(x, \delta)$. Ясно, что функция f непрерывна в точке x тогда и только тогда, когда $\omega_f(x) = 0$.

Задача 3.8. Докажите, что множество непрерывных функций на прямой имеет мощность континуума.

Задача 3.9. Докажете, что множество точек разрыва ограниченной монотонной функции f на отрезке не более чем счётно.

Задача 3.10. а) Докажите, что если функция f, удовлетворяющая уравнению Коши $f(x+y) = f(x) + f(y)$, непрерывна, то $f(x) = cx$, где $c = f(1)$ (Коши, 1821).

б) Докажите, что если функция f, удовлетворяющая уравнению Коши, непрерывна в некоторой точке x_0, то $f(x) = cx$, где $c = f(1)$ (Дарбу, 1875).

Замечание. Если на функцию f, удовлетворяющую уравнению Коши, не накладываются никакие ограничения, то она не обязана иметь вид $f(x) = cx$. Все функции, удовлетворяющие уравнению Коши, можно описать с помощью так называемого *базиса Гамеля*, построенного Георгом Гамелем (1877-1954) в 1905 году. Базис Гамеля — это базис векторного пространства $\mathbb{R}$ над полем $\mathbb{Q}$. Любое вещественное число $x \neq 0$ единственным образом представляется в виде конечной суммы $x = r_1\alpha_1 + \dots + r_n\alpha_n$, где $r_1, \dots, r_n$ — отличные от нуля рациональные числа, $\alpha_1, \dots, \alpha_n$ — некоторые из элементов базиса Гамеля. Тогда $f(x) = r_1 f(\alpha_1) + \dots + r_n f(\alpha_n)$, где значения функции на элементах базиса Гамеля задаются произвольно.

3.3. Теорема о промежуточном значении

Теорема 3.5 (о промежуточном значении). *Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и принимает в его концах значения разных знаков, то $f(x_0) = 0$ для некоторой точки x_0 этого отрезка.*

Доказательство. Пусть для определённости $f(a) > 0$ и $f(b) < 0$. Построим последовательности $\{a_n\}$ и $\{b_n\}$ следующим образом. Положим $a_1 = a$ и $b_1 = b$. Пусть c — середина отрезка $[a, b]$. Если $f(c) < 0$, то положим $a_2 = a_1$ и $b_2 = c$; если же $f(c) > 0$, то положим $a_2 = c$ и $b_2 = b_1$ (мы предполагаем, что $f(c) \neq 0$, поскольку иначе доказательство немедленно завершается). Затем берём середину отрезка $[a_2, b_2]$ и повторяем то же самое и т.д. В результате получим неубывающую последовательность $\{a_n\}$ и невозрастающую последовательность $\{b_n\}$. Эти последовательности ограничены и $\lim_{n\to\infty}(b_n - a_n) = 0$. Поэтому по теореме Вейерштрасса $\lim_{n\to\infty} a_n = x_0 = \lim_{n\to\infty} b_n$. По построению $f(a_n) \geq 0$ для всех n. Функция $f(x)$ непрерывна, поэтому $f(x_0) \geq 0$. С другой стороны, $f(b_n) \leq 0$, поэтому $f(x_0) \leq 0$. Значит, $f(x_0) = 0$. Ясно также, что $a \leq x_0 \leq b$. $\square$

Задача 3.11. а) Пусть f — непрерывная функция, для которой уравнение $f(x) = x$ не имеет вещественных решений. Докажите, что уравнение $f(f(x)) = x$ тоже не имеет вещественных решений.

б) Пусть f и g — непрерывные функции, удовлетворяющие тождеству $f(g(x)) = g(f(x))$. Докажите, что если уравнение $f(x) = g(x)$ не имеет вещественных решений, то уравнение $f(f(x)) = g(g(x))$ тоже не имеет вещественных решений.

Задача 3.12. Функция $f(x)$ непрерывна на отрезке $[0, 1]$ и принимает значения из того же отрезка. Докажите, что $f(x) = x$ для некоторой точки x этого отрезка.

Задача 3.13. Существует ли функция, непрерывная на отрезке $[0, 1]$, которая в рациональных точках принимает иррациональные значения, а в иррациональных — рациональные, и при этом все значения принадлежат отрезку $[0, 1]$?

Задача 3.14. *Непрерывная функция f на отрезке $[0, 1]$ принимает равные значения на его концах. Укажите все числа d, для которых для любой такой функции f существует отрезок длины d, на концах которого функция принимает равные значения.

3.4. Свойства функций, непрерывных на отрезке

Теорема 3.6. *Функция $f(x)$, непрерывная на отрезке $[a, b]$, ограничена на этом отрезке.*

Доказательство. Докажем, что функция $f(x)$ ограничена сверху (ограниченность снизу доказывается аналогично). Предположим, что для любого натурального n на отрезке $[a, b]$ есть точка x_n, для которой $f(x_n) > n$. Согласно теореме 2.6 из ограниченной последовательности $\{x_n\}$ можно выбрать сходящуюся подпоследовательность $\{x_{n_k}\}$. Пусть $\lim_{n\to\infty} x_{n_k} = x_0$ (ясно, что $a \le x_0 \le b$). Тогда $\lim_{n\to\infty} f(x_{n_k}) = f(x_0)$. Но это противоречит тому, что $f(x_{n_k}) > n_k \to \infty$. □

Историческое замечание. Теорему 3.6 Карл Вейерштрасс (1815-1897) доказывал на своих лекциях в 1861 году.

Задача 3.15. Функция непрерывна на интервале (a, b). Верно ли, что она ограничена на этом интервале?

Теорема 3.7 (Вейерштрасс). *Функция $f(x)$, непрерывная на отрезке $[a, b]$, достигает максимума и минимума в некоторых точках этого отрезка.*

Доказательство. Докажем, что на отрезке $[a, b]$ существует точка x_0, для которой $f(x_0) = M$, где M — точная верхняя грань множества чисел $f(x)$ для всех точек x отрезка $[a, b]$. (Для точной нижней грани доказательство аналогично.) Прежде всего заметим, что множество значений функции $f(x)$ ограничено (теорема 3.6), поэтому точная верхняя грань M существует (теорема 1.1). Следовательно, можно выбрать последовательность точек x_n отрезка $[a, b]$ так, что $M - \frac{1}{n} \le f(x_n) \le M$. Выберем из ограниченной последовательности $\{x_n\}$ сходящуюся подпоследовательность $\{x_{n_k}\}$. Если $\lim_{k\to\infty} x_{n_k} = x_0$, то $a \le x_0 \le b$ и $f(x_0) = M$. □

3.5. Логарифм и показательная функция

Определим сначала показательную функцию a^x для рациональных x. (Показательную функцию мы определяем только для положительных a.) Пусть $a > 0$ и $x = p/q$, где p и q — натуральные числа. Определим a^x как $\sqrt[q]{a^p}$, где имеется в виду арифметическое (положительное) значение корня. Ясно, что $\sqrt[nq]{a^{np}} = \sqrt[q]{a^p}$. Действительно, это равенство эквивалентно равенству $(a^{np})^q = (a^p)^{nq}$. При $x < 0$ мы полагаем $a^x = 1/a^{-x}$.

Задача 3.16. а) Пусть $a > 1$. Докажите, что если $x_1 > x_2$, то $a^{x_1} > a^{x_2}$.
б) Пусть $a < 1$. Докажите, что если $x_1 > x_2$, то $a^{x_1} < a^{x_2}$.

Задача 3.17. Пусть $a > 1$. Докажите, что a^x может быть сколь угодно велико, если x достаточно велико.

Задача 3.18. Докажите, что $a^{x_1+x_2} = a^{x_1}a^{x_2}$.

Задача 3.19. Докажите, что $a^{\lambda x} = (a^x)^\lambda$ для любого рационального λ.

Задача 3.20. Пусть a — положительное число, $\{x_n\}$ — последовательность рациональных чисел, причём $\lim_{n\to\infty} x_n = 0$. Докажите, что $\lim_{n\to\infty} a^{x_n} = 1$.

Задача 3.21. Докажите, что если существует предел $\lim_{n\to\infty} x_n = x$, где $\{x_n\}$ — последовательность рациональных чисел, то существует предел $\lim_{n\to\infty} a^{x_n}$, причём этот предел зависит только от x.

Пусть a — положительное число. Определим a^x для произвольного x следующим образом. Пусть $\{x_n\}$ — последовательность рациональных чисел, сходящаяся к x. Положим $a^x = \lim_{n\to\infty} a^{x_n}$.

Задача 3.22. а) Пусть $a > 1$. Докажите, что если $x > y$, то $a^x > a^y$.
б) Пусть $a < 1$. Докажите, что если $x > y$, то $a^x < a^y$.

Задача 3.23. Докажите, что если $\lim_{n\to\infty} x_n = x$, где $\{x_n\}$ — последовательность произвольных (не обязательно рациональных) чисел, то $\lim_{n\to\infty} a^{x_n} = a^x$.

Задача 3.24. Докажите, что $a^{x_1+x_2} = a^{x_1}a^{x_2}$ для произвольных (не обязательно рациональных) чисел x_1 и x_2.

Задача 3.25. Пусть a — положительное число, причём $a \neq 1$. Докажите, что для любого положительного числа x существует единственное число y, для которого $a^y = x$.

Пусть a и x — положительные числа, причём $a \neq 1$. *Логарифм* x по основанию a — это число $y = \log_a x$, для которого $a^y = x$. Для логарифмов по основанию 10 используется обозначение lg, а для логарифмов по основанию e используется обозначение ln.

Задача 3.26. Докажите, что функция $f(x) = \log_a x$ непрерывна.

Задача 3.27. Докажите, что $\log_a(x_1x_2) = \log_a x_1 + \log_a x_2$.

Задача 3.28. Докажите, что $\lim_{x\to\pm\infty} \left(1 + \frac{1}{x}\right)^x = e$.

Задача 3.29. Докажите, что $\lim_{x\to 0} \frac{\ln(1+x)}{x} = 1$.

Задача 3.30. Докажите, что $\lim_{n\to\infty} \left(1 + \frac{a}{n}\right)^n = e^a$ (Эйлер, 1748).

Задача 3.31. Докажите, что $\lim_{x\to 0} \frac{(1+x)^a - 1}{x} = a$ для любого вещественного a.

Задача 3.32. Докажите, что $\lim_{x\to 0} \frac{a^x - 1}{x} = \ln a$ для любого положительного a.

Задача 3.33. Пусть $\{a_n\}$ — последовательность положительных чисел, причём $\lim_{n\to\infty} a_n = a$. Докажите, что $\lim_{n\to\infty} \sqrt[n]{a_1 \ldots a_n} = a$.

Задача 3.34. Докажите, что если $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = a$, то $\lim_{n\to\infty} \sqrt[n]{a_n} = a$.

Задача 3.35. Докажите, что $\lim_{n\to\infty} \frac{\sqrt[n]{n!}}{n} = \frac{1}{e}$.

Задача 3.36. Вычислите $\lim_{n\to\infty} \left(\frac{\sqrt[n]{a} + \sqrt[n]{b}}{2}\right)^n$.

Задача 3.37. Непрерывная функция $f(x)$ определена для всех x, причём $f(x_0) \neq 0$ для некоторого x_0. Докажите, что если выполнено соотношение

$$f(x + y) = f(x)f(y),$$

то $f(x) = a^x$ для некоторого $a > 0$.

Задача 3.38. Непрерывная функция $f(x)$ определена для всех $x > 0$ и удовлетворяет соотношению

$$f(xy) = f(x) + f(y).$$

а) Докажите, что если $f(a) = 1$ для некоторого a, то $f(x) = \log_a x$.
б) Докажите, что если $f(x_0) \neq 0$ для некоторого x_0, то $f(a) = 1$ для некоторого a.

3.6. Гиперболические функции

С помощью показательной функции e^x можно определить *гиперболические функции*, которые параметризуют гиперболу $x^2 - y^2 = 1$ и играют для неё такую же роль, какую играют для окружности тригонометрические функции.

$\operatorname{sh} x = \frac{e^x - e^{-x}}{2}$ (*гиперболический синус*);

$\operatorname{ch} x = \frac{e^x + e^{-x}}{2}$ (*гиперболический косинус*);

$\operatorname{th} x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ (*гиперболический тангенс*);

$\operatorname{cth} x = \frac{e^x + e^{-x}}{e^x - e^{-x}}$ (*гиперболический котангенс*).

Очевидно, что $\operatorname{sh}(-x) = -\operatorname{sh} x$ и $\operatorname{ch}(-x) = \operatorname{ch} x$.

Несложно проверить следующие формулы для производных гиперболических функций:

$$(\operatorname{sh} x)' = \operatorname{ch} x, \quad (\operatorname{ch} x)' = \operatorname{sh} x, \quad (\operatorname{th} x)' = \frac{1}{\operatorname{ch}^2 x}, \quad (\operatorname{cth} x)' = \frac{1}{\operatorname{sh}^2 x}.$$

Задача 3.39. Докажите, что точка с координатами $x = \operatorname{ch} t$, $y = \operatorname{sh} t$ лежит на гиперболе $x^2 - y^2 = 1$.

Задача 3.40. Докажите, что

$$\operatorname{sh}(x \pm y) = \operatorname{sh} x \operatorname{ch} y \pm \operatorname{ch} x \operatorname{sh} y,$$
$$\operatorname{ch}(x \pm y) = \operatorname{ch} x \operatorname{ch} y \pm \operatorname{sh} x \operatorname{sh} y.$$

Обратные гиперболические функции определяются следующим образом:

если $x = \operatorname{sh} y$, то $y = \operatorname{Arsh} x$ (*ареасинус*[1] *гиперболический*);

если $x = \operatorname{ch} y$, то $y = \operatorname{Arch} x$ (*ареакосинус гиперболический*) ;

если $x = \operatorname{th} y$, то $y = \operatorname{Arth} x$ (*ареатангенс гиперболический*) ;

если $x = \operatorname{cth} y$, то $y = \operatorname{Arcth} x$ (*ареакотангенс гиперболический*) .

Задача 3.41. Докажите, что

$$\operatorname{Arsh} x = \ln(x + \sqrt{x^2 + 1});$$
$$\operatorname{Arch} x = \ln(x \pm \sqrt{x^2 - 1}) = \pm \ln(x + \sqrt{x^2 - 1});$$
$$\operatorname{Arth} x = \frac{1}{2} \ln \frac{1 + x}{1 - x}.$$

Гиперболической амплитудой числа x называют угол α $(-\pi/2 < \alpha < \pi/2)$, для которого $\operatorname{sh} x = \operatorname{tg} \alpha$.

Задача 3.42. Докажите следующие свойства гиперболической амплитуды:

а) $\operatorname{ch} x = 1/\cos\alpha$;

б) $\operatorname{th}(x/2) = \operatorname{tg}(\alpha/2)$.

Задача 3.43. *Докажите, что если $0 < y_0 < x_0$ и $x_{n+1} = \frac{x_n + y_n}{2}$ и $y_{n+1} = \sqrt{x_{n+1} y_n}$ при $n \geq 0$, то

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \frac{\sqrt{x_0^2 - y_0^2}}{\operatorname{Arch}(x_0/y_0)} \quad \text{(Борхардт).}$$

Историческое замечание. Гиперболические функции независимо ввели и исследовали в 1760-е годы Винченцо Риккати (1707-1775) и Иоганн Генрих Ламберт (1728-1777).

[1] От латинского area — площадь.

3.7. Равномерная непрерывность. Равномерная сходимость

Пусть множество D содержится в области определения функции $f(x)$. Функцию $f(x)$ называют *равномерно непрерывной* на множестве D, если для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что для любых x_1 и x_2 из D неравенство $|x_1 - x_2| < \delta$ влечёт неравенство $|f(x_1) - f(x_2)| < \varepsilon$. Отличие от обычной непрерывности заключается в том, что число δ одно и то же для всех точек области D. Ясно, что из равномерной непрерывности следует обычная непрерывность. Обратное, вообще говоря, неверно.

Теорема 3.8. *Любая функция $f(x)$, непрерывная на отрезке $[a, b]$, является равномерно непрерывной на этом отрезке.*

Доказательство. Предположим, что функция $f(x)$ на отрезке $[a, b]$ является непрерывной, но не равномерно непрерывной. Тогда существует такое положительное число ε_0, что для любого $\delta > 0$ найдутся точки $x'(\delta)$ и $x''(\delta)$ из отрезка $[a, b]$, обладающие следующими свойствами: $|x'(\delta) - x''(\delta)| < \delta$ и $|f(x'(\delta)) - f(x''(\delta))| \geq \varepsilon_0$. Положим $x'_n = x'(1/n)$ и $x''_n = x''(1/n)$. Из ограниченной последовательности $\{x'_n\}$ можно выбрать подпоследовательность $\{x'_{n_k}\}$, сходящуюся к некоторой точке x_0 отрезка $[a, b]$ (задача 2.6). Неравенство $|x'_n - x''_n| < \frac{1}{n}$ показывает, что подпоследовательность $\{x''_{n_k}\}$ тоже сходится к точке x_0. Из непрерывности функции $f(x)$ следует, что $f(x'_{n_k}) \to f(x_0)$ и $f(x''_{n_k}) \to f(x_0)$. Но это противоречит неравенству $|f(x'(\delta)) - f(x''(\delta))| \geq \varepsilon_0$. $\square$

Историческое замечание. Понятие равномерной непрерывности первым опубликовал Генрих Эдуард Гейне (1821-1881) в 1870 году; он же в 1872 году доказал теорему 3.8. Но уже ранее, в 1854 году, Петер Густав Лежён Дирихле (1805-1859) доказывал эту теорему на своих лекциях. Понятие равномерной непрерывности упоминалось и на лекциях Карла Вейерштрасса (1815-1897) в 1861 году.

Задача 3.44. а) Приведите пример непрерывной функции на множестве всех действительных чисел, которая не является равномерно непрерывной.

б) Приведите пример непрерывной функции на интервале $(0, 1)$, которая не является равномерно непрерывной.

Последовательность функций $f_n(x)$, заданных на некотором множестве D, называют *сходящейся*, если для каждой точки $x_0 \in D$ последовательность $f_n(x_0)$ сходится; предел этой последовательности обозначим $f(x_0)$.

Пример. *Последовательность непрерывных функций $f_n(x) = x^n$ на полуинтервале $(-1, 1]$ сходится к разрывной функции, равной 0 при $-1 < x < 1$ и 1 при $x = 1$.*

Последовательность функций $f_n(x)$ называют *равномерно ограниченной*, если все они ограничены одной и той же константой C, т.е. $|f_n(x)| \leq C$ для всех n и всех x.

Последовательность функций $f_n(x)$ называют *равномерно сходящейся* к предельной функции $f(x)$, если для любого $\varepsilon > 0$ существует такой номер N, что для любой точки $x_0 \in D$ при $n > N$ выполняется неравенство $|f_n(x_0) - f(x_0)| < \varepsilon$.

Историческое замечание. Понятие равномерной сходимости последовательности функций ввёл Карл Вейерштрасс (1815-1897) в 1841 году в письме, опубликованном только в 1894 году. Он указал также на то, что равномерно сходящиеся последовательности функций можно почленно дифференцировать и интегрировать (см. теоремы 5.13 и 6.11).

Задача 3.45. Докажите, что если последовательность функций $f_n(x)$, каждая из которых ограничена, сходится равномерно, то предельная функция ограничена.

Теорема 3.9. *Если последовательность функций $f_n(x)$, каждая из которых непрерывна в точке x_0, сходится равномерно, то предельная функция $f(x)$ непрерывна в точке x_0.*

Доказательство. Воспользовавшись равномерной сходимостью, для данного $\varepsilon > 0$ выберем N так, что

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \tag{3.1}$$

при $n \geq N$ для всех x.

Воспользовавшись непрерывностью функции $f_N(x)$ в точке x_0, выберем $\delta > 0$ так, что $|f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3}$ при $|x - x_0| < \delta$.

Следовательно, если $|x - x_0| < \delta$, то

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Здесь мы дважды воспользовались неравенством (3.1) при $n = N$: для точки x и для точки x_0. □

Задача 3.46. Докажите, что на любом отрезке последовательность функций $f_n(x) = \left(1 + \frac{x}{n}\right)^n$ равномерно сходится к функции e^x.

Теорема 3.10 (Дини). *Пусть последовательность функций f_n, непрерывных на отрезке $I = [a, b]$, возрастающая (т.е. $f_n(x) \leq f_{n+1}(x)$ для любой точки $x \in I$) и поточечно сходится на I к непрерывной функции f. Тогда эта сходимость равномерная.*

Доказательство. Пусть задано число $\varepsilon > 0$. Положим $g_n = f - f_n$ и для каждого n рассмотрим множество $E_n = \{x \in I \,|\, g_n(x) < \varepsilon\}$. Это множество открытое (оно является прообразом открытого множества при непрерывном отображении g_n). Последовательность функций g_n убывающая, поэтому последовательность множеств E_n возрастающая. Последовательность f_n поточечно сходится к f, поэтому каждая точка $x \in I$ содержится в некотором множестве E_n, т.е. множества E_n образуют открытое покрытие отрезка. Из этого покрытия можно выбрать конечное подпокрытие, поэтому $E_N = I$ для некоторого N. Если $n > N$, то для любой точки $x \in I$ выполняется неравенство $|f(x) - f_n(x)| < \varepsilon$. □

Задача 3.47. Докажите, что в теореме Дини нельзя отбросить ни одно из трёх условий: 1) функции f_n непрерывны; 2) функция f непрерывная; 3) область сходимости компактна (отрезок).

Историческое замечание. Улисс Дини (1845-1918) доказал теорему о равномерной сходимости возрастающей последовательности непрерывных функций в 1878 году.

3.8. Липшицевы функции и теорема о неподвижной точке

Функция f удовлетворяет *условию Липшица* с константой L, если $|f(x) - f(y)| \leq L|x - y|$ для всех x, y из области определения функции. Функцию, удовлетворяющую условию Липшица, называют *липшицевой*. Любая липшицева функция равномерно непрерывна и, в частности, непрерывна. Липшицеву функцию с константой $L < 1$ называют *сжимающим отображение*. Сжимающее отображение уменьшает расстояния между точками.

Точку x называют *неподвижной точкой* отображения f, если $f(x) = x$.

Теорема 3.11 (о неподвижной точке). *Сжимающее отображение отрезка в себя имеет единственную неподвижную точку.*

Доказательство. Пусть f — сжимающее отображение отрезка I в себя, $L < 1$ — константа Липшица. Выберем произвольную точку $x_0 \in I$ и положим $x_n = f(x_{n-1})$ для $n \geq 1$. Покажем, что x_n — последовательность Коши. Действительно, пусть $1 \leq m < n$. Тогда

$$\begin{aligned} |x_n - x_m| &< |x_n - x_{n-1}| + |x_{n-1} - x_{n-2}| + \cdots + |x_{m+1} - x_m| = \\ &= |f(x_{n-1}) - f(x_{n-2})| + \cdots + |f(x_m) - f(x_{m-1})|. \end{aligned}$$

Далее, $|f(x_{n-1}) - f(x_{n-2})| \leq L|x_{n-1} - x_{n-2}| \leq L^2|x_{n-2} - x_{n-3}| \leq \cdots \leq L^{n-1}|x_1 - x_0|$, $\ldots$, $|f(x_m) - f(x_{m-1})| \leq L^m|x_1 - x_0|$. Таким образом,

$$|x_n - x_m| \leq (L^{n-1} + L^{n-3} + \cdots + L^m) \cdot |x_1 - x_0| \leq \frac{L^m}{1 - L}|x_1 - x_0|,$$

поскольку $L < 1$ по условию.

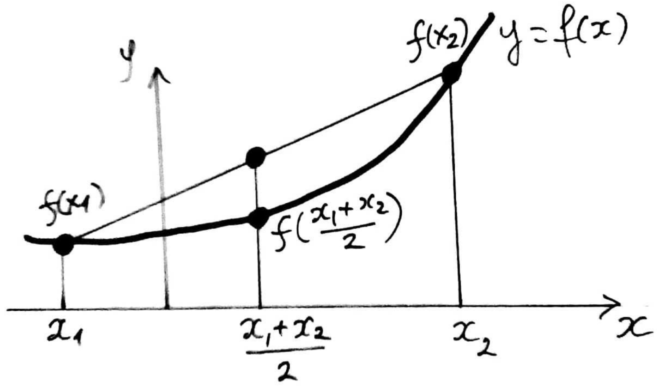

Рис. 3.1.

Последовательность Коши x_n сходится к некоторой точке ξ, причём $\xi \in I$, поскольку отрезок I замкнут. Функция f непрерывна, поэтому

$$f(\xi) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = \lim_{m\to\infty} x_m = \xi,$$

т.е. ξ — неподвижная точка.

Докажем теперь единственность неподвижной точки. Предположим, что $\xi_1 = f(\xi_1)$ и $\xi_2 = f(\xi_2)$. Тогда

$$|\xi_1 - \xi_2| = |f(\xi_1) - f(\xi_2)| \le L|\xi_1 - \xi_2|.$$

По условию $L < 1$, поэтому $\xi_1 = \xi_2$. □

Историческое замечание. Условия Липшица ввёл Рудольф Липшиц (1832-1903) в 1868 году при доказательстве существования решения системы дифференциальных уравнений. Теорему о неподвижной точке для сжимающих отображений доказал Стефан Банах (1892-1945) в 1922 году.

3.9. Выпуклые функции

Функцию $f(x)$, определённую на отрезке $[a, b]$, называют *выпуклой*, если

$$f\left(\frac{x_1 + x_2}{2}\right) \le \frac{f(x_1) + f(x_2)}{2} \qquad (3.2)$$

для всех x_1 и x_2, лежащих на отрезке $[a, b]$.

С геометрической точки зрения неравенство (3.2) означает, что середина любой хорды кривой $y = f(x)$ лежит над этой кривой (рис. 3.1) или на самой кривой.

Функцию $f(x)$ называют *вогнутой*, если функция $-f(x)$ выпукла.

Теорема 3.12. *Функция $f(x)$ выпукла на отрезке $[a, b]$ тогда и только тогда, когда для любого n и любых точек $x_1, \ldots, x_n$ из этого отрезка имеет место неравенство*

$$f\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \le \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n}. \qquad (3.3)$$

Доказательство. Достаточно доказать, что из неравенства (3.2) следует неравенство (3.3) для любого n. Если выполняется неравенство (3.3), то

$$\begin{aligned} 4f\left(\frac{x_1 + x_2 + x_3 + x_4}{4}\right) &\le 2f\left(\frac{x_1 + x_2}{2}\right) + 2f\left(\frac{x_3 + x_4}{2}\right) \le \\ &\le f(x_1) + f(x_2) + f(x_3) + f(x_4). \end{aligned}$$

Аналогично доказывается, что неравенство выполняется для любого $n = 2^m$.

Остаётся доказать, что если неравенство выполняется для n, то оно выполняется и для $n-1$. Пусть даны числа $x_1, \ldots, x_{n-1}$. Положим $x_n = \frac{1}{n-1}(x_1 + \cdots + x_{n-1})$. Тогда

$$f(x_n) = f\left(\frac{(n-1)x_n + x_n}{n}\right) = f\left(\frac{x_1 + \cdots + x_{n-1} + x_n}{n}\right) \leq$$
$$\leq \frac{f(x_1) + \cdots + f(x_{n-1}) + f(x_n)}{n}.$$

Следовательно, $f(x_n) \leq \frac{f(x_1)+\cdots+f(x_{n-1})}{n-1}$, что и требовалось. □

Теорема 3.13. *Пусть $f(x)$ — выпуклая функция, p и q — положительные числа, сумма которых равна 1.*

а) Если числа p и q рациональные, то $f(px_1 + qx_2) \leq pf(x_1) + qf(x_2)$.

б) Если функция $f(x)$ непрерывная, то $f(px_1 + qx_2) \leq pf(x_1) + qf(x_2)$.

Доказательство. а) Если числа p и q рациональные, то $p = \frac{m}{n}$ и $q = \frac{n-m}{n}$, где m, n и $n-m$ — натуральные числа. Применим неравенство из теоремы 3.12 для чисел $x_1, \ldots, x_1, x_2, \ldots, x_2$ (m чисел x_1 и $n-m$ чисел x_2). В результате получим требуемое.

б) Следует из а), поскольку любое число является пределом последовательности рациональных чисел. □

Теорема 3.14 (неравенство Йенсена). *Пусть $f(x)$ — выпуклая функция, $\alpha_1, \ldots, \alpha_n$ — положительные числа, сумма которых равна 1.*

а) Если числа $\alpha_1, \ldots, \alpha_n$ рациональные, то

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n).$$

б) Если функция $f(x)$ непрерывная, то

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n).$$

Доказательство. Доказательство аналогично доказательству теоремы 3.13. □

Историческое замечание. Неравенство из теоремы 3.14 доказал Отто Людвиг Гёльдер (1859-1937) в 1889 году. Йоган Йенсен (1859-1925) в 1906 году доказал это неравенство в интегральной форме, когда α — непрерывная функция, а не конечный набор значений.

Задача 3.48. Докажите, что функция $f(x) = \ln x$ вогнута на интервале $(0, +\infty)$.

Задача 3.49. Докажите неравенство между средним арифметическим и средним геометрическим для положительных чисел $x_1, \ldots, x_n$:

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \cdot \ldots \cdot x_n}.$$

Задача 3.50. Докажите, что если A, B, p и q — положительные числа, причём $\frac{1}{p} + \frac{1}{q} = 1$, то $A^{1/p}B^{1/q} \leq \frac{A}{p} + \frac{B}{q}$.

Задача 3.51. Докажите, что непрерывная функция $f(x)$ выпуклая тогда и только тогда, когда для любых трёх точек $x_1 < x_2 < x_3$ выполняется неравенство

$$\begin{vmatrix} x_1 & f(x_1) & 1 \\ x_2 & f(x_2) & 1 \\ x_3 & f(x_3) & 1 \end{vmatrix} \geq 0.$$

3.10. Функции ограниченной вариации

Пусть f — ограниченная функция на отрезке $[a, b]$. ***Вариация*** $\mathrm{Var}_a^b(f)$ функции f на отрезке $[a, b]$ определяется как точная верхняя грань сумм вида $\sum_{k=0}^{m}|f(x_{k+1})-f(x_k)|$ для всевозможных наборов точек $x_0 = a < x_1 < \dots < x_m < b = x_{m+1}$. Если $\mathrm{Var}_a^b(f) < \infty$, то говорят, что f — ***функция ограниченной вариации*** на отрезке $[a, b]$.

Задача 3.52. Пусть $f(x) = x\sin\frac{1}{x}$ при $x \neq 0$, $f(0) = 0$. Докажите, что на отрезке $[0, 1]$ функция f непрерывна, но не является функцией ограниченной вариации.

Задача 3.53. Докажите, что если функции $f(x)$ и $g(x)$ имеют ограниченную вариацию на отрезке $[a, b]$, то функция $f(x)g(x)$ тоже имеет ограниченную вариацию на этом отрезке.

Задача 3.54. Докажите, что функция f, определённая на отрезке $[a, b]$, является функцией ограниченной вариации тогда и только тогда, когда её можно представить в виде разности двух ограниченных возрастающих функций.

Задача 3.55. Докажите, что для функции ограниченной вариации существуют оба односторонних предела.

Задача 3.56. Пусть f — функция ограниченной вариации на отрезке $[a, b]$. Для $x \in [a, b]$ рассмотрим функцию $V(x) = \mathrm{Var}_a^x(f)$ (предполагается, то $V(a) = 0$). Докажите, что функция f непрерывна в точке $x_0 \in [a, b]$ тогда и только тогда, когда в этой точке непрерывна функция V.

Задача 3.57. Докажите, что непрерывную функцию $f(x)$ ограниченной вариации можно представить в виде разности двух непрерывных неубывающих функций.

Историческое замечание. Понятие функции ограниченной вариации ввёл Камилл Жордан (1838-1922) в 1881 году в связи с исследованием функций, разложимых в ряд Фурье. В 1887 году Жордан обнаружил связь функций ограниченной вариации со спрямляемыми кривыми.

Функцию $f(x)$ на отрезке $[a, b]$ называют *абсолютно непрерывной*, если для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что для любой конечной системы неперекрывающихся интервалов (a_1, b_1), …, (a_n, b_n), сумма длин которых меньше δ, выполняется неравенство $\sum_{k=1}^{n}|f(a_k) - f(b_k)| < \varepsilon$.

Задача 3.58. Докажите, что абсолютно непрерывная функция на отрезке $[a, b]$ является функцией ограниченной вариации.

3.11. Решения задач

3.1. Докажем, например, что для ограниченной возрастающей функции $f(x)$ существует предел $\lim_{x\to x_0+0} f(x)$. Рассмотрим множество значений $f(y)$ для всех $y > x_0$. У этого множества есть точная нижняя грань a. Поэтому для любого $\varepsilon > 0$ можно выбрать $y_0 > x_0$ так, что $0 \leq f(y_0) < a+\varepsilon$. Положим $\delta = y_0 - x_0 > 0$. Тогда для всех x, удовлетворяющих неравенствам $x_0 < x < x_0 + \delta = y_0$, выполняются неравенства $a \leq f(x) \leq f(x_0) < a + \varepsilon$. Поэтому $\lim_{x\to x_0+0} f(x) = a$.

3.2. По условию $f(0) = f(0) + f(0)$, поэтому $f(0) = 0$. Кроме того, $f(x) + f(-x) = f(0)$, поэтому $f(-x) = -f(x)$. Если n — натуральное число, то $f(nx) = f(x + \dots + x) = nf(x)$ и $f(\frac{x}{n}) = \frac{nf(x/n)}{n} = \frac{f(x)}{n}$. В результате получаем, что если r — рациональное число, то $f(rx) = rf(x)$.

Выберем натуральное число $m > \frac{1}{b-a}$. Тогда любое число x можно представить в виде $x = m\alpha+n$, где число n целое и $\alpha \in [a, b]$. Поэтому

$$\lim_{x\to\infty}\frac{f(x)}{x} = \lim_{n\to\infty}\frac{f(m\alpha+n)}{m\alpha+n} = \lim_{n\to\infty}\frac{mf(\alpha)+nf(1)}{m\alpha+n} = f(1),$$

поскольку множества всех чисел вида $mf(\alpha)$ и $m\alpha$ ограничены.

Теперь для произвольного положительного числа x можно воспользоваться тем, что $\frac{f(x)}{x} = \frac{nf(x)}{nx} = \frac{f(nx)}{nx} \to f(1)$ при $n \to \infty$.

3.3. Воспользуемся теоремой 3.3. Ясно, что $\lim_{x\to x_0} x = x_0$, поэтому, применив индукцию по n, получим $\lim_{x\to x_0} x^n = x_0^n$. Значит,

$$\lim_{x\to x_0}(a_n x^n + a_{n-1}x^{n-1} + \dots + a_0) = a_n x_0^n + a_{n-1}x_0^{n-1} + \dots + a_0.$$

Если $Q(x_0) \neq 0$, то $\lim_{x\to x_0} P(x)/Q(x) = P(x_0)/Q(x_0)$.

3.4. Пусть O — центр окружности радиуса 1, A и B — точки этой окружности, для которых $\angle AOB = \alpha$. Рассмотрим также точку C, в которой касательная к окружности в точке B пересекает луч OA. Пусть S_{AOB} — площадь треугольника AOB, S — площадь сектора, высекаемого радиусами AO и OB, S_{BOC} — площадь треугольника BOC. Тогда $S_{AOB} < S < S_{BOC}$. Но $S_{AOB} = \frac{1}{2}\sin\alpha$, $S = \frac{1}{2}\alpha$ и $S_{BOC} = \frac{1}{2}\operatorname{tg}\alpha$.

3.5. Ясно, что

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

Поэтому достаточно доказать, что $\lim_{t\to 0}\sin t = 0$. Но если $|t| < \pi/2$, то $|\sin t| < |t|$ согласно задаче 3.4.

3.6. Согласно задаче 3.4 $\cos\alpha < \frac{\sin\alpha}{\alpha} < 1$ при $0 < \alpha < \pi/2$. Такие же неравенства верны и при $-\pi/2 < \alpha < 0$, поскольку $\cos(-\alpha) = \cos\alpha$ и $\frac{\sin(-\alpha)}{-\alpha} = \frac{\sin\alpha}{\alpha}$. Остаётся заметить, что $\lim_{\alpha\to 0}\cos\alpha = \cos 0 = 1$.

3.7. Ясно, что функция $f(x)$ непрерывна во всех точках $x \neq 0$. Проверим, что функция $f(x)$ непрерывна в точке $x = 0$, т.е. для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что если $|x| < \delta$, то $|f(x)| < \varepsilon$. Но $|f(x)| = |x\sin\frac{1}{x}| \le |x|$. Поэтому можно положить $\delta = \varepsilon$.

3.8. Непрерывная функция задаётся значениями в рациональных точках, поэтому непрерывные функции находятся во взаимно однозначном соответствии с некоторым подмножеством множества $\mathbb{Q}\times\mathbb{R}$, которое имеет мощность континуума согласно задаче 1.16. Поэтому множество непрерывных функций на прямой имеет мощность не больше континуума. Множество непрерывных функций на прямой содержит все постоянные функции, поэтому оно имеет мощность не меньше континуума.

3.9. Пусть для определённости функция f возрастающая. Для монотонной ограниченной функции существуют оба односторонних предела (задача 3.1), поэтому в точке разрыва функции f определено число $f(x_0+0) - f(x_0-0) > 0$ — скачок функции f в точке x_0. Из ограниченности функции f следует, что для любого $\varepsilon > 0$ количество точек разрыва, в которых скачок больше ε, конечно. Пусть H_k — множество точек, в которых скачок больше $1/k$. Множество всех точек разрыва содержится в счётном объединении конечных множеств H_k, $k = 1, 2, \dots$

3.10. а) При решении задачи 3.2 доказано, что если r — рациональное число, то $f(rx) = rf(x)$. В частности, для любого рационального числа r выполняется равенство $f(r) = rf(1) = cr$, где $c = f(1)$. Для любого числа x можно выбрать последовательность рациональных чисел $\{r_n\}$, сходящуюся к x. Из непрерывности функции f следует, что $cr_n = f(r_n) \to f(x)$, поэтому $f(x) = cx$ для любого числа x.

б) Покажем, что функция f непрерывна в любой точке y_0. Рассмотрим последовательность точек $y_n \to y_0$. Тогда $y_n - y_0 + x_0 \to x_0$. Поэтому $f(y_n - y_0 + x_0) \to f(x_0)$. Учитывая, что $f(y_n - y_0 + x_0) = f(y_n) - f(y_0) + f(x_0)$, получаем $f(y_n) - f(y_0) + f(x_0) \to f(x_0)$, т.е. $f(y_n) \to f(y_0)$.

3.11. а) Если f — непрерывная функция и уравнение $f(x) = x$ не имеет вещественных решений, то либо $f(x) > x$ для всех x, либо $f(x) < x$ для всех x. В первом случае $f(f(x)) < f(x) < x$, а во втором случае $f(f(x)) > f(x) > x$.

б) Если f и g — непрерывные функции и уравнение $f(x) = g(x)$ не имеет вещественных решений, то либо $f(x) > g(x)$ для всех x, либо $f(x) < g(x)$ для всех x. В первом случае $f(f(x)) > g(f(x)) = f(g(x)) > g(g(x))$. Во втором случае $f(f(x)) < g(g(x))$.

3.12. Рассмотрим вспомогательную функцию $\varphi(x) = f(x) - x$. Ясно, что $\varphi(0) = f(0) \ge 0$ и $\varphi(1) = f(1) - 1 \le 0$. Поэтому $\varphi(x) = 0$ для некоторой точки x отрезка $[0,1]$. В таком случае $f(x) = x$.

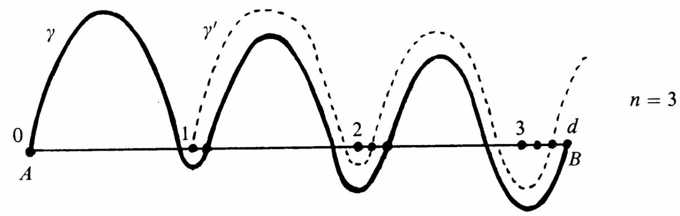

Рис. 3.2. На рисунке нужно заменить d на $1/d$.

3.13. Ответ: нет, не существует. Действительно, если f — требуемая функция, то $f(x) \neq x$ для любой точки x отрезка $[0,1]$. Но согласно задаче 3.12 таких непрерывных функций нет.

3.14. Ответ: $d = \frac{1}{n}$, где n — натуральное число.

Покажем сначала, что если $0 < a < 1$, то функция f принимает равные значения на концах отрезка, длина которого равна a или $1 - a$. Предположим, что на концах всех отрезков длины a и $1 - a$ функция f принимает разные значения. Тогда при сдвиге графика функции $f(x)$ вдоль оси Ox на $\pm a$ или на $\pm(1 - a)$ получается график, не пересекающий исходный график. Сдвинем график функции $f(x)$ на a и на $a - 1$. В результате получим график функции $g(x)$ на отрезке $[a - 1, a + 1]$. Функция g непрерывна: разрыв мог бы быть в точке a, но его нет, поскольку значения функции f на концах отрезка $[0,1]$ равны.

Пусть функция f принимает наибольшее значение в точке $x_{\max}$ и наименьшее значение в точке $x_{\min}$. Функции f и g непрерывны на отрезке с концами $x_{\max}$ и $x_{\min}$, $f(x_{\max}) > g(x_{\max})$ и $f(x_{\min}) < g(x_{\min})$. Полученное противоречие показывает, что функция f принимает равные значения на концах отрезка, длина которого равна a или $1 - a$.

Пусть $d = \frac{1}{n}$, где n — натуральное число. Тогда функция f принимает равные значения на концах отрезка длины $\frac{1}{n}$ или $1 - \frac{1}{n}$. Если функция f принимает равные значения на концах отрезка длины $1 - \frac{1}{n} = \frac{n-1}{n}$, то она принимает равные значения на концах отрезка длины $\frac{1}{n}$ или $\frac{n-1}{n} - \frac{1}{n} = \frac{n-2}{n}$. Продолжая такие рассуждения, получаем, что функция f принимает равные значения на концах отрезка длины $\frac{1}{n}$.

Предположим теперь, что число $\frac{1}{d} > 1$ не целое и n — целая часть этого числа. Построим на отрезке длины $\frac{1}{d}$ непрерывную функцию, принимающую на его концах равные значения и не принимающую равные значения на концах отрезка длины 1. (Из неё легко получить функцию на отрезке длины 1, не принимающую равные значения на концах отрезка длины $\frac{1}{d}$.) Разделим отрезок $[n, \frac{1}{d}]$ на n равных частей. Для $n = 3$ построение требуемого примера изображено на рис. 3.2.

3.15. Ответ: нет, не верно. Функция $f(x) = 1/x$ непрерывна на интервале $(0,1)$, но она не ограничена сверху.

3.16. а) Можно считать, что $x_1 = p_1/q$ и $x_2 = p_2/q$. Тогда $a^{p_1} > a^{p_2}$, поскольку $a > 1$. Для положительных чисел α и β неравенство $\alpha > \beta$ эквивалентно неравенству $\alpha^q > \beta^q$. Поэтому $a^{p_1/q} > a^{p_2/q}$.

б) Решение аналогично, но в этом случае $a^{p_1} < a^{p_2}$, поскольку $a < 1$.

3.17. Положим $a = 1 + \delta$, где $\delta > 0$. Тогда

$$(1+\delta)^n = 1 + n\delta + \frac{n(n-1)}{2}\delta^2 + \dots > n\delta.$$

Поэтому если $n > y/\delta$, то $a^n > y$.

3.18. Можно считать, что $x_1 = p_1/q$ и $x_2 = p_2/q$. Ясно, что $a^{p_1+p_2} = a^{p_1}a^{p_2}$. Поэтому $a^{x_1+x_2} = \sqrt[q]{a^{p_1+p_2}} = \sqrt[q]{a^{p_1}} \cdot \sqrt[q]{a^{p_2}} = a^{x_1}a^{x_2}$.

3.19. Для натурального λ это следует из задачи 3.18. Если $\lambda = p/q$, то достаточно извлечь корень степени q из обеих частей равенства $a^{px} = (a^x)^p$.

3.20. Рассмотрим сначала следующий частный случай: $x_n = 1/n$. В этом случае требуемое утверждение доказано в решении задачи 2.13.

Для исходной последовательности $\{x_n\}$ можно построить последовательность натуральных чисел $k_n \to \infty$ так, что $-\frac{1}{k_n} < x_n < \frac{1}{k_n}$. Тогда $a^{-1/k_n} < a^{x_n} < a^{1/k_n}$ при $a > 1$ и $a^{1/k_n} < a^{x_n} < a^{-1/k_n}$ при $a < 1$. Остаётся заметить, что $\lim_{n\to\infty} a^{1/k_n} = 1 = \lim_{n\to\infty} a^{-1/k_n}$.

Замечание. Не обращаясь к задаче 2.13, равенство $\lim_{n\to\infty} a^{1/n} = 1$ можно доказать следующим образом. Достаточно рассмотреть случай, когда $a > 1$. Согласно задаче 3.16 последовательность $\{a^{1/n}\}$ монотонна. Ясно также, что эта последовательность ограничена, поэтому она имеет некоторый предел $\lim_{n\to\infty} a^{1/n} = c$. Её подпоследовательность $\{a^{1/(2n)}\}$ имеет тот же самый предел, поэтому $c = c^2$, так как $a^{1/n} = (a^{1/(2n)})^2$. Таким образом, $c = 0$ или 1. Но $a^{1/n} \ge 1$, поэтому $c \ge 1$.

3.21. Будем считать, что $a > 1$; случай $a < 1$ разбирается аналогично. Рассмотрим вспомогательные последовательности рациональных чисел $\{x'_n\}$ и $\{x''_n\}$, сходящиеся к x, причём $\{x'_n\}$ монотонно возрастает, а $\{x''_n\}$ монотонно убывает. Согласно задаче 3.16 последовательность $\{a^{x'_n}\}$ монотонно возрастает. Эта последовательность ограничена, поэтому существует предел $\lim_{n\to\infty} a^{x'_n} = c'$. Аналогично существует предел $\lim_{n\to\infty} a^{x''_n} = c''$. Ясно, что $x''_n - x'_n \to 0$, поэтому согласно задаче 3.20 $c''/c' = \lim_{n\to\infty} a^{x''_n - x'_n} = 1$, т.е. $c'' = c' = c$.

Теперь для исходной последовательности $\{x_n\}$ мы можем выбрать последовательность натуральных чисел $k_n \to \infty$ так, что $x'_{k_n} < x_n < x''_{k_n}$. Поэтому $\lim_{n\to\infty} a^{x_n} = c$. Число c зависит только от x.

3.22. а) Выберем рациональные числа p и q так, что $x > p > q > y$. Тогда можно выбрать последовательности рациональных чисел $\{x_n\}$ и $\{y_n\}$ так, что они сходятся к x и y и при этом $x_n \ge p$ и $y_n \le q$ для всех n. Согласно задаче 3.16 имеют место неравенства $a^{x_n} \ge a^p \ge a^q \ge a^{y_n}$, поэтому $a^x \ge a^p \ge a^q \ge a^y$.

б) Решается аналогично.

3.23. Решение аналогично решению задачи 3.21. Мы снова выбираем такие же последовательности $\{x'_n\}$ и $\{x''_n\}$. Согласно задаче 3.22 из неравенств $x'_{k_n} < x_n < x''_{k_n}$ следуют неравенства $a^{x'_{k_n}} < a^{x_n} < a^{x''_{k_n}}$, а потому $\lim_{n\to\infty} a^{x_n} = c = a^x$.

3.24. Соотношение $a^{x_1+x_2} = a^{x_1}a^{x_2}$ для произвольных чисел следует из аналогичного соотношения для рациональных чисел (задача 3.18), поскольку функция $f(x) = a^x$ непрерывна (задача 3.23).

3.25. Функция $f(x) = a^x$ непрерывна (задача 3.23) и монотонна (это легко вывести из утверждения задачи 3.16, воспользовавшись непрерывностью функции f). Кроме того, согласно задаче 3.17 число a^y может быть как сколь угодно велико, так и сколь угодно близко к нулю (для доказательства последнего утверждения нужно заметить, что $a^{-y} = 1/a^y$). Поэтому согласно теореме о промежуточном значении для любого $x > 0$ найдётся такое число y, что $a^y = x$. Единственность числа y следует из монотонности функции f.

3.26. Докажем непрерывность в точке $x_0 = a^{y_0}$, где $y_0 = \log_a x_0$. Для заданного $\varepsilon > 0$ возьмём в качестве δ наименьшее из двух положительных чисел $|a^{y_0} - a^{y_0+\varepsilon}|$ и $|a^{y_0} - a^{y_0-\varepsilon}|$. Функция $g(y) = a^y$ монотонна, поэтому из неравенства $|a^{y_0} - a^y| < \delta$ следует неравенство $|y_0 - y| < \varepsilon$, т.е. из неравенства $|x_0 - x| < \delta$ следует неравенство $|\log_a x_0 - \log_a x| < \varepsilon$.

3.27. Это следует из соответствующего свойства показательной функции: $a^{y_1+y_2} = a^{y_1}a^{y_2}$ (задача 3.24).

3.28. Мы воспользуемся определением числа e как предела $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ (см. с. 26). Для каждого $x \ge 1$ можно выбрать натуральное число n так, что $n \le x < n+1$. Тогда

$$\left(1+\frac{1}{x}\right)^x < \left(1+\frac{1}{n}\right)^{n+1} = \left(1+\frac{1}{n}\right)^n\left(1+\frac{1}{n}\right),$$

$$\left(1+\frac{1}{x}\right)^x > \left(1+\frac{1}{n+1}\right)^n = \left(1+\frac{1}{n+1}\right)^{n+1}\left(1+\frac{1}{n}\right)^{-1}.$$

При $n \to \infty$ пределы правых частей обоих неравенств равны e.

Докажем теперь, что при $x \to -\infty$ предел получается тот же самый. Для этого положим $y = -x$ и заметим, что

$$\left(1 - \frac{1}{y}\right)^{-y} = \left(\frac{y}{y-1}\right)^{y} = \left(1 + \frac{1}{y-1}\right)^{y-1}\left(1 + \frac{1}{y-1}\right).$$

При $y \to \infty$ правая часть стремится к e.

3.29. Согласно задаче 3.28 $\lim_{x\to 0}(1+x)^{1/x} = e$. Из этого, воспользовавшись непрерывностью логарифма, получаем $\lim_{x\to 0} \frac{1}{x}\ln(1+x) = 1$.

3.30. Достаточно доказать, что $\lim_{n\to\infty} n\ln\left(1+\frac{a}{n}\right) = a$. Требуемое равенство следует из задачи 3.29: $n\ln\left(1+\frac{a}{n}\right) = a\left(\frac{n}{a}\ln\left(1+\frac{a}{n}\right)\right) \to a$ при $x = \frac{a}{n} \to 0$.

3.31. Пусть $(1+x)^a = 1+y$. Тогда $y \to 0$ при $x \to 0$. Далее, $a\ln(1+x) = \ln(1+y)$. Поэтому

$$\frac{(1+x)^a - 1}{x} = \frac{y}{x} = \frac{y}{\ln(1+y)} \cdot \frac{a\ln(1+x)}{x}.$$

Воспользовавшись пределом из задачи 3.29, получаем требуемое.

3.32. Пусть $a^x - 1 = y$. Тогда $y \to 0$ при $x \to 0$. Кроме того, $x = \frac{\ln(1+y)}{\ln a}$. Поэтому $\frac{a^x-1}{x} = \frac{y}{\ln(1+y)} \cdot \ln a$. Воспользовавшись пределом из задачи 3.29, получаем требуемое.

3.33. Функция $\ln x$ непрерывна, поэтому $\lim_{n\to\infty}\ln a_n = \ln a$. Значит, согласно задаче 2.2 $\lim_{n\to\infty}\frac{\ln a_1 + \dots + \ln a_n}{n} = \ln a$, т.е. $\lim_{n\to\infty}\ln\sqrt[n]{a_1\dots a_n} = \ln a$. Теперь, воспользовавшись непрерывностью функции e^x, получаем требуемое.

3.34. Рассмотрим последовательность $b_1 = a_1,\ b_n = \frac{a_{n+1}}{a_n}$ при $n > 1$. По условию $\lim_{n\to\infty} b_n = a$. Поэтому согласно задаче 3.33 $\lim_{n\to\infty}\sqrt[n]{b_1\cdots b_n} = a$. Но $b_1\cdots b_n = a_n$.

3.35. Пусть $a_n = \frac{n!}{n^n}$. Тогда $\frac{a_{n+1}}{a_n} = \frac{(n+1)!}{(n+1)^n}\cdot\frac{n^n}{n!} = \frac{1}{\left(1+\frac{1}{n}\right)^n} \to \frac{1}{e}$. Поэтому согласно задаче 3.34 $\lim_{n\to\infty}\sqrt[n]{a_n} = \frac{1}{e}$, т.е. $\lim_{n\to\infty}\frac{\sqrt[n]{n!}}{n} = \frac{1}{e}$.

3.36. Ответ: $\sqrt{ab}$. Легко проверить, что $\frac{\sqrt[n]{a}+\sqrt[n]{b}}{2} = 1 + \frac{c_n}{n}$, где $c_n = \frac{1}{2}(n(\sqrt[n]{a}-1) + n(\sqrt[n]{b}-1)$. Согласно задаче 3.32 $\lim_{n\to\infty} n(\sqrt[n]{a}-1) = \ln a$, поэтому $c = \lim_{n\to\infty} c_n = \ln\sqrt{ab}$. Следовательно, $\lim_{n\to\infty}\left(1+\frac{c_n}{n}\right)^n = \lim_{n\to\infty}\left(1+\frac{c}{n}\right)^n = e^c$ (задача 3.30). Таким образом, искомый предел равен $e^{\ln\sqrt{ab}} = \sqrt{ab}$.

3.37. Равенство $f(x)f(x_0 - x) = f(x_0)$ показывает, что $f(x) \neq 0$ для всех x. Но тогда $f(x) > 0$ для всех x поскольку $f(x) = f(x/2)f(x/2)$.

Индукцией по n легко доказать, что

$$f(nx) = \big(f(x)\big)^n \tag{1}$$

для всех натуральных n. Ясно также, что $f(0) = 1$, поскольку $f(x) = f(x+0) = f(x)f(0)$. Кроме того, $f(x)f(-x) = f(0) = 1$. Поэтому равенство (1) выполняется для всех целых n. В частности, $f(n) = a^n$, где $a = f(1)$.

Докажем теперь, что равенство $f(r) = a^r$ выполняется для любого рационального числа $r = m/n$. Запишем равенство (1) для $x = m/n$: $f(m) = (f(m/n))^n$. Учитывая, что $f(m) = a^m$, получаем $a^m = (f(m/n))^n$, т.е. $f(m/n) = a^{m/n}$.

Равенство $f(x) = a^x$ доказано для всех рациональных x. Поэтому из непрерывности функции f следует, что $f(x) = a^x$ для всех x.

3.38. а) Сделаем замену $u = \log_a x$, т.е. $x = a^u$, и рассмотрим функцию $g(u) = f(x) = f(a^u)$. Функция g тоже непрерывна. Она удовлетворяет соотношению $g(u+v) = g(u)+g(v)$. Действительно, $g(u+v) = f(a^{u+v}) = f(a^u a^v) = f(a^u) + f(a^v) = g(u) + g(v)$. Поэтому согласно задаче 3.10 а) $g(u) = ug(1) = uf(a) = u$. Таким образом, $f(x) = g(u) = u = \log_a x$.

б) Легко проверить, что $f(x_0^n) = nf(x_0)$ для любого целого n. Поэтому функция f принимает значения как больше 1, так и меньше 1.

3.39. Формула для разности квадратов показывает, что

$$\left(\frac{e^t + e^{-t}}{2}\right)^2 - \left(\frac{e^t - e^{-t}}{2}\right)^2 = e^t \cdot e^{-t} = 1.$$

3.40. Ясно, что

$$4\operatorname{sh} x \operatorname{ch} y = e^{x+y} + e^{x-y} - e^{-x+y} - e^{-x-y},$$
$$4\operatorname{ch} x \operatorname{sh} y = e^{x+y} - e^{x-y} + e^{-x+y} - e^{-x-y},$$

Из этого легко получается первое равенство. Второе равенство доказывается аналогично.

3.41. Пусть $y = \operatorname{Arsh} x$. Тогда $x = \operatorname{sh} y = \frac{e^y - e^{-y}}{2}$. Поэтому $e^{2y} - 2xe^y - 1 = 0$. Решая это квадратное уравнение относительно e^y, получаем $e^y = x \pm \sqrt{x^2+1}$. Но $e^y \geq 0$, поэтому $e^y = x + \sqrt{x^2+1}$, т.е. $y = \ln(x + \sqrt{x^2+1})$.

Пусть $y = \operatorname{Arch} x$. Тогда $x = \operatorname{ch} y = \frac{e^y + e^{-y}}{2}$. Поэтому $e^{2y} - 2xe^y + 1 = 0$, а значит, $e^y = x \pm \sqrt{x^2-1}$. Таким образом, $y = \ln(x \pm \sqrt{x^2-1})$. Равенство $(x + \sqrt{x^2-1})(x - \sqrt{x^2-1}) = 1$ показывает, что $\ln(x - \sqrt{x^2-1}) = -\ln(x + \sqrt{x^2-1})$.

Пусть $y = \operatorname{Arth} x$. Тогда $x = \operatorname{th} y = \frac{e^y - e^{-y}}{e^y + e^{-y}}$. Поэтому $e^y(1-x) = e^{-y}(1+x)$, а значит, $e^{2y} = \frac{1+x}{1-x}$. Таким образом, $y = \frac{1}{2}\ln\frac{1+x}{1-x}$.

3.42. а) Ясно, что $1 = \operatorname{ch}^2 x - \operatorname{sh}^2 x = \operatorname{ch}^2 x - \operatorname{tg}^2 \alpha$. Поэтому $\operatorname{ch}^2 x = 1 + \operatorname{tg}^2 \alpha = 1/\cos^2 \alpha$. Кроме того, $\operatorname{ch} x$ и $\cos\alpha$ положительны.

б) Ясно, что $e^x = \operatorname{sh} x + \operatorname{ch} x = \operatorname{tg}\alpha + \frac{1}{\cos\alpha} = \frac{1+\sin\alpha}{\cos\alpha} = \operatorname{tg}\left(\frac{\alpha}{2} + \frac{\pi}{4}\right)$. Поэтому $x = \ln\operatorname{tg}\left(\frac{\alpha}{2} + \frac{\pi}{4}\right)$. Формула для тангенса суммы показывает, что $\operatorname{tg}\left(\frac{\alpha}{2} + \frac{\pi}{4}\right) = \frac{1+\operatorname{tg}(\alpha/2)}{1-\operatorname{tg}(\alpha/2)}$. Таким образом,

$$\frac{x}{2} = \frac{1}{2}\ln\operatorname{tg}\left(\frac{\alpha}{2} + \frac{\pi}{4}\right) = \frac{1}{2}\ln\frac{1+\operatorname{tg}(\alpha/2)}{1-\operatorname{tg}(\alpha/2)} = \operatorname{Arth}\left(\operatorname{tg}\frac{\alpha}{2}\right),$$

что и требовалось.

3.43. Решение аналогично решению задачи 2.26, только теперь мы полагаем $x_0 = y_0 \operatorname{ch}\alpha$, т.е. $\alpha = \operatorname{Arch}(x_0/y_0)$.

3.44. а) Функция $f(x) = x^2$ на множестве всех действительных чисел не является равномерно непрерывной. Действительно, если бы эта функция была равномерно непрерывной, то тогда для любого $\varepsilon > 0$ можно было бы выбрать $\delta > 0$ так, что для любого x выполняется неравенство $\left|\left(x + \frac{\delta}{2}\right)^2 - x^2\right| < \varepsilon$, т.е. $\left|\delta x + \frac{\delta^2}{4}\right| < \varepsilon$. Но если x достаточно велико, то это неравенство не может выполняться.

б) Функция $f(x) = 1/x$ на интервале $(0,1)$ не является равномерно непрерывной. Действительно, предположим, что для любого $\varepsilon > 0$ мы выбрали $\delta > 0$ так, что $\left|\frac{1}{x} - \frac{1}{x+(\delta/2)}\right| < \varepsilon$ для всех x из интервала $\left(0, 1-\frac{\delta}{2}\right)$. Это неравенство можно переписать в виде $\frac{\delta}{2} < \varepsilon\left|x\left(x+\frac{\delta}{2}\right)\right|$. При $x \to 0$ получаем $\delta \leq 0$. Приходим к противоречию.

3.45. По условию $|f_n(x)| \leq C_n$ для всех x. Для некоторого $\varepsilon > 0$ выберем номер N так, что $|f_n(x) - f(x)| < \varepsilon$ при $n \geq N$. В частности, $|f_N(x) - f(x)| < \varepsilon$, поэтому $|f(x)| < |f_N(x)| + \varepsilon < C_N + \varepsilon$.

3.46. Достаточно для любого $a > 0$ доказать равномерную сходимость на отрезке $[-a, a]$. Напомним, что $\lim_{y\to 0}(1+y)^{1/y} = e$ (см. задачу 3.28). Поэтому для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что если $|y| < \delta$, то

$$e(1-\varepsilon) < (1+y)^{1/y} < e(1+\varepsilon).$$

Положим $x = ny$. Тогда $\left(1 + \frac{x}{n}\right)^n = (1+y)^{x/y} = ((1+y)^{1/y})^x$. Следовательно, если $n > |a|/\delta$ и $|x| \leq a$, то

$$e^x(1-\varepsilon)^a \leq e^x(1-\varepsilon)^x < \left(1 + \frac{x}{n}\right)^n < e^x(1+\varepsilon)^x \leq e^x(1+\varepsilon)^a,$$

поэтому

$$e^x((1-\varepsilon)^a - 1) \leq \left(1 + \frac{x}{n}\right)^n - e^x \leq e^x((1+\varepsilon)^a - 1).$$

Из этих неравенств следует равномерная сходимость, поскольку $(1 \pm \varepsilon)^a \to 1$ при $\varepsilon \to 0$.

3.47. 1) Положим $f_n(x) = 1$ для $x \in (0, 1/n)$ и $f_n(x) = 0$ для всех остальных точек $x \in [0,1]$. Последовательность f_n разрывных функций убывающая, сходится к непрерывной функции f, тождественно равной нулю, но сходимость не равномерная на $[0,1]$. Действительно, для любого n для точки $x = \frac{1}{2n}$ выполняется равенство $|f(x) - f_n(x)| = 1$.

2) Положим $f_n(x) = x^n$ для всех $x \in [0,1]$. Последовательность f_n непрерывных функций убывающая, сходится к разрывной функции $f(x)$, но сходимость не равномерная на $[0,1]$. Действительно, если $x \to 1 - 0$, то $|f(x) - f_n(x)| = f_n(x) \to 1$.

3) Положим $f_n(x) = x/n$ для $x \in [0, \infty)$. Последовательность f_n непрерывных функций убывающая, сходится к непрерывной функции f, тождественно равной нулю, но сходимость не равномерная на $[0, \infty)$. Действительно, если $x \to \infty$, то $|f(x) - f_n(x)| = f_n(x) \to \infty$.

3.48. Требуется доказать, что если $x_1, x_2 > 0$, то $\ln\left(\frac{x_1+x_2}{2}\right) \geq \frac{\ln x_1 + \ln x_2}{2}$. Это неравенство эквивалентно неравенству $\frac{x_1+x_2}{2} \geq \exp\left(\frac{\ln x_1+\ln x_2}{2}\right) = \sqrt{x_1x_2}$.

3.49. Функция $f(x) = \ln x$ вогнута (задача 3.48). Поэтому согласно неравенству Йенсена

$$\ln\left(\frac{x_1 + \dots + x_n}{n}\right) \geq \frac{\ln x_1 + \dots + \ln x_n}{n}.$$

Это неравенство эквивалентно требуемому.

Замечание. По поводу других доказательств неравенства между средним арифметическим и средним геометрическим см. задачу 1.9.

3.50. Функция $f(x) = \ln x$ вогнута (задача 3.48). Поэтому согласно неравенству Йенсена

$$\ln\left(\frac{A}{p} + \frac{B}{q}\right) \geq \frac{1}{p}\ln A + \frac{1}{q}\ln B.$$

Это неравенство эквивалентно требуемому.

Замечание. Другое доказательство приведено в решении задачи 1.11.

3.51. Указанное неравенство можно записать в виде

$$(x_3 - x_2)f(x_1) + (x_1 - x_3)f(x_2) + (x_2 - x_1)f(x_3) \geq 0. \tag{1}$$

Предположим сначала, что неравенство (1) выполняется, и запишем это неравенство для $x_2 = \frac{x_1+x_3}{2}$. После сокращения на $x_3 - x_1$ получим неравенство $\frac{f(x_1)+f(x_3)}{2} \geq f\left(\frac{x_1+x_3}{2}\right)$, поэтому функция $f(x)$ выпуклая.

Предположим теперь, что функция $f(x)$ выпуклая. Запишем неравенство из теоремы 3.13 в виде

$$pf(x_1) + (1-p)f(x_3) \geq f(px_1 + (1-p)x_3). \tag{2}$$

Число p выберем так, чтобы выполнялось равенство $x_2 = px_1 + (1-p)x_3$. Домножим обе части неравенства (2) на $(x_3 - x_1)$. Полученное в результате неравенство легко преобразовать в неравенство (1).

3.52. Непрерывность функции $f(x)$ доказана в решении задачи 3.7. Пусть $x_1 = \frac{2}{\pi}$, $x_2 = \frac{1}{\pi}$, $x_3 = \frac{2}{3\pi}$, $x_4 = \frac{1}{2\pi}$, $x_5 = \frac{2}{5\pi}$, $x_6 = \frac{1}{3\pi}$, $x_7 = \frac{2}{7\pi}$, $x_8 = \frac{1}{4\pi}$, … Тогда $f(x_{2k}) = 0$, поэтому

$$\sum_{k=0}^{\infty} |f(x_{k+1}) - f(x_k)| = \frac{2}{\pi} + \frac{2}{3\pi} + \frac{2}{5\pi} + \frac{2}{7\pi} + \dots$$

Последовательность $a_n = 1 + \frac{1}{3} + \frac{1}{7} + \dots + \frac{1}{2n-1} > \frac{1}{2}\left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}\right)$ не ограничена (задача 2.5).

3.53. Неравенство

$$|f(x)g(x) - f(y)g(y)| \leq |f(x)| \cdot |g(x) - g(y)| + |g(y)| \cdot |f(x) - f(y)|$$

показывает, что $\operatorname{Var}_a^b(fg) \leq F \operatorname{Var}_a^b(g) + G \operatorname{Var}_a^b(f)$, где F и G — оценки сверху для функций $|f(x)|$ и $|g(x)|$ на отрезке $[a,b]$.

3.54. Ясно, что если функция g убывающая или возрастающая, то $\mathrm{Var}_a^b(g) = |g(a)-g(b)|$; в частности, g — функция ограниченной вариации. Ясно также, что если g_1 и g_2 — функции ограниченной вариации, то $f = g_1 + g_2$ тоже функция ограниченной вариации.

Рассмотрим сумму $v = \sum_{k=0}^{m} |f(x_{k+1}) - f(x_k)|$. Пусть p — сумма тех слагаемых, для которых $f(x_{k+1})-f(x_k) > 0$, а $(-n)$ — сумма тех слагаемых, для которых $f(x_{k+1})-f(x_k) < 0$. Тогда $v = p+n$ и $f(b) - f(a) = p - n$, поэтому

$$v = 2p + f(a) - f(b) = 2n + f(b) - f(a).$$

Пусть V, P и N — точные верхние грани чисел v, p и n для всех наборов точек $x_0 = a < x_1 < \dots < x_m < b = x_{m+1}$. Аналогично определим числа $V(x)$, $P(x)$ и $N(x)$, заменив отрезок $[a,\, b]$ на отрезок $[a,\, x]$. Ясно, что $P(x)$ и $N(x)$ — возрастающие функции на отрезке $[a,\, b]$. Ясно также, что

$$V(x) = 2P(x) + f(a) - f(x) = 2N(x) + f(x) - f(a),$$

поэтому $f(x) = f(a) + P(x) - N(x)$. Обе функции $f(a) + P(x)$ и $N(x)$ возрастающие.

3.55. Функцию ограниченной вариации можно представить в виде разности ограниченных монотонных функций, а для монотонной функции существуют оба односторонних предела (задача 3.1).

3.56. Предположим сначала, что функция V непрерывна в точке x_0. Функция f является суммой двух монотонных функций, поэтому существуют односторонние пределы $f(x_0+0)$ и $f(x_0-0)$. Из определения функции V следует, что если $x_0 < x \le b$, то

$$|f(x) - f(x_0)| \le V(x) - V(x_0),$$

поэтому $|f(x_0+0) - f(x_0)| \le V(x_0+0) - V(x_0) = 0$, т.е. $f(x_0+0) = f(x_0)$. Аналогично доказывается, что $f(x_0-0) = f(x_0)$.

Предположим теперь, что функция f непрерывна в точке $x_0 \in (a,\, b)$. Тогда для заданного $\varepsilon > 0$ существует такое $\delta > 0$, что если $|x-x_0| < \delta$, то $|f(x)-f(x_0)| < \varepsilon/2$. Для того же самого ε существует такое разбиение отрезка $[x_0,\, b]$, что

$$V(b) - V(x_0) < \sum_{k=0}^{m} |f(x_{k+1}) - f(x_k)| + \varepsilon/2. \tag{1}$$

Выражение, стоящее в правой части неравенства, при измельчении разбиения не уменьшается, поэтому можно считать, что $x_0 < x_1 < x_0 + \delta$. В таком случае

$$\begin{aligned}\sum_{k=0}^{m} |f(x_{k+1}) - f(x_k)| &= |f(x_1) - f(x_0)| + \sum_{k=1}^{m} |f(x_{k+1}) - f(x_k)| \le \\ &\le \varepsilon/2 + (V(b) - V(x_1)). \qquad (2)\end{aligned}$$

Сравнивая неравенства (1) и (2), получаем $V(x_1) - V(x_0) < \varepsilon$. Ясно также, что $V(x_1) \ge V(x_0)$. Следовательно, $V(x_0+0) = V(x_0)$. Аналогично $V(x_0-0) = V(x_0)$. Если $x_0 = a$ или $x_0 = b$, то нужно рассматривать только один из односторонних пределов.

3.57. При решении задачи 3.54 показано, что функцию $f(x)$ ограниченной вариации можно представить в виде разности неубывающих функций $f(a) + P(x)$ и $N(x)$. При этом $V(x) = 2P(x) + f(a) - f(x) = 2N(x) + f(x) - f(a)$, т.е.

$$P(x) = \frac{V(x) + f(x) - f(a)}{2}, \quad N(x) = \frac{V(x) - f(x) + f(a)}{2}.$$

Согласно задаче 3.56 функция $V(x)$ непрерывна, поэтому функции $P(x)$ и $N(x)$ непрерывны.

3.58. Для заданного $\varepsilon > 0$ выберем $\delta > 0$ так, что для любой конечной системы неперекрывающихся интервалов (a_1, b_1), ..., (a_n, b_n), сумма длин которых меньше δ, выполняется неравенство $\sum_{k=1}^{n} |f(a_k) - f(b_k)| < \varepsilon$. Тогда на любом отрезке, длина которого меньше δ, вариация функции меньше ε. Отрезок $[a, b]$ можно представить в виде объединения конечного числа отрезков, длина каждого из которых меньше δ, поэтому вариация функции на отрезке $[a, b]$ конечна.

Глава 4.

Топология вещественных чисел

Непрерывность функции можно определить, введя понятие открытых множеств на прямой. Это упрощает определение и некоторые доказательства и позволяет определить непрерывные отображения в более общей ситуации — для топологических пространств. В этой главе мы обсудим топологические свойства вещественных чисел.

4.1. Открытые и замкнутые множества

Множество $U \subset \mathbb{R}$ называют *открытым*, если для любой точки $x_0 \in U$ можно выбрать $\varepsilon > 0$ так, что все точки x, для которых $|x - x_0| < \varepsilon$ принадлежат U. Пустое множество считается открытым.

Окрестностью точки x называют любое множество V, которое содержит открытое множество U, содержащее точку x: $x \in U \subset V$. *Проколотой окрестностью* точки x называют множество $V \setminus \{x\}$, где V — окрестность точки x.

Задача 4.1. Докажите, что интервал (a, b) — открытое множество.

Множество точек x, для которых $|x - x_0| < \varepsilon$, иногда называют ε-окрестностью точки x_0; мы будем обозначать её $D_\varepsilon(x_0)$.

Следующая теорема проясняет, почему так важно понятие открытого множества.

Теорема 4.1. *Функция $f(x)$, определённая на всей прямой, непрерывна тогда и только тогда, когда прообраз любого открытого множества открыт.*

Доказательство. Сначала переформулируем определение непрерывности функции $f(x)$ в точке x_0 на языке ε-окрестностей: функция $f(x)$ непрерывна в точке x_0 тогда и только тогда, когда для любого $\varepsilon > 0$ существует $\delta > 0$, для которого $D_\delta(x_0)$ содержится в прообразе множества $D_\varepsilon(f(x_0))$.

Предположим, что прообраз любого открытого множества открыт. Тогда, в частности, открыт прообраз множества $D_\varepsilon(f(x_0))$. Поэтому существует $\delta > 0$, для которого $D_\delta(x_0)$ содержится в этом прообразе. Таким образом, функция $f(x)$ непрерывна.

Предположим, что функция $f(x)$ непрерывна. Рассмотрим открытое множество U. Прообраз этого множества состоит из всех точек x_0, для которых $f(x_0) \in U$. Выберем $\varepsilon > 0$ так, что $D_\varepsilon(f(x_0)) \subset U$. Затем выберем $\delta > 0$ так, что $D_\delta(x_0)$ содержится в прообразе множества $D_\varepsilon(f(x_0))$. Тогда $D_\delta(x_0)$ содержится в прообразе множества U, поэтому прообраз этого множества открыт. □

Легко проверить, что открытые множества обладают следующими свойствами.

- Объединение любого набора открытых множеств открыто.
- Пересечение двух (и любого конечного числа) открытых множеств открыто.

Действительно, если точка x_0 принадлежит открытому множеству U, то этому множеству принадлежит её ε-окрестность, и эта же ε-окрестность принадлежит объединению любого набора открытых множеств, в который входит U. А если точка x_0 принадлежит открытым множествам U и V, то этим множествам принадлежат, соответственно, её ε_1-окрестность и ε_2-окрестность. Поэтому пересечению множеств U и V принадлежит ε-окрестность точки x_0, где ε — наименьшее из чисел ε_1 и ε_2. Аналогично можно выбрать наименьшее число из любого конечного набора положительных чисел.

Произвольное множество X с заданным набор подмножеств, называемых *открытыми*, называют *топологическим пространством*, если выполняются оба указанных выше свойства и, кроме того, пустое множество и всё множество X являются открытыми множествам. Отображение топологических пространств называют *непрерывным*, если прообраз любого открытого множества открыт. Отображение топологических пространств называют *непрерывным в точке x*, если прообраз любого открытого множества, содержащего точку x, открыт.

Задача 4.2. Приведите пример бесконечного набора открытых множеств, пересечение которых не является открытым.

Множество называют *замкнутым*, если дополнение этого множества открыто.

Задача 4.3. Докажите, что функция f непрерывна тогда и только тогда, когда прообраз любого замкнутого множества замкнут.

Для объединений и пересечений замкнутых множеств выполняются следующие свойства.

- Пересечение любого набора замкнутых множеств замкнуто.
- Объединение двух (и любого конечного числа) замкнутых множеств замкнуто.

Точку x называют *граничной* точкой множества $X \subset \mathbb{R}$, если любая открытая окрестность точки x содержит точку, принадлежащую X, и точку, не принадлежащую X. Множество всех граничных точек множества X называют *границей* множества X. Границу множества X будем обозначать ∂X.

Все точки прямой можно разбить на три непересекающихся множества:

- Точки, для которых существует открытая окрестность, целиком лежащая в X. Обозначим это множество $\operatorname{Int} X$ и назовём его *внутренностью* множества X.
- Точки, любая открытая окрестность которых содержит как точку в X, так и точку вне X. По определению это множество ∂X.
- Точки, для которых существует открытая окрестность, целиком лежащая вне X. Обозначим это множество $\operatorname{Int} CX$ (внутренность дополнения множества X).

Задача 4.4. Докажите, что множества $\operatorname{Int} X$ и $\operatorname{Int} CX$ открытые.

Задача 4.5. Докажите, что объединение множества и его границы замкнуто.

Объединение множества X и его границы называют *замыканием* множества X. Замыкание множества X будем обозначать $\overline{X}$.

Задача 4.6. а) Докажите, что внутренность множества X — это наибольшее открытое множество, содержащееся в X.

б) Докажите, что замыкание множества X — это наименьшее замкнутое множество, содержащее X.

Задача 4.7. Докажите, что если $X \subset Y$, то $\overline{X} \subset \overline{Y}$ и $\operatorname{Int} X \subset \operatorname{Int} Y$.

Задача 4.8. Докажите, что $\overline{X \cup Y} = \overline{X} \cup \overline{Y}$ и $\operatorname{Int} X \cap \operatorname{Int} Y = \operatorname{Int}(X \cap Y)$.

Задача 4.9. Докажите, что $\overline{X \cap Y} \subset \overline{X} \cap \overline{Y}$ и $\operatorname{Int} X \cup \operatorname{Int} Y \subset \operatorname{Int}(X \cup Y)$. Можно ли эти включения заменить на равенства?

Задача 4.10. Докажите, что множество ∂X замкнутое.

Точку x множества X называют *изолированной*, если некоторая её открытая проколотая окрестность не пересекается с X.

Открытые подмножества любого множества $X \subset \mathbb{R}$ определяются как пересечения множества X со всеми открытыми подмножествами в $\mathbb{R}$. (Чтобы не возникало путаницы, про такие множества говорят, что они открыты в X; например, само множество X всегда открыто в X.) Это позволяет определить непрерывные функции на любом множестве $X \subset \mathbb{R}$: функция f на множестве X *непрерывна*, если прообраз любого открытого множества открыт.

Задача 4.11. Пусть X — объединение конечного набора замкнутых множеств $X_1, \ldots, X_n$. Докажите, что функция f, определённая на X, непрерывна тогда и только тогда, когда непрерывно каждое её ограничение f_i на множество X_i.

Задача 4.12. Пусть X — объединение произвольного набора открытых множеств $\{X_\alpha\}$. Докажите, что функция f, определённая на X, непрерывна тогда и только тогда, когда непрерывны все её ограничения на множества X_α.

Задача 4.13. Докажите, что любое открытое множество на прямой является объединением не более чем счётного набора попарно не пересекающихся интервалов.

Задача 4.14. Докажите, что любое замкнутое множество на прямой является пересечением не более чем счётного множества открытых множеств.

Пусть $A \subset \mathbb{R}$ — произвольное подмножество. Для произвольной точки $x \in \mathbb{R}$ величину $d(x, A) = \inf_{a \in A} |x - a|$ называют *расстоянием* от точки x до множества A.

Задача 4.15. Докажите, что функция $f(x) = d(x, A)$ непрерывна для любого подмножества $A \subset \mathbb{R}$

Задача 4.16. а) Докажите, что если множество A замкнуто, то функция $f(x) = d(x, A)$ для всех $x \notin A$ принимает положительные значения.

б) Докажите, что любое замкнутое множество на прямой является множеством нулей некоторой непрерывной функции.

Задача 4.17. Докажите, что для любых двух непересекающихся замкнутых множеств A и B на прямой существует функция, равная 1 на одном множестве и -1 на другом.

Задача 4.18. На замкнутом множестве $C \subset \mathbb{R}$ задана непрерывная функция f. Докажите, что существует непрерывная функция g, заданная на $\mathbb{R}$, ограничение которой на C совпадает с f.

4.2. Компактные множества

Множество называют *компактным*, если из любого его покрытия открытыми множествами можно выбрать конечное подпокрытие.

Теорема 4.2 (Борель–Лебег). *Отрезок — компактное множество.*

Доказательство. Предположим, что из некоторого открытого покрытия отрезка $[a, b]$ нельзя выбрать конечное подпокрытие. Тогда нельзя выбрать конечное подпокрытие по крайней мере одного из отрезков $[a, c]$ и $[c, b]$, где c — середина отрезка $[a, b]$. Поэтому можно построить последовательность вложенных отрезков, для каждого из которых нельзя выбрать конечное подпокрытие, причём длина n-го отрезка равна $2^{-n}(b - a)$. По теореме Вейерштрасса последовательность левых концов этих отрезков имеет предел α, а последовательность правых концов имеет предел β. Из того, что длина отрезка стремится к нулю, следует, что $\alpha = \beta$. Точка α покрыта одним из открытых множеств, поэтому некоторая ε-окрестность этой точки покрыта этим множеством. Следовательно, если $2^{-n}(b-a) < \varepsilon$, то n-й отрезок покрыт одним из множеств покрытия. Полученное противоречие показывает, что из любого открытого покрытия отрезка можно выбрать конечное подпокрытие. $\square$

Замечание. Эмиль Борель (1871-1956) доказал теорему 4.2 в 1894 году для счётного покрытия, а Пьер Кузен (1867-1933) в 1895 году и Анри Лебег (1875-1941) в 1898 году доказали её для любого покрытия.

Теорема 4.3. *Множество $K \subset \mathbb{R}$ компактно тогда и только тогда, когда оно замкнуто и ограничено.*

Доказательство. Пусть множество $K \subset \mathbb{R}$ компактно. Докажем сначала, что оно замкнуто, т.е. его дополнение CK открыто. Рассмотрим точку $x \in CK$. Для каждого натурального n рассмотрим множество C_n, состоящее из всех $y \in \mathbb{R}$, для которых $|x-y| > \frac{1}{n}$; это множество открытое. Объединение всех множеств C_n — это $\mathbb{R}$ без точки x. Поэтому множество K покрыто открытыми множествами C_n. Из этого покрытия можно выбрать конечное подпокрытие. Поскольку $C_1 \subset C_2 \subset \dots$, множество K содержится в одном из множеств C_N. Поэтому множество K не пересекается с множеством $U = \{z \in \mathbb{R} \,|\, |z-x| < \frac{1}{N}\}$, а его дополнение CK содержит открытую окрестность U точки x. Таким образом, множество CK открыто.

Докажем теперь, что множество K ограничено. Для каждого натурального n рассмотрим открытое множество D_n, состоящее из всех чисел x, для которых $|x| < n$. Из покрытия компактного множества K этими открытыми множествами можно выбрать конечное подпокрытие, поэтому K содержится в некотором множестве D_N, т.е. ограничено.

Пусть теперь множество K замкнуто и ограничено. Предположим, что из некоторого открытого покрытия этого множества нельзя выбрать конечное подпокрытие. Из ограниченности множества K следует, что оно содержится в некотором отрезке $[a,b]$. Как и при доказательстве компактности отрезка, разделим отрезок $[a,b]$ пополам и выберем ту половину, из покрытия пересечения которой с множеством K нельзя выбрать конечное подпокрытие, и т.д. Полученная последовательность вложенных отрезков имеет единственную общую точку x. В каждом из этих вложенных отрезков есть точка множества K, поэтому любая открытая окрестность точки x пересекается с K. Это означает, что точка x принадлежит замыканию множества K, а значит, и самому множеству K. Но тогда точка x принадлежит одному из множеств открытого покрытия, и это множество целиком покрывает один из рассматриваемых вложенных отрезков. Приходим к противоречию. $\square$

Теорема 4.4. *Образ компактного множества при непрерывном отображении компактен.*

Доказательство. Пусть K — компактное множество, $f\colon K \to C$ — непрерывное отображение, причём $C = f(K)$. Рассмотрим открытое покрытие множества C. Прообраз этого открытого покрытия является открытым покрытием компактного множества K. Из него можно выбрать конечное подпокрытие. Множества, соответствующие множествам этого подпокрытия, образуют конечное подпокрытие множества C. $\square$

Задача 4.19. а) Верно ли, что образ замкнутого множества при непрерывном отображении замкнут?

б) Верно ли, что образ ограниченного множества при непрерывном отображении ограничен?

Задача 4.20. Докажите, что непрерывная функция на компактном множестве достигает максимума.

Теорема 4.5. *Функция f, непрерывная на компактном множестве K, равномерно непрерывна.*

Доказательство. Требуется доказать, что для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что если $|x-y| < \delta$ и $x, y \in K$, то $|f(x)-f(y)| < \varepsilon$. Из непрерывности функции f следует, что для любой точки $x \in K$ можно выбрать $\delta_x > 0$ так, что если $|x-y| < \delta_x$ и $y \in K$, то $|f(x)-f(y)| < \varepsilon/2$. Для каждой точки x рассмотрим множество U_x — открытую $\frac{\delta_x}{2}$-окрестность точки x в K. Множества U_x образуют покрытие компактного множества K, поэтому из него можно выбрать конечное подпокрытие $U_{x_1}, \dots, U_{x_n}$. Этим множествам соответствуют числа $\delta_{x_1}, \dots, \delta_{x_n}$, и пусть δ — это половина наименьшего из этих чисел.

Проверим, что что если $|x-y| < \delta$ и $x, y \in K$, то $|f(x)-f(y)| < \varepsilon$. Точка x лежит некотором множестве U_{x_i}. Поэтому $|x-x_i| < \frac{\delta_{x_i}}{2} < \delta_{x_i}$ и

$$|x_i - y| \le |x_i - x| + |x-y| < \frac{\delta_{x_i}}{2} + \delta \le \delta_{x_i}.$$

Из этих неравенств следует, что $|f(x)-f(x_i)| < \varepsilon/2$ и $|f(x_i)-f(y)| < \varepsilon/2$, поэтому $|f(x)-f(y)| < \varepsilon$. $\square$

Теорема 4.6 (Кантор). *Последовательность непустых замкнутых ограниченных подмножеств прямой $F_1 \supset F_2 \supset F_3 \supset \ldots$ имеет общую точку.*

Доказательство. Для каждого натурального n рассмотрим открытое множество G_n — дополнение множества F_n на прямой. Предположим, что пересечение множеств F_n пусто, т.е. объединение множеств G_n совпадает со всей прямой. Тогда из покрытия компактного множества F_1 открытыми множествами G_n можно выбрать конечное подпокрытие. При этом $G_1 \subset G_2 \subset G_3 \subset \ldots$, поэтому $F_1 \subset G_N$ для некоторого N. Это означает, что $F_1 \cap F_N = \varnothing$. Но $F_1 \cap F_N = F_N$, поскольку $F_N \subset F_1$. Приходим к противоречию с тем, что множество F_N непусто. □

Историческое замечание. Георг Кантор (1845-1918) доказал теорему 4.6 в 1874 году.

Семейство множеств называют *центрированным*, если любой конечный набор этих множеств имеет общую точку.

Теорема 4.7. *Центрированное семейство замкнутых подмножеств компактного множества имеет общую точку.*

Доказательство. Предположим, что пересечение центрированного семейства замкнутых подмножеств компактного множества пусто. Тогда семейство дополнений этих множеств является открытым покрытием. Из компактности множества следует, что некоторый конечный набор множеств этого покрытия является покрытием. Пересечение дополнений этих множеств пусто. □

4.3. Связные множества

Множество X называют *связным*, если его нельзя представить в виде объединения двух непересекающихся непустых открытых множеств. Если множество представлено в виде объединения двух непересекающихся непустых открытых множеств, то эти же два множества являются одновременно замкнутыми. Множество связно тогда и только тогда, когда в нём нет множеств (отличных от пустого множества и всего множества), которые одновременно открыты и замкнуты.

Теорема 4.8. *Образ связного множества при непрерывном отображении связен.*

Доказательство. Пусть $f\colon X \to Y$ — непрерывное отображение, причём множество X связно. Без ограничения общности можно считать, что Y — образ отображения f. Предположим, что множество Y можно представить в виде объединения не пересекающихся непустых открытых множеств U и V. Тогда множество X можно представить в виде объединения не пересекающихся непустых открытых множеств — прообразов множеств U и V. Полученное противоречие показывает, что множество Y связно. □

Теорема 4.9. *Отрезок $[a, b]$ — связное множество.*

Доказательство. Предположим, что отрезок $[a, b]$ можно представить в виде объединения замкнутых непересекающихся множеств A и B, причём $a \in A$ и B не пусто. Пусть c — точная нижняя грань множества B. Тогда c — граничная точка замкнутого множества B, поэтому $c \in B$. Следовательно, $c \notin A$. Если $a < x < c$, то $x \in A$, потому что c — нижняя грань множества B. Множество A замкнуто и интервал (a, c) содержится в A, поэтому $c \in A$. Полученное противоречие показывает, что отрезок $[a, b]$ — связное множество. □

Замечание. Интервал и полуинтервал тоже являются связными множествами.

Пример. *Если множество $X \subset \mathbb{R}$ содержит точки a и b и не содержит точку c, где $a < c < b$, то множество X несвязно.*

Доказательство. Пусть $U = \{x \in \mathbb{R} | x < c\}$ и $V = \{x \in \mathbb{R} | x > c\}$. Открытые в X множества $X \cap U$ и $X \cap V$ непусты (они содержат точки a и b соответственно), не пересекаются и их объединение совпадает с X. □

Понятие связности не очень хорошо соотносится с интуитивным представлением о связности. Гораздо лучше с интуитивным представлением соотносится понятие линейной связности. Множество X называют *линейно связным*, если для любых точек $x, y \in X$ существует непрерывное отображение отрезка, переводящее его концы в точки x и y. Это непрерывное отображение называют *путём* в X с концами x и y.

Теорема 4.10. *Линейно связное множество является связным.*

Доказательство. Это непосредственно следует из связности отрезка. Действительно, предположим, что линейно связное множество можно представить в виде объединения двух непустых не пересекающихся открытых множеств U и V, выберем точки $u \in U$ и $v \in V$, и рассмотрим путь с концами u и v. Прообразы множеств U и V дают разбиение отрезка на два непустых не пересекающихся открытых множества. □

Теорема о промежуточном значении (теорема 3.5 на с. 37) следует из теоремы о том, что образ связного множества связен. Действительно, предположим, что непрерывная на отрезке $[a, b]$ функция f принимает в его концах значения разного знака и не принимает значение 0. Тогда образ отрезка $[a, b]$ при отображении f содержит точки разного знака и не содержит точку 0, поэтому образ связного множества $[a, b]$ при непрерывном отображении f несвязен. Полученное противоречие доказывает теорему о промежуточном значении.

4.4. Всюду плотные множества

Множество $A \subset B$ называют всюду плотным в B, если любое непустое открытое в B множество содержит точку множества A. Например, множество рациональных чисел всюду плотно в $\mathbb{R}$.

Историческое замечание. Понятие всюду плотного множества ввёл Георг Кантор (1845-1918) в 1879 году.

Задача 4.21. Докажите, что множество $A \subset B$ всюду плотно в B тогда и только тогда, когда $\overline{A} = B$.

Задача 4.22. Докажите, что множество $A \subset B$ всюду плотно в B тогда и только тогда, когда у его дополнения CA нет внутренних точек.

Задача 4.23. Докажите, что образ всюду плотного множества при сюръективном отображении всюду плотен.

Задача 4.24. Докажите, что непрерывная функция на отрезке полностью задаётся своими значениями на всюду плотном подмножестве отрезка.

Множество $A \subset B$ называют нигде не плотным (в B), если его замыкание $\overline{A}$ не содержит внутренних точек, т.е. $\operatorname{Int} \overline{A} = \varnothing$.

Теорема 4.11. *Множество $A \subset B$ нигде не плотно тогда и только тогда, когда в любом в непустом открытом множестве $U \subset B$ есть непустое открытое подмножество V, не пересекающееся с A.*

Доказательство. Предположим сначала, что множество A нигде не плотно. Пусть $U \subset B$ — непустое открытое множество. Оно не может целиком лежать в $\overline{A}$, поэтому в U есть точка $x \notin \overline{A}$. Выберем окрестность W точки x, не пересекающуюся с $\overline{A}$. Непустое открытое множество $V = U \cap W$ не пересекается с A.

Предположим теперь, что в любом в непустом открытом множестве $U \subset B$ есть непустое открытое подмножество V, не пересекающееся с A. Тогда $\operatorname{Int} \overline{A} = \varnothing$. Действительно, если $\operatorname{Int} \overline{A} \neq \varnothing$, то в непустом открытом множестве $U = \operatorname{Int} \overline{A} \subset A$ есть непустое открытое подмножество V, не пересекающееся с A. □

Задача 4.25. Докажите, что замыкание нигде не плотного множества нигде не плотно.

Задача 4.26. Докажите, что объединение двух нигде не плотных множеств нигде не плотно.

4.5. Совершенные множества

Точку x называют *предельной* точкой множества $X \subset \mathbb{R}$, если любая её открытая проколотая окрестность пересекается с X. Любая точка множества X является либо предельной, либо изолированной.

Задача 4.27. Докажите, что конечное множество не имеет предельных точек.

Задача 4.28. Докажите, что точка x — предельная точка множества X тогда и только тогда, когда $x \in \overline{X \setminus \{x\}}$.

Задача 4.29. Докажите, что замыкание множества X — это объединение множества X и всех его предельных точек.

Производным множеством множества X называют множество X' всех его предельных точек. Второе производное множество — это множество $(X')'$ и т.д.

Согласно задаче 4.29 $\overline{X} = X \cup X'$, поэтому множество X замкнуто тогда и только тогда, когда $X' \subset X$.

Задача 4.30. Докажите, что для $X \subset \mathbb{R}$ производное множество X' замкнуто.

Задача 4.31. Докажите, что предельные точки множества $X \subset \mathbb{R}$ и его замыкания совпадают, т.е. $X' = (\overline{X})'$.

Задача 4.32. Приведите пример множества X, для которого n-е производное множество не пусто, а $(n+1)$-е пусто.

Задача 4.33. Приведите пример множества, у которого все производные множества различны.

Задача 4.34. Докажите, что если одно из производных множеств множества X пусто, то множество X нигде не плотно.

Множество называют *совершенным*, если оно совпадает со своим производным множеством.

Задача 4.35. Докажите, что множество совершенное тогда и только тогда, когда оно замкнутое и не имеет изолированных точек.

Рассмотрим множество C всех точек отрезка $[0,1]$, у которых есть троичная запись $a_1 3^{-1} + a_2 3^{-2} + \ldots$, содержащая только цифры 0 или 2. У такой точки может быть троичная запись, содержащая цифру 1, например, число $1 \cdot 3^{-1} = 2 \cdot 3^{-2} + 2 \cdot 3^{-3} + \ldots$ входит в C. Множество C называют *канторовым множеством*.

Историческое замечание. Понятия предельной точки множества, производного множества и совершенного множества ввёл Георг Кантор (1845-1918) в 1872 году. Канторово множество он описал в 1883 году.

Пусть C_n — множество точек отрезка $[0,1]$, у которых есть троичная запись с цифрой a_n, равной 0 или 2. Это множество замкнуто; например, $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$. Канторово множество замкнуто, поскольку оно является пересечением замкнутых множеств C_n. Ясно также, что канторово множество не имеет изолированных точек, поэтому оно совершенное.

Канторово множество нигде не плотно, поскольку оно замкнуто и не может содержать интервал.

Задача 4.36. а) Докажите, что дополнение открытого всюду плотного множества нигде не плотно.

б) Приведите нетривиальный пример открытого всюду плотного множества на отрезке.

Задача 4.37. *Докажите, что непустое совершенное множество на прямой несчётно.

Задача 4.38. Докажите, что любое счётное замкнутое множество на прямой имеет изолированные точки.

Точку x называют *точкой конденсации* множества $E \subset \mathbb{R}$, если любая открытая окрестность точки x содержит несчётное множество точек E. Обозначим буквой P множество точек конденсации множества $E \subset \mathbb{R}$.

Задача 4.39. Докажите, что множество P замкнуто.

Задача 4.40. *Докажите, что множество $E \setminus P$ не более чем счётно.

Задача 4.41. *Докажите, что если множество P непусто, то оно совершенное и несчётное.

Задача 4.42. *Докажите, что любое замкнутое множество на прямой — объединение совершенного (возможно, пустого) множества и не более чем счётного множества (теорема Кантора–Бендиксона).

Задача 4.43. *Докажите, что непустое ограниченное множество на прямой совершенно тогда и только тогда, когда оно получается из отрезка удалением не более чем счётного множества попарно не пересекающихся интервалов, не имеющих общих граничных точек (Кантор).

4.6. Полунепрерывные функции

Согласно теореме Дини (см. с. 42) монотонная последовательность непрерывных функций на отрезке, сходящаяся к непрерывной функции, сходится равномерно. Несложный пример (см. задачу 3.47) показывает, что монотонная последовательность непрерывных функций может сходится к разрывной функции. Но эта предельная функция, как мы сейчас увидим, не произвольная — она полунепрерывная; более того, предел монотонной сходящейся последовательности полунепрерывных функций — полунепрерывная функция.

Рассмотрим ограниченную функцию $f(x)$, определённую на некотором открытом множестве, содержащем точку x_0. Пусть $M(f, x_0, \varepsilon)$ и $m(f, x_0, \varepsilon)$ — точная верхняя и точная нижняя грани множества значений функции $f(x)$ в ε-окрестности точки x_0. Эти функции от ε монотонны, поэтому при $\varepsilon \to 0$ они сходятся к некоторым пределам. *Верхний предел* функции $f(x)$ в точке x_0 — это предел $\lim_{\varepsilon\to 0} M(f, x_0, \varepsilon)$; этот предел обозначается $\overline{\lim}_{x\to x_0} f(x)$. Аналогично определяется *нижний предел* $\underline{\lim}_{x\to x_0} f(x)$.

Функцию $f(x)$ называют *полунепрерывной сверху* в точке x_0, если $f(x_0) \geq \overline{\lim}_{x\to x_0} f(x)$. Это означает, что для любого $\delta > 0$ можно выбрать $\varepsilon > 0$ так, что $f(x) \leq f(x_0) + \delta$ при $|x - x_0| < \varepsilon$. Для *полунепрерывной снизу* функции должно выполняться неравенство $f(x_0) \leq \underline{\lim}_{x\to x_0} f(x)$.

Историческое замечание. Понятие полунепрерывной функции ввёл Рене Бэр (1874-1932) в 1899 году.

Полунепрерывность сверху функции $f(x)$ в точке x_0 равносильна полунепрерывности снизу функции $-f(x)$ в той же точке.

Из неравенств $m(f, x_0, \varepsilon) \leq f(x_0) \leq M(f, x_0, \varepsilon)$ следуют неравенства $\underline{\lim}_{x\to x_0} f(x) \leq f(x_0) \leq \overline{\lim}_{x\to x_0} f(x)$. Поэтому для полунепрерывной сверху в точке x_0 функции f выполняется равенство $f(x_0) = \overline{\lim}_{x\to x_0} f(x)$.

Функция $f(x)$ непрерывна в точке x_0 тогда и только тогда, когда она в этой точке полунепрерывна сверху и снизу.

Задача 4.44. Докажите, что $\overline{\lim}_{x\to x_0} f(x)$ — это наибольший из пределов сходящихся последовательностей $f(x_n)$, где $x_n \to x_0$.

Задача 4.45. Докажите, что функция $f(x)$ полунепрерывна сверху тогда и только тогда, когда множество $E_c = \{x | f(x) \geq c\}$ замкнуто для любого c.

Характеристическая функция подмножества $A \subset X$ — это функция χ_A на множестве X, принимающая значения 0 и 1: $\chi_A(x) = 0$, если $x \notin A$, и $\chi_A(x) = 1$, если $x \in A$.

Задача 4.46. а) Докажите, что множество $A \subset \mathbb{R}$ замкнуто тогда и только тогда, когда его характеристическая функция полунепрерывна сверху.

б) Докажите, что множество $A \subset \mathbb{R}$ открыто тогда и только тогда, когда его характеристическая функция полунепрерывна снизу.

Задача 4.47. Убывающая последовательность полунепрерывных снизу функций f_n поточечно сходится на отрезке $[a, b]$ к функции f. Докажите, что функция f полунепрерывна снизу.

Задача 4.48. Докажите, что функция на отрезке, полунепрерывная снизу, достигает минимума.

4.7. Теорема Бэра

Множество первой категории — это множество, которое можно представить в виде счётного объединения нигде не плотных множеств. *Множество второй категории* — это множество, которое нельзя представить в виде счётного объединения нигде не плотных множеств.

Теорема 4.12 (Бэр). *а) Дополнение любого множества первой категории на прямой всюду плотно.*

б) Пересечение счётного семейства открытых всюду плотных множеств на прямой всюду плотно.

Доказательство. а) Пусть $A = \cup A_n$, где множества A_n нигде не плотны. Докажем, что любое непустое открытое множество U содержит точку множества CA. Выберем сначала отрезок I_1, содержащийся в $U \setminus A_1$, затем отрезок I_2, содержащийся в $I_1 \setminus A_2$, и т.д. По теореме Кантора (теорема 4.6) последовательность $I_1 \supset I_2 \supset \dots$ имеет непустое пересечение $\cap I_n$. Ясно, что $\cap I_n \subset U$ и $\cap I_n \subset CA$.

б) Дополнение открытого всюду плотного множества нигде не плотно (задача 4.36). Поэтому пересечение дополнений счётного семейства открытых всюду плотных множеств — это дополнение объединения счётного семейства нигде не плотных множеств. □

Историческое замечание. Рене Бэр (1874-1932) ввёл понятия множеств первой и второй категории и доказал теорему 4.12 в 1899 году.

4.8. Предел по фильтру

К разным понятиям предела (предел последовательности, предел функции в точке, односторонний предел) есть единый подход, основанный на понятии фильтра.

Фильтр в множестве X — это множество F его непустых подмножеств, обладающее следующими свойствами: 1) любое подмножество в X, содержащее множество из F, содержится в F; 2) пересечение конечного набора множеств из F содержится в F. В частности, пересечение любого конечного набора множеств из F не пусто.

Примером фильтра в топологическом пространстве X является множество всех окрестностей некоторой фиксированной точки $x \in X$. Обозначим этот фильтр N_x.

Множество G подмножеств в X, все конечные пересечения которых не пусты, порождает фильтр F следующим образом: F состоит из множеств в X, каждое из которых содержит либо множество из G, либо конечное пересечение множеств из G.

Если F — фильтр в множестве X, $g\colon X \to Y$ — некоторое отображение, то множество $\{g(U)\colon U \in F\}$ не всегда является фильтром. Но поскольку $g(U \cap V) \subset g(U) \cap g(V)$, это множество порождает некоторый фильтр; будем обозначать его $g(F)$. Легко проверить, что $V \in g(F)$ тогда и только тогда, когда существует $U \in F$, для которого $g(U) \subset V$.

Фильтр F на топологическом пространстве X сходится к точке $x \in X$ (обозначение: $F \to x$), если F содержит все окрестности точки x. Ясно, что $N_x \to x$.

Теорема 4.13. *Отображение $g\colon X \to Y$ топологических пространств непрерывно в точке $x \in X$ тогда и только тогда, когда для любого фильтра F на X, сходящегося к x, фильтр $g(F)$ на Y сходится к $g(x)$.*

Доказательство. Предположим, что отображение g непрерывно в точке x и $F \to x$. Пусть V — окрестность точки $g(x)$. Из непрерывности следует, что существует окрестность U точки x, для которой $g(U) \subset V$. Но $U \in F$, поэтому $g(U) \in g(F)$. Следовательно, $g(F) \to g(x)$.

Предположим, что если $F \to x$, то $g(F) \to g(x)$. Тогда, в частности, $g(N_x) \to g(x)$. Поэтому $V \in g(F)$ для любой окрестности V точки $g(x)$. Следовательно, существует окрестность $U \in N_x$, для которой $g(U) \subset V$, поэтому отображение g непрерывно в точке x. □

Пусть $f\colon X \to \mathbb{R}$ — некоторая функция, F — некоторый фильтр в множестве X. Число a называют ***пределом функции f по фильтру F***, если для любого $\varepsilon > 0$ существует множество $A \in F$, для которого $|f(x) - a| < \varepsilon$ при $x \in A$. Предел по фильтру обозначают $\lim_F f(x)$.

Чтобы получить определение предела функции в точке, в качестве фильтра можно взять проколотые окрестности этой точки.

Чтобы получить определение одностороннего предела функции, нужно взять односторонние окрестности.

В некоторых случаях для порождения фильтра F не нужно рассматривать конечные пересечения, а можно просто взять все те подмножества в X, каждое из которых содержит какое-то множество из G. Чтобы система образующих G обладала этим свойством, необходимо и достаточно, чтобы пересечение конечного набора множеств из G содержало множество из G. В таком случае систему образующих фильтра называют базисом фильтра.

Базис фильтра множества X — это непустое множество B непустых подмножеств в X, обладающее следующим свойством: пересечение любых двух множество из B содержит некоторое множество из B. Базис фильтра порождает фильтр F, состоящий из всех множеств, каждое из которых содержит какое-то множество из базиса фильтра.

Чтобы получить определение предела последовательности $\{a_n\}$, рассмотрим эту последовательность как отображение $f\colon \mathbb{N} \to \mathbb{R}$, сопоставляющее натуральному числу n число a_n. В качестве базиса фильтра можно взять множества $B_n = \{n, n+1, n+2, \dots\}$.

4.9. Решения задач

4.1. Пусть $x_0 \in (a, b)$, т.е. $a < x_0 < b$. Возьмём наименьшее из двух положительных чисел $x - a$ и $b - x$ и в качестве ε выберем половину этого числа. Тогда все точки x, для которых $|x - x_0| < \varepsilon$ принадлежат (a, b).

4.2. Рассмотрим множества $(0, 1 + \frac{1}{n})$, $n = 1, 2, \dots$ Пересечение этих множеств — полуинтервал $(0, 1]$. Никакая ε-окрестность точки 1 не может целиком содержаться в нём.

4.3. Дополнение прообраза множества A совпадает с прообразом дополнения множества A. Поэтому прообраз любого замкнутого множества замкнут тогда и только тогда, когда прообраз любого открытого множества открыт.

4.4. Пусть $x \in \operatorname{Int} X$. Тогда у точки x есть открытая окрестность U, целиком лежащая в X. Докажем, что U лежит в $\operatorname{Int} X$. Действительно, у каждой точки множества U есть открытая окрестность (а именно, U), целиком лежащая в X. Для множества $\operatorname{Int} CX$ доказательство аналогично.

4.5. Ясно, что $\operatorname{Int} X \subset X \subset \operatorname{Int} X \cup \partial X$, поэтому $X \cup \partial X = \operatorname{Int} X \cup \partial X$. Таким образом, множество $X \cup \partial X$ является дополнением к открытому множеству $\operatorname{Int} CX$, поэтому оно замкнуто.

4.6. а) Если открытое множество U содержится в X, то оно содержится и в $\operatorname{Int} X$. Поэтому объединение всех открытых множеств, содержащихся в X, совпадает с $\operatorname{Int} X$.

б) Замыкание множества X — это дополнение множества $\operatorname{Int} X$, т.е. дополнение наибольшего отрытого множества, содержащего дополнение множества X. Это то же самое, что наименьшее замкнутое множество, содержащее X.

4.7. Замкнутое множество $\overline{Y}$ содержит множество $Y \supset X$, поэтому оно содержит $\overline{X}$.

Открытое множество $\operatorname{Int} X$ содержится в $X \subset Y$, поэтому оно содержится в $\operatorname{Int} Y$.

4.8. Ясно, что $X \subset X \cup Y$ и $Y \subset X \cup Y$, поэтому $\overline{X} \subset \overline{X \cup Y}$ и $\overline{Y} \subset \overline{X \cup Y}$, а значит, $\overline{X} \cup \overline{Y} \subset \overline{X \cup Y}$. Ясно, что $X \subset \overline{X}$ и $Y \subset \overline{Y}$, поэтому $X \cup Y \subset \overline{X} \cup \overline{Y}$, а значит, $\overline{X \cup Y} \subset \overline{\overline{X} \cup \overline{Y}} = \overline{X} \cup \overline{Y}$.

Ясно, что $X \cap Y \subset X$ и $X \cap Y \subset Y$, поэтому $\operatorname{Int}(X \cap Y) \subset \operatorname{Int} X$ и $\operatorname{Int}(X \cap Y) \subset \operatorname{Int} Y$, а значит, $\operatorname{Int}(X \cap Y) \subset \operatorname{Int} X \cap \operatorname{Int} Y$. Ясно, что $\operatorname{Int} X \subset X$ и $\operatorname{Int} Y \subset Y$, поэтому $\operatorname{Int} X \cap \operatorname{Int} Y \subset X \cap Y$ и $\operatorname{Int} X \cap \operatorname{Int} Y = \operatorname{Int}(\operatorname{Int} X \cap \operatorname{Int} Y) \subset \operatorname{Int}(X \cap Y)$.

4.9. Ясно, что $X \cap Y \subset X$ и $X \cap Y \subset Y$, поэтому $\overline{X \cap Y} \subset \overline{X}$ и $\overline{X \cap Y} \subset \overline{Y}$, а значит, $\overline{X \cap Y} \subset \overline{X} \cap \overline{Y}$.

Ясно, что $X \subset X \cup Y$ и $Y \subset X \cup Y$, поэтому $\operatorname{Int}(X) \subset \operatorname{Int}(X \cup Y)$ и $\operatorname{Int}(Y) \subset \operatorname{Int}(X \cup Y)$, а значит, $\operatorname{Int}(X) \cup \operatorname{Int}(Y) \subset \operatorname{Int}(X \cup Y)$.

Доказанные включения нельзя заменить на равенства. Действительно, пусть $X = \mathbb{Q}$ и $Y = \mathbb{R} \setminus \mathbb{Q}$. Тогда $\overline{X \cap Y} = \overline{\varnothing} = \varnothing$ и $\overline{X} \cap \overline{Y} = \mathbb{R} \cap \mathbb{R} = \mathbb{R}$; $\operatorname{Int} X \cup \operatorname{Int} Y = \varnothing \cup \varnothing = \varnothing$ и $\operatorname{Int}(X \cup Y) = \operatorname{Int}(\mathbb{R}) = \mathbb{R}$.

4.10. Дополнение множества ∂X — это объединение двух открытых множеств $\operatorname{Int} X$ и $\operatorname{Int} CX$, поэтому оно открыто.

4.11. Предположим, что отображение f непрерывно и $C \subset Y$ — произвольное замкнутое множество. Тогда множество $f^{-1}(C)$ замкнуто в X, поэтому множество $C_i = f_i^{-1}(C) = f^{-1}(C) \cap X_i$ замкнуто в X_i. Следовательно, отображение f_i непрерывно.

Предположим, что все отображения f_i непрерывны и $C \subset Y$ — произвольное замкнутое множество. Тогда множество $C_i = f_i^{-1}(C) = f^{-1}(C) \cap X_i$ замкнуто в X_i, т.е. существует замкнутое в X множество C_i', для которого $C_i = C_i' \cap X_i$. Оба множества C_i' и X_i замкнуты в X, поэтому множество C_i тоже замкнуто в X. Следовательно, множество $f^{-1}(C) = C_1 \cup \dots \cup C_n$ замкнуто в X, поэтому отображение f_i непрерывно.

4.12. Решение аналогично решению задачи 4.11. Единственное отличие в том, что для открытых множеств объединение любого набора (а не только конечного, как для замкнутых множеств) открыто.

4.13. Для каждой точки x открытого множества $X \subset \mathbb{R}$ рассмотрим все открытые интервалы, целиком лежащие в X, и рассмотрим точную нижнюю грань a их концов и точную верхнюю грань b. Точки a и b не совпадают, не принадлежат X и $(a, b) \subset X$. Полученные открытые интервалы попарно не пересекаются или совпадают. Множество различных таких интервалов не более чем счётно, так как каждый из них содержит рациональное число.

4.14. Согласно задаче 4.13 замкнутое множество на прямой является пересечением не более чем счётного набора множеств вида $\mathbb{R} \setminus (a, b)$. В свою очередь, каждое из этих множеств является пересечением счётного множества открытых множеств $\mathbb{R} \setminus [a - \frac{1}{n}, b + \frac{1}{n}]$.

4.15. Пусть $x, y \in \mathbb{R}$. Тогда $d(x, A) = \inf_{a \in A} |x - a| \leq |x - y| + \inf_{a \in A} |y - a| = |x - y| + d(y, A)$, т.е. $d(x, A) - d(y, A) \leq |x - y|$. Аналогично доказывается, что $d(y, A) - d(x, A) \leq |x - y|$. Следовательно, $|f(x) - f(y)| \leq |x - y|$, поэтому функция f непрерывна.

4.16. а) Если множество A замкнуто, то множество $\mathbb{R} \setminus A$ открыто. Поэтому для любой точки $x_0 \in \mathbb{R} \setminus A$ найдётся такое $\delta > 0$, что открытая δ-окрестность точки x_0 принадлежит множеству $\mathbb{R} \setminus A$. В таком случае $d(x, A) \geq \delta > 0$.

б) Замкнутое множество A является множеством нулей непрерывной функции $f(x) = d(x, A)$.

4.17. Любая точка x не лежит в одном из множеств A и B, поэтому $d(x, A) + d(x, B) > 0$. Положим $f(x) = \frac{d(x,A) - d(x,B)}{d(x,A) + d(x,B)}$. Эта функция непрерывна, $f(x) = 1$ для $x \in A$ и $f(x) = -1$ для $x \in B$.

4.18. Можно считать, что множество C неограниченное. Если, например, C ограничено сверху, то можно рассмотреть точку $x_0 = \sup\{x \mid x \in C\}$ и положить $f(x) = f(x_0)$ при $x > x_0$.

Согласно задаче 4.13 дополнение множества C является объединением не более чем счётного набора попарно не пересекающихся интервалов. Множество C неограниченное, поэтому все эти интервалы ограниченные. На концах каждого из этих интервалов (a_i, b_i) отображение g задано, и его можно продолжить на весь интервал как линейное отображение:

$$g(x) = \frac{f(b_n) - f(a_n)}{b_n - a_n}(x - a_n) + f(a_n)$$

для $x \in (a_n, b_n)$.

4.19. Несложно привести пример взаимно однозначного непрерывного отображения, переводящего замкнутое множество $\mathbb{R}$ в открытый интервал $(0, 1)$. Обратное отображение переводит ограниченное множество в неограниченное.

4.20. Множество значений функции, непрерывной на компактном множестве, компактно, поэтому оно замкнуто и ограничено. Из ограниченности следует, что точная верхняя грань этого множества ограничена, а из замкнутости следует, что точная верхняя грань принадлежит множеству.

4.21. Предположим, что множество $A \subset B$ всюду плотно в B. Пусть $b \notin A$. Тогда любая открытая окрестность точки b содержит как точку, принадлежащую A, так и точку, не принадлежащую A. Поэтому $b \in \partial A$. Таким образом, $B = A \cup \partial A = \overline{A}$.

Предположим, что $\overline{A} = B$. Пусть открытое в B множество U содержит точку $b \notin A$. Тогда $b \in \partial A$, поэтому открытая окрестность U точки b содержит точку, принадлежащую A.

4.22. Предположим, что у множества CA есть внутренняя точка x. Выберем открытое множество $U \ni x$, целиком лежащее в CA. Множество U не пересекается с A, поэтому множество A не является всюду плотным.

Предположим, что множество CA не имеет внутренних точек. Тогда непустое открытое множество U не может целиком лежать в CA, поэтому оно пересекается с A.

4.23. Пусть $f\colon X \to Y$ — сюръективное отображение, $A \subset X$ — всюду плотное множество, $V \subset Y$ — непустое открытое множество. Предположим, что $f(A) \cap V = \varnothing$. Тогда $f^{-1}(f(A) \cap V) = \varnothing$. Но $f^{-1}(f(A) \cap V) = f^{-1}(f(A)) \cap f^{-1}(V)$. Здесь множество $f^{-1}(f(A))$ содержит A, а множество $U = f^{-1}(V)$ непустое (это следует из сюръективности отображения f) и открытое. Поэтому $A \cap U \neq \varnothing$.

4.24. Пусть f — непрерывная функция на отрезке $[a, b]$, $A \subset [a, b]$ — всюду плотное множество. Фиксируем точку $x \in [a, b]$. Для любого натурального n можно выбрать точку $a_n \in A$ так, что $|x - a_n| < \frac{1}{n}$. Тогда $\lim_{n\to\infty} a_n = x$, поэтому $f(x) = \lim_{n\to\infty} f(a_n)$.

4.25. Пусть множество A нигде не плотно. Если открытое множество не пересекается с A, т.е. содержится в $B \setminus A$, то оно содержится и в $B \setminus \overline{A}$.

4.26. Пусть множества A_1 и A_2 нигде не плотны в B. Тогда существуют непустые открытые множества $U_1 \subset B \setminus A_1$ и $U_2 \subset U_1 \setminus A_2$. Поэтому $U_2 \subset B \setminus (A_1 \cup A_2)$.

4.27. Предположим, что x — предельная точка конечного множества X. Среди точек множества X, отличных от точки x, можно выбрать ближайшую к x точку x_1. Тогда открытая окрестность радиуса $|x - x_1|$ с центром x не содержит точек множества X, отличных от точки x.

4.28. Точка x принадлежит множеству $\overline{X \setminus \{x\}}$ тогда и только тогда, когда любая открытая окрестность точки x пересекает множество $X \setminus \{x\}$, т.е. содержит точку, отличную от x и принадлежащую X.

4.29. Предположим сначала, что $x \in \overline{X}$. Если $x \in X$, то доказывать нечего. Если $x \notin X$, то любая открытая окрестность точки x содержит точку, отличную от x и принадлежащую X. Это означает, что x — предельная точка множества X.

Рассмотрим теперь объединение множества X и всех его предельных точек. Если $x \in X$, то $x \in \overline{X}$. Если x — предельная точка множества X, то любая открытая окрестность точки x содержит точку, принадлежащую X (и даже отличную от x). Поэтому $x \in \overline{X}$.

4.30. Пусть x — предельная точка множества X', U — открытая окрестность точки x. Тогда U содержит точку $y \in X'$ (и даже $y \in X' \setminus \{x\}$). Открытая окрестность U точки y содержит точку $a \in X$, отличную от y. Если $a = x$, то рассмотрим открытую окрестность $U \setminus \{x\}$ точки y; она содержит точку $b \in X$, отличную от y и от x. В итоге получаем, что U содержит точку множества X, отличную от x, поэтому $x \in X'$.

4.31. Если $X \subset Y$, то $X' \subset Y'$, поэтому $X' \subset (\overline{X})'$. Докажем обратное включение. Пусть x — предельная точка множества $\overline{X}$, U — открытая окрестность точки x. Тогда U содержит точку $y \in \overline{X} \setminus \{x\}$. Открытая окрестность U точки y содержит точку $a \in X$. Если $a = x$, то рассмотрим открытую окрестность $U \setminus \{x\}$ точки y и найдём в U точку $b \in X$, отличную от x.

4.32. Рассмотрим множество E_n, состоящее из точки 0 и точек вида $\frac{1}{n_1} + \frac{1}{n_2} + \dots + \frac{1}{n_k}$, где n_1, $\dots$, n_k — натуральные числа и $k \leq n$. Легко проверить, что $E_n' = E_{n-1}$ при $n > 1$ и E_1' состоит из точки 0. Таким образом, n-е производное множество множества E_n состоит из точки 0, а $(n+1)$-е пусто.

4.33. Пусть множество X состоит из точек 0 и 1 и точек вида $\frac{1}{2^{n_1}} + \frac{1}{2^{n_2}} + \dots + \frac{1}{2^{n_k}}$, где $1 \leq n_1 < n_2 < \dots < n_k$. Тогда n-е производное множества X состоит из точек 0 и $\frac{1}{2^n}$ и точек вида $\frac{1}{2^{n_1}} + \frac{1}{2^{n_2}} + \dots + \frac{1}{2^{n_k}}$, где $n + 1 \leq n_1 < n_2 < \dots < n_k$.

4.34. Предположим, что множество X не является нигде не плотным, т.е. множество $\overline{X} = X \cup X'$ содержит некоторое непустое открытое множество U. Множество U не может содержать изолированных точек множества X, поэтому $U \subset X'$. Тогда множество U содержится во всех производных множествах множества X.

4.35. Множество X замкнуто тогда и только тогда, когда $X' \subset X$. Множество X не содержит изолированных точек тогда и только тогда, когда $X' \subset X$.

4.36. а) Пусть $A \subset B$ — открытое всюду плотное множества. Тогда любое открытое множество $U \subset B$ содержит некоторую точку a множества A. Множество A открытое, поэтому множество $U \cap A$ открытое. Следовательно, некоторая открытая окрестность W точки a содержится в $U \cap A$. Мы нашли требуемое непустое множество $W \subset U$, не пересекающееся с множеством CA.

б) Дополнение канторова множества — открытое всюду плотное множество.

4.37. Непустое совершенное множество X имеет предельные точки, поэтому оно не может быть конечным. Предположим, что множество X счётно и состоит из точек x_1, x_2, ... Чтобы прийти к противоречию, построим последовательность замкнутых ограниченных множеств

$$\overline{V_1} \supset \overline{V_2} \supset \dots \supset \overline{V_n} \supset \dots,$$

каждое из которых пересекается с X и $x_n \notin \overline{V_{n+1}}$. Тогда согласно теореме 4.6 пересечение замкнутых ограниченных множеств $F_i = X \cap \overline{V_i}$ непусто. С другой стороны, их пересечение содержится в X и не содержит ни одной из точек x_i.

На первом шаге в качестве $\overline{V_1}$ выберем замыкание произвольной r-окрестности V_1 точки x_1. Точка x_1 является предельной точкой множества X, поэтому в V_1 можно выбрать точку $y \in X$, отличную от точки x_1. Затем выбираем открытую окрестность V_2 точки y так, что $V_2 \subset \overline{V_1}$ и $x_1 \notin \overline{V_2}$.

На втором шаге, если $x_2 \neq y$, то выбираем открытую окрестность V_3 точки y так, что $V_3 \subset \overline{V_2}$ и $x_2 \notin \overline{V_3}$. Если же $x_2 = y$, то сначала в V_2 выбираем точку $z \in X$, отличную от точки x_2, а затем выбираем открытую окрестность V_3 точки z так, что $V_3 \subset \overline{V_2}$ и $x_2 \notin \overline{V_3}$. Все последующие шаги аналогичны второму шагу.

4.38. Предположим, что замкнутое множество не имеет изолированных точек. Тогда это множество совершенное (задача 4.35). Но непустое совершенное множество несчётно (задача 4.37).

4.39. Докажем, что дополнение множества P открыто. Пусть $x \notin P$. Тогда некоторая открытая ε-окрестность $D_\varepsilon(x)$ точки x содержит не более чем счётное множество точек из E. Пусть $y \in D_\varepsilon(x)$. Тогда некоторая открытая r-окрестность $D_r(y)$ содержится в $D_\varepsilon(x)$, а потому содержит не более чем счётное множество точек из E. Таким образом, $y \notin P$, т.е. $D_\varepsilon(x) \subset CP$.

4.40. Множество интервалов на прямой с рациональными концами счётно. Занумеруем их: V_1, V_2, ... Пусть W — объединение тех интервалов V_n, для которых множество $E \cap V_n$ не более чем счётно. Покажем, что $P = CW$.

Пусть сначала $x \in P$. Тогда любая открытая r-окрестность $D_r(x)$ точки x содержит несчётное множество точек E. Поэтому если множество $E \cap V_n$ не более чем счётно, то $x \notin V_n$, а значит, $x \notin W$. Пусть теперь $x \in CW$. Тогда если множество $E \cap V_n$ не более чем счётно, то $x \notin V_n$, а значит, если $x \in V_n$, то множество $E \cap V_n$ несчётно. Любая открытая r-окрестность $D_r(x)$ точки x содержит некоторое множество V_n, поэтому множество $E \cap D_r(x)$ тоже несчётно. Следовательно, $x \in P$.

Итак, $P = W$, где W — объединение не более чем счётного набора множество V_n, каждое из которых содержит не более чем счётное множество точек E. Поэтому множество $E \setminus P = CP \cap E = W \cap E$ не более чем счётно.

4.41. Если множество P непусто, то множество E несчётное. Согласно задаче 4.40 множество $E \setminus P$ не более чем счётное, поэтому множество P несчётное.

Согласно задаче 4.39 множество P замкнутое. Замкнутое множество без изолированных точек является совершенным (задача 4.35). Поэтому остаётся доказать, что множество P не имеет изолированных точек. Предположим, что I — изолированная точка множества P. Тогда существует ε-окрестность $D_\varepsilon(I)$ точки I, содержащая несчётное множество точек E, но не содержащая других точек P. Рассмотрим множество интервалов V_n с рациональными концами, как и в задаче 4.40.

Каждая точка множества $D_\varepsilon(I) \cap E$, отличная от точки I, не принадлежит P, поэтому её можно покрыть каким-либо из этих интервалов, содержащим не более чем счётное множество точек E. В результате получим, что множество $D_\varepsilon(I) \cap E$ не более чем счётно, а это противоречит тому, что оно несчётно.

4.42. Рассмотрим замкнутое несчётное множество E на прямой. Пусть P — множество его точек конденсации. Согласно задаче 4.40 множество $E \setminus P$ не более чем счётное, поэтому множество P несчётное. В частности, оно не пустое, поэтому согласно задаче 4.41 множество P совершенное.

Множество E замкнутое, поэтому оно содержит все предельные точки, а все точки конденсации являются также и предельными точками. Поэтому $P \subset E$. Представим множество E в виде

$$E = E \cap (P \cup CP) = (E \cap P) \cup (E \cap CP).$$

Но $P \subset E$, поэтому $E \cap P = P$, а значит, $E = P \cup (E \cap CP)$ — объединение совершенного множества и не более чем счётного множества.

4.43. Рассмотрим сначала непустое ограниченное совершенное множество P на прямой. Пусть a и b — точная нижняя и точная верхняя грани множества P. Точки a и b — предельные точки множества P, поэтому они принадлежат P. Предположим, что множество P не совпадает с отрезком $[a, b]$. Тогда существует точка $x \in [a, b]$, не принадлежащая P. Точка x разделяет множество P на два совершенных множества $P_1 \ni a$ и $P_2 \ni b$. Для множества P_1 находим точную верхнюю грань c, а для множества P_2 находим точную нижнюю грань d. Ясно, что $x \in (c, d)$. Для каждой точки $x \in [a, b]$, не принадлежащей P, находим такой интервал (c, d). Эти интервалы попарно не пересекаются и не имеют общих граничных точек. Каждый такой интервал содержит рациональную точку, поэтому их множество не более чем счётно. Удалив из отрезка $[a, b]$ все такие интервалы, получим множество P.

Рассмотрим теперь множество P, полученное удалением из отрезка $[a, b]$ не более чем счётного множества попарно не пересекающихся интервалов, не имеющих общих граничных точек. Множество P не пустое, потому что концы интервалов не удаляются. Дополнение множества P открыто, поскольку каждая точка отрезка $[a, b]$, не принадлежащая P, является внутренней точкой одного из удалённых отрезков. Изолированной точкой множества P могла бы быть только общая граничная точка удалённых отрезков, поэтому множество P не имеет изолированных точек. Согласно задаче 4.35 множество P совершенное.

4.44. Пусть $\overline{\lim}_{x\to x_0} f(x) = M(x_0)$. Для любого натурального n в $\frac{1}{n}$-окрестности точки x_0 найдётся точка x_n, для которой $M(f, x_0, \frac{1}{n}) - \frac{1}{n} < f(x_n) \le M(f, x_0, \frac{1}{n})$, поэтому $\lim_{n\to\infty} f(x_n) = M(x_0)$. Если же $B > M(x_0)$, то существует $\delta > 0$, для которого $M(f, x_0, \delta) < B$. Для всех x из δ-окрестности точки x_0 выполняются неравенства $f(x) \le M(f, x_0, \delta) < B$, поэтому не существует последовательности $x_n \to x_0$, для которой $f(x_n) \to B$.

4.45. Предположим сначала, что что функция $f(x)$ полунепрерывна сверху. Пусть x_0 — предельная точка множества E_c. Тогда любая ε-окрестность точки x_0 содержит некоторую точку множества E_c, поэтому $M(f, x_0, \varepsilon) \ge c$ для любого $\varepsilon > 0$. Следовательно, $f(x_0) \ge \overline{\lim}_{x\to x_0} f(x) \ge c$, т.е. $x_0 \in E_c$. Таким образом, множество E_c замкнуто.

Предположим теперь, что множество E_c замкнуто для любого c. Возьмём число c, для которого $c > f(x_0)$. Тогда множество E_c замкнуто и не содержит x_0. Поэтому для достаточно малого $\varepsilon > 0$ множество E_c не содержит ε-окрестность точки x_0. Следовательно, $\overline{\lim}_{x\to x_0} f(x) \le M(f, x_0, \varepsilon) \le c$ для любого числа c, для которого $c > f(x_0)$. Поэтому $\overline{\lim}_{x\to x_0} f(x) \le f(x_0)$.

4.46. а) Пусть χ_A — характеристическая функция множества A, $E_c = \{x | \chi_A(x) \ge c\}$. Тогда E_c — пустое множество, всё множество X или множество A. Поэтому функция χ_A полунепрерывна сверху тогда и только тогда, когда множество A замкнуто.

б) Множество $X \setminus A$ замкнуто тогда и только тогда, когда функция $\chi_{X\setminus A} = 1 - \chi_A$ полунепрерывна сверху, т.е. функция χ_A полунепрерывна снизу.

4.47. Функция $f(x)$ полунепрерывна снизу тогда и только тогда, когда множество $E_c(f) = \{x | f(x) \le c\}$ замкнуто для любого c (задача 4.45 с заменой функции f на $-f$). Последовательность функций f_n убывающая, поэтому $E_c(f) = \bigcap_{n=1}^{\infty} E_c(f_n)$.

4.48. Пусть m — точная нижняя грань значений функции f на отрезке $[a,b]$. Тогда существует последовательность точек $x_n \in [a,b]$, для которой $f(x_n) \to m$. Выберем из этой последовательности подпоследовательность, сходящуюся к некоторой точке $x_0 \in [a,b]$. Тогда $f(x_0) \le \underline{\lim}_{x \to x_0} f(x) \le \lim_{n\to\infty} f(x_n) = m$. Неравенство $f(x_0) < m$ невозможно, поэтому $f(x_0) = m$.

Глава 5.

Дифференцируемые функции

5.1. Определение производной

Производная функции $f(x)$ в точке x_0 — это предел $\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} = f'(x_0)$. Если этот предел существует, то говорят, что функция $f(x)$ *дифференцируема* в точке x_0.

Для функции $f(x)$, определённой на отрезке $[a,b]$, производные в точках a и b определяются как односторонние пределы.

Производную функции $g(x) = f'(x)$ называют *второй производной* или *производной второго порядка* функции $f(x)$ и обозначают $f''(x)$. Аналогично определяется производная третьего порядка $f'''(x)$ и т.д. Производную n-го порядка обозначают $f^{(n)}(x)$.

Теорема 5.1. *Пусть функция $f(x)$ дифференцируема в точке x_0, а прямая, проходящая через точки (x_0,y_0) и (x_1,y_1), где $y_0 = f(x_0)$ и $y_1 = f(x_1)$, задаётся уравнением $y - y_0 = k(x_1)(x - x_0)$. Тогда $\lim_{x_1\to x_0} k(x_1) = f'(x_0)$.*

Доказательство. Рассматриваемая прямая задаётся уравнением $y - y_0 = \frac{f(x_1)-f(x_0)}{x_1-x_0}(x - x_0)$, поэтому $k(x_1) = \frac{f(x_1)-f(x_0)}{x_1-x_0}$. Теперь непосредственно из определения производной видно, что $\lim_{x_1\to x_0} k(x_1) = f'(x_0)$. □

Прямую $y - f(x_0) = f'(x_0)(x - x_0)$ называют *касательной* к графику $y = f(x)$ в точке $(x_0, f(x_0))$.

Теорема 5.2. *Если функция $f(x)$ дифференцируема в точке x_0, то она непрерывна в этой точке.*

Доказательство. Запишем тождество

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

При этом $\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} = f'(x_0)$ и $\lim_{x\to x_0}(x - x_0) = 0$. Поэтому $\lim_{x\to x_0}\big(f(x) - f(x_0)\big) = f'(x_0)\cdot 0 = 0$. Это означает, что функция $f(x)$ непрерывна в точке x_0. □

Пример 5.1. *Функция $f(x) = |x|$ всюду непрерывна, но не дифференцируема в точке $x = 0$.*

Теорема 5.3. *Если функции $f(x)$ и $g(x)$ дифференцируемы в точке x_0, то:*
а) $(f+g)'(x_0) = f'(x_0) + g'(x_0)$;
б) $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$, где $(fg)(x) = f(x)g(x)$;
в) $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{(g(x_0))^2}$, если $g(x_0) \neq 0$.

Доказательство. а) Непосредственно следует из свойств предела суммы двух функций.
б) Пусть $h(x) = f(x)g(x)$. Тогда

$$h(x) - h(x_0) = f(x)\big(g(x) - g(x_0)\big) + g(x_0)\big(f(x) - f(x_0)\big).$$

Поделим обе части этого равенства на $x - x_0$ и заметим, что $f(x) \to f(x_0)$ при $x \to x_0$.

в) Пусть $h(x) = f(x)/g(x)$. Тогда

$$\frac{h(x)-h(x_0)}{x-x_0} = \frac{1}{g(x)g(x_0)}\left(g(x_0)\frac{f(x)-f(x_0)}{x-x_0} - f(x_0)\frac{g(x)-g(x_0)}{x-x_0}\right).$$

Устремляя x к x_0, получаем требуемое. □

Композицией функций f и g называют функцию $g \circ f(x) = g(f(x))$.

Теорема 5.4. *Пусть функция f дифференцируема в точке x_0, а функция g дифференцируема в точке $y_0 = f(x_0)$. Предположим, что у точки x_0 есть такая окрестность $U(x_0)$, что если x принадлежит $U(x_0)$ и $x \neq x_0$, то $f(x) \neq f(x_0)$. Тогда функция $g \circ f$ дифференцируема в точке x_0 и $(g \circ f)'(x_0) = g'[f(x_0)]f'(x_0)$.*

Доказательство. Тождество

$$\frac{g\circ f(x) - g\circ f(x_0)}{x-x_0} = \frac{g\circ f(x) - g\circ f(x_0)}{f(x)-f(x_0)}\cdot\frac{f(x)-f(x_0)}{x-x_0}$$

показывает, что функция $g \circ f$ дифференцируема в точке x_0 и $(g\circ f)'(x_0) = g'[f(x_0)]f'(x_0)$. □

Доказательство теоремы 5.4 получилось столь простым из-за предположения о том, что у точки x_0 есть такая окрестность $U(x_0)$, что если x принадлежит $U(x_0)$ и $x \neq x_0$, то $f(x) \neq f(x_0)$. В задаче 5.1 это предположение заменено другим.

Задача 5.1. Пусть функция f непрерывна на отрезке $[a,b]$, дифференцируема в точке $x \in (a,b)$, а функция g определена на некотором отрезке I, содержащем все значения функции f и дифференцируема в точке $y = f(x)$. Докажите, что функция $h(t) = g(f(t))$ дифференцируема в точке x и $h'(x) = g'[f(x)]f'(x)$.

Пусть функция $f(x)$ монотонно возрастает на отрезке $[a,b]$. Тогда каждой точке y отрезка $[f(a), f(b)]$ соответствует единственная точка x отрезка $[a,b]$, для которой $y = f(x)$. Поэтому можно определить *обратную* функцию $g(y) = x$.

Теорема 5.5. *Пусть $f'(x_0) \neq 0$, $g(y)$ — обратная к $f(x)$ функция. Тогда если $y = f(x_0)$, то $g'(y_0) = \frac{1}{f'(x_0)}$.*

Доказательство. Достаточно заметить, что $\frac{f(x)-f(x_0)}{x-x_0} = \frac{y-y_0}{f(y)-f(y_0)}$. □

5.2. Производные элементарных функций

Вычисление производных элементарных функций мы оставим в качестве задач.

Задача 5.2. Докажите, что $(x^n)' = nx^{n-1}$ для любого натурального n.

Задача 5.3. Докажите, что $(x^a)' = ax^{a-1}$ для любого вещественного a и положительного x.

Задача 5.4. Докажите, что $(\sin x)' = \cos x$ и $(\cos x)' = -\sin x$.

Задача 5.5. Докажите, что $(\operatorname{tg} x)' = \frac{1}{\cos^2 x}$ при $x \neq \frac{(2k+1)\pi}{2}$ и $(\operatorname{ctg} x)' = -\frac{1}{\sin^2 x}$ при $x \neq k\pi$.

Задача 5.6. Докажите, что $(a^x)' = a^x \ln a$ для $a > 0$.

Задача 5.7. Докажите, что $(\log_a x)' = \frac{1}{x \ln a}$.

Задача 5.8. Докажите, что $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$ и $(\operatorname{arctg} x)' = \frac{1}{1+x^2}$.

Теперь приведём несколько задач, связанных с применением производных элементарных функций.

Задача 5.9. Пусть функции $u(x)$ и $v(x)$ дифференцируемы, причём функция $u(x)$ положительна. Докажите, что функция $u^v = u(x)^{v(x)}$ дифференцируема и найдите её производную.

Задача 5.10. Вычислите производную функции $f(x) = x^{\sin x}$ (для $x > 0$).

Задача 5.11. Вычислите производную функции $f(x) = \cos^x a - \sin^x a$, где $0 < a < \pi/2$ — постоянный угол.

5.3. Производная многочлена и кратные корни

Корень x_0 многочлена $f(x)$ называют *кратным*, если $f(x) = (x - x_0)^2 g(x)$, где $g(x)$ — некоторый многочлен.

Теорема 5.6. *Многочлен $f(x)$ степени $n \geq 2$ имеет кратный корень тогда и только тогда, когда многочлены $f(x)$ и $f'(x)$ имеют общий корень.*

Доказательство. Предположим, что $f(x) = (x-x_0)^m g(x)$, где $m \geq 2$. Тогда многочлен $f'(x) = m(x - x_0)^{m-1} g(x) + (x - x_0)^m g'(x)$ имеет корень x_0.

Предположим, что $f(x) = (x - x_0)g(x)$, причём $g(x_0) \neq 0$. Тогда $f'(x) = g(x) + (x - x_0)g'(x)$, поэтому $f'(x_0) = g(x_0) \neq 0$. Таким образом, если все корни многочлена $f(x)$ имеют кратность 1, то они не являются корнями многочлена $f'(x)$. $\square$

Задача 5.12. Докажите, что многочлен

$$f_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

не имеет кратных корней.

Задача 5.13. Пусть $P(x)$ — многочлен с целыми коэффициентами. Докажите, что все коэффициенты его n-й производной $P^{(n)}(x)$ делятся на $n!$ для любого натурального n.

Задача 5.14. Докажите, что среднее арифметическое корней многочлена равно среднему арифметическому корней его производной.

Задача 5.15. Пусть f и g — многочлены степени n. Докажите, что $fg^{(n)} - f'g^{(n-1)} + f''g^{(n-2)} - f^{(3)}g^{(n-3)} + \cdots + (-1)^n f^{(n)} g$ — константа.

Задача 5.16. Пусть p и q — вещественные числа. Выясните, сколько вещественных корней имеет кубическое уравнение $x^3 + px + q = 0$ в зависимости от знаков числа p и дискриминанта $D = \frac{q^2}{4} + \frac{p^3}{27}$.

Задача 5.17. Пусть $f(x) = (x - x_1) \cdots (x - x_n)$, где числа $x_1, \ldots, x_n$ попарно различны и отличны от нуля. Докажите, что

$$\sum_{i=1}^{n} \frac{x_i^k}{f'(x_i)} = \begin{cases} 0 & \text{при } 0 \leq k \leq n-2; \\ 1 & \text{при } k = n-1. \end{cases}$$

Задача 5.18. Пусть $P(x) = (x-x_1) \cdots (x-x_n)$, где $x_1, \ldots, x_n$ — вещественные числа. Докажите, что $\bigl(P'(x)\bigr)^2 \geq P(x)P''(x)$ для всех вещественных x.

Задача 5.19. Докажите, что любой многочлен можно представить в виде разности двух монотонно возрастающих многочленов.

Задача 5.20. Докажите, что многочлен $P(x) = a_0 + a_1 x^{k_1} + a_2 x^{k_2} + \cdots + a_n x^{k_n}$ имеет не более n положительных корней.

Задача 5.21. *Функция $f(x)$ бесконечно дифференцируема на всей прямой, причём в каждой точке некоторая производная равна нулю. Докажите, что $f(x)$ — многочлен.

5.4. Касательная и нормаль

Касательная в точке $(x_0, f(x_0))$ к графику $y = f(x)$ дифференцируемой функции задаётся уравнением

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

Производная $f'(x_0)$ — это тангенс угла наклона касательной.

Секущая, проходящая через точки $(x_0, f(x_0))$ и $(x_1, f(x_1))$ задаётся уравнением

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

поэтому касательная — предельное положение секущей, когда $x_1 \to x_0$.

Задача 5.22. Касательная к кривой $y = e^x$ в точке (x_0, y_0) пересекает ось Ox в точке $(x_1, 0)$. Докажите, что разность $x_1 - x_0$ одна и та же для всех точек кривой.

Задача 5.23. На параболе, ось которой параллельна оси Oy, взяты точки A_1, A_2 и A_3. Пусть k_1 — тангенс угла наклона касательной в точке A_1, k_{ij} — тангенс угла наклона секущей A_iA_j. Докажите, что $k_1 = k_{12} + k_{13} - k_{23}$.

Нормаль к кривой $y = f(x)$ в точке (x_0, y_0) — это прямая, которая проходит через точку (x_0, y_0) перпендикулярно касательной в этой точке.

Задача 5.24. Докажите, что нормаль к кривой $y = f(x)$ в точке (x_0, y_0) задаётся уравнением

$$-f'(x_0)(y - y_0) = x - x_0.$$

Задача 5.25. Нормаль к параболе $y = x^2$ в точке (x_0, y_0) пересекает ось Oy в точке $(0, y_1)$. Докажите, что разность $y_1 - y_0$ постоянна для всех точек параболы.

5.5. Функции, дифференцируемые на отрезке

Значение функции, которое является наибольшим или наименьшим, называют *экстремальным*. Точку, в которой функция принимает экстремальное значение, называют *точкой экстремума*.

Теорема 5.7 (Ферма). *Пусть функция $f(x)$ определена на отрезке $[a, b]$, внутренняя точка x_0 этого отрезка — точка экстремума, и в точке x_0 существует производная. Тогда $f'(x_0) = 0$.*

Доказательство. Пусть для определённости $f(x_0) \leq f(x)$ для всех x из отрезка $[a, b]$. Рассмотрим односторонние пределы

$$\lim_{x \to x_0 + 0} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{и} \quad \lim_{x \to x_0 - 0} \frac{f(x) - f(x_0)}{x - x_0}.$$

В обоих пределах числитель неотрицателен. При этом в первом пределе знаменатель положителен, а во втором отрицателен. Значит, первый предел неотрицателен, а второй неположителен. Но оба предела равны $f'(x_0)$. $\square$

Историческое замечание. Общий метод нахождения максимумов и минимумов Пьер Ферма (1601-1665) разработал в 1629 году.

Теорема 5.8 (Ролль). *Пусть функция $f(x)$ дифференцируема на отрезке $[a, b]$, причём $f(a) = f(b)$. Тогда существует внутренняя точка x_0 этого отрезка, для которой $f'(x_0) = 0$.*

Доказательство. Функция $f(x)$ непрерывна на отрезке $[a, b]$, поэтому по теореме Вейерштрасса (теорема 3.7) среди её значений есть наибольшее M и наименьшее m. Если $M = m$, то функция $f(x)$ постоянна, поэтому в качестве x_0 можно взять любую внутреннюю точку отрезка. Если же $M > m$, то одно из этих двух значений достигается не в конце отрезка, потому что по условию $f(a) = f(b)$. Значит, в некоторой внутренней точке x_0 отрезка $[a, b]$ функция $f(x)$ достигает наибольшего или наименьшего значения, поэтому по теореме Ферма (теорема 5.7) $f'(x_0) = 0$. $\square$

Историческое замечание. Мишель Ролль (1652-1719) опубликовал теорему 5.8 для многочленов в 1691 году. Его подход был чисто алгебраический, он был убеждён, что методы анализа бесконечно малых неизбежно приведут к ошибкам.

Теорема 5.9 (Лагранж). *Пусть функция $f(x)$ дифференцируема на отрезке $[a, b]$. Тогда существует внутренняя точка x_0 этого отрезка, для которой*

$$f'(x_0) = \frac{f(a) - f(b)}{a - b}.$$

Теорему Лагранжа, записанную в виде $f(b) - f(a) = f'(x_0)(b - a)$, часто называют *формулой конечных приращений* или *теоремой о среднем значении*

Доказательство. Рассмотрим вспомогательную функцию $F(x) = f(x) - \frac{f(a)-f(b)}{a-b}(x-a)$. Эта функция дифференцируема на отрезке $[a,b]$ и $F(a) = F(b) = f(a)$. Поэтому к функции $F(x)$ можно применить теорему Ролля (теорема 5.8). В результате получим, что существует внутренняя точка x_0 отрезка $[a,b]$, для которой $F'(x_0) = 0$, т.е. $f'(x_0) - \frac{f(a)-f(b)}{a-b} = 0$. □

Историческое замечание. Жозеф-Луи Лагранж (1736-1813) получил формулу конечных приращений в 1797 году как частный случай остаточного члена для формулы Тейлора.

Задача 5.26. Функция $f(x)$ дифференцируема на отрезке $[a,b]$, причём $f'(x) = 0$ для всех точек x отрезка $[a,b]$. Докажите, что функция $f(x)$ постоянна на отрезке $[a,b]$.

Задача 5.27. Функция $f(x)$ дифференцируема на отрезке $[a,b]$.

а) Докажите, что эта функция неубывающая (на этом отрезке) тогда и только тогда, когда $f'(x) \geq 0$ для любой точки x интервала (a,b).

б) Докажите, что если $f'(x) \geq 0$ для любой точки x интервала (a,b) и не существует отрезка $[p,q]$, содержащегося в $[a,b]$, во всех точках которого f' обращается в нуль, то функция $f(x)$ возрастающая.

Задача 5.28. Функции $f(x)$ и $g(x)$ дифференцируемы на отрезке $[a,b]$. Докажите, что если $f(a) = g(a)$ и $f'(x) > g'(x)$ для любой точки x интервала (a,b), то $f(x) > g(x)$ для любой точки x интервала (a,b).

Задача 5.29. Докажите, что если $x > 0$, то $\cos x > 1 - \frac{x^2}{2}$ и $\sin x > x - \frac{x^3}{6}$.

Задача 5.30. Докажите, что если $0 < x < \pi/2$, то $\operatorname{tg} x > x + \frac{1}{3}x^3$.

Замечание. Доказательства неравенств из задач 5.29 и 5.30 основаны на том, что из неравенства для производных следует неравенство для функций. Другими словами, из неравенства для функций следует неравенство для первообразных (интегралов). Подход с применением интегралов (по сути дела, эквивалентный) в некотором смысле более естествен: чтобы получить неравенство, нужно вычислить интеграл. Поэтому здесь мы привели только два неравенства. Более подробно этот метод доказательства неравенств обсуждается в §6.13.

Задача 5.31. а) Пусть $0 < \alpha < 1$ и $x \geq 0$. Докажите, что $x^\alpha - \alpha x \leq 1 - \alpha$.

б) Пусть a, b, p и q — положительные числа, причём $\frac{1}{p} + \frac{1}{q} = 1$. Докажите, что $ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$.

Задача 5.32. Функция $f(x)$ дифференцируема на отрезке $[a,b]$, причём $f''(x) > 0$ для всех $x \in (a,b)$. Докажите, что эта функция выпуклая.

Задача 5.33. Функция f дифференцируема на отрезке $[a,b]$, причём для некоторой константы c для всех $x \in [a,b]$ выполняется неравенство $|f'(x)| \leq c|f(x)|$. Докажите, что если $f(x_0) = 0$ для некоторой точки $x \in [a,b]$, то $f(x) \equiv 0$ на $[a,b]$.

Теорема 5.10 (Коши). *Пусть функции $f(x)$ и $g(x)$ дифференцируемы на отрезке $[a,b]$, причём производная $g'(x)$ не обращается в нуль во внутренних точках этого отрезка. Тогда существует внутренняя точка x_0 отрезка $[a,b]$, для которой*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

Доказательство. Рассмотрим вспомогательную функцию

$$F(x) = \big(f(b)-f(a)\big)\big(g(x)-g(a)\big) - \big(f(x)-f(a)\big)\big(g(b)-g(a)\big).$$

Ясно, что

$$F'(x) = \big(f(b)-f(a)\big)g'(x) - f'(x)\big(g(b)-g(a)\big)$$

и $F(a) = F(b) = 0$. Поэтому к функции $F(x)$ можно применить теорему Ролля (теорема 5.8). В результате получим, что существует внутренняя точка x_0, для которой

$$\big(f(b)-f(a)\big)g'(x_0) - f'(x_0)\big(g(b)-g(a)\big) = 0. \tag{1}$$

По условию $g'(x_0) \neq 0$. Легко также видеть, что $g(b) - g(a) \neq 0$, поскольку иначе по теореме Ролля нашлась бы внутренняя точка x_1, для которой $g'(x_1) = 0$. Поэтому равенство (1) можно поделить на $g'(x_0)\big(g(b)-g(a)\big)$ и получить требуемое. □

5.6. Неравенства

Приведём ещё несколько задач на доказательство неравенств с помощью производных.

Задача 5.34. Докажите, что если $0 < \alpha < \beta < \pi/2$, то $\alpha \sin\beta < \beta \sin\alpha$.

Задача 5.35. Докажите, что если $0 < \alpha < \beta < \pi/2$, то $\alpha \operatorname{tg}\beta > \beta \operatorname{tg}\alpha$.

Задача 5.36. Докажите, что если $0 < \alpha < \pi/2$, то $2\sin\alpha + \operatorname{tg}\alpha > 3\alpha$.

Задача 5.37. а) Докажите, что $e^x > 1 + x$ для любого $x \neq 0$.
б) Докажите, что $\frac{1}{n+1} < \ln\frac{n+1}{n} < \frac{1}{n}$ для любого натурального n.

Задача 5.38. Пусть $x > 0,\ \ x \neq 1$. Докажите, что: а) $\ln x < x - 1$; б) $\ln x > \frac{x-1}{x}$.

Задача 5.39. Докажите, что $\ln x < n(x^{1/n} - 1) < x^{1/n}\ln x$ для любого положительного числа $x \neq 1$.

Задача 5.40. Докажите, что $\lim_{n\to\infty} n\ln\left(1 + \frac{x}{n}\right) = x$ при $x > 0$.

Задача 5.41. Докажите, что $e^x > x^e$ для любого положительного $x \neq e$.

Задача 5.42. Пусть a и b — положительные числа. Докажите, что $b\cdot 2^a + a\cdot 2^{-b} \geq a + b$.

Задача 5.43. Пусть $a > b > 0$. Докажите, что

$$\sqrt{ab} < \frac{a-b}{\ln a - \ln b} < \frac{a+b}{2}.$$

Задача 5.44. *Докажите, что

$$\sin x + \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x + \cdots + \frac{1}{n}\sin nx > 0$$

при $0 < x < \pi$.

Задача 5.45. Пусть $0 < x < \pi/4$. Докажите, что

$$(\cos x)^{\cos^2 x} > (\sin x)^{\sin^2 x} \quad \text{и} \quad (\cos x)^{\cos^4 x} < (\sin x)^{\sin^4 x}.$$

Задача 5.46. Докажите, что если $x > -1$ и $x \neq 0$, то

$$\frac{2|x|}{2+x} < |\ln(1+x)| < \frac{|x|}{\sqrt{1+x}}.$$

5.7. Правило Лопиталя

Правило Лопиталя позволяет вычислять некоторые пределы с помощью производной.

Теорема 5.11 (правило Лопиталя). *Пусть функции $f(x)$ и $g(x)$ удовлетворяют условиям теоремы Коши (теорема 5.10) и, кроме того, $f(a) = g(a) = 0$. Тогда если $\lim_{x\to a+0}\frac{f'(x)}{g'(x)} = A$, то $\lim_{x\to a+0}\frac{f(x)}{g(x)} = A$.*

Доказательство. Фиксируем точку x, где $a < x \leq b$, и применим теорему Коши к отрезку $[a, x]$. В результате получим, что внутри этого отрезка есть точка x_1, для которой

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(x_1)}{g'(x_1)}.$$

Если $x \to a$, то $x_1 \to a$. Из этого следует требуемое. $\square$

Историческое замечание. Гийом де Лопиталь (1661-1704) привёл правило нахождения пределов функций в своём учебнике анализа, изданном в 1696 году. Но этот учебник составлен на основе лекций, которые ему читал Иоганн Бернулли (1667-1748). В 1692 году Бернулли уже знал это правило.

Задача 5.47. Вычислите с помощью правила Лопиталя $\lim_{x\to 0} \frac{x-\sin x}{x^3}$.

Задача 5.48. Вычислите предел $\lim_{x\to a} \frac{a^x-x^a}{x-a}$.

Задача 5.49. Вычислите предел $\lim_{x\to 0} \frac{\operatorname{tg} x-x}{x-\sin x}$.

Задача 5.50. Пусть $f(x)$ и $g(x)$ — многочлены, не имеющие общих корней, причём $\deg f < \deg g$ и многочлен $g(x)$ не имеет кратных корней. Докажите, что

$$\frac{f(x)}{g(x)} = \sum_{i=1}^{n} \frac{A_i}{x-a_i},$$

где $a_1, \dots a_n$ — корни многочлена g и $A_i = f(a_i)/g'(a_i)$.

5.8. Алгебраические и трансцендентные функции

Функцию $f(x)$ называют *алгебраической*, если существуют многочлены $P_0(x), \dots, P_n(x)$, для которых

$$P_0(x)\bigl(f(x)\bigr)^n + P_1(x)\bigl(f(x)\bigr)^{n-1} + \dots + P_n(x) = 0,$$

причём многочлен $P_0(x)$ не равен тождественно нулю. В противном случае функцию $f(x)$ называют *трансцендентной*.

Задача 5.51. Докажите, что функция $f(x) = \sin x$ трансцендентная.

Задача 5.52. Докажите, что функция $f(x) = e^x$ трансцендентная.

5.9. Формула Тейлора

Проще всего формула Тейлора выглядит для многочленов; сформулируем соответствующее утверждение в виде задачи.

Задача 5.53. а) Пусть a — фиксированное число. Докажите, что любой многочлен $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ можно записать в виде

$$f(x) = A_0 + A_1(x-a) + A_2(x-a)^2 + \dots + A_n(x-a)^n,$$

где $A_0, A_1, \dots, A_n$ — константы.

б) Докажите, что $A_0 = f(a)$, $A_1 = f'(a)$, $A_2 = \frac{f''(a)}{2!}, \dots, A_n = \frac{f^{(n)}(a)}{n!}$.

Теорема 5.12 (формула Тейлора). *Пусть a — фиксированное число, $f(x)$ — функция, имеющая производные до порядка $n+1$ включительно для любого x между a и b (для некоторого b). Тогда если число x заключено между a и b и*

$$T(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

то

$$f(x) - T(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!}(x-a)^{n+1}$$

для некоторого θ между a и x.

Доказательство. Ясно, что $T(a) = f(a)$. Далее,

$$\begin{aligned} T'(x) &= f'(a) + \frac{f''(a)}{1!}(x-a) + \dots + \frac{f^{(n)}(a)}{(n-1)!}(x-a)^{n-1},\\ T''(x) &= f''(a) + \frac{f'''(a)}{1!}(x-a) + \dots + \frac{f^{(n)}(a)}{(n-2)!}(x-a)^{n-2},\\ &\dots\dots\\ T^{(n)}(x) &= f^{(n)}(a),\\ T^{(n+1)}(x) &= 0. \end{aligned}$$

Поэтому $T'(a) = f'(a)$, $T''(a) = f''(a)$, $\dots$, $T^{(n)}(a) = f^{(n)}(a)$ и $T^{(n+1)}(x) = 0$.

Рассмотрим вспомогательную функцию $\varphi(x) = f(x) - T(x)$. Для неё $\varphi(a) = \varphi'(a) = \varphi''(a) = \dots = \varphi^{(n)}(a) = 0$ и $\varphi^{(n+1)}(x) = f^{(n+1)}(x)$. Рассмотрим ещё вспомогательную функцию $\psi(x) = \frac{(x-a)^{n+1}}{(n+1)!}$. Для неё $\psi(a) = \psi'(a) = \psi''(a) = \dots = \psi^{(n)}(a) = 0$ и $\psi^{(n+1)}(x) = 1$. Более того, ни сама функция ψ, ни её производные до $(n+1)$-й включительно не обращаются в нуль в точках, отличных от a.

Фиксируем точку $x \neq a$. Из равенств $\varphi(a) = \psi(a) = 0$ следует, что $\frac{\varphi(x)}{\psi(x)} = \frac{\varphi(x)-\varphi(a)}{\psi(x)-\psi(a)}$. Поэтому по теореме Коши (теорема 5.10) существует точка x_1 между a и x, для которой $\frac{\varphi(x)}{\psi(x)} = \frac{\varphi'(x_1)}{\psi'(x_1)}$. Из равенств $\varphi'(a) = \psi'(a) = 0$ следует, что $\frac{\varphi'(x)}{\psi'(x)} = \frac{\varphi'(x)-\varphi'(a)}{\psi'(x)-\psi'(a)}$. Поэтому по теореме Коши существует точка x_2 между a и x_1 (а значит, между a и x), для которой $\frac{\varphi'(x_1)}{\psi'(x_1)} = \frac{\varphi''(x_2)}{\psi''(x_2)}$. Продолжая эти рассуждения, получаем

$$\frac{\varphi(x)}{\psi(x)} = \frac{\varphi'(x_1)}{\psi'(x_1)} = \frac{\varphi''(x_2)}{\psi''(x_2)} = \dots = \frac{\varphi^{(n+1)}(x_{n+1})}{\psi^{(n+1)}(x_{n+1})},$$

где точка x_{n+1} лежит между a и x.

Напомним, что $\varphi^{(n+1)}(x) = f^{(n+1)}(x)$ и $\psi^{(n+1)}(x) = 1$. Пусть $\theta = x_{n+1}$. Тогда $\frac{\varphi(x)}{\psi(x)} = f^{(n+1)}(\theta)$, т.е. $f(x) - T(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!}(x-a)^{n+1}$. $\square$

Разность $f(x) - T(x)$ называют *остаточным членом*. Остаточный член вида $\frac{f^{(n+1)}(\theta)}{(n+1)!}(x-a)^{n+1}$ называют *остаточным членом в форме Лагранжа*.

Историческое замечание. Брук Тейлор (1685-1731) пришёл к представлению функций с помощью формулы Тейлора в 1715 году. Но Джеймс Грегори (1638-1675) уже знал эту формулу в 1671 году и получил разложения нескольких важных функций. Жозеф-Луи Лагранж (1736-1813) получил остаточный член в формуле Тейлора в 1797 году. Он первым оценил точность приближения функции суммой конечного числа членов ряда Тейлора.

Задача 5.54. а) Докажите, что разность между $\sin x$ и $x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ не превосходит $\frac{|x|^{2n+2}}{(2n+2)!}$.

б) Докажите, что разность между $\cos x$ и $1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!}$ не превосходит $\frac{|x|^{2n+1}}{(2n+1)!}$.

в) Докажите, что разность между e^x и $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!}$ не превосходит $e^x \frac{|x|^{n+1}}{(n+1)!}$.

Замечание. Приводимые в задаче 5.54 следствия из формулы Тейлора можно получить и более простыми средствами. По этому поводу см. §6.13.

Историческое замечание. Формулы $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$ и $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$ предложил Исаак Ньютон (1643-1727) в 1669 году.

Задача 5.55. Докажите, что если $(n+1)$-я производная функции f тождественно равна нулю, то f — многочлен степени не выше n.

5.10. Равномерная сходимость дифференцируемых функций

Начнём с двух задач.

Задача 5.56. Докажите, что поточечно сходящаяся последовательность функций, дифференцируемых на отрезке, может сходится к не дифференцируемой функции.

Задача 5.57. Пусть последовательность дифференцируемых функций f_n сходится на отрезке к дифференцируемой функции f. Докажите, что при этом последовательность f_n' может не сходиться к f'.

Но при определённых условиях последовательность дифференцируемых функций сходится к дифференцируемой функции и при этом последовательность производных сходится к производной.

Теорема 5.13. *Если последовательность непрерывно дифференцируемых функций f_n на интервале (a, b) сходится в некоторой точке x_0, а последовательность их производных f_n' равномерно сходится на этом интервале, то последовательность функций f_n равномерно сходится на интервале к непрерывно дифференцируемой функции f, причём $f'(x) = \lim_{n\to\infty} f'(x)$ для всех $x \in (a, b)$*

Доказательство. Фиксируем точку $c \in (a, b)$ и положим $g_n(x) = f_n'(c)$ при $x = c$ и $g_n(x) = \frac{f_n(x)-f_n(c)}{x-c}$ при $x \neq c$. Тогда

$$f_n(x) = f_n(c) + (x - c)g_n(x) \tag{5.1}$$

для всех $x \in (a, b)$. Покажем, что для любого c последовательность g_n сходится равномерно на интервале (a, b). Пусть задано $\varepsilon > 0$. Согласно теореме о среднем значении (теорема 5.9) для точки $x \in (a, b)$, $x \neq c$, можно выбрать точку ξ между x и c так, что

$$g_n(x) - g_m(x) = \frac{f_n(x) - f_m(x) - (f_n(c) - f_m(c))}{x - c} = f_n'(\xi) - f_m'(\xi).$$

Последовательность f_n' сходится равномерно на (a, b), поэтому можно выбрать N так, что если $n, m \geq N$, то $|g_n(x) - g_m(x)| < \varepsilon$. Для $x = c$ это неравенство тоже выполняется, поскольку $g_n(c) = f_n'(c)$. Равномерная сходимость последовательности g_n доказана.

Докажем теперь, что последовательность f_n сходится равномерно на интервале (a, b). Запишем равенство (5.1) для $c = x_0$:

$$f_n(x) = f_n(x_0) + (x - x_0)g_n(x).$$

Из этого равенства следует равномерная сходимость последовательности f_n, поскольку последовательность чисел $f_n(x_0)$ сходится (по условию) и последовательность функций g_n сходится равномерно.

Снова фиксируем точку $c \in (a, b)$ и положим $f(x) = \lim_{n\to\infty} f_n(x)$ и $g(x) = \lim_{n\to\infty} g_n(x)$. Требуется доказать, что

$$f'(c) = \lim_{n\to\infty} f_n'(c). \tag{5.2}$$

Последовательность g_n сходится равномерно и каждая функция g_n непрерывна в точке c, поэтому функция g непрерывна в точке c. Кроме того, $g_n(c) = f_n'(c)$. Следовательно, правую часть равенства (5.2) можно записать в следующем виде:

$$\lim_{n\to\infty} f_n'(c) = \lim_{n\to\infty} g_n(c) = g(c) = \lim_{x\to c} g(x).$$

Ясно также, что

$$\frac{f(x) - f(c)}{x - c} = \lim_{n\to\infty} \frac{f_n(x) - f_n(c)}{x - c} = \lim_{n\to\infty} g_n(x) = g(x).$$

Поэтому левую часть равенства (5.2) можно записать в следующем виде:

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

□

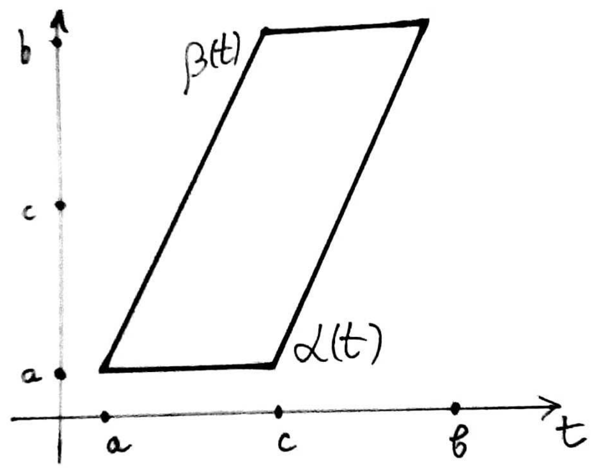

Рис. 5.1.

5.11. Промежуточные значения производной

Производная всюду дифференцируемой функции может не быть непрерывной. Рассмотрим, например, функцию $f(x) = x^2 \sin\frac{1}{x}$ при $x \neq 0$, $f(0) = 0$. Если $x \neq 0$, то $f'(x) = 2x\sin\frac{1}{x} + x^2\left(\frac{1}{x}\right)'\cos\frac{1}{x} = 2x\sin\frac{1}{x} - \cos\frac{1}{x}$. Для вычисления производной в нуле можно воспользоваться непосредственно определением производной. Ясно, что $\left|\frac{f(t)-f(0)}{t-0}\right| = \left|t\sin\frac{1}{t}\right| \leq |t|$ при $t \neq 0$, поэтому $f'(0) = 0$. Таким образом, функция $f'(x)$ всюду определена, но в точке 0 она не непрерывна, поскольку функция $\cos\frac{1}{x}$ не имеет предела при $x \to 0$.

Тем не менее, производная функции, дифференцируемой на отрезке, принимает все промежуточные значения.

Теорема 5.14 (Дарбу). *Пусть функция $f(x)$ дифференцируема на отрезке $[a, b]$ и $f'(a) < \lambda < f'(b)$. Тогда $f'(x) = \lambda$ для некоторой точки $x \in (a, b)$.*

Аналогичное утверждение верно и в том случае, когда $f'(a) > f'(b)$.

Доказательство. Пусть точка c — середина отрезка $[a, b]$. Рассмотрим на отрезке $[a, b]$ две непрерывные кусочно линейные функции $\alpha(t)$ и $\beta(t)$: если $a \leq t \leq c$, то $\alpha(t) = a$ и $\beta(t) = 2t - a$, а если $c \leq t \leq b$, то $\alpha(t) = 2t - b$ и $\beta(t) = b$ (рис. 5.1). Ясно, что при $a < t < b$ выполняются неравенства $a \leq \alpha(t) < \beta(t) \leq b$. Поэтому можно рассмотреть непрерывную функцию

$$g(t) = \frac{f(\beta(t)) - f(\alpha(t))}{\beta(t) - \alpha(t)}$$

на интервале (a, b). Ясно, что $g(t) \to f'(a)$ при $t \to a$ и $g(t) \to f'(b)$ при $t \to b$. Поэтому функцию $g(t)$ можно доопределить до функции, непрерывной на отрезке $[a, b]$, и из теоремы о промежуточном значении (теорема 3.5) следует, что на интервале (a, b) существует точка t_0, для которой $g(t_0) = \lambda$, т.е.

$$\frac{f(\beta(t_0)) - f(\alpha(t_0))}{\beta(t_0) - \alpha(t_0)} = \lambda.$$

Поэтому по формуле конечных приращений (теорема 5.9) существует точка $x \in (\alpha(t_0), \beta(t_0))$, для которой $f'(x) = \lambda$. $\square$

Историческое замечание. Гастон Дарбу (1842-1917) доказал теорему 5.14 в 1875 году.

5.12. Многочлены Чебышева

Определение многочленов Чебышева основано на том, что $\cos n\varphi$ полиномиально выражается через $\cos\varphi$, т.е. существует такой многочлен $T_n(x)$, что $T_n(x) = \cos n\varphi$ при $x = \cos\varphi$. В самом деле, можно положить $T_0(x) = 1$ и $T_1(x) = x$, а затем воспользоваться формулой

$$\cos(n+1)\varphi + \cos(n-1)\varphi = 2\cos\varphi\cos n\varphi.$$

Эта формула показывает, что многочлены $T_n(x)$, определённые при $n \geq 1$ рекуррентным соотношением

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$

обладают требуемым свойством. Эти многочлены $T_n(x)$ называют *многочленами Чебышева.*

Непосредственно из того, что $T_n(x) = \cos n\varphi$ при $x = \cos\varphi$, следует, что $|T_n(x)| \leq 1$ при $x \leq 1$. А из рекуррентного соотношения следует, что $T_n(x) = 2^{n-1}x^n + a_1x^{n-1} + \cdots + a_n$, где $a_1, \ldots, a_n$ — целые числа.

Наиболее важное свойство многочленов Чебышева заключается в том, что многочлен $\frac{1}{2^{n-1}}T_n(x)$ — наименее уклоняющийся от нуля на интервале $[-1, 1]$ многочлен степени n со старшим коэффициентом 1. Это свойство тесно связано с другим важным свойством многочленов Чебышева: $T_n\big(\cos(k\pi/n)\big) = \cos k\pi = (-1)^k$ при $k = 0, 1, \ldots, n$, т.е. все локальные максимумы функции T_n равны 1, все локальные минимумы равны -1, и общее количество локальных максимумов и минимумов равно $n+1$.

Теорема 5.15. *Пусть $P_n(x) = x^n + \cdots$ — многочлен степени n со старшим коэффициентом 1, причём $|P_n(x)| \leq \frac{1}{2^{n-1}}$ при $|x| \leq 1$. Тогда $P_n(x) = \frac{1}{2^{n-1}}T_n(x)$.*

Доказательство. Рассмотрим многочлен $Q(x) = \frac{1}{2^{n-1}}T_n(x) - P_n(x)$. Его степень не превосходит $n-1$, поскольку старшие члены многочленов $\frac{1}{2^{n-1}}T_n(x)$ и $P_n(x)$ равны. Из того, что $|P_n(x)| \leq \frac{1}{2^{n-1}}$ при $|x| \leq 1$, следует, что в точке $x_k = \cos(k\pi/n)$ либо $Q(x_k) = 0$, либо знак числа $Q(x_k)$ совпадает со знаком числа $T_n(x_k)$.

Предположим сначала, что $Q(x_k) \neq 0$ для всех k. В таком случае в концах каждого отрезка $[x_{k+1}, x_k]$ многочлен $Q(x)$ принимает значения разного знака, поэтому у многочлена $Q(x)$ на этом отрезке есть корень. Количество отрезков $[x_{k+1}, x_k]$ равно n, поэтому многочлен $Q(x)$ имеет по крайней мере n корней. Для многочлена степени не более $n-1$ это означает, что он тождественно равен нулю, т.е. $P_n(x) = \frac{1}{2^{n-1}}T_n(x)$.

Предположим теперь, что $Q(x_k) = 0$ для некоторого k. В этом случае либо x_k — кратный корень многочлена Q, либо в окрестности точки x_k многочлен Q принимает как положительные, так и отрицательные значения. Во втором случае заменим точку x_k на близкую к ней точку x'_k, для которой знак числа $Q(x'_k)$ такой же, как у числа $(-1)^k$. Мы снова получаем n отрезков, каждому из которых соответствует по крайней мере один корень многочлена Q. Действительно, либо хотя бы в одном из концов отрезка расположен кратный корень многочлена Q, либо в концах отрезка многочлен Q принимает значения разного знака. В первом случае каждому из отрезков с концом в кратном корне можно сопоставить свой корень многочлена Q. □

Историческое замечание. Пафнутий Львович Чебышев (1821-1894) ввёл многочлены Чебышева и доказал теорему 5.15 в 1854 году.

5.13. Интерполяционные многочлены Лагранжа и Эрмита

Пусть $x_1, \ldots, x_{n+1}$ — попарно различные точки прямой. Тогда существует ровно один многочлен $P(x)$ степени не выше n, принимающий в точке x_i заданное значение a_i. Действительно, единственность многочлена P следует из того, что разность двух таких многочленов обращается в нуль в точках $x_1, \ldots, x_{n+1}$ и имеет при этом степень не выше n. Ясно также, что следующий многочлен обладает всеми требуемыми свойствами:

$$\begin{aligned} P(x) &= \sum_{k=1}^{n+1} a_k \frac{(x-x_1)\cdots(x-x_{k-1})\cdot(x-x_{k+1})\cdots(x-x_{n+1})}{(x_k-x_1)\cdots(x_k-x_{k-1})\cdot(x_k-x_{k+1})\cdots(x_k-x_{n+1})} = \\ &= \sum_{k=1}^{n+1} a_k \frac{\omega(x)}{(x-x_k)\omega'(x_k)}, \end{aligned}$$

где $\omega(x)=(x-x_1)\cdot\dots\cdot(x-x_{n+1})$. Этот многочлен $P(x)$ называют *интерполяционным многочленом Лагранжа*.

Пусть $x_1,\dots,x_n$ — попарно различные точки прямой, $\alpha_1,\dots,\alpha_n$ — натуральные числа, сумма которых равна $m+1$. Предположим, что в каждой точке x_i заданы числа $y_i^{(0)},y_i^{(1)},\dots,y_i^{(\alpha_i-1)}$. Тогда существует единственный многочлен $H_m(x)$ степени не выше m, для которого выполняются равенства

$$H_m(x_i)=y_i^{(0)},H'_m(x_i)=y_i^{(1)},\dots,H_m^{(\alpha_i-1)}(x_i)=y_i^{(\alpha_i-1)},$$

$i=1,\dots,n$. Иными словами, в точке x_i многочлен H_m имеет заданные значения производных до порядка α_i-1 включительно. Такой многочлен H_m называют *интерполяционным многочленом Эрмита*.

Единственность интерполяционного многочлена Эрмита достаточно очевидна. Действительно, если $G(x)$ — разность двух интерполяционных многочленов Эрмита, то $\deg G\le m$ и $G(x)$ делится на $(x-x_1)^{\alpha_1}\cdot\dots\cdot(x-x_n)^{\alpha_n}$.

Пусть $\Omega(x)=(x-x_1)^{\alpha_1}\cdot\dots\cdot(x-x_n)^{\alpha_n}$. Чтобы построить интерполяционный многочлен Эрмита, достаточно указать многочлены $\varphi_{ik}(x)$ ($i=1,\dots,n$ и $k=0,1,\dots,\alpha_i-1$), обладающие следующими свойствами:

1) $\deg\varphi_{ik}\le m$;

2) $\varphi_{ik}(x)$ делится на многочлен $\Omega(x)/(x-x_i)^{\alpha_i}$, т.е. $\varphi_{ik}(x)$ делится на $(x-x_j)^{\alpha_j}$ при $j\ne i$;

3) разложение $\varphi_{ik}(x)$ по степеням $(x-x_i)$ начинается с $\frac{1}{k!}(x-x_i)^k+(x-x_i)^{\alpha_i}$.

Действительно, $\varphi_{ik}^{(0)}(x_j)=\dots=\varphi_{ik}^{(\alpha_i-1)}(x_j)=0$ при $j\ne i$, $\varphi_{ik}^{(k)}(x_i)=1$ и $\varphi_{ik}^{(l)}(x_i)=0$ при $0\le l\le\alpha_i-1$, $l\ne k$. Поэтому можно положить

$$H_m(x)=\sum_{i=1}^{n}\sum_{k=0}^{\alpha_i-1}y_i^{(k)}\varphi_{ik}(x).$$

Функция $\frac{1}{k!}\frac{(x-x_i)^{\alpha_i}}{\Omega(x)}$ регулярна в точке x_i, поэтому в окрестности точки x_i ее можно разложить в ряд Тейлора:

$$\frac{1}{k!}\frac{(x-x_i)^{\alpha_i}}{\Omega(x)}=\sum_{s=0}^{\infty}a_{iks}(x-x_i)^s=l_{ik}(x)+\sum_{s=\alpha_i-k}^{\infty}a_{iks}(x-x_i)^s.$$

Здесь $l_{ik}(x)$ — многочлен степени не выше α_i-k-1, являющийся начальной частью ряда Тейлора. Несложно проверить, что многочлен

$$\varphi_{ik}(x)=\frac{\Omega}{(x-x_i)^{\alpha_i}}\left(l_{ik}(x)(x-x_i)^k\right)$$

обладает всеми требуемыми свойствами. Свойства 1 и 2 очевидны, а свойство 3 доказывается следующим образом:

$$\begin{aligned}\varphi_{ik}(x)&=\frac{l_{ik}(x)(x-x_i)^k}{k!l_{ik}(x)+a(x-x_i)^{\alpha_i-k}+\cdots}=\\&=\frac{(x-x_i)^k}{k!}\left(1+b(x-x_i)^{\alpha_i-k}+\cdots\right).\end{aligned}$$

Записывая в явном виде начальный участок ряда Тейлора $l_{ik}(x)$, получаем

$$H_m(x)=\sum_{i=1}^{n}\sum_{k=0}^{\alpha_i-1}\sum_{s=0}^{\alpha_i-k-1}y_i^{(k)}\frac{1}{k!}\frac{1}{s!}\left(\frac{(x-x_i)^{\alpha_i}}{\Omega(x)}\right)^{(s)}_{x=x_i}\frac{\Omega(x)}{(x-x_i)^{\alpha_i-k-s}}.$$

Историческое замечание. Жозеф-Луи Лагранж (1736-1813) опубликовал формулу интерполяционного многочлена в 1795 году. Шарль Эрмит (1822-1901) обобщил интерполяционный многочлен Лагранжа в 1878 году.

5.14. Формула Фаа ди Бруно

Формула Фаа ди Бруно — это обобщение формулы производной сложной функции на производные высшего порядка. У этой формулы есть разные записи. Мы приведём только комбинаторную запись. Отметим, что наше доказательство следует [Joh2].

Теорема 5.16 (Фаа ди Бруно). *Пусть функции f и g дифференцируемы достаточно много раз. Тогда*

$$\frac{d^m}{dt^m} g(f(t)) =$$
$$= \sum \frac{m!}{b_1!b_2!\dots b_m!} g^{(k)}(f(t)) \left(\frac{f'(t)}{1!}\right)^{b_1} \left(\frac{f''(t)}{2!}\right)^{b_2} \dots \left(\frac{f^{(m)}(t)}{m!}\right)^{b_m},$$

где суммирование ведётся по всем наборам целых неотрицательных чисел $b_1, \dots, b_m$, для которых $b_1 + 2b_2 + \dots + mb_m = m$; при этом $k = b_1 + \dots + b_m$.

Доказательство. Ясно, что $(g(f(t)))' = g'(f(t))f'(t)$ и $(g(f(t)))'' = g'(f(t))f''(t) + g''(f(t))(f'(t))^2$. Предположим теперь, что $(g(f(t)))^{(m)}$ является суммой слагаемых вида

$$g^{(k)}(f(t))(f'(t))^{b_1} \dots (f^{(m)}(t))^{b_m} \tag{5.3}$$

где $b_1 + 2b_2 + \dots + mb_m = m$ и $b_1 + \dots + b_m = k$. Производная функции (5.3) равна

$$g^{(k+1)}(f(t))(f'(t))^{b_1+1} \dots (f^{(m)}(t))^{b_m} + g^{(k)}(f(t))\big((f'(t))^{b_1} \dots (f^{(m)}(t))^{b_m}\big)'.$$

Второе слагаемое является суммой m слагаемых; при этом

$$\big((f^{(i)}(t))^{b_i}\big)' = b_i(f^{(i)}(t))^{b_i-1} f^{(i+1)}(t),$$

т.е. b_i заменяется на $b_i - 1$, а b_{i+1} заменяется на $b_{i+1} + 1$. Ясно также, что $b_1 + 2b_2 + \dots + i(b_i - 1) + (i+1)(b_{i+1} + 1) + \dots = m + 1$ и $b_1 + \dots + (b_i - 1) + (b_{i+1} + 1) + \dots = k$. Таким образом, $(g(f(t)))^{(m+1)}$ тоже является суммой слагаемых такого же вида (только с заменой m на $m + 1$).

Доказываемая формула имеет комбинаторный смысл. Чтобы прояснить его, сопоставим разбиению чисел от 1 до m на блоки выражение вида (5.3) следующим образом: количество блоков — это порядок производной функции g, а каждому блоку из b чисел сопоставляется множитель $f^{(b)}(t)$. Например, разбиению $\{1, 2, 3\}$ сопоставляется $g'(f(t))f'''(t)$, каждому из разбиений $\{1, 2\}, \{3\}$; $\{1, 3\}, \{2\}$ и $\{2, 3\}, \{1\}$ сопоставляется $g''(f(t))f'(t)f''(t)$, а разбиению $\{1\}, \{2\}, \{3\}$ сопоставляется $g'''(f(t))(f'(t))^3$.

Это сопоставление позволяет доказать, что

$$\frac{d^m}{dt^m} g(f(t)) = \sum g^{(k)}(f(t)) \, (f'(t))^{b_1} \, (f''(t))^{b_2} \dots \left(f^{(m)}(t)\right)^{b_m},$$

где суммирование ведётся по всем разбиениям чисел от 1 до m на блоки, причём для каждого разбиения число k — это количество блоков, а b_i — это количество блоков, содержащих ровно i чисел.

Действительно, применим индукцию по m. Каждое разбиение чисел от 1 до $m+1$ единственным образом получается из разбиения чисел от 1 до m присоединением числа $m + 1$. Если мы присоединяем $m + 1$ в качестве отдельного блока, то общее количество блоков увеличивается на 1, и количество блоков, содержащих ровно одно число, увеличивается на 1. Это соответствует применению дифференцирования к $g^{(k)}(f(t))$, в результате чего получается $g^{(k+1)}(f(t))f'(t)$. Если же мы добавляем $m + 1$ к уже существующему блоку, состоящему из i чисел, то количество блоков такого размера уменьшается на 1, а количество блоков размера $i + 1$ увеличивается на 1; такая операция применяется к любому из b_i блоков. Это в точности соответствует применению дифференцирования к $(f^{(i)}(t))^{b_i}$, в результате чего получается $b_i(f^{(i)}(t))^{b_i-1}f^{(i+1)}(t)$; в самом деле, количество блоков длины i становится равным $b_i - 1$ (множитель $(f^{(i)}(t))^{b_i-1}$) и добавляется один блок длины $i + 1$ (множитель $f^{(i+1)}(t)$).

Остаётся доказать, что количество разбиений числа m на b_1 блоков длины 1, b_2 блоков длины 2 и т.д. равно

$$\frac{m!}{(1!)^{b_1}\dots(m!)^{b_m}b_1!\dots b_m!}.$$

Рассмотрим $m!$ перестановок чисел от 1 до m. Каждой такой перестановке можно сопоставить разбиений на блоки: сначала идут b_1 блоков длины 1, потом b_2 блоков длины 2 и т.д. Но при этом в каждом из b_i блоков длины i числа можно переставлять, от этого разбиений на блоки не изменяется; это соответствует множителю $(i!)^{b_i}$ в знаменателе. Кроме того, сами b_i блоков длины i тоже можно переставлять; это соответствует множителю $b_i!$ в знаменателе. $\square$

Историческое замечание. Франческо Фаа ди Бруно (1825-1888) получил формулу для высших производных сложной функции в виде определителя в 1855 году. Но первым такую формулу опубликовал Луи Арбогас (1759-1803) в 1800 году.

5.15. Решения задач

5.1. По определению производной

$$\begin{aligned} f(t)-f(x) &= (t-x)[f'(x)+u(t)], \\ g(s)-g(y) &= (s-y)[g'(y)+v(s)], \end{aligned}$$

где $t\in[a,b]$, $s\in I$ и $u(t)\to 0$ при $t\to x$, $v(s)\to 0$ при $s\to y$. Пусть $s=f(t)$. Тогда

$$\begin{aligned} h(t)-h(x) &= g(f(t))-g(f(x)) = \\ &= [f(t)-f(x)]\cdot[g'(y)+v(s)] = \\ &= (t-x)\cdot[f'(x)+u(t)]\cdot[g'(y)+v(s)]. \end{aligned}$$

При $t\neq x$ получаем

$$\frac{h(t)-h(x)}{t-x}=[g'(y)+v(s)]\cdot[f'(x)+u(t)].$$

Правая часть стремится к $g'(y)f'(x)$ при $t\to x$, поскольку из непрерывности функции f следует, что $s\to y$ при $t\to x$.

5.2. Ясно, что

$$\frac{f(x)-f(x_0)}{x-x_0}=\frac{x^n-x_0^n}{x-x_0}=x^{n-1}+x^{n-2}x_0+\dots+x_0^{n-1}.$$

При $x\to x_0$ каждое из n слагаемых стремится к x_0^{n-1}.

5.3. Легко проверить, что

$$\frac{x^a-x_0^a}{x-x_0}=x_0^a\frac{\left(\frac{x^a}{x_0^a}\right)-1}{x-x_0}=x_0^{a-1}\frac{(1+y)^a-1}{y},$$

где $y=\frac{x-x_0}{x_0}$. Ясно, что $y\to 0$ при $x\to x_0$. Кроме того, $\lim_{y\to 0}\frac{(1+y)^a-1}{y}=a$ (задача 3.31).

5.4. Ясно, что

$$\sin x-\sin x_0=2\sin\frac{x-x_0}{2}\cos\frac{x+x_0}{2},$$
$$\cos x-\cos x_0=-2\sin\frac{x-x_0}{2}\sin\frac{x+x_0}{2}.$$

Остаётся заметить, что $\frac{1}{x-x_0}\sin\frac{x-x_0}{2}\to\frac{1}{2}$ при $x\to x_0$ (задача 3.6) и функции $\sin x$ и $\cos x$ непрерывны (задача 3.5).

5.5. Согласно задаче 5.3 в)

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

Для $(\operatorname{ctg} x)'$ вычисления аналогичны.

5.6. Ясно, что $a^x - a^{x_0} = a^{x_0}(a^{x-x_0} - 1)$. Остаётся заметить, что $\lim_{x\to x_0} \frac{a^{x-x_0}-1}{x-x_0} = \ln a$ согласно задаче 3.32.

5.7. Поскольку $\log_a x = \frac{\ln x}{\ln a}$, достаточно проверить, что $(\ln x)' = \frac{1}{x}$. Равенство $\ln x - \ln x_0 = \ln\frac{x}{x_0} = \ln\left(1 + \frac{x-x_0}{x_0}\right)$ показывает, что $\frac{\ln x - \ln x_0}{x-x_0} = \frac{1}{x_0}\frac{\ln(1+y)}{y}$, где $y = \frac{x-x_0}{x_0}$. Остаётся заметить, что $\lim_{y\to 0}\frac{\ln(1+y)}{y} = 1$ (задача 3.29).

5.8. Если $x = \sin y$, то $\arcsin x = y$. Поэтому arcsin — функция, обратная к sin. Значит, согласно задаче 5.5,

$$(\arcsin x)' = \frac{1}{(\sin y)'} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-x^2}}.$$

Если $x = \operatorname{tg} y$, то $\operatorname{arctg} x = y$, поэтому

$$(\operatorname{arctg} x)' = \frac{1}{(\operatorname{tg} y)'} = \cos^2 y = \frac{1}{1+\operatorname{tg}^2 y} = \frac{1}{1+x^2}.$$

5.9. Ясно, что если $y = u^v$, то $\ln y = v\ln u$. Дифференцируя это равенство, получаем $\frac{y'}{y} = v'\ln u + v\frac{u'}{u}$. Поэтому $y' = y\left(v'\ln u + \frac{vu'}{u}\right)$.

5.10. Согласно задаче 5.9 $f'(x) = x^{\sin x}\left(\frac{\sin x}{x} + \cos x\cdot\ln x\right)$.

5.11. Согласно задаче 5.9 $f'(x) = \cos^x a\cdot\ln\cos a - \sin^x a\cdot\ln\sin a = \cos^x a(\ln\cos a - \operatorname{tg}^x a\cdot\ln\sin a)$.

5.12. Предположим, что у многочлена $f_n(x)$ есть кратный корень. Тогда многочлены $f_n(x)$ и $f_n'(x) = f_{n-1}(x)$ имеют общий корень x_0. Следовательно, $f_n(x_0) - f_{n-1}(x_0) = 0$. Но $f_n(x) - f_{n-1}(x) = \frac{x^n}{n!}$, поэтому $x_0 = 0$. Приходим к противоречию, поскольку $f_n(0) \neq 0$.

5.13. Требуемое утверждение достаточно доказать для монома x^m. Ясно, что если $n \leq m$, то

$$\begin{aligned}(x^m)^{(n)} &= m(m-1)\cdot\ldots\cdot(m-n+1)x^{m-n} = \\ &= n!\frac{m!}{n!(m-n)!}x^{m-n} = n!\binom{m}{n}x^{m-n}.\end{aligned}$$

Число $n!\binom{m}{n}$, очевидно, делится на $n!$, поскольку число $\binom{m}{n}$ целое. Если же $n > m$, то $(x^m)^{(n)} = 0$.

5.14. Пусть $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Тогда по теореме Виета сумма корней многочлена f равна $-\frac{a_{n-1}}{a_n}$, поэтому их среднее арифметическое равно $-\frac{a_{n-1}}{na_n}$. Сумма корней производной $f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$ равна $-\frac{(n-1)a_{n-1}}{na_n}$, поэтому их среднее арифметическое равно $-\frac{(n-1)a_{n-1}}{(n-1)na_n} = -\frac{a_{n-1}}{na_n}$.

5.15. Производная рассматриваемого выражения равна $fg^{(n+1)} + f'g^{(n)} - f'g^{(n)} - f''g^{(n-2)} + f''g^{(n-2)} + \cdots + (-1)^n f^{(n)}g' + (-1)^n f^{(n+1)}g$. Все промежуточные члены взаимно сокращаются, и остаются только $fg^{(n+1)} \pm f^{(n+1)}g$. Но f и g — многочлены степени n; поэтому их $(n+1)$-е производные равны нулю.

5.16. Положим $f(x) = x^3 + px + q$. Тогда $f'(x) = 3x^2 + p$. Если $p \geq 0$, то функция f монотонна, и её график пересекает ось x ровно в одной точке. В этом случае уравнение имеет ровно один вещественный корень. Отметим, что в этом случае $D \geq 0$.

Пусть теперь $p < 0$. Тогда функция f возрастает на участке от $-\infty$ до $-\sqrt{-p/3}$, затем убывает на участке от $-\sqrt{-p/3}$ до $\sqrt{-p/3}$, а после этого возрастает на участке от $\sqrt{-p/3}$ до ∞. Ясно, что

$f(x)f(-x) = q^2 - x^2(x^2+p)^2$. При $x = \sqrt{-p/3}$ выражение в правой части превращается в $4D$. Поэтому если $D < 0$, то график функции f пересекает ось x на участке от $-\sqrt{-p/3}$ до $\sqrt{-p/3}$, на котором функция убывает. В этом случае уравнение имеет три вещественных корня. Если $D > 0$, то уравнение имеет один вещественный корень. Если же $D = 0$, то уравнение имеет два вещественных корня.

5.17. Многочлен $\sum_{i=1}^n \frac{g(x_i)f(x)}{f'(x_i)(x-x_i)}$ принимает в точках $x_1, \ldots, x_n$ значения $g(x_1), \ldots, g(x_n)$. Поэтому если $g(x)$ — многочлен степени не выше $n-1$, то $g(x) = \sum_{i=1}^n \frac{g(x_i)f(x)}{f'(x_i)(x-x_i)}$, т.е. $\frac{g(x)}{f(x)} = \sum_{i=1}^n \frac{g(x_i)}{f'(x_i)(x-x_i)}$.

Положим $g(x) = x^{k+1}$, где $0 \le k \le n-2$. В результате получим $\frac{x^{k+1}}{f(x)} = \sum_{i=1}^n \frac{x_i^{k+1}}{f'(x_i)(x-x_i)}$. По условию $f(0) \ne 0$, поэтому, положив $x = 0$, получим требуемое равенство при $0 \le k \le n-2$.

Положим теперь $g(x) = x^n - f(x)$. В результате получим $\frac{x^n - f(x)}{f(x)} = \sum_{i=1}^n \frac{x_i^n}{f'(x_i)(x-x_i)}$, поскольку $f(x_i) = 0$. Положив $x = 0$, получим требуемое равенство для $k = n-1$.

5.18. Если $x = x_i$, то неравенство очевидно. Поэтому будем считать, что x — не корень многочлена P. Тогда

$$\frac{P'(x)}{P(x)} = \frac{1}{x-x_1} + \cdots + \frac{1}{x-x_n}.$$

Продифференцировав обе части этого равенства, получаем

$$\frac{P''(x)P(x) - \big(P'(x)\big)^2}{\big(P(x)\big)^2} = -\frac{1}{(x-x_1)^2} - \cdots - \frac{1}{(x-x_n)^2} < 0.$$

5.19. Пусть f — данный многочлен. Рассмотрим многочлены F и G, которые обладают следующими свойствами: $F' = \dfrac{(f')^2 + f' + 1}{2}$, $G' = \dfrac{(f')^2 - f' + 1}{2}$, $F(0) = f(0)$ и $G(0) = 0$. Тогда $f = F - G$, причём $F' > 0$ и $G' > 0$ (т.е. F и G монотонно возрастают), поскольку $(f')^2 \pm f' + 1 = \big(f' \pm \frac{1}{2}\big)^2 + \frac{3}{4} \ge \frac{3}{4}$.

5.20. Применим индукцию по n. При $n = 1$ получаем многочлен $a_0 + a_1x^k$, который имеет не более одного положительного корня. Предположим, что многочлен $P(x)$ имеет более n положительных корней. Между любыми двумя положительными корнями многочлена есть по крайней мере один положительный корень его производной. Поэтому многочлен $P'(x) = b_1x^{k_1-1} + b_2x^{k_2-1} + \cdots + b_nx^{k_n-1}$ имеет более $n-1$ положительных корней. Можно считать, что $k_1 < k_2 < \cdots < k_n$. Тогда многочлен $P'(x)$ можно сократить на x^{k_1-1} (при этом мы потеряем только корень $x = 0$). В результате получим, что многочлен $b_1 + b_2x^{k_2-k_1} + \cdots + b_nx^{k_n-k_1}$ имеет более $n-1$ положительных корней. Это противоречит предположению индукции.

5.21. Рассмотрим замкнутое множество $F_n = \{x \mid f^{(n)}(x) = 0\}$ и его дополнение — открытое множество $G_n = \mathbb{R} \setminus F_n$. Предположим, что функция f не является многочленом на некотором интервале Δ. Тогда для любого n открытое множество $G_n \cap \Delta$ непусто. Действительно, если это множество пусто, то f является многочленом степени не более $n-1$ на Δ. Напомним, что множество $G_n \cap \Delta$, как любое открытое множество на прямой, является объединением не более чем счётного набора попарно не пересекающихся интервалов (задача 4.13). Докажем, что функция f не является многочленом на одном из этих интервалов.

Предположим, что для некоторого натурального n и некоторого интервала Δ функция f — не многочлен на Δ, но её ограничение на любой интервал δ, из которых состоит множество $G_n \cap \Delta$, является многочленом. Рассмотрим наибольший интервал $\sigma(\delta)$, содержащий δ и содержащийся в Δ, на котором функция f остаётся многочленом.

Если ограничения функции f на два интервала (a, b) и (b, c) с общей граничной точкой b — многочлены, то f — многочлен на интервале (a, c). Действительно, для некоторого m непрерывная функция $f^{(m)}$ равна нулю на интервалах (a, b) и (b, c), поэтому она равна нулю и в общей граничной точке b. Из этого следует, что изолированная точка b множества F_n, расположенная внутри Δ, не может быть граничной точкой интервала $\sigma(\delta)$. Действительно, точка b является общей граничной точкой двух интервалов $\delta_1 = (a, b)$ и $\delta_2 = (b, c)$, на каждом из которых функция f — многочлен.

Кроме того, согласно предположению функция f — не многочлен на Δ, поэтому $\sigma(\delta)$ не может совпасть с Δ. Таким образом, по крайней мере одна граничная точка интервала $\sigma(\delta)$ не является изолированной точкой множества F_n.

Пусть a — граничная точка интервала $\sigma(\delta)$, которая не является изолированной точкой множества F_n. Тогда существует сходящаяся к a последовательность точек $x_k \in F_n$, отличных от a. Для точки $x_k \in F_n$ выполняется равенство $f^{(n)}(x_k) = 0$, поэтому $f^{(n)}(a) = 0$. Между двумя точками, в которых обращается в нуль функция $f^{(m)}$, есть точка, в которой обращается в нуль функция $f^{(m+1)}$, поэтому $f^{(m)}(a) = 0$ для всех $m \geq n$.

На интервале $\sigma(\delta)$ с граничной точкой a функция f является многочленом. Значения этого многочлена и его производных в точке a совпадают со значениями в этой точке функции f. Если для некоторого многочлена P для всех $m \geq n$ выполняется равенство $P^{(m)}(a) = 0$, то степень этого многочлена меньше n, поэтому $f^{(n)}$ обращается в нуль тождественно на $\sigma(\delta)$. В частности, $f^{(n)}$ обращается в нуль тождественно на δ. Но $\delta \subset G_n$, т.е. $f^{(n)}(x) \neq 0$ для всех $x \in \delta$. Приходим к противоречию.

Вернёмся теперь к исходной задаче. Предположим, что функция f не является многочленом. Тогда она не является многочленом на некотором интервале Δ. Найдём интервал $(a_1, b_1) \subset G_1 \cap \Delta$, на котором функция f не является многочленом. Затем найдём интервал $(c_2, d_2) \subset G_2 \cap (a_1, b_1)$, на котором функция f не является многочленом, и рассмотрим отрезок $[a_2, b_2] \subset (c_2, d_2)$, и т.д. В результате получим последовательность вложенных отрезков. В общей точке этих отрезков все производные функции f отличны от нуля, что противоречит условию.

5.22. Касательная к кривой $y = f(x)$ в точке (x_0, y_0) задаётся уравнением $y - y_0 = f'(x_0)(x - x_0)$. Поэтому для x_1 получаем уравнение $-y_0 = e^{x_0}(x_1 - x_0)$. Учитывая, что $y_0 = e^{x_0}$, получаем $x_0 - x_1 = 1$.

5.23. Числа k_1 и k_{ij} не изменяются при параллельном переносе осей координат, поэтому можно считать, что мы имеем дело с параболой $y = ax^2$. Пусть $A_i = (x_i, y_i)$. Тогда $k_1 = 2ax_1$ и

$$k_{ij} = \frac{y_i - y_j}{x_i - x_j} = a\frac{x_i^2 - x_j^2}{x_i - x_j} = a(x_i + x_j).$$

Теперь требуемое равенство легко проверяется.

5.24. Если прямая задаётся уравнением $\frac{y-y_0}{x-x_0} = k$, то перпендикулярная ей прямая, проходящая через точку (x_0, y_0), задаётся уравнением $\frac{y-y_0}{x-x_0} = -\frac{1}{k}$, поскольку $\operatorname{tg}(\varphi + 90^\circ) = -\frac{1}{\operatorname{tg}(\varphi)}$. Касательная к кривой $y = f(x)$ в точке (x_0, y_0) задаётся уравнением $\frac{y-y_0}{x-x_0} = f'(x_0)$, поэтому нормаль в точке (x_0, y_0) задаётся уравнением $\frac{y-y_0}{x-x_0} = -\frac{1}{f'(x_0)}$.

5.25. Нормаль к параболе $y = x^2$ в точке (x_0, y_0) задаётся уравнением $\frac{y-y_0}{x-x_0} = -\frac{1}{2x_0}$. Чтобы найти y_1, полагаем $x = 0$. В результате получаем $\frac{y_1-y_0}{-x_0} = -\frac{1}{2x_0}$, т.е. $y_1 - y_0 = \frac{1}{2}$.

5.26. Пусть $a < x \leq b$. Применим теорему Лагранжа (теорема 5.9) к функции f на отрезке $[a, x]$. В результате получим, что $f(x) - f(a) = f'(x_0)(x - a)$ для некоторой точки x_0 отрезка $[a, x]$. По условию $f'(x_0) = 0$, поэтому $f(x) = f(a)$.

5.27. а) Непосредственно из определения производной видно, что $f'(x) \geq 0$ для неубывающей функции $f(x)$.

Предположим теперь, что $f'(x) \geq 0$ для любой точки x интервала (a, b). Возьмём на отрезке $[a, b]$ две точки x и y так, что $x < y$. Применим теорему Лагранжа (теорема 5.9) к отрезку $[x, y]$. В результате получим, что $f(y) - f(x) = f'(z)(y - x)$ для некоторой точки z интервала (x, y). По условию $f'(z) \geq 0$, поэтому $f(y) \geq f(x)$.

б) Согласно задаче а) функция $f(x)$ неубывающая. Выберем на отрезке $[a, b]$ две точки x и y так, что $x < y$. Тогда для любой точки z отрезка $[x, y]$ имеют место неравенства $f(x) \leq f(z) \leq f(y)$. Поэтому если $f(x) = f(y)$, то функция f постоянна на отрезке $[x, y]$. Но тогда f' обращается в нуль на этом отрезке, что противоречит условию.

5.28. Согласно задаче 5.27 функция $f(x) - g(x)$ монотонно возрастает на отрезке $[a, b]$.

5.29. Пусть $f(x) = \cos x$ и $g(x) = 1 - \frac{x^2}{2}$. Тогда $f(0) = g(0)$ и $f'(x) = -\sin x > -x = g'(x)$ при $x > 0$, поскольку $\sin x < x$ (задача 3.4). Поэтому $f(x) > g(x)$ при $x > 0$ согласно задаче 5.28.

Пусть теперь $f(x) = \sin x$ и $g(x) = x - \frac{x^3}{6}$. Тогда $f'(x) = \cos x$ и $g'(x) = 1 - \frac{x^2}{2}$. Неравенство $\cos x > 1 - \frac{x^2}{2}$ только что было доказано.

5.30. Достаточно доказать, что $(\operatorname{tg} x)' = \frac{1}{\cos^2 x} > 1 + x^2$, т.е. $\frac{1}{\cos^2 x} - 1 = \operatorname{tg}^2 x > x^2$. Неравенство $\operatorname{tg} x > x$ доказано в решении задачи 3.4.

5.31. а) Пусть $f(x) = x^\alpha - \alpha x$. Тогда $f'(x) = \alpha(x^{\alpha-1} - 1)$. Поэтому $f'(x) > 0$ при $0 < x < 1$ и $f'(x) < 0$ при $x > 1$. Следовательно, $f(x)$ возрастает при $0 < x < 1$ и убывает при $x > 1$. Таким образом, для неотрицательных x наибольшее значение $f(x)$ равно $f(1) = 1 - \alpha$.

б) Непосредственно следует из а). Действительно, положим $\alpha = \frac{1}{p}$ и $x = \frac{a^p}{b^q}$. В результате получим $\left(\frac{a^p}{b^q}\right)^{1/p} \le \frac{1}{p}\frac{a^p}{b^q} + \frac{1}{q}$. Остаётся умножить обе части неравенства на $b^{1+\frac{q}{p}} = b^q$.

Замечание. По поводу других доказательств см. задачи 1.11 и 3.50.

5.32. Рассмотрим функцию

$$g(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a) - f(x).$$

Применив теорему Лагранжа к функции f, получим, что $g'(x) = f'(c) - f'(x)$ для некоторой точки $c \in (a, b)$. Применив теорему Лагранжа к функции f', получим, что $g'(x) = f''(d)(c - x)$ для некоторой точки d между c и x. Поэтому g строго возрастает на $[a, c]$ и строго убывает на $[c, b]$. А так как $g(a) = 0 = g(b)$, то $g(x) > 0$ при $x \in (a, b)$. При $x = \frac{a+b}{2}$ получаем неравенство

$$f\left(\frac{a+b}{2}\right) < \frac{f(a) + f(b)}{2}.$$

Аналогично получаем такое же неравенство и для любых точек $x_1, x_2 \in [a, b]$.

5.33. Можно считать, что $c > 0$, потому что при $c = 0$ утверждение очевидно. Положим $\varepsilon = \frac{1}{2c}$ и рассмотрим точку x_1, в которой достигает максимума функция $|f(x)|$ на ε-окрестности точки x_0. Тогда

$$\begin{aligned} |f(x_1)| &= |f(x_1) - f(x_0)| \le |x_1 - x_0| \cdot |f'(x_1)| \le \\ &\le c|x_1 - x_0| \cdot |f(x_1)| \le c\varepsilon|f(x_1)| \le \frac{1}{2}|f(x_1)|. \end{aligned}$$

Следовательно, $f(x_1) = 0$ и $f(x) \equiv 0$ на ε-окрестности точки x_0. Применим это же утверждение, взяв вместо x_0 один из концов этой ε-окрестности, и т.д.

5.34. Пусть $0 < x < \pi/2$. Тогда

$$\left(\frac{\sin x}{x}\right)' = \frac{x\cos x - \sin x}{x^2} = \frac{\cos x(x - \operatorname{tg} x)}{x^2} < 0.$$

Поэтому если $0 < \alpha < \beta < \pi/2$, то $\frac{\sin\alpha}{\alpha} > \frac{\sin\beta}{\beta}$.

5.35. Пусть $0 < x < \pi/2$. Тогда

$$\left(\frac{\operatorname{tg} x}{x}\right)' = \frac{x - \sin x\cos x}{x^2\cos^2 x} > 0,$$

поскольку $\sin x < x$ и $\cos x < 1$.

5.36. Рассмотрим функцию $f(x) = 2\sin x + \operatorname{tg} x - 3x$. Ясно, что

$$\begin{aligned} f'(x) = 2\cos x + \frac{1}{\cos^2 x} - 3 = \frac{2\cos^3 x - 3\cos^2 x + 1}{\cos^2 x} = \\ = \frac{(1 - \cos x)(1 + \cos x - 2\cos^2 x)}{\cos^2 x}. \end{aligned}$$

Если $0 < \alpha < \pi/2$, то $1 - \cos x > 0$ и $1 + \cos x - 2\cos^2 x = 1 - \cos^2 x + \cos x(1 - \cos x) > 0$. Поэтому если $0 < \alpha < \pi/2$, то $f(\alpha) > f(0) = 0$.

5.37. а) Пусть $f(x) = e^x - x$. Тогда $f'(x) = e^x - 1$. Поэтому $f'(x) > 0$ при $x > 0$ и $f'(x) < 0$ при $x < 0$. Значит, $f(x) > f(0) = 1$ для любого $x \neq 0$.

б) Из неравенства $e^x > 1+x$ следует, что $x > \ln(1+x)$ при $1+x > 0$, $x \neq 0$. Поэтому $\ln\left(1 + \frac{1}{n}\right) < \frac{1}{n}$ и $\ln\left(1 - \frac{1}{n+1}\right) < -\frac{1}{n+1}$, т.е. $\ln\frac{n+1}{n} < \frac{1}{n}$ и $\ln\frac{n}{n+1} < -\frac{1}{n+1}$.

5.38. а) Пусть $f(x) = x - \ln x$. Тогда $f'(x) = 1 - \frac{1}{x}$. Поэтому $f'(x) < 0$ при $x < 1$ и $f'(x) > 0$ при $x > 1$. Следовательно, если $x > 0$, $x \neq 1$, то $f(x) > f(1) = 0$. Это и есть требуемое неравенство.

б) В неравенстве $\ln t < t - 1$ положим $t = \frac{1}{x}$. В результате получим $-\ln x < \frac{1}{x} - 1$, т.е. $\ln x > \frac{x-1}{x}$.

5.39. Положим $y = x^{1/n}$. Тогда после сокращения на n требуемые неравенства запишутся в виде $\ln y < y - 1 < y\ln y$. Такие неравенства доказаны в решении задачи 5.38.

5.40. Согласно задаче 5.38

$$\frac{\frac{x}{n}}{1 + \frac{x}{n}} < \ln\left(1 + \frac{x}{n}\right) < \frac{x}{n},$$

поэтому $\frac{x}{1+\frac{x}{n}} < n\ln\left(1 + \frac{x}{n}\right) < x$.

5.41. Требуемое неравенство эквивалентно неравенству $x > e\ln x$, т.е. $\ln x/x < 1/e$. Рассмотрим функцию $f(x) = \ln x/x$. Легко проверить, что $f'(x) = (1 - \ln x)/x^2$. Производная функции f обращается в нуль лишь в точке $x = e$. Кроме того, $\lim_{x\to+0}(\ln x/x) = -\infty$ и $\lim_{x\to+\infty}(\ln x/x) = 0$. Поэтому $f(x) \leq f(e) = \ln e/e = 1/e$, причём равенство достигается лишь при $x = e$.

5.42. Рассмотрим функцию $f(x) = b\cdot x^a + a\cdot x^{-b}$, где $x \geq 1$. Ясно, что $f'(x) = ab(x^{a-1} - x^{-b-1}) > 0$ при $x > 1$, поскольку $x^{a-1}/x^{-b-1} = x^{a+b} > 1$. Значит, функция f возрастающая и $f(2) \geq f(1)$, т.е. $b \cdot 2^a + a \cdot 2^{-b} \geq a + b$.

5.43. Положим $a = bx$, где $x > 1$. Требуемое неравенство запишется тогда в виде

$$\sqrt{x} < \frac{x-1}{\ln x} < \frac{x+1}{2}, \text{ т.е. } 2\frac{x-1}{x+1} < \ln x < \frac{x-1}{\sqrt{x}}.$$

Если $x > 1$, то $\left(\sqrt{x} - \frac{1}{\sqrt{x}}\right)' = \frac{1}{2}x^{-1}\left(x^{1/2} + x^{-1/2}\right) > x^{-1} = (\ln x)'$ и $\left(2\frac{x-1}{x+1}\right)' = \frac{4}{(x+1)^2} < \frac{1}{x}$. Остаётся заметить, что если $f(1) = g(1)$ и $f'(x) > g'(x)$ при $x > 1$, то $f(x) > g(x)$ при $x > 1$.

5.44. Применим индукцию по n. При $n = 1$ неравенство очевидно. При $n = 2$ получаем $\sin x + \frac{1}{2}\sin 2x = \sin x(1 + \cos x)$. Ясно, что $\sin x > 0$ и $1 + \cos x > 0$ при $0 < x < \pi$.

Предположим, что $f_{n-1}(x) = \sin x + \frac{1}{2}\sin 2x + \dots + \frac{1}{n-1}\sin(n-1)x > 0$ при $0 < x < \pi$. Покажем, что тогда $f_n(x) = f_{n-1}(x) + \frac{1}{n}\sin nx > 0$ при $0 < x < \pi$. Пусть x_0 — точка отрезка $[0, \pi]$, в которой функция $f_n(x)$ принимает минимальное значение. Предположим, что $f_n(x_0) \leq 0$, причём $x_0 \neq 0$ и π. Тогда $f_n'(x_0) = 0$. Но

$$f_n'(x_0) = \cos x_0 + \cos 2x_0 + \dots + \cos nx_0 = \frac{\sin\left(n + \frac{1}{2}\right)x_0 - \sin\frac{x_0}{2}}{\sin\frac{x_0}{2}}.$$

Поэтому $\sin\left(n + \frac{1}{2}\right)x_0 = \sin\frac{x_0}{2}$, а значит, $\left|\cos\left(n + \frac{1}{2}\right)x_0\right| = \cos\frac{x_0}{2}$. Далее,

$$\begin{aligned} f_n(x_0) - f_{n-1}(x_0) &= \frac{1}{n}\sin nx_0 = \\ &= \frac{1}{n}\left(\sin\left(n + \frac{1}{2}\right)x_0\cos\frac{x_0}{2} - \cos\left(n + \frac{1}{2}\right)x_0\sin\frac{x_0}{2}\right). \end{aligned}$$

Полученное выражение равно 0 или $\frac{2}{n}\sin\frac{x_0}{2}\cos\frac{x_0}{2} = \frac{1}{n}\sin x_0 > 0$. Таким образом, $f_n(x_0) - f_{n-1}(x_0) \geq 0$, а значит, $f_{n-1}(x_0) \leq f_n(x_0) \leq 0$. Получено противоречие.

5.45. Фиксируем число x $(0 < x < \pi/4)$ и рассмотрим функцию $f(y) = \cos^y x - \sin^y x$ для $y \geq 0$. Ясно, что $f(0) = 0$, $f(y) > 0$ при $y > 0$ и $f(y) \to 0$ при $y \to \infty$. Далее, согласно задаче 5.11

$$f'(y) = \cos x^y(\ln\cos x - \operatorname{tg}^y x\ln\sin x).$$

Функция $g(y) = \operatorname{tg}^y x$ монотонна, поэтому равенство $f'(y) = 0$ выполняется для единственного положительного числа y. Равенство $f(2) = f(2)(\cos^2 x + \sin^2 x) = f(4)$ показывает, что это число заключено между 2 и 4, поэтому $f'(2) > 0$ и $f'(4) < 0$.

Неравенство $f'(2) > 0$ записывается следующим образом: $\cos^2 x \ln\cos x > \sin^2 x \ln\sin x$, т.е. $\ln\bigl((\cos x)^{\cos^2 x}\bigr) > \ln\bigl((\sin x)^{\sin^2 x}\bigr)$. Так мы получаем первое требуемое неравенство. Второе неравенство получается из неравенства $f'(4) < 0$.

Замечание. Фактически мы доказали, что $(\cos x)^{\cos^y x} > (\sin x)^{\sin^y x}$ при $0 < y \leq 2$ и $(\cos x)^{\cos^y x} < (\sin x)^{\sin^y x}$ при $y \geq 4$.

5.46. Рассмотрим функции

$$f(x) = \ln(1+x) - \frac{2x}{2+x},$$
$$g(x) = \ln(1+x) - \frac{x}{\sqrt{1+x}}.$$

Легко проверить, что $f(0) = g(0) = 0$ и $g'(0) \leq 0 \leq f'(0)$, причём при $x \neq 0$ неравенства строгие. Действительно,

$$f(x)' = \frac{1}{1+x} - \frac{4}{(2+x)^2} = \left(\frac{x}{2+x}\right)^2,$$
$$g(x)' = \frac{1}{1+x} - \frac{2+x}{2(1+x)^{3/2}} = -\frac{(1-\sqrt{1+x})^2}{2(1+x)^{3/2}}.$$

Таким образом, $f(x) > 0$ при $x > 0$ и $f(x) < 0$ при $x < 0$. Из этого следует неравенство в правой части. Неравенство в левой части получается аналогично.

5.47. Применив правило Лопиталя дважды, получим

$$\lim_{x\to 0}\frac{x - \sin x}{x^3} = \lim_{x\to 0}\frac{1-\cos x}{3x^2} = \lim_{x\to 0}\frac{\sin x}{6x} = \frac{1}{6}.$$

5.48. По правилу Лопиталя

$$\begin{aligned}\lim_{x\to a}\frac{a^x - x^a}{x-a} &= \lim_{x\to a}(a^x - x^a)' = \lim_{x\to a}(\ln a\cdot a^x - ax^{a-1}) = \\ &= \ln a\cdot a^a - a^a = a^a(\ln a - 1).\end{aligned}$$

5.49. Ответ: 2. Преобразуем отношение производных:

$$\frac{\frac{1}{\cos^2 x} - 1}{1-\cos x} = \frac{1}{\cos^2 x}\cdot\frac{1-\cos x^2}{1-\cos x} = \frac{1+\cos x}{\cos^2 x}.$$

Последнее выражение стремится к 2 при $x \to 0$.

5.50. Рациональная функция f/g обладает разложением требуемого вида (см., например, доказательство теоремы 6.3); нужно лишь найти коэффициенты A_i. Ясно, что $A_i = \lim_{x\to a_i}\frac{(x-a_i)f(x)}{g(x)}$. При $x = a_i$ производная функции $(x-a_i)f(x)$ равна $f(a_i)$, поэтому согласно правилу Лопиталя $\lim_{x\to a_i}\frac{(x-a_i)f(x)}{g(x)} = \frac{f(a_i)}{g'(a_i)}$.

5.51. Предположим, что $P_0(x)\sin^n x + P_1(x)\sin^{n-1} x + \dots + P_n(x) = 0$, причём $P_0(x) \neq 0$ и n — наименьшее натуральное число, для которого имеет место тождество такого вида. Из минимальности n следует, что $P_n(x) \neq 0$. С другой стороны, если $x = k\pi$, где k — целое число, то $\sin x = 0$, поэтому $P_n(k\pi) = 0$ для всех целых k. Но многочлен P_n не может иметь бесконечно много корней. Получено противоречие.

5.52. Первое решение. Предположим, что

$$P_0(x)e^{nx} + P_1(x)e^{(n-1)x} + \dots + P_n(x) = 0, \tag{1}$$

причём $P_0(x) = a_0x^k + a_1x^{k-1} + \dots + a_k$, $a_0 \neq 0$. Рассмотрим функцию $G(x) = \frac{1}{x^k e^{nx}}\bigl(P_0(x)e^{nx} + P_1(x)e^{(n-1)x} + \dots + P_n(x)\bigr)$. С одной стороны, $G(x) = 0$ для всех x. С другой стороны, $G(x) = a_0 + \frac{a_1}{x} + \dots + \frac{a_x}{x^k} + \frac{1}{e^x}\frac{P_1(x)}{x^k} + \dots + \frac{1}{e^{nx}}\frac{P_n(x)}{x^k}$, поэтому $\lim_{x\to\infty} G(x) = a_0 \neq 0$, поскольку $\lim_{x\to\infty} \frac{x^m}{e^x} = 0$. Получено противоречие.

В т о р о е   р е ш е н и е. Предположим, что выполняется тождество (1), причём из всех тождеств такого вида выбрано то, для которого n минимально. Из минимальности n следует, что многочлен $P_n(x)$ отличен от нуля. Поделив обе части тождества (1) на $P_n(x)$, получим

$$R_0(x)e^{nx} + R_1(x)e^{(n-1)x} + \dots + R_{n-1}(x)e^x + 1 = 0,$$

где $R_0(x)$, $\dots$, $R_{n-1}(x)$ — рациональные функции, причём $R_0(x) \neq 0$. Продифференцировав это тождество и поделив обе части на e^x, получим тождество

$$(R_0' + nR_0)e^{(n-1)x} + (R_1' + (n-1)R_1)e^{(n-2)x} + \dots + (R_{n-1}' + R_{n-1}) = 0. \tag{2}$$

Из этого тождества легко получить тождество для многочленов: достаточно умножить обе части на общий знаменатель рациональных функций. Поэтому из минимальности n следует, что $R_0'+nR_0 = 0$, $\dots$, $R_{n-1}' + R_{n-1} = 0$. Пусть $R_0 = P/Q$, где P и Q — взаимно простые многочлены (мы знаем, что $P \neq 0$). Из равенства $R_0' + nR_0 = 0$ следует, что $P'Q - PQ' + nPQ = 0$. Многочлены $P'Q$ и PQ делятся на Q, поэтому PQ' делится на Q, а значит, Q' делится на Q. Это возможно лишь в том случае, когда Q — константа. Аналогично доказывается, что P — константа. Поэтому $R_0' = 0$, а значит, $R_0 = -R_0'/n = 0$. Получено противоречие.

5.53. а) Достаточно заметить, что

$$x^k = \bigl(a + (x-a)\bigr)^k = \sum_{i=0}^{k} \binom{k}{i} a^i (x-a)^{k-i}.$$

б) Чтобы найти A_0, положим $x = a$. В результате получим $f(a) = A_0$. Продифференцировав выражение для $f(x)$, получим

$$f'(x) = A_1 + 2A_2(x-a) + \dots + nA_n(x-a)^{n-1}.$$

Положив $x = a$, получаем $f'(a) = A_1$. Теперь дифференцируем выражение для $f'(x)$:

$$f''(x) = 1\cdot 2A_2 + 2\cdot 3A_3(x-a) + \dots + (n-1)nA_n(x-a)^{n-2}.$$

Снова подставляя $x = a$, получаем $f''(a) = 1\cdot 2A_2$. Продолжая эти рассуждения дальше, будем получать $A_k = \frac{f^{(k)}(a)}{k!}$ для $k \leq n$.

5.54. Эти утверждения непосредственно следуют из формулы Тейлора для $a = 0$ и соответствующих функций $f(x)$.

5.55. Для такой функции остаточный член в форме Лагранжа тождественно равен нулю.

5.56. Последовательность функций $f_n(x) = x^n$ на отрезке $[0,1]$ сходится к разрывной функции.

5.57. Рассмотрим функции $f_n(x) = x^n/n$ на отрезке $[0,1]$. Последовательность этих функций сходится к функции $f(x) = 0$. Ясно, что $f_n'(x) = x^{n-1}$, поэтому $f_n'(1) = 1$.

Глава 6.

Интегрирование

6.1. Неопределённый интеграл

Функцию $F(x)$, для которой $F'(x) = f(x)$, называют *первообразной* или *неопределённым интегралом* функции $f(x)$. Для функции, заданной на отрезке (или на всей вещественной прямой), её первообразная определена с точностью до прибавления некоторой константы C. Действительно, если $F_1(x)$ и $F_2(x)$ — первообразные функции $f(x)$, а $\varphi(x) = F_1(x) - F_2(x)$, то по условию $\varphi'(x) = 0$. Поэтому согласно задаче 5.26 функция $\varphi(x)$ постоянна.

Первообразную функции $f(x)$ обозначают $\int f(x)\,dx$.

Задача 6.1. Докажите, что $\int \frac{dx}{x^2-a^2} = \frac{1}{2a} \ln \frac{x-a}{x+a} + C$ на любом отрезке, не содержащем точек $\pm a$.

Теорема 6.1. *Пусть $F(x)$ — первообразная функции $f(x)$ на отрезке $[a,b]$, $\varphi(y)$ — дифференцируемая функция на отрезке $[p,q]$, причём $a \le \varphi(y) \le b$ для всех y из отрезка $[p,q]$ и у любой точки y_0 из отрезка $[p,q]$ есть такая окрестность $U(y_0)$, что если $y \in U(y_0)$ и $y \ne y_0$, то $\varphi(y) \ne \varphi(y_0)$. Тогда $F(\varphi(y))$ — первообразная функции $f(\varphi(y))\varphi'(y)$.*

Доказательство. Формула для производной композиции функций (теорема 5.4) показывает, что $\big(F(\varphi(y))\big)' = F'(\varphi(y))\varphi'(y)$. Ясно также, что $F'(\varphi(y)) = f(\varphi(y))$. $\square$

Теорему 6.1 можно сформулировать следующим образом: если $\int f(x)\,dx = F(x) + C$, то $\int f(\varphi(y))\varphi'(y)\,dy = F(\varphi(y)) + C$. В связи с этим удобно ввести обозначение $\varphi'(y)\,dy = d\varphi(y)$.

Задача 6.2. Вычислите $\int \frac{dx}{x(1+\ln^2 x)}$; $\int \frac{e^x\,dx}{1+e^{2x}}$; $\int \frac{\cos x\,dx}{1+\sin^2 x}$.

Задача 6.3. Вычислите $\int \operatorname{tg} x\,dx$.

Задача 6.4. Докажите, что $\int f(x)\,dx$ можно вычислить следующим образом. Пусть $x = \varphi(y)$ — дифференцируемая функция, имеющая обратную функцию $y = \psi(x)$. Вычислим $\int f(\varphi(y))\varphi'(y)\,dy = F(y)$ и положим $y = \psi(x)$. Тогда $F(\psi(x))$ — искомый интеграл.

Задача 6.5. Вычислите $\int \sqrt{1-x^2}dx$ и $\int \sqrt{a^2-x^2}dx$.

Задача 6.6. Вычислите $\int \sqrt{x^2+1}dx$, $\int \sqrt{x^2-1}dx$ и $\int \sqrt{x^2+A}dx$.

Задача 6.7. Вычислите $\int \frac{dx}{1+\sqrt{x}}$.

Теорема 6.2 (формула интегрирования по частям). *Если у функций $f(x)$ и $g(x)$ есть первообразные $f'(x)$ и $g'(x)$, то*

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

Доказательство. Требуется доказать, что

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

Это немедленно следует из формулы для производной произведения двух функций. □

Задача 6.8. Вычислите с помощью интегрирования по частям: а) $\int \ln x\,dx$; б) $\int x^3 \ln x\,dx$; в) $\int \operatorname{arctg} x\,dx$; г) $\int x\cos x\,dx$.

Задача 6.9. Вычислите а) $\int xe^x dx$; б) $\int x^m e^{-x} dx$, где m — натуральное число.

Задача 6.10. Вычислите $\int \frac{\ln x}{x} dx$.

Задача 6.11. Вычислите: а) $\int e^{ax}\sin bx\,dx$; б) $\int e^{ax}\cos bx\,dx$.

Задача 6.12. Докажите, что

$$\int fg^{(n+1)}dx = fg^{(n)} - f'g^{(n-1)} + f''g^{(n-2)} - \dots +$$
$$+ (-1)^n f^{(n)}g + (-1)^{n+1}\int f^{(n+1)}g\,dx.$$

Задача 6.13. Пусть $P(x)$ — многочлен. Вычислите: а) $\int P(x)e^{ax}dx$; б) $\int P(x)\sin ax\,dx$; в) $\int P(x)\cos ax\,dx$.

6.2. Вычисление неопределённых интегралов

Обсудим теперь некоторые общие методы вычисления неопределённых интегралов. Начнём с вычисления интеграла от рациональной функции $R(x)$.

Теорема 6.3. *Если $R(x)$ — рациональная функция, то $\int R(x)\,dx$ есть сумма рациональной функции и некоторого числа слагаемых вида $c_i \ln(x - a_i)$.*

Доказательство. Достаточно доказать, что рациональную функцию $R(x)$ можно представить в виде

$$A(x) + \sum_{i,k} \frac{c_{ik}}{(x-a_i)^k}\,,$$

где $A(x)$ — многочлен.

Пусть $R(x) = P(x)/Q(x)$, где P и Q — многочлены. Поделив P на Q с остатком, можно перейти к дроби S/Q, где $\deg S < \deg Q$. Пусть $Q = Q_1Q_2$, где Q_1 и Q_2 — взаимно простые многочлены. Тогда существуют такие многочлены a и b, что $a(x)\,Q_1(x) + b(x)\,Q_2(x) = 1$. Поэтому

$$\frac{S}{Q_1Q_2} = \frac{a_1SQ_1 + a_2SQ_2}{Q_1Q_2} = \frac{a_1S}{Q_2} + \frac{a_2S}{Q_1}\,.$$

В полученных дробях нужно снова поделить с остатком числитель на знаменатель. После нескольких таких операций придём к сумме многочлена $A(x)$ и нескольких дробей вида $p(x)\,(x-a)^{-n}$, где $\deg p(x) < n$. Остается записать многочлен $p(x)$ в виде

$$p(x) = b_1(x-a)^{n-1} + b_2(x-a)^{n-2} + \dots + b_n.$$

□

Вычислим теперь с помощью замены переменной интеграл вида $\int R(x, \sqrt{ax^2+bx+c})\,dx$, где $R(x,y)$ — рациональная функция от двух переменных. Эту замену переменных предложил Леонард Эйлер (1707-1783) в 1768 году.

Наиболее простой случай — когда квадратный трёхчлен вырождается в линейный, т.е. когда интеграл имеет вид $\int R(x, \sqrt{ax+b})\,dx$. В этом случае можно положить $y = t$ и $x = \frac{t^2-b}{a}$, поэтому

$$\int R(x, \sqrt{ax+b})\,dx = \int R\left(\frac{t^2-b}{a}, t\right)\frac{2t\,dt}{a}.$$

Если квадратный трёхчлен под знаком радикала имеет два вещественных корня, то общий случай можно свести к этому вырожденному случаю. Действительно, пусть $y^2 = ax^2 + bx + c = a(x-\alpha)(x-\beta)$. Положим $x = \frac{1}{x_1} + \alpha$ и $y = \frac{y_1}{x_1}$. Тогда $y_1^2 = a(1 - (\alpha - \beta)x_1)$, поэтому

$$\int R(x,y)dx = \int R\left(\frac{1}{x_1} + \alpha, \frac{y_1}{x_1}\right) d\left(\frac{1}{x_1} + \alpha\right) = \int R_1(x_1, y_1)dx_1,$$

где R_1 — некоторая рациональная функция.

Несложно указать и замену, позволяющую вычислить интеграл $\int R(x, \sqrt{ax^2+bx+c})\, dx$ непосредственно, не сводя его к вырожденному случаю. Действительно, положим $t^2 = a\frac{x-\beta}{x-\alpha}$. Тогда через t рационально выражаются как $x = \frac{a\beta - \alpha t^2}{a - t^2}$, так и $\sqrt{ax^2+bx+c} = (x-\alpha)t$.

Отметим, что замена переменной, аналогичная сведению интеграла от $R(x, \sqrt{ax^2+bx+c})$ к интегралу от $R(x, \sqrt{Ax+B})$, позволяет свести интеграл с корнем из многочлена степени $2n$ к интегралу с корнем из многочлена степени $2n-1$. Прежде всего покажем, как с помощью замены переменной можно преобразовать кривую вида $y^2 = P_{2n}(x)$ в кривую вида $y_1^2 = P_{2n-1}(x_1)$ (здесь P_m обозначает многочлен степени m). Пусть α — корень многочлена P_{2n}. Положим $x = \frac{1}{x_1} + \alpha$ и $y = \frac{y_1}{x_1^n}$. Тогда уравнение $y^2 = P_{2n}(x)$ запишется в виде

$$\begin{aligned}
\frac{y_1^2}{x_1^{2n}} &= P_{2n}\left(\frac{1}{x_1} + \alpha\right) = \\
&= \frac{b_{2n}}{x_1^{2n}} + \frac{b_{2n-1}}{x_1^{2n-1}} + \cdots + \frac{b_1}{x_1} + P_{2n}(\alpha) = \\
&= \frac{b_{2n}}{x_1^{2n}} + \frac{b_{2n-1}}{x_1^{2n-1}} + \cdots + \frac{b_1}{x_1},
\end{aligned}$$

поскольку $P_{2n}(\alpha) = 0$. Поэтому $y_1^2 = b_{2n} + b_{2n-1}x_1 + \cdots + b_1 x^{2n-1}$, что и требовалось. Теперь, применяя эту замену переменной к интегралу $\int R(x,y)dx$, где R — некоторая рациональная функция, получаем:

$$\int R(x,y)dx = \int R\left(\frac{1}{x_1} + \alpha, \frac{y_1}{x_1^n}\right) d\left(\frac{1}{x_1} + \alpha\right) = \int R_1(x_1, y_1)dx_1,$$

где R_1 — некоторая рациональная функция.

Если квадратный трёхчлен под знаком радикала имеет мнимые корни, то рассмотренная выше замена переменной приводит к выражению, содержащему комплексные числа. Чтобы избежать их, можно применить общий подход, основанный на рациональной параметризации кривой второго порядка (коники). Фиксируем на конике некоторую точку (x_0, y_0) и проведём через нее прямую $y = y_0 + t(x - x_0)$. Она пересекает конику ещё в одной точке. Вычислим координаты этой точки. Для этого нужно в уравнение коники вместо y подставить $y_0 + t(x - x_0)$; в результате получим уравнение вида $A(t)x^2 + B(t)x + C(t) = 0$, где A, B и C — многочлены от t. У этого уравнения известен один корень x_0. По теореме Виета второй корень x_1 равен $-x_0 - \frac{B(t)}{A(t)} = \frac{P(t)}{A(t)}$. Кроме того,

$$y_1 = y_0 + t(x_1 - x_0) = y_0 + t\left(-2x_0 - \frac{B(t)}{A(t)}\right) = \frac{Q(t)}{A(t)}.$$

Следовательно, для кривой второго порядка существуют такие многочлены $A(t), P(t)$ и $Q(t)$, что при изменении t от $-\infty$ до $+\infty$ точка с координатами $\left(\frac{P(t)}{A(t)}, \frac{Q(t)}{A(t)}\right)$ движется по данной кривой; если при этом добавить точку, соответствующую бесконечному значению t, то в результате получим все точки кривой. Такое представление кривой с помощью многочленов A, P и Q называют *рациональной параметризацией.*

Кривая $y^2 = ax^2 + bx + c$ допускает рациональную параметризацию $(x(t), y(t))$. Поэтому

$$\int R(x, \sqrt{ax^2+bx+c})\, dx = \int R(x(t), y(t))\, dx(t) = \int Q(t)\, dt,$$

где Q — некоторая рациональная функция.

Рассмотрим теперь интеграл от бинома $\int x^m(a+bx^n)^p dx$, где $m,\ n,\ p \in \mathbb{Q}$. Этот интеграл выражается в элементарных функциях в следующих трёх случаях.

Случай 1. *p — целое число.*

Если l — наименьшее общее кратное чисел m и n, то рассматриваемый интеграл имеет вид $\int R(\sqrt[l]{x})dx$. Его можно вычислить с помощью подстановки $t = \sqrt[l]{x}$.

Случай 2. *$\dfrac{m+1}{n}$ — целое число.*

После подстановки $z = x^n$ получаем интеграл $\displaystyle\int \frac{1}{n}(a+bz)^p z^{\frac{m+1}{n}-1}dz$. По условию $\frac{m+1}{n}$ — целое число, поэтому полученный интеграл имеет вид $\int R(z, \sqrt[k]{a+bz})dz$, где k — знаменатель числа p. Этот интеграл можно вычислить с помощью подстановки $t = \sqrt[k]{a+bz} = \sqrt[k]{a+bx^n}$.

Случай 3. *$\dfrac{m+1}{n} + p$ — целое число.*

Пусть $\frac{m+1}{n} - 1 = q$. После подстановки $z = x^n$ получаем интеграл

$$\int (a+bz)^p z^q dz = \int \left(\frac{a+bz}{z}\right)^p z^{p+q} dz.$$

По условию $p+q$ — целое число, поэтому полученный интеграл имеет вид $\int R(z, \sqrt[k]{\frac{a+bz}{z}})dz$, где k — знаменатель числа p. Этот интеграл можно вычислить с помощью подстановки $t = \sqrt[k]{\frac{a+bz}{z}} = \sqrt[k]{ax^{-n}+b}$.

Историческое замечание. Интеграл от бинома в указанных трёх случаях вычислил Исаак Ньютон (1643-1727) в 1676 году. Он полагал, что в остальных случаях интеграл $\int x^m(a+bx^n)^p dx$ нельзя выразить через элементарные функции. Но доказать это смог лишь Пафнутий Львович Чебышев (1821-1894) в 1853 году.

При интегрировании рациональных функций от $\sin x$ и $\cos x$ можно воспользоваться заменой переменной $\operatorname{tg}\frac{x}{2} = t$. Действительно, $\sin x = \frac{2t}{1+t^2}$, $\cos x = \frac{1-t^2}{1+t^2}$ и $dx = 2d\operatorname{arctg} t = \frac{2dt}{1+t^2}$. Поэтому

$$\int R(\sin x, \cos x)dx = \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2} = \int R_1(t)dt,$$

где R_1 — некоторая рациональная функция.

6.3. Интеграл Римана

Пусть на отрезке $[a,b]$ задана функция $f(x)$. Рассмотрим *разбиение* этого отрезка на части точками $x_0 = a < x_1 < x_2 < \dots < x_{n-1} < b = x_n$. На каждом отрезке $[x_k, x_{k+1}]$ выберем произвольную точку ξ_k и рассмотрим так называемую *интегральную сумму* $\sum_{k=0}^{n-1} f(\xi_k)(x_{k+1}-x_k)$. Оказывается, что если функция $f(x)$ непрерывна и наибольшая длина отрезка разбиения стремится к нулю, то интегральные суммы имеют конечный предел, который не зависит от выбора разбиения и от выбора точек ξ_k. Этот предел называют *определённым интегралом* функции $f(x)$ по отрезку $[a,b]$ и обозначают $\int_a^b f(x)\,dx$.

Функцию $f(x)$ называют *интегрируемой по Риману* на отрезке $[a,b]$, если для неё существует указанный предел интегральных сумм. Для неограниченной функции интегральные суммы не ограничены, поэтому интегрируемой по Риману может быть только ограниченная функция.

Историческое замечание. В 1854 году Бернхард Риман (1826-1866) дал определение интеграла, которое годилось и для некоторых функций, имеющих бесконечно много точек разрыва. С такими функциями от встретился при исследовании рядов Фурье, и для них не годились более ранние определения, например, определение Коши.

Пример 6.1. *Рассмотрим на отрезке $[0,1]$ функцию, которая равна 1 в рациональных точках и 0 в иррациональных точках. Эта функция не интегрируема по Риману.*

Доказательство. Для любого разбиения отрезка в качестве точек ξ_k можно выбрать рациональные точки. В этом случае интегральная сумма равна 1. Если же в качестве точек ξ_k выбрать иррациональные точки, то интегральная сумма равна 0. □

Историческое замечание. Функцию из примера 6.1 предложил Петер Густав Лежён Дирихле (1805-1859) в 1829 году. Её называют *функцией Дирихле*.

Интегрируемыми могут быть не только непрерывные функции, но в начале этой главы нас в основном будут интересовать непрерывные функции. Чтобы доказать интегрируемость по Риману непрерывных функций, введём следующие вспомогательные понятия. Пусть на отрезке $[a, b]$ задана ограниченная функция $f(x)$ и задано некоторое разбиение этого отрезка. Для каждого отрезка $[x_k, x_{k+1}]$ рассмотрим M_k и m_k — наибольшее и наименьшее значения функции $f(x)$ на этот отрезке. Суммы

$$\sum_{k=0}^{n-1} M_k(x_{k+1} - x_k) \quad \text{и} \quad \sum_{k=0}^{n-1} m_k(x_{k+1} - x_k)$$

назовём *верхней* и *нижней интегральными суммами*; они зависят только от разбиения и не зависят от выбора точек ξ_k. Ясно, что каждая интегральная сумму заключена между верхней и нижней интегральными суммами для того же самого разбиения.

Историческое замечание. Верхнюю и нижнюю интегральные суммы ввёл Гастон Дарбу (1842-1917) в 1875 году.

Лемма 6.1. *При добавлении новых точек разбиения нижняя интегральная сумма не уменьшается, а верхняя не увеличивается.*

Доказательство. Достаточно рассмотреть случай, когда добавляется одна точка деления. Пусть на отрезке $[c, d]$ взята точка y и наибольшие значения функции $f(x)$ на отрезках $[c, d]$, $[c, y]$ и $[y, d]$ равны M, M_1 и M_2. Требуется доказать, что

$$M_1(y - c) + M_2(d - y) \leq M(d - c).$$

Ясно, что $M_1 \leq M$ и $M_2 \leq M$. Поэтому

$$M_1(y - c) + M_2(d - y) \leq M(y - c) + M(d - y) = M(d - c).$$

Для нижних интегральных сумм доказательство аналогично. □

Лемма 6.2. *Любая нижняя интегральная сумма не больше любой верхней интегральной суммы.*

Доказательство. Пусть s_1 — нижняя интегральная сумма для одного разбиения, S_2 — верхняя интегральная сумма для другого разбиения. Разобьём отрезок всеми точками, входящими в эти разбиения. Пусть s и S — нижняя и верхняя интегральная суммы для полученного разбиения. Согласно лемме 6.1 $s_1 \leq s$ и $S \leq S_2$. Остаётся заметить, что $s \leq S$. □

Лемма 6.3. *Пусть I — точная верхняя грань нижних интегральных сумм, σ — произвольная интегральная сумма, s и S — нижняя и верхняя интегральные суммы, соответствующие тому же разбиению, что и σ. Тогда*

$$|\sigma - I| \leq S - s.$$

Доказательство. Прежде всего заметим, что точная верхняя грань нижних интегральных сумм существует, поскольку согласно лемме 6.2 любая нижняя интегральная сумма s не превосходит некоторой фиксированной интегральной суммы S_0. Из неравенства $s \leq S_0$ следует, что $I \leq S_0$. Это неравенство выполняется для произвольной верхней интегральной суммы S_0. Таким образом, $s \leq I \leq S$. Ясно также, что $s \leq \sigma \leq S$. Из этих двух неравенств следует требуемое неравенство. □

Теорема 6.4. *Любая непрерывная функция $f(x)$ на отрезке $[a, b]$ интегрируема по Риману.*

Доказательство. Согласно лемме 6.3 достаточно доказать, что если наибольшая длина отрезка разбиения стремится к нулю, то разность $S - s$ стремится к нулю.

Функция $f(x)$ равномерно непрерывна на отрезке $[a, b]$ (теорема 3.8), поэтому для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что если расстояние между точками x' и x'' отрезка $[a, b]$ меньше δ, то $|f(x') - f(x'')| < \varepsilon$. Поэтому если наибольшая длина отрезка разбиения меньше δ, то $M_k - m_k < \varepsilon$, а значит,

$$S - s = \sum_{k=0}^{n-1}(M_k - m_k)(x_{k+1} - x_k) < \varepsilon \sum_{k=0}^{n-1}(x_{k+1} - x_k) = \varepsilon(b - a).$$

Таким образом, если наименьшая длина отрезка разбиения стремится к нулю, то разность $S - s$ стремится к нулю. □

6.4. Теорема о среднем

Теорема 6.5 (о среднем). *Пусть $f(x)$ — непрерывная функция на отрезке $[a, b]$. Тогда на отрезке $[a, b]$ существует точка ξ, для которой $\int_a^b f(x)\,dx = f(\xi)(b - a)$.*

Доказательство. Пусть M и m — наибольшее и наименьшее значения функции $f(x)$ на отрезке $[a, b]$. Тогда каждая интегральная сумма заключена между $m(b - a)$ и $M(b - a)$. Поэтому $\int_a^b f(x)\,dx = c(b - a)$, где число c заключено между m и M. Непрерывная функция на отрезке принимает все значения между m и M, поэтому $c = f(\xi)$ для некоторой точки ξ отрезка $[a, b]$. □

Задача 6.14. Докажите, что если непрерывная функция $f(x)$ неотрицательна на отрезке $[a, b]$, то $\int_a^b f(x)\,dx \geq 0$. Более того, если $f(x_0) > 0$ для некоторой точки x_0 отрезка $[a, b]$, то $\int_a^b f(x)\,dx > 0$.

Задача 6.15. Пусть $f(x)$ и $\varphi(x)$ — непрерывные функции на отрезке $[a, b]$, причём функция $\varphi(x)$ неотрицательна. Докажите, что

$$\int_a^b f(x)\varphi(x)dx = f(\xi)\int_a^b \varphi(x)dx$$

для некоторого $\xi \in (a, b)$.

6.5. Формула Ньютона–Лейбница

Теорема 6.6. *Пусть $f(x)$ — непрерывная функция на отрезке $[a, b]$ и $F(x) = \int_a^x f(t)\,dt$. Тогда $F'(x) = f(x)$, т.е. $F(x)$ — первообразная функции $f(x)$.*

Доказательство. Пусть точки x и $x + \Delta x$ лежат на отрезке $[a, b]$. Тогда $F(x + \Delta x) - F(x) = \int_x^{x+\Delta x} f(t)\,dt$. Согласно теореме 6.5 этот интеграл равен $f(\xi)\Delta x$, где ξ — некоторая точка между x и $x + \Delta x$. Следовательно, $\lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = f(x)$, поскольку $\xi \to x$ при $\Delta x \to 0$. □

Теорема 6.7 (формула Ньютона–Лейбница). *Пусть $f(x)$ — непрерывная функция на отрезке $[a, b]$. Тогда*

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

где $F(x)$ — первообразная функции $f(x)$.

Для разности $F(b) - F(a)$ мы будем использовать обозначение $F(x)|_a^b$.

Доказательство. Рассмотрим функцию $\Phi(x) = \int_a^x f(t)\,dt$. Ясно, что $\Phi(b) = \int_a^b f(x)\,dx$. Кроме того, согласно теореме 6.6 $\Phi'(x) = f(x)$. Поэтому $\Phi(x)$ тоже первообразная, а значит, $\Phi(x) = F(x) + C$, где C — некоторая константа. Заметим, что $\Phi(a) = 0$. Поэтому $F(a) = -C$. Следовательно,

$$\int_a^b f(x)\,dx = \Phi(b) = F(b) + C = F(b) - F(a).$$

□

Историческое замечание. Первые публикации о связи интегрирования и дифференцирования принадлежат Джеймсу Грегори (1638-1675) и Исааку Барроу (1630-1677). В 1668 году Грегори опубликовал первый учебник анализа. В этом учебнике было показано, что метод касательных (дифференцирование) и метод квадратур (интегрирование) взаимно обратны. В 1670 году Барроу опубликовал свои лекции по геометрии, посвящённые методу касательных. Он установил, что задача о вычислении площади под кривой и задача о построении касательной взаимно обратны. К публикации лекции Барроу готовили его ученики, в их числе Исаак Ньютон (1643-1727). Завершили оформление этих геометрических идей независимо Ньютон и Готфрид Вильгельм Лейбниц (1646-1716).

Для положительных a и b из формулы Ньютона–Лейбница следует, что $\int_a^b x^n dx = \frac{b^{n+1}-a^{n+1}}{n+1}$ при $n \neq -1$ и $\int_a^b x^{-1}dx = \ln b - \ln a$. Отметим, что $\lim_{n\to -1} \frac{b^{n+1}-a^{n+1}}{n+1} = \lim_{x\to 0} \frac{b^x - a^x}{x} = \lim_{x\to 0} \frac{(b^x-1)-(a^x-1)}{x} = \ln b - \ln a$ поскольку $\lim_{x\to 0} \frac{a^x-1}{x} = \ln a$ для любого положительного a (задача 3.32).

Задача 6.16. Функция f непрерывна на некотором отрезке, числа a и b положительны, причём $f(x+a)-f(x) > 0$ и $f(x+b)-f(x) < 0$ на всей области определения. Докажите, что длина отрезка, на котором определена функция f, меньше $a+b$.

Задача 6.17. Для натурального числа $n > 2$ приведите пример функции f, непрерывной на отрезке $[0, 2n-1]$, для которой $f(x+n) < f(x)$ и $f(x+n+1) > f(x)$ на всей области определения.

Задача 6.18. *Для взаимно простых натуральных чисел m и n $(m > n+1 > 2)$ приведите пример функции f, непрерывной на отрезке $[0, m+n-2]$, для которой $f(x+n) < f(x)$ и $f(x+m) > f(x)$ на всей области определения.

6.6. Формула замены переменной

Пусть функция φ имеет непрерывную производную на отрезке $[\alpha, \beta]$, $\varphi(\alpha) = a$, $\varphi(\beta) = b$ и функция f непрерывна на области значений функции φ. Тогда имеет место формула замены переменной для определённого интеграла:

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

Действительно, рассмотрим интегралы с переменными верхними пределами $F(x) = \int_a^x f(y)dy$ и $\Phi(t) = \int_\alpha^t f(\varphi(z))\varphi'(z)dz$. Положим $x = \varphi(t)$ и вычислим производную сложной функции $F(\varphi(t)) = \int_a^{\varphi(t)} f(y)dy$:

$$\frac{dF(\varphi(t))}{dt} = \frac{dF(x)}{dx}\frac{d\varphi(t)}{dt} = f(x)\varphi'(t) = f(\varphi(t))\varphi'(t).$$

Таким образом, функции $F(\varphi(t))$ и $\Phi(t)$ имеют на отрезке $[\alpha, \beta]$ равные производные и $F(\varphi(\alpha)) = 0 = \Phi(\alpha)$. Поэтому $F(\varphi(t)) = \Phi(t)$ для всех $t \in [\alpha, \beta]$. При $t = \beta$ получаем требуемое равенство.

Часто бывает удобно рассматривать определённые интегралы $\int_b^a f(x)\,dx$ для $a < b$. А именно, мы полагаем $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$. При таких соглашениях остаются верными формулы замены переменной и соотношение

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

Формула замены переменной свидетельствует о том, что под знаком интеграла не случайно стоит не функция $f(x)$, а некоторый другой объект $f(x)dx$. При замене переменной $x = \varphi(t)$ он заменяется на $f(\varphi(t))\frac{dx}{dt}dt = f(\varphi(t))\varphi'(t)dt$. Этот объект (дифференциальная 1-форма) обсуждается в § 6.21.

6.7. Остаточный член в интегральной форме

Напомним, что остаточный член $R_n(x)$ — это разность $f(x)-T(x)$, где $T_n(x)=\sum_{k=0}^{n}\frac{f^k(a)}{k!}(x-a)^k$. Остаточный член в форме Лагранжа (см. с. 76) — это выражение вида $R_n(x)=\frac{f^{(n+1)}(\theta)}{(n+1)!}(x-a)^{n+1}$, где $\theta\in[a,x]$. В остаточный член в форме Лагранжа входит неопределённая величина θ, но можно записать остаточный член в форме, не содержащей неопределённых величин. Таков, например, *остаточный член в интегральной форме*

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

Для доказательства этой формулы воспользуемся обобщённой формулой интегрирования по частям (задача 6.12):

$$\begin{aligned}\int fg^{(n+1)}dt &= fg^{(n)}-f'g^{(n-1)}+f''g^{(n-2)}-\dots+\\ &\quad +(-1)^n f^{(n)}g+(-1)^{n+1}\int f^{(n+1)}g\,dt.\end{aligned}$$

Положим $g(t)=(x-t)^n$ и рассмотрим интегралы от a до x. Ясно, что $g^{(k)}(t)=(-1)^k\frac{n!}{(n-k)!}(x-t)^{n-k}$ при $k\le n$ и $g^{(n+1)}=0$. Поэтому $f(t)g^{(n)}(t)\big|_a^x=(-1)^n n!f(t)\big|_a^x=(-1)^n n!(f(x)-f(a))$ и

$$\begin{aligned}(-1)^k f^{(k)}(t)g^{(n-k)}(t)\Big|_a^x &= (-1)^{k+1}f^{(k)}(t)(-1)^{n-k}\frac{n!}{k!}(x-t)^k\Big|_a^x=\\ &= (-1)^{n+1}f^{(k)}(t)(-1)^{n-k}\frac{n!}{k!}(x-a)^k\end{aligned}$$

при $k=1,\ldots,n$. В результате получаем

$$\begin{aligned}f(x) &= f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\dots+\\ &\quad +\frac{f^{(n)}(a)}{n!}(x-a)^n+\frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n dt.\end{aligned}$$

6.8. Вычисление определённых интегралов

Приведём несколько задач на вычисление определённых интегралов.

Задача 6.19. Пусть $U_n=\int_0^{\pi/2}\cos^n x\,dx$.

а) Докажите, что если n нечётно, то $U_n=\frac{(n-1)!!}{n!!}$, а если n чётно, то $U_n=\frac{(n-1)!!}{n!!}\frac{\pi}{2}$, где $n!!$ обозначает произведение всех натуральных чисел, не превосходящих n и имеющих с n одинаковую чётность.

б) Докажите, что $\pi=\lim_{n\to\infty}\frac{1}{n}\left(\frac{(2n)!!}{(2n-1)!!}\right)^2$ (*формула Валлиса*).

Задача 6.20. *Докажите, что $\int_0^{\pi/2}\ln(\sin x)dx=-\frac{\pi}{2}\ln 2$.

Задача 6.21. *Вычислите интеграл $\int_0^{\pi}\ln(1-2\alpha\cos x+\alpha^2)dx$.

6.9. Вычисление площадей

Пусть функция $f(x)$ на отрезке $[a,b]$ неотрицательна. Член $f(\xi_k)(x_{k-1}-x_k)$ интегральной суммы заключён между $m(x_{k-1}-x_k)$ и $M(x_{k-1}-x_k)$, где m и M — наибольшее и наименьшее значения функции $f(x)$ на отрезке $[x_k,x_{k+1}]$. Таким образом, этот член заключён между площадью прямоугольника, содержащегося в фигуре под графиком функции $y=f(x)$, и площадью прямоугольника, содержащего часть этой фигуры, расположенную над отрезком $[x_k,x_{k+1}]$. Поэтому определённый интеграл $\int_a^b f(x)\,dx$ — это *площадь* фигуры, ограниченной графиком $y=f(x)$ и прямыми $y=0$, $x=a$ и $x=b$.

Задача 6.22. Вычислите площадь круга радиуса 1 с помощью определённого интеграла.

Задача 6.23. Вычислите площадь под графиком функции $y = \sin x$ на отрезке от 0 до π.

Задача 6.24. а) Вычислите площадь сегмента гиперболы $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, отсечённого прямой $x = x_0$, где $x_0 > a > 0$.

б) Докажите, что площадь сектора, ограниченного дугой гиперболы, осью Ox и лучом, идущим из начала координат в точку (x_0, y_0), в которой указанная прямая пересекает гиперболу, равна

$$\frac{1}{2}ab\ln\left(\frac{x_0}{a} + \frac{y_0}{b}\right).$$

6.10. Вычисление объёмов

Пусть тело расположено в пространстве с прямоугольными координатами $Oxyz$, причём проекция этого тела на ось Ox — это отрезок $[a, b]$. Предположим, что плоскость, проходящая через точку x отрезка $[a, b]$ перпендикулярно оси Ox, высекает на этом теле фигуру площади $S(x)$. Тогда *объём* этого тела определяется как $\int_a^b S(x)\,dx$.

Задача 6.25. Вычислите с помощью определённого интеграла объём V конуса с радиусом R и высотой h.

Задача 6.26. Вычислите с помощью определённого интеграла объём V шара радиуса R.

Задача 6.27. Найдите объём фигуры, образованной при пересечении двух прямых круговых цилиндров радиуса R, оси которых перпендикулярны и пересекаются.

Задача 6.28. Найдите объём фигуры, отсекаемой от цилиндра плоскостью, проходящей через диаметр его основания. Известны радиус цилиндра R и высота полученной фигуры h.

6.11. Длина кривой

Рассмотрим кривую $y = f(x)$, где $f(x)$ — непрерывная функция на отрезке $[a, b]$. Пусть $x_0 = a < x_1 < \dots < x_{n-1} < b = x_n$ — точки на этом отрезке. Рассмотрим ломаную с вершинами $A_k = \bigl(x_k, f(x_k)\bigr)$, $k = 0, 1, \dots, n$. Если длина этой ломаной стремится к конечному пределу l, когда длина наибольшего звена ломаной стремится к нулю, то это число l называют *длиной* кривой.

Теорема 6.8. *Если функция $f(x)$ на отрезке $[a, b]$ имеет непрерывную производную, то кривая $y = f(x)$ имеет длину $\int_a^b \sqrt{1 + f'^2(x)}\,dx$.*

Доказательство. Ясно, что $A_kA_{k+1} = \sqrt{(x_{k+1} - x_k)^2 + \bigl(f(x_{k+1}) - f(x_k)\bigr)^2}$. По теореме Лагранжа

$$f(x_{k+1}) - f(x_k) = f'(\xi_k)(x_{k+1} - x_k)$$

для некоторой точки ξ_k отрезка $[x_k, x_{k+1}]$. Поэтому длина ломаной равна

$$\sum_{k=0}^{n-1} \sqrt{1 + f'^2(\xi_k)}(x_{k+1} - x_k),$$

т.е. она равна интегральной сумме для непрерывной функции $\sqrt{1 + f'^2(x)}$. Остаётся заметить, что длина отрезка $[x_k, x_{k+1}]$ не превосходит A_kA_{k+1}, поэтому длина наибольшего отрезка разбиения стремится к нулю, когда стремится к нулю наибольшая длина звена ломаной. $\square$

Задача 6.29. Вычислите с помощью определённого интеграла длину окружности радиуса R.

Задача 6.30. Вычислите длину дуги параболы $2y = x^2$ от точки $(0, 0)$ до точки $\left(x_0, \frac{x_0^2}{2}\right)$.

Задача 6.31. Вычислите длину дуги кривой $y = \operatorname{ch} x$, заключённой между точками (x_1, y_1) и (x_2, y_2).

Теорема 6.9. *Пусть график функции $y = f(x)$ на отрезке $[a, b]$ параметризован параметром t, т.е. задана монотонно возрастающая функция $x(t)$, причём $a = x(t_0)$ и $b = x(t_1)$, и мы полагаем $y(t) = f(x(t))$. Тогда длина этого графика равна $\int_{t_0}^{t_1} \sqrt{x'^2 + y'^2}dt$, где $x' = \frac{dx}{dt}$ и $y' = \frac{dy}{dt}$.*

Доказательство. По формуле замены переменной

$$\int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \int_{t_0}^{t_1} \sqrt{1 + \frac{y'^2}{x'^2}} x' dt = \int_{t_0}^{t_1} \sqrt{x'^2 + y'^2} dt;$$

при записи последнего равенства мы воспользовались тем, что $x' > 0$. □

Задача 6.32. Вычислите длину *астроиды*, заданной уравнением $x^{2/3} + y^{2/3} = a^{2/3}$.

Задача 6.33. Вычислите длину ветви *циклоиды*

$$x = a(t - \sin t), \quad y = a(1 - \cos t), \quad \text{где} \quad 0 \le t \le 2\pi.$$

Рассмотрим следующую параметризацию эллипса: $x(t) = a \cos t$, $y(t) = b \sin t$. Тогда $x'^2 + y'^2 = a^2 \sin^2 t + b^2 \cos^2 t$. Положив $a^2 - b^2 = e^2 a^2$, получим, что длина дуги эллипса выражается интегралом

$$a \int \sqrt{1 - e^2 \cos^2 t} dt.$$

Сделав замену $u = \cos t$, приходим к интегралу

$$a \int \frac{\sqrt{1 - e^2 u^2}}{\sqrt{1 - u^2}} du = a \int \frac{1 - e^2 u^2}{\sqrt{(1 - u^2)(1 - e^2 u^2)}} du.$$

Лемнискату можно представить параметрически следующим образом: $x = \frac{t(t^2+1)}{1+t^4}$, $y = \frac{t(t^2-1)}{1+t^4}$. В этом случае $x'^2 + y'^2 = \frac{2}{1+t^4}$.

Воспользуемся полярными координатами $x = r \cos \varphi$, $y = r \sin \varphi$. В полярных координатах $ds^2 = dx^2 + dy^2 = (\cos \varphi dr - r \sin \varphi d\varphi)^2 + (\sin \varphi dr + r \cos \varphi d\varphi)^2 = dr^2 + r^2 d\varphi^2$. В декартовых координатах лемниската задаётся уравнением $(x^2 + y^2)^2 = x^2 - y^2$. В полярных координатах она задаётся уравнением $r^2 = \cos 2\varphi$, поэтому $r \frac{dr}{d\varphi} = -\sin 2\varphi$, $r^4 + r^2 \left(\frac{dr}{d\varphi}\right)^2 = 1$, поэтому $ds = \frac{dr}{\sqrt{1 - r^4}}$.

6.12. Площадь поверхности вращения

Пусть $f(x)$ — непрерывная положительная функция на отрезке $[a, b]$, $x_0 = a < x_1 < \dots < x_{n-1} < b = x_n$ — некоторые точки на отрезке $[a, b]$. Будем вращать кривую $y = f(x)$ вокруг оси Ox. Ломаная с вершинами $A_k = (x_k, f(x_k))$ заметает при этом некоторую поверхность, состоящую из усечённых конусов. Если сумма площадей боковых поверхностей этих конусов имеет предел, когда длина наибольшего из отрезков $A_k A_{k+1}$ стремится к нулю, то этот предел называют *площадью поверхности вращения*, заметаемой кривой $y = f(x)$.

Теорема 6.10. *Если положительная функция $f(x)$ на отрезке $[a, b]$ имеет непрерывную производную, то площадь поверхности вращения, заметаемой кривой $y = f(x)$, равна*

$$2\pi \int_a^b f(x) \sqrt{1 + (f'(x))^2} dx.$$

Доказательство. Площадь боковой поверхности усечённого конуса равна произведению образующей на длину среднего сечения. Площадь боковой поверхности конуса, образованного вращением звена $A_k A_{k+1}$, равна

$$\pi\big(f(x_k) + f(x_{k+1})\big) \sqrt{(x_{k+1} - x_k)^2 + \big(f(x_{k+1}) - f(x_k)\big)^2}.$$

По формуле Лагранжа
$$f(x_{k+1}) - f(x_k) = f'(\xi_k)(x_{k+1} - x_k)$$
для некоторой точки ξ_k отрезка $[x_k, x_{k+1}]$. Поэтому сумма площадей боковых поверхностей конусов равна
$$\pi \sum_{k=0}^{n-1} \big(f(x_k) + f(x_{k+1})\big)\sqrt{1 + (f'(\xi_k))^2}(x_{k+1} - x_k).$$
Нужно сравнить эту сумму с интегральной суммой
$$2\pi \sum_{k=0}^{n-1} f(\xi_k)\sqrt{1 + (f'(\xi_k))^2}(x_{k+1} - x_k).$$
Это сравнение несложно получить, основываясь на равномерной непрерывности функции $f(x)$ и ограниченности функции $\sqrt{1 + (f'(x))^2}$. □

Задача 6.34. Докажите, что площадь поверхности шара радиуса R, заключённой между двумя параллельными плоскостями (пересекающими шар), равна $2\pi Rh$, где h — расстояние между этими плоскостями.

6.13. Неравенства

Интегрирование можно применять при доказательстве неравенств. Приведём несколько задач такого рода.

Задача 6.35. Пусть $t > 0$.
а) Докажите, что $t - \frac{t^3}{6} \le \sin t \le t$ и $1 - \frac{t^2}{4} \le \cos t \le 1 - \frac{t^2}{2} + \frac{t^4}{24}$.
б) Докажите, что
$$t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots - \frac{t^{4n+3}}{(4n+3)!} \le \sin t \le t - \frac{t^3}{3!} + \frac{t^5}{5!} - \dots + \frac{t^{4n+1}}{(4n+1)!},$$
$$1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots - \frac{t^{4n+2}}{(4n+2)!} \le \cos t \le 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \dots + \frac{t^{4n}}{(4n)!}.$$

Задача 6.36. Докажите, что если $0 < t \le \pi/2$, то $\left(\frac{\sin t}{t}\right)^3 > \cos t$.

Задача 6.37. Докажите, что если $0 < t \le \pi/2$, то $\frac{1}{\sin^2 t} \le \frac{1}{t^2} + 1 - \frac{4}{\pi^2}$.

Задача 6.38. Докажите, что $3\sin t < 2t + t\cos t$ для любого $t > 0$.

Задача 6.39. Пусть $t > 0$. Докажите, что
$$t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \dots - \frac{t^{2n}}{2n} < \ln(1+t) < t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \dots + \frac{t^{2n+1}}{2n+1}.$$

Задача 6.40. Докажите, что для любого $t > 0$
$$t - \frac{t^3}{3} + \frac{t^5}{5} - \frac{t^7}{7} + \dots - \frac{t^{4n-1}}{4n-1} < \operatorname{arctg} t < t - \frac{t^3}{3} + \frac{t^5}{5} - \frac{t^7}{7} + \dots + \frac{t^{4n+1}}{4n+1}.$$

Задача 6.41. а) Докажите, что если $0 \le t \le a$, то
$$1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots + \frac{t^n}{n!} \le e^t \le 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots + \frac{t^{n-1}}{(n-1)!} + e^a\frac{t^n}{n!}.$$

б) Докажите, что если $0 \ge t \ge a$, то при нечётном n имеют место такие же неравенства, а при чётном n знаки неравенств заменяются на обратные.

* * *

Задача 6.42. Функция $f(x)$ дифференцируема на отрезке $[a, b]$, длина которого равна π. Докажите, что на отрезке $[a, b]$ есть точка x, для которой $f'(x) - \big(f(x)\big)^2 < 1$.

6.14. Вычисление пределов с помощью интегралов

Задача 6.43. Докажите, что $\lim_{n\to\infty}\frac{\sqrt[n]{n!}}{n}=\frac{1}{e}$.

Задача 6.44. Вычислите предел $\lim_{n\to\infty}\sum_{k=0}^{n-1}\frac{n}{n^2+k^2}$.

Задача 6.45. Вычислите предел $\lim_{n\to\infty}\sum_{k=0}^{n-1}\frac{\sqrt{n^2-k^2}}{n^2}$.

Задача 6.46. Вычислите $\lim_{n\to\infty}\sum_{k=n}^{2n}\frac{1}{k}$.

Задача 6.47. Вычислите $\lim_{n\to\infty}\sum_{k=0}^{n-1}\frac{1}{\sqrt{n^2-k^2}}$.

6.15. Несобственные интегралы

Интеграл $\int_a^b f(x)\,dx$ называют *несобственным* в двух случаях:

1) если функция $f(x)$ определена лишь на интервале (a,b), но не на всём отрезке $[a,b]$;

2) если $a=-\infty$ и/или $b=\infty$.

Несобственный интеграл понимается как предел интеграла $\int_u^v f(x)\,dx$, когда $u\to a+0$ (соответственно, $u\to-\infty$) и $v\to b-0$ (соответственно, $v\to+\infty$). Имеется в виду, что u и v стремятся к указанным пределам независимо друг от друга.

Задача 6.48. Вычислите несобственный интеграл $\int_0^1 x^a dx$, где $-1<a<0$.

Задача 6.49. Вычислите несобственный интеграл $\int_1^\infty x^{-a}dx$, где $a>1$.

Задача 6.50. Докажите, что $\int_0^\infty x^k e^{-x}dx=k!$ для любого целого неотрицательного k.

Если несобственный интеграл функции $f(x)$ сходится, то его называют сходящимся *абсолютно* или *условно*, в зависимости от того, сходится или расходится интеграл функции $|f(x)|$.

Пример 6.2. *Интеграл $\int_1^\infty \frac{\cos x}{x^2}dx$ сходится абсолютно.*

Доказательство. Ясно, что $\left|\frac{\cos x}{x^2}\right|\le\frac{1}{x^2}$ и интеграл $\int_1^\infty\frac{1}{x^2}dx$ сходится (задача 6.49). □

Пример 6.3. *Интеграл $\int_{\pi/2}^\infty\frac{\sin x}{x}dx$ сходится условно.*

Доказательство. Интегрируя по частям, получаем

$$\int_{\pi/2}^\infty\frac{\sin x}{x}dx=-\left.\frac{\cos x}{x}\right|_{\pi/2}^\infty+\int_{\pi/2}^\infty\frac{\cos x}{x^2}dx=\int_{\pi/2}^\infty\frac{\cos x}{x^2}dx.$$

Последний интеграл сходится, поэтому первый интеграл тоже сходится.

Если $b>\pi/2$, то

$$\int_{\pi/2}^\infty\left|\frac{\sin x}{x}\right|dx\ge\int_{\pi/2}^\infty\frac{\sin^2 x}{x}dx=\frac{1}{2}\int_{\pi/2}^\infty\frac{dx}{x}-\frac{1}{2}\int_{\pi/2}^\infty\frac{\cos 2x}{x}dx.$$

Сходимость интеграла $\frac{1}{2}\int_{\pi/2}^\infty\frac{\cos 2x}{x}dx$ доказывается интегрированием по частям. Интеграл $\int_{\pi/2}^\infty\frac{dx}{x}$ расходится. □

Если на отрезке $[a,b]$ есть внутренняя точка c, в окрестности которой функция f неограниченная, то интеграл $\int_a^b f(x)dx$ можно представить в виде суммы двух несобственных интегралов $\int_a^c f(x)dx+\int_c^b f(x)dx$. Но часто в таких случаях рассматривают так называемый *интеграл в смысле главного значения*

$$P\int_a^b f(x)dx=\lim_{\delta\to+0}\left(\int_a^{c-\delta}f(x)dx+\int_{c+\delta}^b f(x)dx.\right)$$

Например, интеграл $\int_{-1}^1\frac{dx}{x}$ расходится, а интеграл в смысле главного значения $P\int_{-1}^1\frac{dx}{x}$ равен нулю.

6.16. Равномерная сходимость интегрируемых функций

Начнём с двух задач.

Задача 6.51. Докажите, что последовательность функций, интегрируемых на отрезке, может сходиться к функции, не интегрируемой на этом отрезке.

Задача 6.52. Пусть последовательность функций f_n, интегрируемых на отрезке $[a, b]$, сходится на этом отрезке к интегрируемой функции f. Докажите, что при этом равенство $\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$ может не выполняться.

Но при определённых условиях последовательность интегрируемых функций сходится к интегрируемой функции и при этом последовательность интегралов сходится к интегралу.

Теорема 6.11. *Если последовательность функций f_n, интегрируемых на отрезке $[a, b]$, равномерно сходится на этом отрезке к функции f, то функция f интегрируема и $\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx$. Более того, последовательность функций $F_n(x) = \int_a^x f_n(t)dt$ равномерно сходится на отрезке $[a, b]$ к функции $F(x) = \int_a^x f(t)dt$.*

Доказательство. Функции f_n ограничены, поэтому предельная функция f тоже ограничена (задача 3.45). Докажем, что функция f интегрируема на отрезке $[a, b]$. Пусть задано $\varepsilon > 0$. Выберем натуральное N так, что

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3(b-a)}$$

при $n \geq N$. Из этого неравенства вытекают следующие неравенства для верхней интегральной суммы S и нижней интегральной суммы s (для любого разбиения отрезка):

$$S(f - f_N) \leq \frac{\varepsilon}{3}, \quad s(f - f_N) \geq \frac{\varepsilon}{3}.$$

Функция f_N интегрируема, поэтому разбиение отрезка можно выбрать так, что

$$S(f_N) - s(f_N) \leq \frac{\varepsilon}{3}.$$

В таком случае

$$\begin{aligned} S(f) - s(f) &\leq S(f - f_N) + S(f_N) - s(f_N) - s(f - f_N) < \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Это означает, что функция f интегрируема на отрезке $[a, b]$. Кроме того,

$$\left|\int_a^x f_n(t)dt - \int_a^x f(t)dt\right| \leq \int_a^x |f_n(t) - f(t)|dt \leq \frac{\varepsilon(x-a)}{3(b-a)} < \varepsilon$$

для всех $x \in [a, b]$ и всех $n \geq N$. □

С помощью теоремы 6.11 можно вычислить несобственный интеграл Эйлера–Пуассона.

Пример 6.4. $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$ *(интеграл Эйлера–Пуассона).*

Доказательство. Воспользовавшись вычислением интеграла $U_n = \int_0^{\pi/2} \cos^n x\,dx$ и формулой Валлиса (задача 6.19), несложно получить, что $\lim_{n\to\infty} \sqrt{n} \int_0^{\pi/2} \cos^{2n+1} x\,dx = \frac{1}{2}\sqrt{\pi}$. Действительно,

$$\sqrt{n}\frac{(2n)!!}{(2n+1)!!} = \frac{\sqrt{n}}{2n+1} \cdot \frac{(2n)!!}{(2n-1)!!} = \frac{n}{2n+1} \cdot \left(\frac{1}{\sqrt{n}} \cdot \frac{(2n)!!}{(2n-1)!!}\right) \to \frac{1}{2}\sqrt{\pi}.$$

Сделаем замену переменной $u = \sin x$. Тогда $du = \cos x\,dx$ и $\int_0^{\pi/2} \cos^{2n+1} x\,dx = \int_0^1 (1-u^2)^n du$, поэтому $\lim_{n\to\infty} \sqrt{n} \int_0^1 (1-u^2)^n du = \frac{1}{2}\sqrt{\pi}$. Сделаем ещё одну замену переменной: $t = \sqrt{n}u$. В результате получим $\lim_{n\to\infty} \int_0^{\sqrt{n}} \left(1 - \frac{t^2}{n}\right)^n dt = \frac{1}{2}\sqrt{\pi}$. Фиксируем число $a > 1$ и будем рассматривать числа

n, для которых $\sqrt{n} > a$. Тогда интеграл от 0 до $\sqrt{n}$ можно представить в виде суммы интеграла от 0 до a и интеграла от a до $\sqrt{n}$:

$$\int_0^{\sqrt{n}} \left(1 - \frac{t^2}{n}\right)^n dt = \int_0^a \left(1 - \frac{t^2}{n}\right)^n dt + \int_a^{\sqrt{n}} \left(1 - \frac{t^2}{n}\right)^n dt.$$

Согласно задаче 3.46 на любом отрезке последовательность функций $f_n(x) = \left(1 + \frac{x}{n}\right)^n$ равномерно сходится к функции e^x. Поэтому $\lim_{n\to\infty} \int_0^a \left(1 - \frac{t^2}{n}\right)^n dt = \int_0^a e^{-t^2} dt$.

Пусть $t > a > 1$. Тогда $e^{-t^2} < e^{-t}$. Согласно задаче 5.37 а) $e^x > 1 + x$ для любого $x \neq 0$, поэтому $1 - \frac{t^2}{n} < e^{-\frac{t^2}{n}}$ и $\left(1 - \frac{t^2}{n}\right)^n < e^{-t^2}$. Следовательно,

$$0 < \int_a^{\sqrt{n}} \left(1 - \frac{t^2}{n}\right)^n dt < \int_a^{\sqrt{n}} e^{-t^2} dt < \int_a^{\sqrt{n}} e^{-t} dt = e^{-a} - e^{-\sqrt{n}} < e^{-a}.$$

Устремляя a к бесконечности, получаем $\int_0^\infty e^{-x^2} dx = \frac{1}{2}\sqrt{\pi}$. Функция e^{-x^2} чётная, поэтому $\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}$. □

6.17. Ортогональные многочлены

В пространстве многочленов степени не выше $n - 1$ можно задать скалярное произведение следующим образом: скалярное произведение многочленов $P(x)$ и $Q(x)$ равно $\int_a^b P(x)Q(x)dx$. Все свойства скалярного произведения (симметричность, линейность, положительность) при этом очевидны. Отметим, что непосредственным следствием того, что формула такого вида задаёт скалярное произведение, является неравенство Буняковского.

Теорема 6.12 (неравенство Буняковского). *Пусть $f(x)$ и $g(x)$ — непрерывные функции на отрезке $[a, b]$. Тогда*

$$\left(\int_a^b f(x)g(x)dx\right)^2 \leq \int_a^b (f(x))^2 dx \int_a^b (g(x))^2 dx.$$

Доказательство. Рассмотрим квадратный трёхчлен

$$\int_a^b (tf(x) + g(x))^2 dx = At^2 + Bt + C,$$

где $A = \int_a^b (f(x))^2 dx$, $B = 2\int_a^b f(x)g(x)dx$ и $C = \int_a^b (g(x))^2 dx$. Этот квадратный трёхчлен неотрицателен при всех t, поэтому его дискриминант $B^2 - 4AC$ неположителен. Следовательно, $(B/2)^2 \leq AC$, а это и есть требуемое неравенство. □

Замечание. Виктор Яковлевич Буняковский (1804-1889) доказал это неравенство для интегралов в 1859 году. Иногда это неравенство называют неравенством Шварца, хотя Герман Шварц (1843-1921) доказал его лишь в 1885 году. В 1821 году Огюстен Луи Коши (1789-1857) доказал аналогичное неравенство для чисел (задача 1.8).

Скалярное произведение позволяет применить процесс ортогонализации и построить систему многочленов $P_0(x) = 1$, $P_1(x)$, $P_2(x)$, ... ($\deg P_n = n$), ортогональных относительно этого скалярного произведения.

Найдём многочлен $P_n(x)$, обладающий следующим свойством: $\int_a^b P_n(x)Q(x)dx = 0$ для любого многочлена $Q(x)$, степень которого меньше n. Будем искать многочлен $P_n(x)$ в виде n-й производной многочлена R степени $2n$: $P_n(x) = \frac{d^n R(x)}{dx^n}$. При этом P_n не изменится, если к многочлену R добавить произвольный многочлен степени меньше n, поэтому можно считать, что $R(a) = R'(a) = \cdots =$

$R^{(n-1)}(a) = 0$. Интегрируя по частям, получаем

$$\int_a^b Q(x)\frac{d^nR(x)}{dx^n}dx =$$
$$\left(Q(x)\frac{d^{n-1}R(x)}{dx^{n-1}} - \frac{dQ(x)}{dx}\frac{d^{n-2}R(x)}{dx^{n-2}} + \dots \pm \frac{d^{n-1}Q(x)}{dx^{n-1}}R(x)\right)\Big|_a^b =$$
$$= Q(b)\frac{d^{n-1}R(b)}{dx^{n-1}} - \frac{dQ(b)}{dx}\frac{d^{n-2}R(b)}{dx^{n-2}} + \dots \pm \frac{d^{n-1}Q(b)}{dx^{n-1}}R(b).$$

Это выражение должно быть равно нулю для любых $Q(b)$, $\frac{dQ(b)}{dx}$, ..., $\frac{d^{n-1}Q(b)}{dx^{n-1}}R(b)$, поэтому $R(b) = R'(b) = \dots = R^{(n-1)}(b) = 0$. Следовательно, $R(x) = C(x-a)^n(x-b)^n$, где C — некоторая константа, и

$$P_n(x) = C\frac{d^n}{dx^n}((x-a)^n(x-b)^n).$$

Если $a = -1$ и $b = 1$, то многочлен P_n называют *многочленом Лежандра*.

Скалярное произведение многочленов $P(x)$ и $Q(x)$ можно определить не только формулой $\int_a^b P(x)Q(x)dx$, но и формулой $\int_a^b P(x)Q(x)w(x)dx$, где $w(x)$ — непрерывная строго положительная функция на отрезке $[a, b]$, называемая *весовой функцией*. Для скалярного произведения с весовой функцией $w(x)$ можно построить систему многочленов $\varphi_0(x) = 1$, $\varphi_1(x)$, $\varphi_2(x)$, ... $(\deg\varphi_n = n)$, ортогональных относительно этого скалярного произведения.

Для любой весовой функции на отрезке $[a, b]$ ортогональные многочлены $\varphi_n(x)$ обладают некоторыми общими свойствами. Докажем два таких свойства: теоремы 6.13 и 6.14.

Теорема 6.13. *Внутри отрезка $[a, b]$ многочлен $\varphi_n(x)$ имеет ровно n некратных корней.*

Доказательство. Пусть x_1, ..., x_m — корни многочлена $\varphi_n(x)$ на отрезке $[a, b]$, причём каждый корень берётся столько раз, какова его кратность. Предположим, что либо $m < n$, либо $m = n$, но среди данных корней есть совпадающие. В первом случае рассмотрим многочлен $\psi(x)$, равный $(x - x_1)\dots(x - x_m)$, а во втором случае поделим этот многочлен на $(x - x_i)^2$, где x_i — кратный корень. Многочлен $\psi(x)$ обладает двумя свойствами: 1) на всём отрезке $[a, b]$ многочлен $\varphi_n(x)\psi(x)$ либо неотрицательный, либо неположительный; 2) степень многочлена $\psi(x)$ строго меньше n. Из свойства 1) следует, что $(\varphi_n, \psi) \neq 0$ (ясно, что многочлен $\varphi_n(x)\psi(x)$ ненулевой). Из свойства 2) следует, что $\psi = \lambda_0\varphi_0 + \dots + \lambda_{n-1}\varphi_{n-1}$, где λ_0, ..., λ_{n-1} — константы. Поэтому $(\varphi_n, \psi) = 0$. Приходим к противоречию. □

Теорема 6.14. *Любые три последовательных ортогональных многочлена связаны соотношением*

$$\varphi_{n+1}(x) = (a_nx + b_n)\varphi_n(x) + c_n\varphi_{n-1}(x), \quad n = 1, 2, \dots,$$

где a_n, b_n и c_n — некоторые константы.

Доказательство. Пусть $\varphi_n(x) = k_nx^n + k'_nx^{n-1} + \dots$. Коэффициент многочлена $\varphi_{n+1}(x) - a_nx\varphi_n(x)$ при x^{n+1} равен $k_{n+1} - a_nk_n$, поэтому если $a_n = k_{n+1}/k_n$, то этот многочлен имеет степень не выше n, а значит,

$$\varphi_{n+1}(x) - a_nx\varphi_n(x) = \lambda_0\varphi_0(x) + \dots + \lambda_n\varphi_n(x)$$

для некоторых констант λ_0, ..., λ_n.

Из условия ортогональности следует, что $\big(\varphi_m(x), x^k\big) = 0$ при $k \leq m-1$. Поэтому $\big(x\varphi_n(x), x^k\big) = \big(\varphi_n(x), x^{k+1}\big) = 0$ при $k \leq n-2$. Таким образом,

$$\big(\lambda_0\varphi_0(x) + \dots + \lambda_n\varphi_n(x), x^k\big) = 0$$

при $k \leq n-2$. Любой многочлен φ_0, ..., φ_{n-2} линейно выражается через 1, x, ..., x^{n-2}, поэтому

$$\big(\lambda_0\varphi_0(x) + \dots + \lambda_n\varphi_n(x), \varphi_k(x)\big) = 0$$

при $k \leq n-2$. Следовательно, $\lambda_0 = \dots = \lambda_{n-2} = 0$. В итоге получаем

$$\varphi_{n+1}(x) - a_nx\varphi_n(x) = \lambda_{n-1}\varphi_{n-1}(x) + \lambda_n\varphi_n(x).$$

Остаётся положить $b_n = \lambda_n$ и $c_n = \lambda_{n-1}$. □

Например, любые три последовательных многочлена Лежандра связаны соотношением

$$P_{n+1}(x) = \frac{2n+1}{n+1}xP_n(x) - \frac{n}{n+1}P_{n-1}(x), \quad n \geq 1.$$

В самом деле, $P_n(x) = k_n x^n + \dots$, где $k_n = \frac{(2n)!}{2^n(n!)^2}$. Поэтому согласно теореме 6.14 $P_{n+1}(x) - a_n x P_n(x) = b_n P_n(x) + c_n P_{n-1}(x)$, где $a_n = \frac{k_{n+1}}{k_n} = \frac{2n+1}{n+1}$. Чтобы найти константы b_n и c_n, мы воспользуемся тем, что $P_n(1) = 1$ и $P_n(-1) = (-1)^n$. Эти равенства следуют из того, что

$$\begin{aligned} P_n(x) &= \frac{1}{2^n n!}\frac{d^n}{dx^n}\big((x-1)^n(x+1)^n\big) = \\ &= \frac{1}{2^n n!}\sum_{k=0}^{n}\binom{n}{k}\frac{n!}{(n-k)!}(x-1)^{n-k}\frac{n!}{k!}(x+1)^k. \end{aligned}$$

Для b_n и c_n при $x = \pm 1$ получаем уравнения $1 - \frac{2n+1}{n+1} = b_n + c_n$, $1 - \frac{2n+1}{n+1} = -b_n + c_n$, откуда $b_n = 0$ и $c_n = -\frac{n}{n+1}$.

Историческое замечание. Адриен-Мари Лежандр (1752-1833) ввёл многочлены P_n в 1782 году при исследовании притяжения тел вращения. Он доказал некоторые свойства этих многочленов, в том числе ортогональность. Формулу $P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$ получил Бенжамен Олинд Родриг (1795-1851) в 1816 году.

Теорема 6.15. *Многочлены Чебышева образуют ортогональную систему многочленов на отрезке* $[-1,1]$ *с весовой функцией* $w(x) = \dfrac{1}{\sqrt{1-x^2}}$.

Доказательство. Сделав замену $x = \cos\varphi$, получим

$$\begin{aligned} \int_{-1}^{1} T_n(x)T_m(x)\frac{dx}{\sqrt{1-x^2}} &= \int_0^\pi \cos n\varphi \cos m\varphi \, d\varphi = \\ &= \int_0^\pi \frac{\cos(m+n)\varphi + \cos(m-n)\varphi}{2} d\varphi. \end{aligned}$$

Ясно также, что $\int\limits_0^\pi \cos k\varphi \, d\varphi = 0$ при $k \neq 0$. □

Скалярное произведение можно задавать не только собственным интегралом, но и несобственным. Например, можно рассмотреть скалярное произведение $(f,g) = \int_{-\infty}^{\infty} e^{-x^2} f(x)g(x)\,dx$.

Многочлены $H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}$ называют *многочленами Эрмита*. Прежде всего заметим, что $H_n(x)$ — многочлен степени n. Действительно, $\left(f(x)e^{-x^2}\right)' = f'(x)e^{-x^2} - 2xf(x)e^{-x^2}$, поэтому $H_{n+1}(x) = -H_n'(x) + 2xH_n(x)$. Ясно также, что $H_0(x) = 1$.

При доказательстве теоремы 6.16 нам понадобится интеграл Эйлера–Пуассона $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$, вычисленный в примере 6.4.

Теорема 6.16. $\int_{-\infty}^{\infty} e^{-x^2} H_n(x)H_m(x)\,dx = 2^n n! \sqrt{\pi}\delta_{mn}$.

Доказательство. Можно считать, что $m \leq n$. По определению

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x)H_m(x)\,dx = (-1)^n \int_{-\infty}^{\infty} H_m(x)\varphi^{(n)}(x)\,dx,$$

где $\varphi(x) = e^{-x^2}$. Интегрируя по частям, получаем

$$\begin{aligned} &(-1)^n \int_{-\infty}^{\infty} H_m(x)\varphi^{(n)}(x)\,dx = \\ &\qquad = (-1)^n \, H_m(x)\varphi^{(n-1)}(x)\Big|_{-\infty}^{\infty} + (-1)^{n+1}\int_{-\infty}^{\infty} H_m'(x)\varphi^{(n-1)}(x)\,dx. \end{aligned}$$

Ясно, что $H_m(x)\varphi^{(n-1)}(x)$ представляет собой произведение многочлена на e^{-x^2}, поэтому при $x = \pm\infty$ это выражение обращается в нуль. Повторяя интегрирование по частям, приходим к $(-1)^{n+m}\int_{-\infty}^{\infty} H_m^{(m)}(x)\varphi^{(n-m)}(x)\,dx$. Если $m < n$, то можно ещё раз проинтегрировать по частям, и мы получим 0, поскольку $H_m^{(m+1)}(x) = 0$. Если $m = n$, то мы получаем $n!2^n\int_{-\infty}^{\infty} e^{-x^2}dx = n!2^n\sqrt{\pi}$, поскольку $H_m^{(m)}(x) = (2^n x^n + \dots)^{(n)} = n!2^n$. $\square$

Теорема 6.17. *Любые три последовательных многочлена Эрмита связаны соотношением*

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \quad n \geq 1.$$

Доказательство. Пусть $\varphi(x) = e^{-x^2}$. Тогда $\varphi'(x) = -2xe^{-x^2}$, поэтому $\varphi'(x) + 2x\varphi(x) = 0$. Дифференцируя это равенство, получаем $\varphi''(x) + 2x\varphi'(x) + 2\varphi(x) = 0$, затем $\varphi'''(x) + 2x\varphi''(x) + 4\varphi'(x) = 0$ и т.д. Индукцией по n легко доказать, что соответствующее равенство имеет вид $\varphi^{(n+1)}(x) + 2x\varphi^{(n)}(x) + 2n\varphi^{(n-1)}(x) = 0$, что эквивалентно требуемому равенству. $\square$

Следствие. *Для любого натурального n выполняется равенство $H_n'(x) = 2nH_{n-1}(x)$.*

Доказательство. Достаточно сравнить полученное вначале рекуррентное соотношение $H_{n+1}(x) = -H_n'(x) + 2xH_n(x)$ с соотношением из теоремы. $\square$

6.18. Среднее значение длины проекции

Для функции f, интегрируемой на отрезке $[a, b]$, можно определить её *среднее значение* $M = \frac{1}{b-a}\int_a^b f(x)dx$. Для постоянной функции её среднее значение совпадает с её значением во всех точках, поэтому функция не может быть всюду больше среднего значения и не может быть всюду меньше среднего значения.

Рассмотрим на плоскости отрезок длины a. Пусть прямая, на которой он лежит, образует угол α с осью Ox. Тогда длина проекции этого отрезка на прямую l, образующую угол φ с осью Ox, равна $a|\cos(\alpha - \varphi)|$. Длина проекции отрезка является функцией от φ, поэтому можно рассмотреть среднее значение этой функции на отрезке $[0, \pi]$. Это будет среднее значение по всем возможным направлениям прямой l. Таким образом, на плоскости среднее значение проекции отрезка длины a на прямую равно $\frac{a}{\pi}\int_0^{\pi}|\cos(\alpha - \varphi)|d\varphi = \frac{a}{\pi}\int_0^{\pi}|\cos\psi|d\psi = \frac{2a}{\pi}$. Как и следовало ожидать, среднее значение длины отрезка пропорционально длине и не зависит от его положения на плоскости (т.е. не зависит от угла α).

Отношение среднего значения длины проекции к длине отрезка можно найти также из следующих геометрических соображений. Рассмотрим правильный многоугольник с большим числом сторон, вписанный в окружность радиуса R. Сумма длин его сторон приблизительно равна $2\pi R$, а сумма длин их проекций приблизительно равна удвоенному диаметру, т.е. $4R$. Поэтому отношение среднего значения длины проекции к длине отрезка равно $\frac{4R}{2\pi R} = \frac{2}{\pi}$.

Одно из приложений среднего значения длины проекции связано с так называемыми кривыми постоянной ширины. Ширина кривой — это длина её проекции на прямую l. Замкнутую кривую называют *кривой постоянной ширины*, если длина проекции не зависит от выбора прямой l. Одним из примеров кривой постоянной ширины является окружность. Но есть и много других кривых постоянной ширины. Например, кривую постоянной ширины несложно построить, проведя дуги окружностей с вершинами в вершинах равностороннего треугольника, радиус которых равен его стороне. Тем не менее, для любой выпуклой кривой отношение её периметра к ширине такое же, как для окружности. (Замкнутую кривую называют *выпуклой*, если она ограничивает выпуклую фигуру.)

Теорема 6.18. *Если длина проекции замкнутой выпуклой кривой на любую прямую равна 1, то её длина равна π.*

Доказательство. Длина кривой — это предел периметров вписанных в неё многоугольников. Рассмотрим вписанный многоугольник с периметром P и длиной проекции на прямую l, равной $d(l)$. Пусть $1 - \varepsilon < d(l) < 1$ для всех прямых l. Многоугольник можно подобрать так, чтобы ε было сколь угодно мало. Так как многоугольник выпуклый, сумма длин проекций его сторон на прямую l равна $2d(l)$. Среднее значение величины $2d(l)$ равно $\frac{2P}{\pi}$, поэтому $2 - 2\varepsilon < \frac{2P}{\pi} < 2$, т.е. $\pi - \pi\varepsilon < P < \pi$. Устремляя ε к нулю, получаем, что длина кривой равна π. $\square$

Среднее значение длины проекции связано также с *задачей об игле Бюффона*. Предположим, что на плоскость, разделённую на полосы ширины d, бросают иглу длины $l < d$. Какова вероятность того, что игла пересекает границу какой-то полосы?

Удобно убрать ограничение $l < d$ и сформулировать задачу по-другому: чему равно среднее число $E(l)$ точек, в которых игла пересекает границы полос? При $l < d$ это будет прежняя задача. Но при такой постановке задачи ясно, что $E(x+y) = E(x) + E(y)$ при $x > 0$ и $y > 0$. Ясно также, что $E(x) \geq 0$ при $x \geq 0$, поэтому функция E возрастающая. При решении уравнения Коши мы показали, что $E(r) = cr$ для положительных рациональных r. Поэтому из монотонности функции E следует, что $E(x) = cx$ для всех положительных x.

Рассмотрим прямую a, перпендикулярную границам полос. Границы полос пересекают прямую a в точках, делящих её на отрезки длины d; отметим эти точки. Задача об игле Бюффона эквивалентна следующей задаче. На прямую a бросают отрезок, длина которого равна длине проекции иглы на прямую a. Какова вероятность того, что этот отрезок содержит отмеченную точку? Пусть среднее число отмеченных точек на отрезке длины x равно $E_1(x)$. Тогда $E_1(x) = cx$ и $E_1(d) = 1$, поэтому $E_1(x) = \frac{x}{d}$. Нас интересует среднее число отмеченных точек на отрезке, длина x которого равна средней длине проекции отрезка длины l, т.е. $x = \frac{2l}{\pi}$. Поэтому вероятность того, что игла пересекает границу какой-то полосы, равна $\frac{2l}{\pi d}$.

Покажем теперь, как можно найти коэффициент c и искомую вероятность без интегрирования. Вместо иглы длины l можно бросать многоугольник, периметр которого равен l. Формула $E(x) = cx$ показывает, что среднее число точек пересечения при этом будет то же самое. Но если бросать окружность диаметра d, то число точек пересечения всегда равно 2. Поэтому $2 = E(\pi d) = c\pi d$, т.е. $c = \frac{2}{\pi d}$. Таким образом, искомая вероятность равна $\frac{2l}{\pi d}$.

Историческое замечание. Жорж-Луи Леклерк де Бюффон (1707-1788) предложил задачу о бросании иглы в 1777 году. Решение задачи Бюффона без интегрирования предложил Жозеф-Эмиль Барбье (1839-1889) в 1860 году. В той же статье он доказал теорему 6.18 о кривых постоянной ширины.

Задача 6.53. Сумма длин нескольких векторов на плоскости равна L. Докажите, что из этих векторов можно выбрать некоторое число векторов (может быть, только один) так, что длина их суммы будет не меньше L/π.

Задача 6.54. Докажите, что если длины всех сторон и диагоналей выпуклого многоугольника меньше d, то его периметр меньше πd.

6.19. Преобразование Лежандра

Рассмотрим функцию $f(x)$, область определения D которой — отрезок или интервал, возможно, бесконечный с одной или с обеих сторон. Функцию $f^*(p) = \sup_{x \in D}[xp - f(x)]$, определённую там, где этот супремум конечен, называют *преобразованием Лежандра* функции f.

Непосредственно из определения следует, что $xp \leq f(x) + f^*(p)$ для всех x и p.

Пример 6.5. *Пусть $f(x) = x^a$, где $a > 1$, и $D = (0, \infty)$. Тогда $f^*(p) = \left(\frac{p}{a}\right)^{\frac{a}{a-1}} \cdot (a-1)$.*

Доказательство. Фиксируем p и продифференцируем по x выражение $xp - f(x) = xp - x^a$. В результате получим $p - ax^{a-1}$. Эта величина обращается в нуль при $x_0 = \left(\frac{p}{a}\right)^{\frac{1}{a-1}}$. Поэтому при изменении x от 0 до ∞ величина $xp - f(x)$ сначала возрастает от 0 до $x_0p - f(x_0) = \left(\frac{p}{a}\right)^{\frac{a}{a-1}} \cdot (a-1)$, а затем убывает до $-\infty$. $\square$

Ясно, что числа a и $b = \frac{a}{a-1}$ связаны соотношением $\frac{1}{a} + \frac{1}{b} = 1$. Несложно проверить, что если $f(x) = \frac{\lambda}{a}x^a$, то $f^*(p) = \frac{\mu}{b}x^b$, где $\mu^a\lambda^b = 1$.

Пример 6.6. *Пусть $f(x) = e^x$ и $D = \mathbb{R}$. Тогда $f^*(p) = p(\ln p - 1)$.*

Доказательство. Фиксируем p и продифференцируем по x выражение $xp - f(x) = xp - e^x$. В результате получим $p - e^x$. Эта величина обращается в нуль при $x_0 = \ln p$. Поэтому при изменении x от $-\infty$ до ∞ величина $xp - f(x)$ сначала возрастает от $-\infty$ до $x_0p - f(x_0) = p\ln p - p$, а затем убывает до $-\infty$. $\square$

Свойства преобразования Лежандра наиболее просты в том случае, когда $D = (0, \infty)$, $f''(x) > 0$ для всех $x \in D$ и функция $f'(x)$ строго монотонно возрастает от 0 до ∞ при изменении x от 0 до ∞. В таком случае при фиксированном $p > 0$ функция $xp - f(x)$ имеет единственный максимум в точке x_0, в которой $f'(x_0) = p$. Действительно, производная этой функции равна $p - f'(x)$. Поэтому производная положительна при $0 < x < x_0$ и отрицательна при $x > x_0$.

Положим $g(x) = f'(x)$. Тогда $f^*(p) = xp - f(x)$, где $g(x) = p$, т.е. $x = g^{-1}(p)$. Таким образом, $f^*(p) = xp - f(x)$, где $x = g^{-1}(p)$. Поэтому

$$\frac{df^*(p)}{dp} = \frac{d}{dp}(xp - f(x)) = x + p\frac{dx}{dp} - \frac{df}{dx} \cdot \frac{dx}{dp} = x,$$

поскольку $p = \frac{df}{dx}$. В результате получаем, что функция f и её преобразование Лежандра f^* связаны следующими соотношениями: $p = f'(x)$ и $x = (f^*)'(p)$.

Равенство $x = (f^*)'(p)$ означает, что $(f^*)'(p) = x = g^{-1}(p)$. Таким образом, $f(x) = \int_0^x g(t)dt$ и $f^*(p) = \int_0^p g^{-1}(t)dt$. Функция g^{-1} обладает такими же свойствами, как и функция $g = f'$: её производная положительна и функция g^{-1} строго монотонно возрастает от 0 до ∞ при изменении x от 0 до ∞. Поэтому $f^{**} = f$.

Историческое замечание. Адриен-Мари Лежандр (1752-1833) ввёл преобразование Лежандра в 1787 году. Но уже ранее, в 1770 году, это преобразование использовал Леонард Эйлер (1707-1783).

6.20. Среднее арифметико-геометрическое

Пусть $a_0 > b_0 > 0$, $a_{n+1} = \frac{a_n + b_n}{2}$ и $b_{n+1} = \sqrt{a_n b_n}$. Тогда обе последовательности a_n и b_n сходятся к одному и тому же числу $\mu(a, b)$, которое называется *средним арифметико-геометрическим* чисел $a = a_0$ и $b = b_0$ (см. задачу 2.24).

Гаусс показал, что среднее арифметико-геометрическое тесно связано с интегралом

$$G = \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}}.$$

Эта связь основана на замене переменной

$$\sin \varphi = \frac{2a \sin \theta}{(a + b) + (a - b) \sin^2 \theta}.$$

Если θ изменяется от 0 до $\frac{\pi}{2}$, то φ изменяется в тех же пределах. Кроме того, производная функции $\varphi(\theta)$ положительна, поскольку

$$\cos \varphi \, d\varphi = 2a \frac{(a + b) - (a - b) \sin^2 \theta}{((a + b) + (a - b) \sin^2 \theta)^2} \cos \theta \, d\theta.$$

Легко также проверить, что

$$\cos \varphi = \sqrt{1 - \sin^2 \varphi} = \frac{\sqrt{(a + b)^2 - (a - b)^2 \sin^2 \theta}}{(a + b) + (a - b) \sin^2 \theta} \cos \theta$$

и

$$\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi} = a \frac{(a + b) - (a - b) \sin^2 \theta}{((a + b) + (a - b) \sin^2 \theta)^2},$$

поэтому

$$\frac{d\varphi}{\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}} = \frac{d\theta}{\sqrt{\left(\frac{a+b}{2}\right)^2 \cos^2 \theta + ab \sin^2 \theta}}.$$

Положим $a_1 = \frac{a+b}{2}$ и $b_1 = \sqrt{ab}$. Тогда

$$G = \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{a^2 \cos^2 \varphi + b^2 \sin^2 \varphi}} = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{a_1^2 \cos^2 \theta + b_1^2 \sin^2 \theta}}$$

(*формула Гаусса*). Применяя формулу Гаусса повторно, получаем, что

$$G = \int_0^{\frac{\pi}{2}} \frac{d\theta}{\sqrt{a_n^2 \cos^2\theta + b_n^2 \sin^2\theta}}$$

для всех n. Следовательно, $G = \frac{\pi}{2\mu(a,b)}$ и $\mu(a,b) = \frac{\pi}{2G}$.

6.21. Прямая как дифференцируемое многообразие

Напомним формулу замены переменной для определённого интеграла. Если функция φ имеет непрерывную производную на отрезке $[\alpha,\beta]$, $\varphi(\alpha) = a$, $\varphi(\beta) = b$ и функция f непрерывна на области значений функции φ, то

$$\int_a^b f(y)dy = \int_\alpha^\beta f(\varphi(x))\varphi'(x)dx.$$

Эта формула побуждает задуматься: а что же именно мы интегрируем? При замене переменной функция $f(x)$ заменяется на функцию $f(\varphi(t))$, без множителя $\varphi'(t)$. Поэтому мы интегрируем не функцию, а какой-то другой объект.

Пока нас интересует только интеграл по прямой или отрезку, этот вопрос преждевременный. На прямой есть естественная параметризация, и она позволяет определить интеграл от функции. Но этот вопрос очень важен для кривых, поверхностей и их обобщений — многообразий. На них естественной параметризации нет, поэтому необходимо, чтобы интеграл не зависел от параметризации. А уже для прямой видно, что при изменении параметризации изменяется функция, которую нужно интегрировать. Не зависящий от параметризации объект, который интегрируется, — это дифференциальная форма.

Для прямой введение дифференциальных форм во многих отношениях — чрезмерное усложнение. Но в случае многообразий без дифференциальных форм уже не обойтись. Поэтому есть смысл обсудить дифференциальные формы уже в простейшей ситуации — на прямой.

Рассмотрим замену переменной $y = \varphi(x)$. Это означает, что задано отображение $\varphi\colon M \to N$, где M и N — подмножества в $\mathbb{R}$. На множестве M есть координата x, а на множестве N — координата y. При этом функции переносятся в обратном направлении, с N на M: функции $f(y)$ на N соответствует функция $(\varphi^* f)(x) = f(\varphi(x))$ на M. Звёздочка вверху здесь символизирует то, что объекты переносятся в направлении, противоположном направлению отображения φ.

Введём теперь объект, который переносится в направлении отображения φ. Этот объект — *касательный вектор* в точке $a \in M$. Рассмотрим параметризованную кривую $\gamma\colon (-\varepsilon,\varepsilon) \to M$, $\gamma(0) = a$. Для функции $f(x)$ определим в точке a производную в направлении этой кривой: $\partial_\gamma f(a) = \left.\frac{df(\gamma(t))}{dt}\right|_{t=0}$. Эта производная равна $\frac{df}{dx}(a) \cdot \frac{d\gamma}{dt}(0)$. Касательный вектор v в точке a — это класс кривых, для которых число $\frac{d\gamma}{dt}(0)$ одно и то же. Производная в направлении всех этих кривых одна и та же; будем считать её производной в направлении данного вектора. Производную функции f в направлении вектора v будем обозначать $v(f)$. Число $\frac{d\gamma}{dt}(0)$ — это *координата* вектора v в системе координат x.

Например, кривой $\gamma(t) = a+t$ соответствует вектор v, координата которого равна 1. Этот вектор действует на функции следующим образом: $v(f) = \frac{df}{dx}(a)$.

Сделаем теперь замену координат $y = \varphi(x)$. Кривые при этом переносятся в направлении отображения φ: кривой $\gamma = \gamma(t)$, для которой $\gamma(0) = a \in M$ соответствует кривая $\varphi_*\gamma = \varphi(\gamma(t))$, для которой $\varphi_*\gamma(0) = \varphi(a) \in N$. Соответствующие этим кривым векторы тоже переносятся в направлении отображения φ: вектор v в системе координат x переходит в вектор $\varphi_* v$ в системе координат $y = \varphi(x)$.

Функции $f(y)$ соответствует функция $(\varphi^* f)(x) = f(\varphi(x))$. Производная функции $\varphi^* f$ в направлении кривой $\varphi_*\gamma$ равна

$$\left.\frac{df(\varphi(\gamma(t)))}{dt}\right|_{t=0} = \frac{df}{d\varphi}(\varphi(a)) \cdot \frac{d\varphi}{dx}(a) \cdot \frac{d\gamma}{dt}(0) = \frac{df}{dx}(a) \cdot \frac{d\gamma}{dt}(0).$$

Таким образом, вектор $\varphi_* v$ действует на функцию $\varphi^* f$ точно так же, как вектор v действует на функцию f, т.е. $\varphi_* v(\varphi^* f) = v(f)$. Но координата у вектора $\varphi_* v$ другая: она равна $\frac{d\varphi}{dx}(a) \cdot \frac{d\gamma}{dt}(0)$, а не $\frac{d\gamma}{dt}(0)$.

Дифференциальная 1-форма ω на прямой (или на отрезке) в каждой точке a представляет собой линейную функцию от вектора v, зависящую от a: $\omega_a(v) = \beta(a)v$. Замена координат φ действует на дифференциальную 1-форму следующим образом: $(\varphi^*\omega)(v) = \omega(\varphi_* v)$.

Примером дифференциальной 1-формы является форма dx, принимающая в точке a значение 1 на векторе v, заданном кривой $\gamma(t) = t + a$. Напомним, что координата этого вектора равна 1.

Рассмотрим замену координат $y = \varphi(x)$ и вычислим 1-форму $dy = \varphi^* dx$. По определению для вектора v, заданного кривой $\gamma(t) = t + a$, получаем:

$$(\varphi^* dx)(v) = dx(\varphi_* v) = \frac{d\varphi}{dx}(a) = \frac{d\varphi}{dx}(a) dx(v).$$

Таким образом, $dy = \frac{d\varphi}{dx}(a)dx$ и $f(y)dy = f(\varphi(x))\varphi'(x)dx$. Это в точности согласуется с формулой замены переменной, и для 1-формы $\omega = f(x)dx$ мы можем положить

$$\int_{[a,b]} \omega = \int_a^b f(x)dx.$$

Формула замены переменной показывает, что выражение в правой части не зависит от параметризации отрезка $[a, b]$.

Отметим, что интеграл от дифференциальной 1-формы ω по отрезку $[a, b]$ равен пределу интегральных сумм $\sum_{i=1}^n \omega_{x_i}(b_i - a_i)$, где $[a_i, b_i]$ — i-й отрезок разбиения, x_i — точка этого отрезка ($i = 1, \ldots, n$). Интеграл зависит от ориентации, при изменении ориентации отрезка интеграл меняет знак.

Дифференцируемой функции f можно сопоставить её *дифференциал* — дифференциальную 1-форму df, которая в точке x равна $\frac{df}{dx}(x)dx$. При этом $\int_{[a,b]} df = f(a) - f(b)$.

6.22. Решения задач

6.1. Ясно, что $\left(\ln \frac{x-a}{x+a}\right)' = \left(\ln(x-a) - \ln(x+a)\right)' = \frac{1}{x-a} - \frac{1}{x+a} = \frac{2a}{x^2-a^2}$.

6.2. Согласно теореме 6.1 и задаче 5.8

$$\int \frac{dx}{x(1+\ln^2 x)} = \int \frac{d\ln x}{1+\ln^2 x} = \int \frac{dy}{1+y^2} = \operatorname{arctg} y + C = \operatorname{arctg} \ln x + C;$$
$$\int \frac{e^x dx}{1+e^{2x}} = \int \frac{de^x}{1+e^{2x}} = \int \frac{dy}{1+y^2} = \operatorname{arctg} y + C = \operatorname{arctg} e^x + C;$$
$$\int \frac{\cos x\, dx}{1+\sin^2 x} = \int \frac{d\sin x}{1+\sin^2 x} = \int \frac{dy}{1+y^2} = \operatorname{arctg} y + C = \operatorname{arctg} \sin x + C.$$

6.3. Согласно теореме 6.1

$$\int \operatorname{tg} x\, dx = \int -\frac{d\cos x}{\cos x} = -\ln\cos x + C.$$

6.4. Согласно теореме 6.1 $\int f(\varphi(y))\varphi'(y)\, dy = \int f(\varphi(y))\, d\varphi(y) = \int f(x)\, dx$.

6.5. Положим $x = \sin y$. Тогда, согласно задаче 6.4

$$\int \sqrt{1-x^2}dx = \int \cos^2 y\, dy = \int \frac{1+\cos 2y}{2}dy = \frac{y}{2} + \frac{1}{4}\int \cos 2y\, d(2y) =$$
$$= \frac{y}{2} + \frac{\sin 2y}{4} + C = \frac{\arcsin x}{2} + \frac{x\sqrt{1-x^2}}{2} + C.$$

Положим $y = \frac{x}{a}$. Тогда $\int \sqrt{1-y^2}dy = \frac{1}{a^2}\int \sqrt{a^2-x^2}dx$, поэтому

$$\int \sqrt{a^2-x^2}dx = \frac{a^2}{2}\arcsin\frac{x}{a} + \frac{x}{2}\sqrt{a^2-x^2} + C.$$

6.6. Положим $x = \operatorname{sh} y$. Тогда, согласно теореме 6.4

$$\int \sqrt{1+x^2}dx = \int \operatorname{ch}^2 y\,dy = \int \frac{1+\operatorname{ch} 2y}{2}dy = \frac{y}{2} + \frac{1}{4}\int \operatorname{ch} 2y\,d(2y) = \\ = \frac{y}{2} + \frac{\operatorname{sh} 2y}{4} + C = \frac{\operatorname{Arsh} x}{2} + \frac{x\sqrt{1+x^2}}{2} + C = \\ = \frac{1}{2}\ln(x+\sqrt{1+x^2}) + \frac{x\sqrt{1+x^2}}{2} + C.$$

Аналогично, положив $x = \operatorname{ch} y$, получаем

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

Положив $y = \frac{x}{\sqrt{A}}$ при $A > 0$ и $y = \frac{x}{\sqrt{-A}}$ при $A < 0$, получим

$$\int \sqrt{x^2+A} = \frac{A}{2}\ln(x+\sqrt{x^2+A}) + \frac{x\sqrt{x^2+A}}{2} + C.$$

6.7. Положим $x = y^2$. Тогда, согласно теореме 6.4

$$\int \frac{dx}{1+\sqrt{x}} = \int \frac{2y\,dy}{1+y} = 2\int \left(1 - \frac{1}{1+y}\right)dy = \\ = 2y - 2\ln(1+y) + C = 2\sqrt{x} - 2\ln(1+\sqrt{x}) + C.$$

6.8. а) Положим $f(x) = \ln x$ и $g(x) = x$. По формуле интегрирования по частям получаем $\int \ln x\,dx = x\ln x - \int \frac{1}{x}x\,dx = x\ln x - \int dx = x\ln x - x + C$.

б) $\int x^3 \ln x\,dx = \int \ln x d\left(\frac{x^4}{4}\right) = \ln x \cdot \frac{x^4}{4} - \int \frac{x^4}{4}d(\ln x) = \ln x \cdot \frac{x^4}{4} - \int \frac{x^4}{4}\cdot\frac{1}{x}dx = \frac{x^4}{4}\ln x - \frac{x^4}{16} + C$.

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

г) $\int x\cos x\,dx = \int x\,d(\sin x) = x\sin x - \int \sin x\,dx = x\sin x + \cos x + C$.

д) $\int xe^x dx = \int x\,d(e^x) = xe^x - \int e^x dx = xe^x - e^x + C$.

6.9. а) $\int xe^x dx = \int x\,d(e^x) = xe^x - \int e^x dx = xe^x - e^x + C$.

б) Пусть $u_m(x) = \int x^m e^{-x}dx$ и $\alpha_m(x) = -x^m e^{-x}$. Интегрируя по частям, получаем $\int x^m e^{-x} dx = -\int x^m d(e^{-x}) = -x^m e^{-x} + m\int x^{m-1}e^x dx$, т.е. $u_m = \alpha_m + mu_{m-1}$. Сложим равенства $u_1 = \alpha_1 + u_0$, $u_2 = \alpha_2 + 2u_1, \ldots, u_m = \alpha_m + mu_{m-1}$, предварительно умножив их на $1, 1/2!, \ldots, 1/m!$. В результате получим

$$\frac{u_m}{m!} = u_0 + \alpha_1 + \frac{\alpha_2}{2!} + \cdots + \frac{\alpha_m}{m!}.$$

Здесь $u_0 = -e^{-x} + C$, поэтому

$$\int x^m e^{-x} dx = -m!e^{-x}\left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^m}{m!}\right) + C.$$

6.10. Применим формулу интегрирования по частям для $f(x) = g(x) = \ln x$. В результате получим

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

Первообразная определена с точностью до константы, поэтому это равенство нужно понимать так: $I = (\ln x)^2 - (I + C)$, где $I = \int \frac{\ln x}{x}dx$. Таким образом, $\int \frac{\ln x}{x}dx = \frac{(\ln x)^2}{2} + C$ (константа C здесь другая).

6.11. а) Интегрируя по частям, получаем

$$\int e^{ax}\sin bx\,dx = -\frac{1}{b}e^{ax}\cos bx + \frac{a}{b}\int e^{ax}\cos bx\,dx = \\ = -\frac{1}{b}e^{ax}\cos bx + \frac{a}{b^2}e^{ax}\sin bx - \frac{a^2}{b^2}\int e^{ax}\sin bx\,dx.$$

Для $I = \int e^{ax} \sin bx\, dx$ мы получили соотношение

$$I = \frac{e^{ax}}{b^2}(-b\cos bx + a\sin bx) - \frac{a^2}{b^2}(I + C).$$

Следовательно,

$$\int e^{ax} \sin bx\, dx = \frac{e^{ax}}{a^2+b^2}(a\sin bx - b\cos bx) + C.$$

б) Аналогично получаем

$$\int e^{ax} \cos bx\, dx = \frac{e^{ax}}{a^2+b^2}(a\cos bx + b\sin bx) + C.$$

6.12. Воспользуемся формулой из теоремы 6.2, заменив в ней g на $g^{(n)}$:

$$\int fg^{(n+1)}dx = fg^{(n)} - \int f'g^{(n)}dx.$$

Аналогично получаем

$$\int f'g^{(n)}dx = f'g^{(n-1)} - \int f''g^{(n-1)}dx,$$
$$\int f''g^{(n-1)}dx = f''g^{(n-2)} - \int f'''g^{(n-2)}dx,$$
$$\dots\dots$$
$$\int f^{(n)}g'dx = f^{(n)}g - \int f^{(n+1)}g\, dx.$$

Из этих формул легко следует требуемое.

6.13. а) Воспользуемся формулой из задачи 6.12. Если $g^{(n+1)} = e^{ax}$, то $g^{(n)} = \frac{e^{ax}}{a}$, $g^{(n-1)} = \frac{e^{ax}}{a^2}$, ... Поэтому

$$\int P(x)e^{ax}dx = e^{ax}\left(\frac{P}{a} - \frac{P'}{a^2} + \frac{P''}{a^3} - \dots\right) + C.$$

б) Если $g^{(n+1)} = \sin ax$, то $g^{(n)} = -\frac{\cos ax}{a}$, $g^{(n-1)} = -\frac{\sin ax}{a^2}$, $g^{(n-2)} = \frac{\cos ax}{a^3}$, ... Поэтому

$$\int P(x)\sin ax\, dx = \sin ax\left(\frac{P'}{a^2} - \frac{P'''}{a^4} + \dots\right) - \cos ax\left(\frac{P}{a} - \frac{P''}{a^3} + \dots\right) + C.$$

в) Аналогично решению задачи б) получаем

$$\int P(x)\cos ax\, dx = \sin ax\left(\frac{P}{a} - \frac{P''}{a^3} + \dots\right) + \cos ax\left(\frac{P'}{a^2} - \frac{P'''}{a^4} + \dots\right) + C.$$

6.14. Если $f(x) \geq 0$ для всех точек отрезка $[a, b]$, то все интегральные суммы неотрицательны, поэтому их предел тоже неотрицателен.

Предположим теперь, что $f(x_0) = c > 0$. Можно считать, что x_0 — внутренняя точка отрезка $[a, b]$. Выберем $\delta > 0$ так, что $f(x) > c/2$ при $|x - x_0| < \delta$. Можно считать, что $a_0 + \delta \leq x_0 \leq b - \delta$. Тогда

$$\int_a^b f(x)\, dx = \int_a^{x_0-\delta} f(x)\, dx + \int_{x_0-\delta}^{x_0+\delta} f(x)\, dx + \int_{x_0+\delta}^{b} f(x)\, dx.$$

Два крайних интеграла неотрицательны, а средний интеграл больше $2\delta \cdot \frac{c}{2} = c\delta > 0$.

6.15. Пусть m и M — наименьшее и наибольшее значения функции f на отрезке $[a, b]$. Тогда $m\varphi(x) \leq f(x)\varphi(x) \leq M\varphi(x)$ для всех $x \in [a, b]$. Поэтому

$$m\int_a^b \varphi(x)dx \leq \int_a^b f(x)\varphi(x)dx \leq M\int_a^b \varphi(x)dx.$$

Следовательно, $\int_a^b f(x)\varphi(x)dx = P\int_a^b \varphi(x)dx$ для некоторого числа $P \in [m, M]$. На отрезке $[a, b]$ непрерывная функция f принимает все промежуточные значения, поэтому $f(c) = P$ для некоторого $c \in [a, b]$.

6.16. Предположим, что непрерывная функция $f(x)$ определена на отрезке $[0, a+b]$. Тогда определены оба интеграла $\int_0^a (f(x+b) - f(x))dx$ и $\int_0^b (f(x+a) - f(x))dx$. Покажем, что эти интегралы равны. Действительно, пусть $F(x)$ — первообразная функции $f(x)$. Тогда

$$\begin{aligned}\int_0^a (f(x+b) - f(x))dx &= \int_b^{a+b} f(x)dx - \int_0^a f(x)dx = \\ &= F(a+b) - F(b) - F(a) + F(0).\end{aligned}$$

Аналогично

$$\int_0^b (f(x+a) - f(x))dx = F(a+b) - F(b) - F(a) + F(0).$$

Из условия задачи следует, что $\int_0^a (f(x+b) - f(x))dx > 0$ и $\int_0^b (f(x+a) - f(x))dx < 0$, а это противоречит равенству интегралов.

6.17. Рассмотрим функцию g, которая равна 1 на отрезках $[0, n-1]$ и $[n, 2n-1]$, а на интервале $(n-1, n)$ равна $-n+\frac{1}{2}$. Положим $f(x) = \int_0^x g(t)dt$. Тогда $f(x+n) - f(x)$ — это интеграл от функции g по отрезку длины n, а $f(x+n+1) - f(x)$ — это интеграл по отрезку длины $n+1$. На всей области определения на любом отрезке длины n функция g равна $-n+\frac{1}{2}$ на интервале длины 1, а на остальной части отрезка, имеющей длину $n-1$, функция равна 1. Поэтому интеграл равен $n-1-n+\frac{1}{2} = -\frac{1}{2} < 0$. На любом отрезке длины $n+1$ функция g равна $-n+\frac{1}{2}$ на интервале длины 1, а на остальной части отрезка, имеющей длину n, функция равна 1. Поэтому интеграл равен $n-n+\frac{1}{2} = \frac{1}{2} > 0$.

6.18. Отметим середины $m+n-2$ отрезков, на которые точки $1, 2, \ldots, m+n-3$ делят отрезок $[0, m+n-2]$. Это будут вершины графа. Рёбра графа соединяют вершины, расстояние между которыми равно m или n. Из первых $n-2$ вершин справа и слева выходит по два ребра (одно длины n, другое длины m). Из средних $m-n-2$ вершин тоже выходит по два ребра (оба длины n). Из четырёх остальных вершин выходит по одному ребру длины n. Если из четырёх вершин графа выходит по одному ребру, а из остальных вершин выходит по два ребра, то такой граф состоит из двух компонент. Вершинам из одной компоненты сопоставим некоторое число u, а вершинам из другой компоненты — некоторое число v. Рассмотрим функцию g на отрезке $[0, m+n-2]$, которая на интервале вида $(k, k+1)$, где число k целое, принимает значение, сопоставленное середине интервала. Значения в целых точках можно взять произвольные. Положим $f(x) = \int_0^x g(t)dt$. Тогда на всей области определения $f(x+n) - f(x) = au + bv$, где a и b — постоянные натуральные числа, причём $a+b = n$, и $f(x+m) - f(x) = cu + dv$, где $c+d = m$. Остаётся подобрать числа u и v так, что $au + bv < 0$ и $cu + dv > 0$. Это можно сделать, если $\frac{a}{b} \neq \frac{c}{d}$. Предположим, что $\frac{a}{b} = \frac{c}{d}$. Тогда $\frac{a+b}{b} = \frac{c+d}{d}$, т.е. $\frac{n}{b} = \frac{m}{d}$. Но это противоречит взаимной простоте чисел m и n.

6.19. а) Ясно, что $U_0 = \frac{\pi}{2}$ и $U_1 = \sin x|_0^{\pi/2} = 1$. Пусть теперь $n > 1$. Тогда, интегрируя по частям, получаем

$$U_n = \int_0^{\pi/2} \cos^{n-1} x \, d\sin x = \sin x \cos^{n-1} x|_0^{\pi/2} - \int_0^{\pi/2} \sin x \, d\cos^{n-1} x.$$

Первый член равен нулю, поэтому

$$U_n = (n-1)\int_0^{\pi/2} \cos^{n-2} x \sin^2 x \, dx = (n-1)\int_0^{\pi/2} \cos^{n-2} x(1-\cos^2 x)\, dx,$$

т.е. $U_n = (n-1)(U_{n-2} - U_n)$. Таким образом, $U_n = \frac{n-1}{n} U_{n-2}$. Значит, $U_{n-2} = \frac{n-3}{n-2} U_{n-4}$ и т.д. Последовательно применяя эту формулу, получаем требуемое.

б) Если $0 \leq x \leq \pi/2$, то

$$\cos^{2n+2} x \leq \cos^{2n+1} x \leq \cos^{2n} x,$$

поэтому $U_{2n+2} < U_{2n+1} < U_{2n}$. Используя формулы из задачи а), получаем

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

Следовательно,

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

Учитывая, что $\lim_{n\to\infty}\frac{2n+1}{2n+2}=1$ и $\lim_{n\to\infty}\frac{n}{2n+1}=\frac{1}{2}$, получаем требуемое.

6.20. Первое решение. Сначала докажем тождество

$$\sin\frac{\pi}{2n}\sin\frac{2\pi}{2n}\cdots\sin\frac{(n-1)\pi}{2n}=\frac{\sqrt{n}}{2^{n-1}}.$$

Корни $2n$-й степени из единицы, отличные от ±1, имеют вид $\cos\frac{k\pi}{n}\pm i\sin\frac{k\pi}{n}$, где $k=1, 2, \ldots, n-1$. Кроме того,

$$\left(x-\cos\frac{k\pi}{n}-i\sin\frac{k\pi}{n}\right)\left(x-\cos\frac{k\pi}{n}+i\sin\frac{k\pi}{n}\right)=x^2-2x\cos\frac{k\pi}{n}+1.$$

Следовательно,

$$x^{2n}-1=(x^2-1)\left(x^2-2x\cos\frac{\pi}{n}+1\right)\left(x^2-2x\cos\frac{2\pi}{n}+1\right)\cdots$$
$$\cdots\left(x^2-2x\cos\frac{(n-1)\pi}{n}+1\right).$$

Воспользуемся этим равенством при $x=1$, предварительно заметив, что

$$\frac{x^{2n}-1}{x^2-1}=x^{2n-2}+x^{2n-4}+\cdots+x^2+1.$$

В результате получим

$$2^{n-1}\left(1-\cos\frac{\pi}{n}\right)\left(1-\cos\frac{2\pi}{n}\right)\cdots\left(1-\cos\frac{(n-1)\pi}{n}\right)=n.$$

Заметим теперь, что $1-\cos\alpha=2\sin^2\frac{\alpha}{2}$. Ясно также, что рассматриваемое произведение синусов положительно.

Из доказанного тождества следует, что

$$\sum_{k=1}^{n-1}\frac{\pi}{2n}\ln\sin\frac{k\pi}{2n}=\frac{\pi}{2}\cdot\frac{-(n-1)\ln 2+\frac{1}{2}\ln n}{n}\to-\frac{\pi}{2}\ln 2,$$

поскольку $\frac{1}{n}\ln n\to 0$ (задача 2.10 б). Но указанная сумма — это интегральная сумма для разбиения отрезка $[0,\pi/2]$ на n равных отрезков.

Второе решение. Заметим, что

$$\int_0^{\pi/2}\ln(\sin x)dx=\int_0^{\pi/2}\ln(\cos x)dx=\frac{1}{2}\int_0^{\pi/2}\ln\left(\frac{\sin 2x}{2}\right)dx,$$

поскольку $\ln(\sin x)+\ln(\cos x)=\ln(\sin x\cos x)=\ln\left(\frac{\sin 2x}{2}\right)$. Далее,

$$\int_0^{\pi/2}\ln\left(\frac{\sin 2x}{2}\right)dx=\int_0^{\pi/2}\ln(\sin 2x)dx-\frac{\pi}{2}\ln 2,$$

а

$$\int_0^{\pi/2}\ln(\sin 2x)dx=\int_0^{\pi}\ln(\sin y)\frac{dy}{2}=\int_0^{\pi/2}\ln(\sin y)dy.$$

Таким образом, для $a=\int_0^{\pi/2}\ln(\sin x)dx$ получаем соотношение $a=\frac{a}{2}-\frac{\pi}{4}\ln 2$, откуда $a=-\frac{\pi}{2}\ln 2$.

6.21. Ответ: 0 при $\alpha^2 \le 1$ и $\pi \ln \alpha^2$ при $\alpha^2 > 1$.

Разделим отрезок $[0, \pi]$ на n равных частей и вычислим интегральную сумму, беря значения функции в левых концах отрезков. В результате получим

$$\frac{\pi}{n} \ln \left((1 - 2\alpha + \alpha^2)(1 - 2\alpha \cos \frac{\pi}{n} + \alpha^2) \cdot \ldots \cdot (1 - 2\alpha \cos \frac{(n-1)\pi}{n} + \alpha^2) \right).$$

Группируя сопряжённые корни $2n$-й степени из единицы, несложно проверить равенство

$$\frac{\alpha^{2n} - 1}{\alpha^2 - 1} = (1 - 2\alpha \cos \frac{\pi}{n} + \alpha^2) \cdot \ldots \cdot (1 - 2\alpha \cos \frac{(n-1)\pi}{n} + \alpha^2).$$

Поэтому интегральная сумма равна

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

При $\alpha^2 \le 1$ эта интегральная сумма стремится к 0, а при $\alpha^2 > 1$ — к $\pi \ln \alpha^2$.

6.22. Площадь круга в 2 раза больше площади полукруга, ограниченного графиком $y = \sqrt{1 - x^2}$ и прямой $y = 0$. Далее, площадь полукруга равна

$$\int_{-1}^{1} \sqrt{1 - x^2}\, dx = \left. \frac{\arcsin x}{2} \right|_{-1}^{1} + \left. \frac{x\sqrt{1 - x^2}}{2} \right|_{-1}^{1}$$

(см. задачу 6.5). Второе слагаемое обращается в нуль, а первое слагаемое равно $\frac{\pi}{2}$.

6.23. Искомая площадь равна $\int_0^\pi \sin x\, dx = -\cos x|_0^\pi = 2$.

6.24. а) Искомая площадь сегмента равна $2\int_a^{x_0} \frac{b}{a}\sqrt{x^2 - a^2} dx$. Неопределённый интеграл $\int \sqrt{x^2 + A} dx$ вычислен в решении задачи 6.6. Воспользовавшись этим вычислением, получаем, что искомая площадь равна

$$\frac{bx_0}{a}\sqrt{x_0^2 - a^2} - ab \ln \frac{x_0 + \sqrt{x_0^2 - a^2}}{a}.$$

б) Чтобы получить искомую площадь сектора, нужно из площади треугольника, ограниченного прямой $x = x_0$, осью Ox и лучом, идущим из начала координат в точку (x_0, y_0), вычесть половину площади сегмента. Поскольку $y_0 = \frac{b}{a}\sqrt{x_0^2 - a^2}$, площадь этого треугольника равна $\frac{bx_0}{2a}\sqrt{x_0^2 - a^2}$. Следовательно, искомая площадь сектора равна

$$\frac{1}{2} ab \ln \frac{x_0 + \sqrt{x_0^2 - a^2}}{a} = \frac{1}{2} ab \ln \left(\frac{x_0}{a} + \frac{y_0}{b} \right).$$

6.25. Ответ: $\frac{1}{3}\pi R^2 h$. Выберем в качестве оси Ox ось конуса и поместим начало координат в вершину конуса. Тогда $V = \int_0^h \pi \left(\frac{R}{h} x \right)^2 dx = \frac{1}{3}\pi R^2 h$.

6.26. Ответ: $\frac{4}{3}\pi R^3$. Плоскость, удалённая от центра шара на расстояние x, высекает на шаре круг радиуса $\sqrt{R^2 - x^2}$. Площадь этого круга равна $\pi(R^2 - x^2)$, поэтому

$$V = \int_{-R}^{R} \pi(R^2 - x^2) dx = 2\pi R^3 - \left. \pi \frac{x^3}{3} \right|_{-R}^{R} = 2\pi R^3 - \frac{2\pi R^3}{3} = \frac{4\pi R^3}{3}.$$

6.27. Ответ: $\frac{16}{3} R^3$. Введём систему координат, направив оси Ox и Oy по осям цилиндров. Пересечение цилиндров задаётся неравенствами $x^2 + z^2 \le R$ и $y^2 + z^2 \le R$. Его сечение плоскостью $z = z_0$ задаётся неравенствами $x_0 \le R^2 - z_0^2$ и $y_0 \le R^2 - z_0^2$. Таким образом, сечение — квадрат со стороной $2\sqrt{R^2 - z_0^2}$. Его площадь равна $4(R^2 - z_0^2)$. Поэтому объём рассматриваемой фигуры равен

$$\int_{-R}^{R} 4(R^2 - z^2) dz = 8R^3 - \left. 4\frac{z^3}{3} \right|_{-R}^{R} = 8R^3 - \frac{8}{3} R^3 = \frac{16}{3} R^3.$$

6.28. О т в е т: $\frac{2}{3}hR^2$. Рассматриваемая фигура задаётся неравенствами $x^2+y^2\le R^2$ и $0\le z\le \frac{h}{R}x$. Сечение этой фигуры плоскостью $y=y_0$ представляет собой прямоугольный треугольник с катетами $\sqrt{R^2-y_0^2}$ и $\frac{h}{R}\sqrt{R^2-y_0^2}$. Его площадь равна $\frac{1}{2}\frac{h}{R}(R^2-y_0^2)$. Поэтому объём рассматриваемой фигуры равен

$$\int_{-R}^{R}\frac{1}{2}\frac{h}{R}(R^2-y^2)dy=\frac{h}{2R}\left(2R^3-\frac{y^3}{3}\Big|_{-R}^{R}\right)=\frac{h}{2R}\cdot\frac{4R^3}{3}.$$

6.29. Длина окружности равна удвоенной длине полуокружности $y=\sqrt{R^2-x^2}$. Ясно, что $y'=\frac{-x}{\sqrt{R^2-x^2}}$, поэтому длина полуокружности равна

$$\lim_{\varepsilon\to 0}\int_{-R+\varepsilon}^{R-\varepsilon}\sqrt{1+y'^2}\,dx=\lim_{\varepsilon\to 0}\int_{-R+\varepsilon}^{R-\varepsilon}\frac{R\,dx}{\sqrt{R^2-x^2}}=\lim_{\varepsilon\to 0}R\arcsin\frac{x}{R}\Big|_{-R+\varepsilon}^{R-\varepsilon}=\pi.$$

6.30. Длина рассматриваемой дуги параболы равна

$$\int_0^{x_0}\sqrt{1+y'^2}\,dx=\int_0^{x_0}\sqrt{1+x^2}\,dx=\frac{x_0\sqrt{1+x_0^2}}{2}+\frac{1}{2}\ln(x_0+\sqrt{1+x_0^2})$$

(см. задачу 6.6).

6.31. Ясно, что $y'=\operatorname{sh}x$ и $\sqrt{1+(y')^2}=\operatorname{ch}x$. Поэтому длина рассматриваемой дуги равна $\int_{x_1}^{x_2}\operatorname{ch}x\,dx=\operatorname{sh}x_2-\operatorname{sh}x_1$ (предполагается, что $x_1<x_2$).

6.32. О т в е т: $6a$. Астроиду можно задать параметрически: $x=a\sin^3 t,\ \ y=a\cos^3 t$. Достаточно вычислить длину четверти астроиды, для которой x и y неотрицательны. Ясно, что $x'=3a\cos t\sin^2 t$ и $y'=-3a\sin t\cos^2 t$. Поэтому $\sqrt{x'^2+y'^2}=3a\sin t\cos t$, а значит, длина четверти астроиды равна

$$3a\int_0^{\pi/2}\sin t\cos t\,dt=\frac{3a}{2}\sin^2 t\Big|_0^{\pi/2}=\frac{3a}{2}.$$

Длина всей астроиды равна $6a$.

6.33. О т в е т: $8a$. Ясно, что

$$\sqrt{x'^2+y'^2}=a\sqrt{(1-\cos t)^2+\sin^2 t}=2a\sin\frac{t}{2}.$$

Поэтому длина ветви циклоиды равна

$$2a\int_0^{2\pi}\sin\frac{t}{2}dt=-4a\cos\frac{t}{2}\Big|_0^{2\pi}=8a.$$

6.34. Мы рассматриваем поверхность шара как фигуру, образованную при вращении кривой $y=\sqrt{R^2-x^2}$. Ясно, что $y'=\frac{-x}{\sqrt{R^2-x^2}}$ и $\sqrt{1+y'^2}=\frac{R}{\sqrt{R^2-x^2}}$. Поэтому рассматриваемая площадь равна

$$2\pi\int_a^{a+h}\sqrt{R^2-x^2}\frac{R}{\sqrt{R^2-x^2}}dx=2\pi R\int_a^{a+h}dx=2\pi Rh.$$

6.35. Из очевидного неравенства $\cos x\le 1$ следует, что $\int_0^t\cos x\,dx\le\int_0^t 1\,dx$, т.е. $\sin t\le t$. Из полученного неравенства следует, что $\int_0^t\sin x\,dx\le\int_0^t x\,dx=\frac{t^2}{2}$, т.е. $1-\cos t\le\frac{t^2}{2}$. Затем из неравенства $1-\frac{t^2}{2}\le\cos t$ аналогично получаем $t-\frac{t^3}{6}\le\sin t$ и т.д.

6.36. Воспользуемся неравенствами $t-\frac{t^3}{6}\le\sin t$ и $\cos t\le 1-\frac{t^2}{2}+\frac{t^4}{24}$ (задача 6.35). Из первого неравенства следует, что

$$\left(\frac{\sin t}{t}\right)^3\ge\left(1-\frac{t^2}{6}\right)^3=1-\frac{t^2}{2}+\frac{t^4}{12}-\frac{t^6}{216}.$$

Поэтому достаточно доказать, что $1-\frac{t^2}{2}+\frac{t^4}{12}-\frac{t^6}{216}>1-\frac{t^2}{2}+\frac{t^4}{24}$. Это неравенство эквивалентно неравенству $\frac{t^4}{24}-\frac{t^6}{216}>0$, т.е. $t^2<9$. При $0<t\le\pi/2$ это неравенство выполняется.

6.37. Рассмотрим функцию $f(t)=\frac{1}{\sin^2 t}-\frac{1}{t^2}$. Ясно, что $f\left(\frac{\pi}{2}\right)=1-\frac{4}{\pi^2}$. Кроме того,

$$f'(t)=\frac{2}{t^3}-\frac{2\cos t}{\sin^3 t}=\frac{2}{\sin^3 t}\left(\left(\frac{\sin t}{t}\right)^3-\cos t\right).$$

Поэтому согласно задаче 6.36 $f'(t)>0$, т.е. на рассматриваемом интервале функция $f(t)$ монотонно возрастает. Значит, $f(t)\le f\left(\frac{\pi}{2}\right)$, т.е. $\frac{1}{\sin^2 t}-\frac{1}{t^2}\le 1-\frac{4}{\pi^2}$.

6.38. Согласно задаче 6.35

$$3\sin t\le 3t-\frac{t^3}{2}+\frac{t^5}{40},$$
$$2t+t\cos t\ge 3t-\frac{t^3}{2}+\frac{t^5}{24}-\frac{t^7}{720}.$$

Поэтому если $\frac{t^5}{24}-\frac{t^7}{720}>\frac{t^5}{40}$, т.е. $t^2<12$, то требуемое неравенство выполняется. С другой стороны, требуемое неравенство очевидным образом выполняется, если $2t>3$, т.е. $t>3/2$.

6.39. Если $x>0$, то

$$1-x+x^2-x^3+\dots-x^{2n-1}<\frac{1}{1+x}<1-x+x^2-x^3+\dots+x^{2n}.$$

Действительно,

$$(1-x+x^2-x^3+\dots-x^{2n-1})(1+x)=1-x^{2n}<1,$$
$$(1-x+x^2-x^3+\dots+x^{2n})(1+x)=1+x^{2n+1}>1.$$

Требуемые неравенства следуют из того, что $\int_0^t\frac{dx}{1+x}=\ln(1+t)$ и $\int_0^t x^k dx=\frac{t^k}{k}$.

6.40. Мы воспользуемся тем, что $\int_0^t\frac{dx}{1+x^2}=\operatorname{arctg} t$. Для $\frac{1}{1+x^2}$ имеем неравенства

$$1-x^2+x^4-x^6+\dots-t^{4n-2}<\frac{1}{1+x^2}<1-x^2+x^4-x^6+\dots+t^{4n}.$$

Интегрируя эти неравенства, получаем требуемое.

6.41. а) Так как $e>1$, то $1\le e^x\le e^a$ при $0\le x\le t$. Поэтому $\int_0^t 1\,dx\le\int_0^t e^x dx\le\int_0^t e^a dx$, т.е. $t\le e^t-1\le e^a t$. Из неравенства $1+x\le e^x\le 1+e^a x$ аналогично получаем $1+t+\frac{t^2}{2}\le e^t\le 1+t+e^a\frac{t^2}{2}$ и т.д.

б) Доказательство аналогично. При $0\ge x\ge t$ получаем $1\ge e^x\ge e^a$. Поэтому $\int_t^0 1\,dx\ge\int_t^0 e^x dx\ge\int_t^0 e^a dx$, т.е. $-t\ge 1-e^t\ge -e^a t$. Последнее неравенство переписывается в виде $t\le e^t-1\le e^a t$. Из неравенства $1+x\le e^x\le 1+e^a x$ аналогично получаем $1+t+\frac{t^2}{2}\ge e^t\ge 1+t+e^a\frac{t^2}{2}$ и т.д.

6.42. Предположим, что если $a\le x\le b$, то $f'(x)-\big(f(x)\big)^2\ge 1$, т.е. $\frac{f'(x)}{1+\left(f(x)\right)^2}\ge 1$. Рассмотрим функцию $F(x)=\operatorname{arctg}\big(f(x)\big)$. Тогда $F'(x)=\frac{f'(x)}{1+\left(f(x)\right)^2}\ge 1$, поскольку $(\operatorname{arctg} y)'=\frac{1}{1+y^2}$. Значит, $\int_a^b F'(x)\,dx\ge b-a=\pi$, т.е. $F(b)-F(a)\ge\pi$. Но арктангенс принимает лишь значения, заключённые между $-\pi/2$ и $\pi/2$, поэтому $F(b)-F(a)<\pi$. Получено противоречие.

6.43. Ясно, что

$$\ln\frac{\sqrt[n]{n!}}{n}=\frac{1}{n}\ln n!-\ln n=$$
$$=\frac{1}{n}\sum_{k=1}^n\ln k-\frac{1}{n}\sum_{k=1}^n\ln n=$$
$$=\sum_{k=1}^n\ln\left(\frac{k}{n}\right)\frac{1}{n}.$$

При $n\to\infty$ последнее выражение превращается в $\int_0^1\ln(x)\,dx=-1$.

Решение этой задачи без использования интегралов приведено в решении задачи 3.35.

6.44. О т в е т: $\pi/4$. Рассматриваемую сумму можно записать в виде $\sum_{k=0}^{n-1}\frac{1}{n}\frac{1}{1+(k/n)^2}$. При $n\to\infty$ эта сумма стремится к $\int_0^1\frac{dx}{1+x^2}=\frac{\pi}{4}$.

6.45. О т в е т: $\pi/4$. Рассматриваемую сумму можно записать в виде $\sum_{k=0}^{n-1}\frac{1}{n}\sqrt{1-\left(\frac{k}{n}\right)^2}$. При $n\to\infty$ эта сумма стремится к $\int_0^1\sqrt{1-x^2}\,dx=\frac{\pi}{4}$.

6.46. О т в е т: $\ln 2$. Заметим, что рассматриваемая сумма равна

$$\frac{1}{n}\left(1+\frac{1}{1+1/n}+\frac{1}{1+2/n}+\dots+\frac{1}{1+n/n}\right)\to\int_1^2\frac{dx}{1+x}=\ln 2.$$

Замечание. Другое доказательство можно найти в решении задачи 7.16 а).

6.47. О т в е т: $\pi/2$. Рассматриваемую сумму можно записать в виде $\sum_{k=0}^{n-1}\frac{1}{n}\frac{1}{\sqrt{1-(k/n)^2}}$. Эта сумма стремится к $\int_0^1\frac{dx}{\sqrt{1-x^2}}=\arcsin x\big|_0^1=\frac{\pi}{2}$.

6.48. Ясно, что $\int_0^1 x^a dx=\frac{1}{a+1}x^{a+1}\Big|_0^1$. Если $a>-1$, то $a+1>0$. Поэтому функция x^{a+1} обращается в нуль при $x=0$. Таким образом, $\int_0^1 x^a dx=\frac{1}{a+1}$ при $-1<a<0$.

6.49. Ясно, что $\int_1^\infty x^{-a}dx=\frac{1}{1-a}x^{1-a}\Big|_1^\infty$. Если $a>1$, то $1-a<0$. Поэтому функция x^{1-a} обращается в нуль при $x=\infty$. Таким образом, $\int_1^\infty x^{-a}dx=\frac{1}{a-1}$ при $a>1$.

6.50. Интегрируя по частям, получаем

$$\int_0^\infty x^k e^{-x}dx=-x^k e^{-x}\big|_0^\infty+k\int_0^\infty x^{k-1}e^{-x}dx.$$

Функция $x^k e^{-x}$ обращается в нуль как при $x=0$, так и при $x=\infty$. Поэтому для $J_k=\int_0^\infty x^k e^{-x}dx$ получаем рекуррентное соотношение $J_k=kJ_{k-1}$. Ясно также, что $\int_0^\infty e^{-x}dx=-e^{-x}\big|_0^\infty=1$.

6.51. Определим на отрезке $[0,1]$ функцию f_n следующим образом: если число x можно представить в виде дроби, знаменатель которой не превосходит n, то $f_n(x)=1$; во всех остальных точках функция равна 0. Последовательность функций f_n сходится к функции, которая равна 1 в рациональных точках и 0 в иррациональных точках. Эта функция не интегрируема (пример 6.1 на с. 94).

6.52. Пусть f_n, где $n\geq 2$, — функция на отрезке $[0,1]$, график которой соединяет последовательно точки $(0,0)$, $(1/n,n)$, $(2/n,0)$ и $(1,0)$. Интегралы от этих функций равны 1, а функции поточечно сходятся к функции, тождественно равной нулю.

6.53. Наибольшее значение функции не меньше среднего, поэтому найдётся такая прямая l, что сумма длин проекций данных векторов на неё не меньше $2L/\pi$. Зададим на прямой l направление. Тогда либо сумма длин положительных проекций на это направление, либо сумма длин отрицательных проекций не меньше L/π. Поэтому либо длина суммы векторов, дающих положительные проекции, либо длина суммы векторов, дающих отрицательные проекции, не меньше L/π.

6.54. Среднее значение суммы длин проекций сторон выпуклого многоугольника равно $2P/\pi$, где P — периметр. Среднее значение не превосходит максимального, поэтому $\frac{2P}{\pi}<2d$, т.е. $P<\pi d$.

Глава 7.

Ряды

Пусть $\{a_n\}$ — некоторая последовательность, $S_n = a_1 + a_2 + \cdots + a_n$. Если последовательность $\{S_n\}$ имеет предел, то говорят, что *ряд* $\sum_{n=1}^{\infty} a_n$ *сходится* (в противном случае — *расходится*). Число $\lim_{n\to\infty} S_n$ называют при этом *суммой ряда*. Числа $S_n = a_1 + a_2 + \cdots + a_n$ называют *частичными суммами* ряда $\sum_{n=1}^{\infty} a_n$.

7.1. Ряды с положительными членами

Начнём с наиболее простого признака сходимости ряда. Ряд вида $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$, где $a_n > 0$, называют *знакочередующимся*. Для знакочередующегося ряда, абсолютные величины членов которого монотонно убывают, имеет место *признак сходимости Лейбница*

Теорема 7.1 (Лейбниц). *Если для знакочередующегося ряда $a_{n+1} \leq a_n$ для всех n и $a_n \to 0$ при $n \to \infty$, то такой ряд сходится.*

Доказательство. Последовательность $S_n = (a_1 - a_2) + \cdots + (a_{2n-1} - a_{2n})$ неубывающая, поскольку $a_{2k-1} - a_{2k} \geq 0$. Кроме того, эта последовательность ограниченная, поскольку $S_n = a_1 - (a_2 - a_3) - \cdots - (a_{2n-2} - a_{2n-1}) - a_{2n} \leq a_1$. Поэтому последовательность S_n имеет предел S. Последовательность $S_n + a_{2n+1}$ имеет тот же самый предел S. $\square$

Историческое замечание. Признак сходимости знакочередующегося ряда сформулировал Готфрид Вильгельм Лейбниц (1646-1716) в 1705 году. В 1713 году он сформулировал определение сходящегося ряда.

Перейдём теперь к рядам с положительными членами. Будем называть ряд $\sum a_n$ с положительными членами a_n *положительным*. Сумма первых n членов положительного ряда монотонно возрастает, поэтому такой ряд сходится тогда и только тогда, когда последовательность чисел S_n ограничена.

Замечание. Для доказательства сходимости положительного ряда достаточно оценить его сверху сходящимся рядом, а для доказательства расходимости — оценить снизу расходящимся положительным рядом.

Задача 7.1. Докажите, что ряд $\sum_{n=2}^{\infty} \frac{1}{n \ln n}$ расходится (Абель).

Теорема 7.2. *Пусть ряды $\sum a_n$ и $\sum b_n$ положительные и, начиная с некоторого n, выполняется неравенство $\frac{b_{n+1}}{b_n} \leq \frac{a_{n+1}}{a_n}$. Тогда если ряд $\sum a_n$ сходящийся ряд, то $\sum b_n$ тоже сходящийся.*

Доказательство. Пусть $\frac{b_{n+1}}{b_n} \leq \frac{a_{n+1}}{a_n}$ при $n \geq p$. Если все члены ряда умножить на одно и то же число, то его сходимость не изменится. Поэтому можно считать, что $b_p = a_p$. Тогда $b_{p+1} \leq a_{p+1}$, поэтому $b_{p+2} \leq a_{p+2}$, и т.д. $\square$

Простейший ряд, которым можно воспользоваться для сравнения с другими рядами, — геометрическая прогрессия $\sum k^n$. Этот ряд сходится при $k < 1$.

Теорема 7.3 (признак сходимости Даламбера). *Пусть все члены ряда $\sum a_n$, начиная с некоторого, положительны и отношение каждого следующего к предыдущему не превосходит некоторого числа $k < 1$. Тогда этот ряд сходящийся.*

Доказательство. Сравнение отношений последовательных членов ряда $\sum a_n$ с отношениями последовательных членов сходящейся геометрической прогрессии доказывает сходимость ряда $\sum a_n$. □

Историческое замечание. Жан ле Рон Даламбер (1717-1783) предложил этот признак сходимости рядов в 1768 году.

Теорема 7.4 (признак сходимости Коши). *Пусть все члены ряда $\sum a_n$, начиная с некоторого, положительны и для некоторого $k < 1$ выполняется неравенство $\sqrt[n]{a_n} \le k$. Тогда этот ряд сходящийся.*

Доказательство. Из неравенства $\sqrt[n]{a_n} \le k$ следует, что $a_n \le k^n$. Поэтому, сравнивая члены ряда $\sum a_n$ с членами сходящейся геометрической прогрессии, получаем сходимость ряда. □

Историческое замечание. Огюстен Луи Коши (1789-1857) предложил этот признак сходимости в 1821 году.

Признак сходимости Коши можно применить в том случае, когда $\overline{\lim}_{n\to\infty} \sqrt[n]{a_n} < 1$. Если же $\overline{\lim}_{n\to\infty} \sqrt[n]{a_n} > 1$, то a_n не стремится к нулю, поэтому ряд $\sum a_n$ расходящийся.

Задача 7.2. Пусть $0 < r < 1$, α — некоторое число. Докажите с помощью признака Коши, что ряд $\sum r^n |\sin n\alpha|$ сходящийся. Можно ли доказать сходимость этого ряда, применив признак Даламбера?

Теорема 7.5 (Коши). *Пусть функция $f(x)$ монотонно убывает и стремится к нулю при $x \to +\infty$. Тогда ряд $\sum f(a+n)$ сходящийся, если $\int_a^l f(x)dx$ стремится к пределу при $l \to +\infty$. Если же этот интеграл неограниченно возрастает, то ряд расходящийся.*

Доказательство. При $a+p-1 \le x \le a+p$ выполняется неравенство $f(a+p-1) \ge f(x) \ge f(a+p)$, поэтому

$$f(a+p-1) \ge \int_{a+p-1}^{a+p} f(x)dx \ge f(a+p).$$

Просуммировав такие равенства для $p = 1, 2, \ldots, n$, получим:

$$\sum_{i=0}^{n-1} f(a+i) \ge \int_a^{a+n} f(x)dx \ge \sum_{i=1}^{n} f(a+i).$$

Поэтому если интеграл $\int_a^l f(x)dx$ стремится к пределу при $l \to +\infty$, то ряд сходящийся, а если этот интеграл неограниченно возрастает, то ряд расходящийся. □

Задача 7.3. Докажите, что ряд $\sum \frac{1}{n^s}$ сходящийся при $s > 1$ и расходящийся при $s \le 1$.

Задача 7.4. Докажите, что положительный ряд $\sum_{n=1}^{\infty} a_n$ с монотонно убывающими членами сходится тогда и только тогда, когда сходится ряд $\sum_{k=1}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \ldots$ (Коши, 1821).

Для положительных рядов есть также следующий признак сходимости Куммера, частными случаями которого являются многие другие известные признаки сходимости.

Теорема 7.6 (Куммер). *а) Положительный ряд $\sum a_n$ сходится тогда и только тогда, когда существует положительный ряд $\sum p_n$ и вещественное число $c > 0$, для которых $p_n \frac{a_n}{a_{n+1}} - p_{n+1} \ge c$.*

б) Положительный ряд $\sum a_n$ расходится тогда и только тогда, когда существует расходящийся положительный ряд $\sum \frac{1}{p_n}$, для которого $p_n \frac{a_n}{a_{n+1}} - p_{n+1} \le 0$.

Доказательство. Докажем сначала достаточность существования требуемого положительного ряда $\sum p_n$.

а) Пусть $p_n \frac{a_n}{a_{n+1}} - p_{n+1} \geq c > 0$. Тогда

$$\begin{aligned} a_1 + (a_2 + \cdots + a_N) &\leq a_1 + \frac{1}{c}(p_1 a_1 - p_2 a_2) + \cdots + \frac{1}{c}(p_N a_N - p_{N+1} a_{N+1}) = \\ &= a_1 + \frac{1}{c}(p_1 a_1 - p_{N+1} a_{N+1}) \leq a_1 + \frac{p_1 a_1}{c}, \end{aligned}$$

поэтому ряд $\sum a_n$ сходится.

б) Пусть положительный ряд $\sum \frac{1}{p_n}$ расходится и $p_n \frac{a_n}{a_{n+1}} - p_{n+1} \leq 0$. Тогда

$$a_2 \geq \frac{p_1 a_1}{p_2}, \quad a_3 \geq \frac{p_2 a_2}{p_3} \geq \frac{p_1 a_1}{p_3}, \ldots$$

Следовательно,

$$\sum_{k=1}^{\infty} a_k \geq a_1 + p_1 a_1 \sum_{k=2}^{\infty} \frac{1}{p_k},$$

поэтому ряд $\sum a_n$ расходится.

Докажем теперь необходимость существования требуемого положительного ряда $\sum p_n$.

а) Пусть положительный ряд $\sum a_n$ сходится. Построим положительный ряд $\sum p_n$, для которого $p_n \frac{a_n}{a_{n+1}} - p_{n+1} \geq 1$. Положим $p_n = (M - \sum_{i=1}^{n} a_i)/a_n$, где $M = \sum_{i=1}^{\infty} a_i$. Ясно, что $p_n > 0$ и

$$\begin{aligned} p_n \frac{a_n}{a_{n+1}} - p_{n+1} &= \frac{1}{a_{n+1}} \left(\left(M - \sum_{i=1}^{n} a_i \right) - \left(M - \sum_{i=1}^{n+1} a_i \right) \right) = \\ &= \frac{a_{n+1}}{a_{n+1}} = 1. \end{aligned}$$

б) Пусть положительный ряд $\sum a_n$ расходится. Положим $p_n = \frac{1}{a_n} \sum_{i=1}^{n} a_i$. Ясно, что $p_n > 0$ и

$$p_n \frac{a_n}{a_{n+1}} - p_{n+1} = \frac{1}{a_{n+1}} \left(\sum_{i=1}^{n} a_i - \sum_{i=1}^{n+1} a_i \right) = -\frac{a_{n+1}}{a_{n+1}} = -1.$$

Докажем, что ряд $\sum \frac{1}{p_n}$ расходится. Для этого достаточно доказать, что для любого натурального m можно выбрать $n > m$ так, что $\sum_{k=m}^{n} \frac{1}{p_k} > \frac{1}{2}$. Выберем $n > m$ так, что $a_m + \cdots + a_n > a_1 + \cdots + a_{m-1}$. Тогда

$$\begin{aligned} \sum_{k=m}^{n} \frac{1}{p_k} &= \frac{a_m}{a_1 + \cdots + a_m} + \cdots + \frac{a_n}{a_1 + \cdots + a_n} > \frac{a_m + \cdots + a_n}{a_1 + \cdots + a_n} = \\ &= \left(\frac{a_1 + \cdots + a_n}{a_m + \cdots + a_n} \right)^{-1} = \left(\frac{a_1 + \cdots + a_{m-1}}{a_m + \cdots + a_n} + 1 \right)^{-1} > \\ &> (1+1)^{-1} = \frac{1}{2}. \end{aligned}$$

□

Следствие. *Предположим, что последовательность чисел $K_n = c_n \frac{a_n}{a_{n+1}} - c_{n+1}$ имеет предел K. Тогда при $K > 0$ ряд $\sum a_n$ сходится, а при $K < 0$ расходится.*

Историческое замечание. Эрнст Эдуард Куммер (1810-1893) предложил достаточность этого признака сходимости в 1835 году с некоторым дополнительным ограничением. Улисс Дини (1845-1918) устранил это ограничение в 1867 году. Необходимость этого признака сформулирована и доказана в [Ji].

Из признака сходимости Куммера можно вывести другие важные признаки сходимости положительного ряда $\sum a_n$. Например, при $p_n = 1$ получаем уже известный нам признак сходимости Даламбера.

Положим теперь $p_n = n$ и заметим, что ряд $\sum \frac{1}{p_n} = \sum \frac{1}{n}$ расходится. Для такой последовательности чисел p_n признак сходимости Куммера превращается в следующий *признак сходимости Раабе*. Рассмотрим последовательность чисел $R_n = n\left(\frac{a_n}{a_{n+1}} - 1\right)$. Если существует $r > 1$, для которого $R_n \geq r$ при всех достаточно больших n, то ряд $\sum a_n$ сходится. Если же $R_n \leq 1$ при всех достаточно больших n, то ряд $\sum a_n$ расходится.

Историческое замечание. Йозеф Людвиг Раабе (1801-1859) доказал этот признак сходимости в 1832 году.

Положим, наконец, $p_n = n \ln n$, $(n \geq 2)$. Ряд $\sum_{n=2}^{\infty} 1/p_n = \sum_{n=2}^{\infty} \frac{1}{n \ln n}$ расходится (задача 7.1). Для такой последовательности чисел p_n получаем:

$$\begin{aligned} K_n &= n \ln n \frac{a_n}{a_{n+1}} - (n+1)\ln(n+1) = \\ &= \ln n \left[n\left(\frac{a_n}{a_{n+1}} - 1\right) - 1\right] - \ln\left(1 + \frac{1}{n}\right)^{n+1} = \\ &= B_n - \ln\left(1 + \frac{1}{n}\right)^{n+1}, \end{aligned}$$

где $B_n = \ln n \left[n\left(\frac{a_n}{a_{n+1}} - 1\right) - 1\right]$. Ясно, что $\lim_{n\to\infty} \ln\left(1 + \frac{1}{n}\right)^{n+1} = \ln e = 1$. Поэтому последовательность K_n имеет предел K тогда и только тогда, когда последовательность B_n имеет предел B, и при этом $K = B - 1$. В этом случае признак сходимости Куммера превращается в следующий *признак сходимости Бертрана*: если последовательность B_n имеет предел B (конечный или бесконечный), то при $B > 1$ ряд $\sum a_n$ сходится, а при $B < 1$ расходится.

Историческое замечание. Жозеф Бертран (1822-1900) предложил этот признак сходимости в 1842 году.

Из признаков Даламбера, Раабе и Бертрана можно вывести следующий *признак сходимости Гаусса*

Теорема 7.7 (Гаусс). *Пусть $\sum a_n$ — положительный ряд и*

$$\frac{a_n}{a_{n+1}} = \lambda + \frac{\mu}{n} + \frac{\theta_n}{n^2},$$

где λ и μ — константы, $|\theta_n| \leq C$. Тогда:

а) если $\lambda > 1$ или $\lambda = 1$ и $\mu > 1$, то ряд $\sum a_n$ сходится.

б) если $\lambda < 1$ или $\lambda = 1$ и $\mu \leq 1$, то ряд $\sum a_n$ расходится.

Доказательство. При $\lambda \neq 1$ требуемое утверждение следует из признака Даламбера, поскольку $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \frac{1}{\lambda}$.

Рассмотрим теперь случай, когда $\lambda = 1$. В этом случае

$$R_n = n\left(\frac{a_n}{a_{n+1}} - 1\right) = \mu + \frac{\theta_n}{n}$$

и $\lim_{n\to\infty} R_n = \mu$. Поэтому при $\mu \neq 1$ требуемое утверждение следует из признака Раабе. Если же $\mu = 1$, то

$$B_n = \ln n \left[n\left(\frac{a_n}{a_{n+1}} - 1\right) - 1\right] = (R_n - 1)\ln n = \frac{\ln n}{n}\theta_n.$$

Последовательность θ_n ограничена и $\lim_{n\to\infty} \frac{\ln n}{n} = 0$, поэтому $B = \lim_{n\to\infty} = 0$. Следовательно, по признаку Бертрана ряд $\sum a_n$ расходится. □

Историческое замечание. Гаусс использовал этот признак сходимости в 1812 году при исследовании сходимости гипергеометрического ряда.

Задача 7.5. а) Пусть a_1, a_2, a_3, … — возрастающая последовательность натуральных чисел, в десятичной записи которых не встречается цифра 1. Докажите, что ряд $\sum_{n=1}^{\infty} 1/a_n$ сходится.

б) Пусть a_1, a_2, a_3, … — возрастающая последовательность натуральных чисел, в десятичной записи которых не встречается подряд сто единиц. Докажите, что ряд $\sum_{n=1}^{\infty} 1/a_n$ сходится.

Задача 7.6. Пусть a_1, a_2, a_3, … — возрастающая последовательность натуральных чисел, причём ряд $\sum_{n=1}^{\infty} 1/a_n$ расходится. Докажите, что число $0, a_1a_2a_3\ldots$ иррационально.

7.2. Абсолютно сходящиеся ряды

Ряд $\sum a_n$ называется *абсолютно сходящимся*, если ряд $\sum |a_n|$ сходящийся. Абсолютно сходящийся ряд сходится. Доказать это можно, например, воспользовавшись признаком Коши сходимости последовательности: при достаточно большом n сумма $\sum_{i=n}^{n+p} |a_i|$ сколь угодно мала, поэтому величина $|\sum_{i=n}^{n+p} a_i| \leq \sum_{i=n}^{n+p} |a_i|$ тоже сколь угодно мала.

Теорема 7.8 (о перестановке членов ряда). *Ряд, полученный при перестановке членов абсолютно сходящегося ряда, абсолютно сходится. Сумма абсолютно сходящегося ряда не меняется при перестановке его членов.*

Доказательство. Рассмотрим сначала случай, когда члены ряда неотрицательны. Пусть a_{i_1}, $a_{i_2}, \ldots$ — некоторая перестановка членов ряда $a_1, a_2, \ldots$ Тогда $a_{i_1} + a_{i_2} + \cdots + a_{i_n} \leq a_1 + a_2 + \cdots + a_N$, где N — наибольшее из чисел $i_1, \ldots, i_n$. Поэтому при перестановке членов ряда сумма ряда не увеличивается. Но ряд $\sum a_k$ тоже можно получить из ряда $\sum a_{i_k}$ перестановкой, поэтому при перестановке членов ряда сумма ряда не уменьшается.

Член a_k произвольного абсолютно сходящегося ряда можно представить в виде $a_k = b_k - c_k$, где одно из числе b_k и c_k равно 0, а другое неотрицательно. Ряды $\sum b_k$ и $\sum c_k$ сходятся, поэтому $\sum a_{i_k} = \sum b_{i_k} - \sum c_{i_k} = \sum b_k - \sum c_k = \sum a_k$. □

Историческое замечание. Теорему 7.8 доказал Петер Густав Лежён Дирихле (1805-1859) в 1837 году.

Сходящийся, но не абсолютно сходящийся, ряд $\sum a_n$ называется *условно сходящимся*

Пример 7.1. *Ряд $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$ сходится условно.*

Доказательство. Ряд $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$ сходится согласно признаку Лейбница. Ряд $\sum_{n=1}^{\infty} \frac{1}{n}$ расходится. □

Задача 7.7. Сопоставим числу a_n числа $p_n = \frac{|a_n|+a_n}{2}$ и $q_n = \frac{|a_n|-a_n}{2}$ (одно из этих чисел равно 0, а другое $|a_n|$). Докажите, что если ряд $\sum a_n$ сходится условно, то оба ряда $\sum p_n$ и $\sum q_n$ расходятся.

Теорема 7.9 (Риман). *Сумма ряда, полученного при перестановке членов условно сходящегося ряда, может быть произвольным числом.*

Доказательство. Нулевые члены ряда не влияют на сходимость; при перестановке нулевые члены можно ставить на нечётные места (пока они не закончатся). Поэтому можно считать, что все члены ряда ненулевые. Расположим по порядку отдельно положительные члены $p_1, p_2, \ldots$ и отрицательные $q_1, q_2, \ldots$ Из условной сходимости ряда следует, что $\sum p_n = \infty$, $\sum q_n = -\infty$ и $\lim_{n\to\infty} p_n = \lim_{n\to\infty} q_n = 0$.

Пусть S — данное число. Сначала будем брать по порядку положительные члены, пока не добьёмся выполнения неравенства $p_1 + \cdots + p_k > S$, затем добавим к ним отрицательные члены, чтобы сумма $p_1 + \cdots + p_k + q_1 + \ldots q_l$ стала меньше S, затем добавим положительные члены, чтобы сумма стала больше S, и т.д. (мы всегда сразу останавливаемся, как только требуемое неравенство выполняется). В результате получим требуемую перестановку. Действительно, частичные суммы S_k, $\ldots$, S_{k+l-1} отличаются от S не больше чем на p_k, следующие частичные суммы отличаются от S не больше чем на q_l и т.д. □

Историческое замечание. Бернхард Риман (1826-1866) доказал теорему об условно сходящихся рядах в 1853 году.

7.3. Признак Абеля сходимости рядов

Лемма (Абель). *Пусть $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_p$ — убывающая последовательность положительных чисел, $a_0, a_1, \ldots, a_p$ — некоторая последовательность чисел. Тогда если все суммы $S_k = \sum_{i=0}^{k} a_i$, $k = 0, 1, \ldots, p$, заключены между двумя числами A и B, то сумма $S = \sum_{i=0}^{p} \varepsilon_i a_i$ заключена между $A\varepsilon_0$ и $B\varepsilon_0$.*

Доказательство. Ясно, что $a_0 = S_0$, $a_1 = S_1 - S_0$, $\dots$, $a_p = S_p - S_{p-1}$. Поэтому

$$S = S_0(\varepsilon_0 - \varepsilon_1) + S_1(\varepsilon_1 - \varepsilon_2) + \dots + S_{p-1}(\varepsilon_{p-1} - \varepsilon_p) + S_p\varepsilon_p.$$

Все разности $\varepsilon_0 - \varepsilon_1$, $\varepsilon_1 - \varepsilon_2$, $\dots$, $\varepsilon_{p-1} - \varepsilon_p$ положительны, поэтому каждое из чисел S_i можно заменить на A или на B и получить требуемую оценку для S. □

Теорема 7.10 (Абель). *Пусть ряд $\sum a_i$ сходящийся, а последовательность положительных чисел ε_i монотонно убывает и стремится к нулю. Тогда ряд $\sum \varepsilon_i a_i$ сходящийся.*

Историческое замечание. Нильс Генрик Абель (1802-1829) доказал теорему 7.10 в 1826 году. В 1862 году Дирихле доказал более общий признак сходимости рядов (теорема 7.11).

Теорема 7.11 (Дирихле). *Пусть сумма первых n членов ряда $\sum a_i$ ограничена некоторой константой (не зависящей от n), а последовательность положительных чисел ε_i монотонно убывает и стремится к нулю. Тогда ряд $\sum \varepsilon_i a_i$ сходящийся.*

Доказательство. Пусть $|\sum_{i=1}^{n} a_i| < A$ для всех n. Тогда $|\sum_{i=n+k}^{n+l} a_i| < 2A$, поэтому согласно лемме Абеля $|\sum_{i=n+k}^{n+l} \varepsilon_i a_i| < 2A\varepsilon_{n+k}$. По условию при достаточно большом n число ε_{n+k} достаточно мало. □

Задача 7.8. Докажите, что ряд $\sum(-1)^n a_n$, где a_n — последовательность положительных чисел, монотонно стремящаяся к нулю, сходится.

Задача 7.9. Последовательность положительных чисел ε_n монотонно убывает и стремится к нулю. Докажите, что ряд $\sum \varepsilon_n \sin n\alpha$ сходится.

Задача 7.10. Последовательность положительных чисел ε_n монотонно убывает и стремится к нулю. Докажите, что при $\alpha \neq 2k\pi$ ряд $\sum \varepsilon_n \cos n\alpha$ сходится.

Задача 7.11. Последовательность положительных чисел ε_n монотонно убывает и стремится к нулю. Докажите, что ряд $\sum(-1)^n \varepsilon_n \sin n\alpha$ сходится и при $\alpha \neq 2k\pi$ ряд $\sum(-1)^n \varepsilon_n \cos n\alpha$ сходится.

Задача 7.12. Пусть $a_0, a_1, \dots$ и $b_0, b_1, \dots$ — произвольные последовательности чисел, $A_n = a_0 + a_1 + \dots + a_n$. Тогда при $n \geq 0$ и $k \geq 1$ имеет место равенство

$$\sum_{i=n+1}^{n+k} a_i b_i = \sum_{i=n+1}^{n+k} A_i(b_i - b_{i+1}) - A_n b_{n+1} + A_{n+k} b_{n+k+1}.$$

Замечание. Формулу из задачи 7.12 называют *формулой суммирования Абеля* или *формулой суммирования по частям*.

Лемма Абеля применяется не только в теории рядов. Например, её можно применить для доказательства следующей теоремы о среднем для интегралов.

Задача 7.13. *а) Пусть $f(x)$ и $\varphi(x)$ — непрерывные функции на отрезке $[a, b]$, причём функция $\varphi(x)$ неотрицательная убывающая. Докажите, что

$$\int_a^b f(x)\varphi(x)dx = \varphi(a)\int_a^c f(x)dx$$

для некоторого $c \in (a, b)$ (теорема Бонне о среднем).

б) Пусть $f(x)$ и $\varphi(x)$ — непрерывные функции на отрезке $[a, b]$, причём функция $\varphi(x)$ неотрицательная возрастающая. Докажите, что

$$\int_a^b f(x)\varphi(x)dx = \varphi(b)\int_d^b f(x)dx$$

для некоторого $d \in (a, b)$.

в) Пусть $f(x)$ и $\varphi(x)$ — непрерывные функции на отрезке $[a, b]$, причём функция $\varphi(x)$ монотонная. Докажите, что

$$\int_a^b f(x)\varphi(x)dx = \varphi(a)\int_a^c f(x)dx + \varphi(b)\int_c^b f(x)dx$$

для некоторого $c \in (a, b)$

7.4. Произведение Коши двух рядов

Произведением Коши рядов $\sum_{i=0}^{\infty} a_i$ и $\sum_{j=0}^{\infty} b_j$ называют ряд $\sum_{k=0}^{\infty} c_k$, где $c_k = \sum_{i+j=k} a_i b_j$.

Теорема 7.12 (Коши, 1821). *Если оба ряда $\sum_{i=0}^{\infty} a_i$ и $\sum_{j=0}^{\infty} b_j$ сходятся абсолютно, то их произведение Коши сходится абсолютно, причём сумма произведения Коши этих рядов равна произведению их сумм.*

Доказательство. Занумеруем произвольно произведения $a_i b_j$ и составим из них ряд $\sum_{k=0}^{\infty} p_k$. Покажем, что этот ряд сходится абсолютно. Пусть $A = \sum_{i=0}^{\infty} |a_i|$ и $B = \sum_{j=0}^{\infty} |b_j|$. Среди индексов i и j для произведений $a_i b_j$, равных $p_0, p_1, \ldots, p_N$, выберем наибольший; пусть он равен M. Тогда

$$\sum_{i=0}^{N} |p_k| \leq \left(\sum_{i=0}^{M} |a_i|\right)\left(\sum_{j=0}^{M} |b_j|\right) \leq AB.$$

Из абсолютной сходимости ряда $\sum_{k=0}^{\infty} p_k$ следует, что ряд $\sum_{k=0}^{\infty} c_k$ сходится абсолютно и $\sum_{k=0}^{\infty} c_k = \sum_{k=0}^{\infty} p_k$. Вычислять сумму ряда $\sum_{k=0}^{\infty} p_k$ можно для произвольной нумерации чисел p_k. Их можно последовательно занумеровать так, чтобы для всех натуральных n выполнялось равенство

$$\sum_{k=0}^{(n+1)^2-1} p_k = \left(\sum_{i=0}^{n} a_i\right)\left(\sum_{j=0}^{n} b_j\right).$$

Устремляя n к бесконечности, получаем требуемое. □

Следующий пример показывает, что произведение Коши условно сходящихся рядов может быть расходящимся рядом.

Пример 7.2. *Пусть $a_0 = b_0 = 0$ и $a_n = b_n = \frac{(-1)^{n-1}}{\sqrt{n}}$ при $n \geq 1$. Тогда ряды $\sum_{i=0}^{\infty} a_i$ и $\sum_{j=0}^{\infty} b_j$ сходятся (по признаку Лейбница), но их произведение Коши $\sum_{k=0}^{\infty} c_k$ — расходящийся ряд.*

Доказательство. Ясно, что $c_0 = c_1 = 0$ и при $n \geq 2$

$$c_n = (-1)^n \left(\frac{1}{\sqrt{1\cdot(n-1)}} + \frac{1}{\sqrt{2(n-2)}} + \cdots + \frac{1}{\sqrt{(n-1)\cdot 1}}\right).$$

Знаменатель каждой из $n-1$ дробей меньше $n-1$, поэтому $|c_n| > 1$. □

Но если один из двух рядов сходится абсолютно, то справедлива следующая теорема.

Теорема 7.13 (Мертенс). *Если хотя бы один из двух сходящихся рядов $\sum_{i=0}^{\infty} a_i = A$ и $\sum_{j=0}^{\infty} b_j = B$ сходится абсолютно, то их произведение Коши $\sum_{k=0}^{\infty} c_k$ сходится, причём $\sum_{k=0}^{\infty} c_k = AB$.*

Доказательство. Для определённости будем считать, что абсолютно сходится ряд $\sum_{i=0}^{\infty} a_i$. Пусть $A_n = \sum_{i=0}^{n} a_i$, B_n и C_n определены аналогично; положим также $\beta_n = B_n - B$. Тогда

$$\begin{aligned} C_n &= a_0 + (a_0b_1 + a_1b_0) + \cdots + (a_0b_n + \cdots + a_nb_0) = \\ &= a_0B_n + a_1B_{n-1} + \cdots + a_nB_0 = \\ &= a_0(B+\beta_n) + a_1(B+\beta_{n-1}) + \cdots + a_n(B+\beta_0) = \\ &= A_nB + a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0. \end{aligned}$$

Ясно, что $\lim_{n\to\infty} A_nB = AB$, поэтому достаточно доказать, что $\lim_{n\to\infty} \gamma_n = 0$, где $\gamma_n = a_0\beta_n + a_1\beta_{n-1} + \cdots + a_n\beta_0$.

Пусть $\alpha = \sum_{i=0}^{\infty} |a_i|$. Для заданного $\varepsilon > 0$ выберем N так, что $|\beta_n| \leq \varepsilon$ при $n \geq N$. Для таких n получаем

$$\begin{aligned} |\gamma_n| &\leq |a_n\beta_0 + \cdots + a_{n-N}\beta_N| + |a_{n-N-1}\beta_{N+1} + \cdots + a_0\beta_n| \leq \\ &\leq |a_n\beta_0 + \cdots + a_{n-N}\beta_N| + \varepsilon\alpha. \end{aligned}$$

Ясно также, что $\lim_{n\to\infty} |a_n\beta_0 + \cdots + a_{n-N}\beta_N| = 0$. □

Историческое замечание. Франц Мертенс (1840-1927) доказал теорему 7.13 в 1875 году.

Замечание. В 1826 году Нильс Генрик Абель (1802-1829) доказал, что если два ряда сходятся и их произведение Коши тоже сходится, то сумма произведения Коши этих рядов равна произведению их сумм (теорема 7.19).

7.5. Гармонический ряд

Ряд $\sum_{n=1}^{\infty} \frac{1}{n}$ называют *гармоническим*. Этот ряд расходящийся (задача 7.3; другое доказательство даёт задача 7.14). Мы будем использовать обозначение

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n}.$$

Задача 7.14. Докажите, что для любого натурального n

$$\ln(n+1) < H_n < 1 + \ln n.$$

Задача 7.15. Докажите, что ряд $1 + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \dots$ расходится.

Задача 7.16. а) Пусть $a_n = \frac{1}{n} + \frac{1}{n+1} + \dots + \frac{1}{2n}$. Вычислите предел $\lim_{n\to\infty} a_n$.
б) Пусть $b_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots - \frac{1}{2n}$. Докажите, что $\lim_{n\to\infty} b_n = \ln 2$.

Задача 7.17. а) Докажите, что

$$\frac{1}{1\cdot 2} + \frac{1}{3\cdot 4} + \frac{1}{5\cdot 6} + \frac{1}{7\cdot 8} + \dots = \ln 2.$$

б) Докажите, что

$$\frac{1}{2\cdot 3} + \frac{1}{4\cdot 5} + \frac{1}{6\cdot 7} + \dots = 1 - \ln 2.$$

Задача 7.18. Докажите, что

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \dots = \frac{1}{2}\ln 2.$$

Задача 7.19. Пусть $a_n = \frac{1}{2n+1} + \frac{1}{2n+3} + \frac{1}{2n+5} + \dots + \frac{1}{4n-1}$. Докажите, что $\lim_{n\to\infty} a_n = \frac{1}{2}\ln 2$.

Задача 7.20. Докажите, что

$$\frac{1}{1\cdot 2\cdot 3} + \frac{1}{5\cdot 6\cdot 7} + \frac{1}{9\cdot 10\cdot 11} + \dots = \frac{1}{4}\ln 2.$$

Задача 7.21. Докажите, что

$$\frac{1}{1\cdot 2\cdot 3} + \frac{1}{3\cdot 4\cdot 5} + \frac{1}{5\cdot 6\cdot 7} + \dots = \ln 2 - \frac{1}{2}.$$

Задача 7.22. а) Пусть $a > 0$ и $A_n = \prod_{k=1}^{n}\left(1 + \frac{a}{k}\right)$. Докажите, что $\lim_{n\to\infty} \sqrt[n]{A_n} = 1$.
б) Пусть число $b > 0$ не целое и $B_n = \prod_{k=1}^{n}\left|1 - \frac{b}{k}\right|$. Докажите, что $\lim_{n\to\infty} \sqrt[n]{B_n} = 1$.

Задача 7.23. Докажите, что

$$\sum_{k=1}^{n} \frac{x^k}{k} = H_n + \sum_{k=1}^{n} \binom{n}{k} \frac{(x-1)^k}{k}.$$

Задача 7.24. Пусть $p_1 = 2$, $p_2 = 3$, ...— последовательные простые числа.
а) Докажите, что

$$\left(1 - \frac{1}{2}\right)^{-1}\left(1 - \frac{1}{3}\right)^{-1} \cdots \left(1 - \frac{1}{p_m}\right)^{-1} > \sum_{n=1}^{p_m} \frac{1}{n}.$$

б) Докажите, что $1 + \sum_{n=1}^{m} \frac{1}{p_n} > \ln\ln p_m$.

Задача 7.25. Докажите, что число $0{,}12357111317\dots$ (подряд записываются последовательные простые числа) иррационально.

Задача 7.26. Докажите, что $H_n = \int_0^1 \frac{1-(1-t)^n}{t}\,dt$.

Постоянная Эйлера

Задача 7.27. Докажите, что существует предел

$$\lim_{n\to\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}-\ln n\right).$$

Предел $\gamma = 0,5772157\ldots$ из задачи 7.27 называют *постоянной Эйлера*.

Следующая задача не имеет прямого отношения к постоянной Эйлера, но её формулировка и решение близки к задаче 7.27.

Задача 7.28. Докажите, что существует предел

$$\lim_{n\to\infty}\left(1+\frac{1}{\sqrt{2}}+\cdots+\frac{1}{\sqrt{n}}-2\sqrt{n}\right).$$

Задача 7.29. Докажите, что $\displaystyle\prod_{n=1}^{\infty}\frac{e^{1/n}}{1+\frac{1}{n}} = e^{\gamma}$, где γ — постоянная Эйлера.

7.6. Степенные ряды

Помимо числовых рядов часто рассматривают и ряды функций. Говорят, что ряд $\sum_{k=1}^{\infty} u_k(x)$ *равномерно сходится* на множестве A, если последовательность функций $f_n(x) = \sum_{k=1}^{n} u_k(x)$ равномерно сходится на множестве A к некоторой функции $f(x)$. Напомним, что равномерно сходящийся ряд непрерывных функций сходится к непрерывной функции (теорема 3.9).

Историческое замечание. Понятие равномерно сходящегося ряда функций ввёл Георг Габриэль Стокс (1819-1903) в 1848 году.

Теорема 7.14 (Вейерштрасс). *Пусть $|u_k(x)| \le M_k$ для всех $x \in A \subset \mathbb{R}$ при $k \ge k_0$ и ряд $\sum M_k$ сходится. Тогда ряд $\sum u_k(x)$ равномерно сходится на множестве A.*

Доказательство. Ясно, что ряд $\sum u_k(x)$ при фиксированном $x \in A$ сходится; пусть $s(x)$ — сумма этого ряда. Если $s_n(x) = \sum_{k=1}^{n} u_k(x)$, то

$$|s(x) - s_n(x)| = |u_{n+1}(x) + u_{n+2}(x) + \cdots| \le M_{n+1} + M_{n+2} + \cdots$$

Для любого $\varepsilon > 0$ существует такое число $N \ge k_0$, что $M_{n+1} + M_{n+2} + \cdots < \varepsilon$ при $n \ge N$. Поэтому ряд $\sum u_k(x)$ равномерно сходится. $\square$

Этот признак сходимости ряда функций иногда называют *M-тестом Вейерштрасса*.

Задача 7.30. Докажите, что постоянная Эйлера γ равна пределу $\lim_{x\to 1+0}\left(\sum_{n=1}^{\infty}\frac{1}{n^x}-\frac{1}{x-1}\right)$.

Степенной ряд — это ряд функций вида $\sum_{n=0}^{\infty} a_n x^n$.

Теорема 7.15. *Если степенной ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится при $x = x_0$, то он абсолютно сходится при $|x| < |x_0|$ и равномерно сходится на любом отрезке, содержащемся в интервале $(-|x_0|, |x_0|)$.*

Доказательство. По условию ряд $\sum_{n=0}^{\infty} a_n x_0^n$ сходится, поэтому $a_n x_0^n \to 0$ при $n \to \infty$. Выберем M так, что $|a_n x_0^n| < M$ для всех n. Тогда $|a_n x^n| < M\left|\frac{x}{x_0}\right|^n$ при $|x| < |x_0|$. Поэтому каждый член ряда $\sum_{n=0}^{\infty} |a_n x^n|$ меньше соответственного члена сходящегося ряда $\sum_{n=0}^{\infty} M\left|\frac{x}{x_0}\right|^n$.

Равномерная сходимость следует из теоремы 7.14. $\square$

Теорема 7.16. *Пусть $\mu = \overline{\lim}_{n\to\infty}\sqrt[n]{|a_n|}$. Если $\mu = 0$, то степенной ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится абсолютно при всех x. Если $\mu = \infty$, то степенной ряд расходится при $x \ne 0$. Если $0 < \mu < \infty$, то степенной ряд сходится абсолютно при $|x| < \mu^{-1}$ и расходится при $|x| > \mu^{-1}$.*

Доказательство. Пусть $\mu < \infty$. Тогда для любого $\varepsilon > 0$ неравенство $\sqrt[n]{|a_n|} > \mu + \varepsilon$ выполняется лишь для конечного множества чисел n, поэтому $|a_n| \le (\mu+\varepsilon)^n$ для всех достаточно больших n. Поэтому ряд $\sum_{n=0}^{\infty} |a_n x^n|$ сходится при $|x| < \frac{1}{\mu+\varepsilon}$. При $\mu > 0$ получаем, что при $|x| < \mu^{-1}$ ряд сходится абсолютно. При $\mu = 0$ получаем, что для любого $\varepsilon > 0$ ряд сходится абсолютно при $|x| < 1/\varepsilon$, поэтому ряд сходится абсолютно при всех x.

Пусть $0 < \mu < \infty$. Тогда для любого $\varepsilon > 0$ для всех достаточно больших n выполняется неравенство $|a_n| \ge (\mu-\varepsilon)^n$. Поэтому если $|x| > \frac{1}{\mu-\varepsilon}$, то $|a_n x^n| > 1$ для достаточно больших n. Поэтому ряд $\sum_{n=0}^{\infty} a_n x^n$ расходится.

Пусть $\mu = \infty$. Тогда для любого числа положительного числа M для всех достаточно больших n выполняется неравенство $|a_n| \ge M^n$, поэтому при $|x| > \frac{1}{M}$ выполняется неравенство $|a_n x^n| \ge |Mx|^n > 1$. $\square$

Число $R = (\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|})^{-1}$ (или ∞) называют *радиусом сходимости* степенного ряда.

Историческое замечание. Теорему 7.16 доказал Огюстен Луи Коши (1789-1857) в 1821 году, но на неё долго не обращали должного внимания. В 1888 году Жак Адамар (1865-1963) заново открыл эту теорему и показал, что она имеет много важных приложений.

Пример 7.3. *Радиус сходимости ряда $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ равен ∞, радиус сходимости ряда $\sum_{n=0}^{\infty} n! x^n$ равен 0. Радиус сходимости ряда $\sum_{n=0}^{\infty} \frac{x^n}{n^s}$ (при $s > 0$) равен 1. При $s > 1$ этот ряд сходится при $|x| = 1$; при $0 < s \le 1$ ряд сходится при $x = -1$ и расходится при $x = 1$; при $s \le 0$ ряд расходится при $|x| = 1$.*

Доказательство. Первые два утверждения следуют из того, что $\lim_{n\to\infty} \sqrt[n]{n!} = \infty$ (задача 2.15). Утверждение о радиусе сходимости последнего ряда следует из того, что $\lim_{n\to\infty} \sqrt[n]{n} = 1$ (задача 2.14). Согласно задаче 7.3 ряд $\sum \frac{1}{n^s}$ сходящийся при $s > 1$ и расходящийся при $s \le 1$. Сходимость ряда $\sum_{n=0}^{\infty} \frac{(-1)^n}{n^s}$ при $0 < s \le 1$следует из признака Лейбница. При $s \le 0$ и $|x| = 1$ члены ряда не стремятся к нулю. $\square$

Теорема 7.17. *Радиус сходимости ряда $\sum_{n=0}^{\infty} a_n x^n$ равен радиусу сходимости ряда $\sum_{n=1}^{\infty} n a_n x^{n-1}$, полученного почленным дифференцированием.*

Доказательство. Пусть радиус сходимости ряда $\sum_{n=0}^{\infty} a_n x^n$ равен R и $|x| < R$. Выберем R_1 так, что $|x| < R_1 < R$. Тогда ряд $\sum_{n=0}^{\infty} a_n R_1^n$ сходится абсолютно и $a_n R_1^n \to 0$ при $n \to \infty$. Выберем M так, что $|a_n| < M R_1^{-n}$ для всех n. Тогда все члены ряда $\sum_{n=1}^{\infty} n|a_n x|^{n-1}$ меньше соответствующих членов ряда $\frac{M}{R_1} \sum_{n=0}^{\infty} \frac{n|x|^{n-1}}{R_1^{n-1}}$. Последний ряд сходится, поскольку радиус сходимости ряда $\sum_{n=0}^{\infty} n x^n$ равен 1. Если же $|x| > R$, то $a_n x^n$ не стремится к нулю и, тем более, $n a_n x^{n-1}$ не стремится к нулю. $\square$

Следствие. *Радиус сходимости ряда $\sum_{n=0}^{\infty} a_n x^n$ равен радиусу сходимости ряда $\sum_{n=0}^{\infty} \frac{1}{n+1} a_n x^{n+1}$, полученного почленным интегрированием.*

Пусть радиус сходимости ряда $\sum_{n=0}^{\infty} a_n x^n$ равен R. Тогда при $|x| < R$ определены функции $f(x) = \sum_{n=0}^{\infty} a_n x^n$, $g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$ и $h(x) = \sum_{n=0}^{\infty} \frac{1}{n+1} a_n x^{n+1}$. Из равномерной сходимости этих трёх рядов на любом отрезке, содержащемся в интервале $(-R, R)$ следует, что на этом интервале функция $f(x)$ непрерывна и выполняются равенства $g(x) = f'(x)$ (теорема 5.13) и $h(x) = \int_0^x f(t)dt$ (теорема 6.11). Радиус сходимости ряда, который сходится к функции $g'(x) = f''(x)$, тоже равен R, и т.д. Таким образом, функция $f(x)$ бесконечно дифференцируема при $|x| < R$.

Задача 7.31. а) Докажите, что если радиус сходимости ряда $\sum_{n=0}^{\infty} a_n x^n = f(x)$ положителен, то

$$\lim_{x\to 0} \frac{f(x) - (a_0 + a_1 x + \cdots + a_n x^n)}{x^{n+1}} = a_{n+1}.$$

б) Докажите, что если радиусы сходимости рядов $\sum_{n=0}^{\infty} a_n x^n = f(x)$ и $\sum_{n=0}^{\infty} b_n x^n = g(x)$ положительны и $f(x_n) = g(x_n)$ для некоторой последовательности отличных от нуля чисел x_n, сходящейся к нулю, то $a_n = b_n$ для всех n.

Задача 7.32. Докажите, что если радиус сходимости ряда $\sum_{n=0}^{\infty} a_n x^n = f(x)$ положителен, то $f^{(n)}(0) = n! a_n$.

Теорема 7.18 (Абель). *Пусть ряд $\sum_{n=0}^{\infty} a_n$ сходится. При $|x| < 1$ положим $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Тогда $\lim_{x\to 1-} f(x) = \sum_{n=0}^{\infty} a_n$.*

Доказательство. Пусть $s_n = a_0 + a_1 + \cdots + a_n$ и $s_{-1} = 0$. Тогда

$$\sum_{n=0}^{m} a_n x^n = \sum_{n=0}^{m} (s_n - s_{n-1}) x^n = (1-x) \sum_{n=0}^{m-1} s_n x^n + s_m x^m.$$

Сходящаяся последовательность $\{s_m\}$ ограничена, поэтому при $|x| < 1$ получаем

$$f(x) = (1-x) \sum_{n=0}^{\infty} s_n x^n.$$

Пусть $s = \lim_{n\to\infty} s_n$. Для данного $\varepsilon > 0$ выберем N так, что при $n > N$ выполняется неравенство $|s - s_n| < \frac{\varepsilon}{2}$. Тогда, воспользовавшись соотношением $(1-x)\sum_{n=0}^{\infty} x^n = 1$ ($|x| < 1$), получаем

$$|f(x) - s| = \left| (1-x) \sum_{n=0}^{\infty} (s_n - s) x^n \right| \le (1-x) \sum_{n=0}^{N} |s_n - s| \cdot |x|^n + \frac{\varepsilon}{2}.$$

Если число $1 - x$ достаточно мало, то в последнем выражении первое слагаемое также меньше $\frac{\varepsilon}{2}$. $\square$

Историческое замечание. Нильс Генрик Абель (1802-1829) доказал теорему 7.18 в 1826 году.

Абель применил доказанную им предельную теорему для доказательства одного свойства произведения Коши. Напомним, что произведение Коши рядов $\sum_{i=0}^{\infty} a_i$ и $\sum_{j=0}^{\infty} b_j$ — это ряд $\sum_{k=0}^{\infty} c_k$, где $c_k = \sum_{i+j=k} a_i b_j$. Несложная проверка показывает, что произведение Коши степенных рядов $\sum a_k x^k$ и $\sum b_k x^k$ — это степенной ряд $\sum c_k x^k$, где $c_k = \sum_{i+j=k} a_i b_j$. Для степенных рядов произведение Коши получается формальным перемножением рядом и группировкой мономов одной степени.

Теорема 7.19 (Абель). *Если два ряда $\sum a_k$ и $\sum b_k$ сходятся и их произведение Коши $\sum c_k$ тоже сходится, то сумма произведения Коши этих рядов равно произведению их сумм.*

Доказательство. Сопоставим рядам $\sum a_k$, $\sum b_k$ и $\sum c_k$ степенные ряды $\sum a_k x^k = a(x)$, $\sum b_k x^k = b(x)$ и $\sum c_k x^k = c(x)$. Эти степенные ряды сходятся абсолютно при $0 \le x < 1$. Кроме того, произведение Коши степенных рядов $\sum a_k x^k$ и $\sum b_k x^k$ — это степенной ряд $\sum c_k x^k$. Поэтому, применив теорему Коши (теорема 7.12), получим, что при $0 \le x < 1$ выполняется соотношение $a(x)b(x) = c(x)$. Согласно предельной теореме Абеля $\lim_{x\to 1-} a(x) = \sum_{n=0}^{\infty} a_n$ и аналогичные равенства выполняются для функций $b(x)$ и $c(x)$. Таким образом, переходя к пределу $x \to 1-$ в равенстве $a(x)b(x) = c(x)$, получаем требуемое. $\square$

7.7. Ряд для логарифма

Покажем, что если $-1 < x < 1$, то

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots = \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n}.$$

Из тождества

$$\frac{1}{1+t} = 1 - t + t^2 - \cdots + (-1)^{n-1} t^{n-1} + \frac{(-1)^n t^n}{1+t}$$

следует, что если $x > -1$, то

$$\ln(1+x) = \int_0^x \frac{dt}{1+t} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1} \frac{x^n}{n} + (-1)^n R_n,$$

где $R_n = \int_0^x \frac{t^n}{1+t}dt$. Если $0 \le x \le 1$, то $0 \le R_n \le \int_0^x t^n dt = \frac{x^n}{n+1} \le \frac{1}{n+1} \to 0$ при $n \to \infty$. Если $-1 \le x < 0$, то сделаем замену $x = -y$. Тогда $R_n = (-1)^{n+1} \int_0^y \frac{t^n}{1-t}dt$, поэтому $|R_n| \le \frac{1}{1-y} \int_0^y t^n dt = \frac{y^n}{(1-y)(n+1)} \to 0$ при $n \to \infty$.

Радиус сходимости ряда $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}x^n}{n}$ равен 1.

Историческое замечание. Ряд для натурального логарифма предложил Николаус Меркатор (1620-1687) в 1668 году. Ряд из задачи 7.33 предложил Джеймс Грегори (1638-1675) в том же году.

Задача 7.33. Докажите, что если $-1 < x < 1$, то

$$\ln \frac{1+x}{1-x} = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \dots\right).$$

Задача 7.34. а) Докажите, что для любого натурального n

$$\ln(n+1) = \ln n + 2\left(\frac{1}{2n+1} + \frac{1}{3(2n+1)^3} + \frac{1}{5(2n+1)^5} + \dots\right).$$

б) Докажите, что

$$\ln 2 = \frac{2}{3}\left(1 + \frac{1}{3 \cdot 9} + \frac{1}{5 \cdot 9^2} + \frac{1}{7 \cdot 9^3} + \dots\right).$$

Задача 7.35. Пусть $H_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n}$. Докажите, что $\frac{1}{1-x} \ln\left(\frac{1}{1-x}\right) = \sum_{n=1}^{\infty} H_n x^n$ при $|x| < 1$.

Формула Стирлинга

Задача 7.36. *а) Докажите, что

$$1 < \left(n + \frac{1}{2}\right) \ln\left(1 + \frac{1}{n}\right) < 1 + \frac{1}{12n(n+1)}.$$

б) Докажите, что последовательность $a_n = \frac{n! e^n}{n^{n+\frac{1}{2}}}$ имеет (конечный) предел a.

в) Докажите, что $n! = a n^{n+\frac{1}{2}} e^{-n+\frac{\theta}{12n}}$ для некоторого θ между 0 и 1.

Задача 7.37. *Докажите, что число a из задачи 7.36 равно $\sqrt{2\pi}$, т.е.

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/12n}$$

(*формула Стирлинга*).

7.8. Бином Ньютона

Если $f(x) = (1+x)^{\alpha}$, где $\alpha \in \mathbb{R}$, то $f^{(n)}(x) = \alpha(\alpha-1)\dots(\alpha-n+1)(1+x)^{\alpha-n}$, поэтому $f^{(n)}(0) = \alpha(\alpha-1)\dots(\alpha-n+1)$ и

$$(1+x)^{\alpha} = 1 + \sum_{n=1}^{m} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n + R_m(x),$$

где $R_m(x)$ — остаточный член.

Если α — целое неотрицательное число, то степенной ряд $1 + \sum_{n=1}^{\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n = \sum_{n=0}^{\infty} a_n x^n$ вырождается в многочлен. Исследуем сходимость этого степенного ряда, когда число α не является целым неотрицательным; в таком случае $a_n \ne 0$ для всех n. Ясно, что $\frac{a_{n+1}x^{n+1}}{a_n x^n} = \frac{x(\alpha-n)}{n+1}$, поэтому ряд сходится при $|x| < 1$ и расходится при $|x| > 1$. На любом отрезке, содержащемся в интервале $(-1, 1)$ биномиальный ряд сходится равномерно к некоторой функции $g(x)$. Убедимся, что $g(x) = (1+x)^{\alpha}$.

Проверим сначала соотношение $(1+x)g'(x) = \alpha g(x)$. Проверка этого соотношения сводится к проверке следующего соотношения между коэффициентами биномиального ряда: $na_n + (n+1)a_{n+1} = \alpha a_n$, т.е. $\frac{a_{n+1}}{a_n} = \frac{\alpha - n}{n+1}$.

Из соотношения $(1+x)g'(x) = \alpha g(x)$ следует, что $\frac{d}{dx}\left(\frac{g(x)}{(1+x)^\alpha}\right) = 0$, поскольку

$$\begin{aligned} \frac{d}{dx}\left(\frac{g(x)}{(1+x)^\alpha}\right) &= \frac{g'(x)(1+x)^\alpha - \alpha g(x)(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} = \\ &= \frac{g'(x)(1+x) - \alpha g(x)}{(1+x)^{\alpha+1}}. \end{aligned}$$

Поэтому $g(x) = C(1+x)^\alpha$, где C — некоторая константа. Ясно также, что $g(0) = 1$. В итоге получаем $g(x) = (1+x)^\alpha$.

Например, при $\alpha = 1/2$ и при $\alpha = -1/2$ биномиальный ряд имеет вид

$$(1+x)^{1/2} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 - \frac{5}{128}x^4 + \frac{7}{256}x^5 - \cdots$$

и

$$(1+x)^{-1/2} = 1 - \frac{1}{2}x + \frac{3}{8}x^2 - \frac{5}{16}x^3 + \frac{35}{128}x^4 - \frac{63}{256}x^5 + \cdots$$

Историческое замечание. Исаак Ньютон (1643-1727) около 1665 года обнаружил формулу бинома для всех вещественных α. До него формула бинома была известна только для целых неотрицательных α. Первым доказал сходимость биномиального ряда при $|x| < 1$ Нильс Генрик Абель (1802-1829) в 1826 году.

7.9. Ряды для числа π

В этом параграфе мы приведём два самых знаменитых ряда для числа π. Отметим, что в 1739 году Эйлер обобщил теорему 7.21 и вычислил сумму $\sum \frac{1}{n^{2k}}$ (см. теорему 7.35).

Теорема 7.20 (Лейбниц, 1682). *Сумма ряда* $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + (-1)^n \frac{1}{2n+1}$ *равна* $\frac{\pi}{4}$.

Доказательство. Пусть $a_n = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots + (-1)^n \frac{1}{2n+1}$. Ясно, что $\lim_{n\to\infty}(a_{2k-1} - a_{2k}) = 0$, поэтому достаточно доказать, что $\frac{1}{2(4k+1)} \le \frac{\pi}{4} - a_{2k-1} \le \frac{1}{4k+1}$.

Из тождества $\frac{1-t^{4k}}{1+t^2} = 1 - t^2 + t^4 - t^6 + \cdots - t^{4k-2}$ следует, что

$$\begin{aligned} \frac{\pi}{4} = \operatorname{arctg} 1 &= \int_0^1 \frac{1}{1+t^2}dt = \int_0^1 \frac{1-t^{4k}}{1+t^2}dt + \int_0^1 \frac{t^{4k}}{1+t^2}dt = \\ &= \int_0^1 (1 - t^2 + t^4 - t^6 + \cdots - t^{4k-2})dt + R_k = \\ &= 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots - \frac{1}{4k-1} + R_k, \end{aligned}$$

где $R_k = \int_0^1 \frac{t^{4k}}{1+t^2}dt$. Но $\frac{1}{2}t^{4k} \le \frac{t^{4k}}{1+t^2} \le t^{4k}$ при $0 \le t \le 1$, поэтому $\frac{1}{2(4k+1)} \le R_k \le \frac{1}{4k+1}$. □

Теорема 7.21 (Эйлер, 1735). *Сумма ряда* $1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots$ *равна* $\frac{\pi^2}{6}$.

Доказательство. Сначала докажем, что числа $\operatorname{ctg}^2 \frac{\pi}{2n+1}$, $\operatorname{ctg}^2 \frac{2\pi}{2n+1}$, $\ldots$, $\operatorname{ctg}^2 \frac{n\pi}{2n+1}$ являются корнями многочлена

$$\binom{2n+1}{1}x^n - \binom{2n+1}{3}x^{n-1} + \binom{2n+1}{5}x^{n-2} - \cdots + (-1)^n. \tag{7.1}$$

Действительно, согласно формуле Муавра $\cos(2n+1)\alpha + i\sin(2n+1)\alpha = (\cos\alpha + i\sin\alpha)^{2n+1}$. Поэтому

$$\sin(2n+1)\alpha = $$

$$= \binom{2n+1}{1}\cos^{2n}\alpha\sin\alpha - \binom{2n+1}{3}\cos^{2n-2}\alpha\sin^3\alpha + \cdots + (-1)^n\sin^{2n+1}\alpha. \tag{7.2}$$

Из формулы (7.2) следует, что

$$\sin(2n+1)\alpha = \\ = \sin^{2n+1}\alpha\left(\binom{2n+1}{1}\operatorname{ctg}^{2n}\alpha - \binom{2n+1}{3}\operatorname{ctg}^{2n-2}\alpha + \dots + (-1)^n\right).$$

Поэтому числа $\pm\operatorname{ctg}\frac{\pi}{2n+1}, \ldots, \pm\operatorname{ctg}\frac{n\pi}{2n+1}$ являются корнями многочлена

$$\binom{2n+1}{1}y^{2n} - \binom{2n+1}{3}y^{2n-2} + \dots + (-1)^n.$$

После замены $x = y^2$ получаем требуемое.

Теперь мы можем вычислить следующие суммы:

$$\text{а) } \operatorname{ctg}^2\frac{\pi}{2n+1} + \operatorname{ctg}^2\frac{2\pi}{2n+1} + \dots + \operatorname{ctg}^2\frac{n\pi}{2n+1};$$
$$\text{б) } \frac{1}{\sin^2\frac{\pi}{2n+1}} + \frac{1}{\sin^2\frac{2\pi}{2n+1}} + \dots + \frac{1}{\sin^2\frac{n\pi}{2n+1}}.$$

Начнём с суммы а). Рассматриваемая сумма равна сумме корней многочлена (7.1), поэтому она равна

$$\binom{2n+1}{3}\Bigg/\binom{2n+1}{1} = \frac{n(2n-1)}{3}.$$

Тождество $\frac{1}{\sin^2\alpha} = 1 + \operatorname{ctg}^2\alpha$ показывает, что сумма б) равна $\frac{n(2n-1)}{3} + n = \frac{2n(n+1)}{3}$.

Теперь мы можем перейти непосредственно к доказательству равенства $1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \dots = \frac{\pi^2}{6}$. Напомним, что согласно задаче 3.4 $\operatorname{ctg}^2\alpha < \frac{1}{\alpha^2} < \frac{1}{\sin^2\alpha}$ для $0 < \alpha < \pi/2$. Просуммируем такие неравенства для $\alpha = \frac{\pi}{2n+1}, \frac{2\pi}{2n+1}, \ldots, \frac{n\pi}{2n+1}$. Воспользовавшись формулами для сумм а) и б), получим

$$\frac{n(2n-1)}{3} < \left(\frac{2n+1}{\pi}\right)^2\left(1 + \frac{1}{2^2} + \dots + \frac{1}{n^2}\right) < \frac{2n(n+1)}{3},$$

т.е.

$$\frac{\pi^2 n(2n-1)}{3(2n+1)^2} < 1 + \frac{1}{2^2} + \dots + \frac{1}{n^2} < \frac{\pi^2 2n(n+1)}{3(2n+1)^2}.$$

Остаётся заметить, что $\lim_{n\to\infty}\frac{n(2n-1)}{(2n+1)^2} = \lim_{n\to\infty}\frac{2n(n+1)}{(2n+1)^2} = \frac{1}{2}$. $\square$

Задача 7.38. *Докажите, что

$$(a_1 + a_2 + \dots + a_n)^2 < \frac{\pi^2}{6}(a_1^2 + 2^2a_2^2 + \dots + n^2a_n^2).$$

7.10. Производящие функции

В комбинаторике часто бывают полезны производящие функции различных последовательностей. *Производящая функция* последовательности $a_0, a_1, \ldots$ — это степенной ряд $\sum_{n=0}^{\infty} a_n x^n$. *Экспоненциальная производящая функция* последовательности $a_0, a_1, \ldots$ — это степенной ряд $\sum_{n=0}^{\infty} a_n\frac{x^n}{n!}$.

Например, согласно задаче 7.35 производящая функция последовательности $H_n = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n}$ равна $\frac{1}{1-x}\ln\left(\frac{1}{1-x}\right)$.

Производящие функции можно рассматривать не только для числовых последовательностей, но и для последовательностей функций. Вычислим производящие функции для многочленов Лежандра и для многочленов Эрмита. Во втором случае производящая функция экспоненциальная.

Рассмотрим биномиальный ряд для $(1+y)^{-1/2}$ и положим $y = -2xr + r^2$. Биномиальный ряд сходится при $|y| < 1$, поэтому при фиксированном x функцию $H(x,r) = \frac{1}{\sqrt{1-2xr+r^2}}$ можно разложить в степенной ряд по r, сходящийся при достаточно малых r:

$$H(x,r) = P_0(x) + P_1(x)r + P_2(x)r^2 + \ldots$$

В частности, $P_0(x) = 1$, $P_1(x) = x$ и $P_2(x) = \frac{1}{2}(3x^2 - 1)$.

Выражение $y^m = (r^2 - 2xr)^m$ состоит из мономов вида $r^{2k}(xr)^{m-k} = x^{m-k}r^{m+k}$. В таком мономе степень по x не превосходит степени по r, поэтому P_n — многочлен степени n. Кроме того, в каждом таком мономе степень по x имеет ту же чётность, что и степень по r, поэтому в P_n есть только мономы, степень которых имеет ту же чётность, что и n.

Ясно, что

$$\frac{\partial H(x,r)}{\partial r} = \frac{\partial}{\partial r}(1 - 2xr + r^2)^{-1/2} = (x - r)(1 - 2xr + r^2)^{-3/2},$$

поэтому

$$(1 - 2rx + r^2)\frac{\partial H}{\partial r} - (x - r)H = 0. \tag{7.3}$$

В выражении в левой части коэффициент при r^n равен

$$(n+1)P_{n+1}(x) - 2nxP_n(x) + (n-1)P_{n-1}(x) - xP_n(x) + P_{n-1}(x),$$

поэтому получаем следующее рекуррентное соотношение между многочленами P_n:

$$(n+1)P_{n+1}(x) - x(2n+1)P_n(x) + nP_{n-1}(x) = 0.$$

Оно совпадает с рекуррентным соотношением между многочленами Лежандра (см. с. 106). Многочлены $P_0(x) = 1$ и $P_1(x) = x$ тоже совпадают с многочленами Лежандра. Поэтому все многочлены P_n совпадают с многочленами Лежандра.

Сравнив производные $H(x,r)$ по r и по x, можно получить ещё одно соотношение между многочленами Лежандра. Действительно, дифференцируя $H(x,r)$ по x, получаем соотношение

$$(1 - 2rx + r^2)\frac{\partial H}{\partial x} - rH = 0.$$

Сравнивая это соотношение с соотношением (7.3), получаем

$$r\frac{\partial H}{\partial r} - (x - r)\frac{\partial H}{\partial x} = 0.$$

Вычислив коэффициент при r^n в выражении в левой части, приходим к соотношению

$$nP_n(x) - xP_n'(x) + P_{n-1}'(x) = 0.$$

Перейдём теперь к производящей функции для многочленов Эрмита. Рассмотрим функцию

$$H(x,t) = e^{2xt-t^2} = H_0(x) + H_1(x)t + \frac{1}{2!}H_2(x)t^2 + \frac{1}{3!}H_3(x)t^3 + \dots$$

Легко проверить, что $\frac{\partial H}{\partial t} = 2(x - t)H$. Из этого следует соотношение

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x).$$

Именно такому рекуррентному соотношению удовлетворяют многочлены Эрмита (см. с. 107). Многочлены $H_0(x) = 1$ и $H_2(x) = 2x$ тоже совпадают с многочленами Эрмита.

7.11. Двойные ряды

Пусть для любых натуральных i и j задано число a_{ij}. Положим $S_{m,n} = \sum_{i \le m,\, j \le n} a_{ij}$. *Двойной ряд* $\sum_{i,j} a_{ij}$ сходится к сумме S, если для любого $\varepsilon > 0$ можно выбрать натуральные числа M и N так, что $|S_{m,n} - S| < \varepsilon$ при $m > M$ и $n > N$.

Двойной ряд $\sum_{i,j} a_{ij}$ *сходится абсолютно*, если сходится двойной ряд $\sum_{i,j} |a_{ij}|$.

Будем считать, что $S_{m,n} =$, если $m \le 0$ или $n \le 0$. Тождество

$$a_{mn} = (S_{m,n} - S_{m,n-1}) - (S_{m-1,n} - S_{m-1,n-1}) \tag{7.4}$$

показывает, что если двойной ряд $\sum_{i,j} a_{ij}$ сходится, то для любого $\varepsilon > 0$ можно выбрать N так, что $|a_{mn}| < \varepsilon$, когда $m > N$ и $n > N$.

Задача 7.39. Докажите, что существует единственный двойной ряд с заданными суммами $S_{m,n}$.

Теорема 7.22. *Двойной ряд* $\sum_{i,j} a_{ij}$ *сходится тогда и только тогда, когда для любого* $\varepsilon > 0$ *можно выбрать* M *и* N *так, что*

$$|S_{m+i,n+j} - S_{m,n}| < \varepsilon,$$

когда $m > M$, $n > N$, *а* i *и* j *принимают произвольные неотрицательные значения.*

Доказательство. Воспользуемся критерием Коши сходимости последовательности. Необходимость указанного условия сходимости ряда непосредственно следует из критерия Коши. Чтобы доказать его достаточность, заметим, что $|S_{m+i,m+i} - S_{m,m}| < \varepsilon$ при $m > M + N$ и $i \geq 0$. Поэтому последовательность $\{S_{m,m}\}$ имеет предел; обозначим его S. Пусть $m > M$ и $n > N$. Устремим i и j к бесконечности так, чтобы выполнялось соотношение $m + i = n_j$. Тогда из неравенства $|S_{m+i,n+j} - S_{m,n}| < \varepsilon$ следует, что $|S - S_{m,n}| \leq \varepsilon$. Поэтому двойной ряд сходится. □

Теорему 7.22 можно применить для доказательства того, что абсолютно сходящийся двойной ряд сходится. Положим $T_{m,n} = \sum_{i\leq m,\, j\leq n} |a_{ij}|$. Из абсолютной сходимости двойного ряда $\sum_{i,j} a_{ij}$ следует, что для любого $\varepsilon > 0$ можно выбрать M и N так, что $T_{m+i,n+j} - T_{m,n} = |T_{m+i,n+j} - T_{m,n}| < \varepsilon$, когда $m > M$, $n > N$, $i \geq 0$ и $j \geq 0$. Поэтому при таких условиях выполняются неравенства $|S_{m+i,n+j} - S_{m,n}| \leq T_{m+i,n+j} - T_{m,n} < \varepsilon$, а значит, двойной ряд $\sum_{i,j} a_{ij}$ сходится.

Для двойного ряда $\sum_{i,j} a_{ij}$ можно рассмотреть *сумму по строкам*

$$\sum_{i=1}^{\infty}\left(\sum_{j=1}^{\infty} a_{ij}\right) = \lim_{m\to\infty}\lim_{n\to\infty} S_{m,n}$$

и *сумму по столбцам*

$$\sum_{j=1}^{\infty}\left(\sum_{i=1}^{\infty} a_{ij}\right) = \lim_{n\to\infty}\lim_{m\to\infty} S_{m,n}.$$

Пример 7.4. *Для двойного ряда с суммами* $S_{m,n} = \frac{m-n}{m+n}$ *сумма по строкам равна* -1, *сумма по столбцам равна* 1, *а сумма* S *не существует.*

Теорема 7.23 (Прингсгейм). *Если двойной ряд* $\sum_{i,j} a_{ij}$ *сходится к сумме* S *и существуют суммы по строкам и по столбцам, то обе эти суммы равны* S.

Доказательство. Выберем M так, что $|S_{m,n} - S| < \varepsilon$ при $m > M$ и $n > M$. Сумма по строкам существует, поэтому существует предел $\lim_{n\to\infty} S_{m,n}$ и $|\lim_{n\to\infty} S_{m,n} - S| < \varepsilon$ при $m > M$. Следовательно, сумма по строкам равна S. Для суммы по столбцам доказательство аналогично. □

Теорема 7.24. *Занумеруем члены абсолютно сходящегося двойного ряда* $\sum_{i,j} a_{ij}$ *в произвольном порядке и составим из них обычный ряд* $\sum a_i$. *Этот ряд сходится и его сумма не зависит от нумерации.*

Доказательство. Пусть сумма двойного ряда $\sum_{i,j} |a_{ij}|$ равна T. Положим $T_{m,n} = \sum_{i\leq m,\, j\leq n} |a_{ij}|$. Для заданного $\varepsilon > 0$ выберем m_0 и n_0 так, что $T - T_{m,n} < \frac{\varepsilon}{2}$ при $m > m_0$ и $n > n_0$. Пусть $M \geq m_0$ и $M \geq n_0$. Определим число N следующим образом: первые N членов ряда $\sum a_i$ содержат все слагаемые суммы $S_{M+1,M+1}$. Пусть сумма этих первых N членов равна S_N. Тогда разность $T_N - S_{M+1,M+1}$ состоит только из членов a_{ij}, для которых $i > m_0$ и $j > n_0$, поэтому

$$|T_N - S_{M+1,M+1}| \leq T - T_{M+1,M+1} < \frac{\varepsilon}{2}.$$

Разность $S - S_{M+1,M+1}$ тоже состоит только из членов a_{ij}, для которых $i > m_0$ и $j > n_0$, поэтому

$$|S - S_{M+1,M+1}| \leq T - T_{M+1,M+1} < \frac{\varepsilon}{2}.$$

Следовательно, $|S - T_N| < \varepsilon$. Для заданного $\varepsilon > 0$ мы выбрали N так, что $|S - T_n| < \varepsilon$ при $n \geq N$. □

Историческое замечание. Теорию двойных рядов разработал Альфред Прингсгейм (1850-1941) в 1897 году.

7.12. Подстановка ряда в ряд

Пусть ряд $\sum_{n=0}^{\infty} a_n x^n = f(x)$ имеет положительный радиус сходимости R. Положим $a_n^{(2)} = a_0 a_n + a_1 a_{n-1} + \dots + a_n a_0$. Тогда при $|x| < R$ ряд $\sum_{n=0}^{\infty} a_n^{(2)} x^n$ сходится, причём его сумма равна $f^2(x)$. Аналогично можно определить ряд $\sum_{n=0}^{\infty} a_n^{(m)} x^n$, который при $|x| < R$ сходится к $f^m(x)$.

Пусть также задан ряд $\sum_{m=0}^{\infty} b_m y^m = g(y)$, который имеет положительный радиус сходимости r, причём $|a_0| < r$. Тогда для достаточно малых x выполняется неравенство $|f(x)| < r$, поэтому можно рассмотреть композицию функций $g(f(x))$.

Подставим в ряд $\sum_{m=0}^{\infty} b_m y^m = g(y)$ вместо y^m $(m \geq 1)$ ряд $\sum_{n=0}^{\infty} a_n^{(m)} x^n$ и сгруппируем члены с одинаковыми степенями x. В результате получим некоторый ряд $\sum_{k=0}^{\infty} c_k x^k$.

Теорема 7.25. *Радиус сходимости ряда $\sum_{k=0}^{\infty} c_k x^k$ положителен и сумма этого ряда равна $g(f(x))$.*

Доказательство. По условию $|a_0| < r$, поэтому при достаточно малых x выполняется неравенство $\sum_{n=0}^{\infty} |a_n x^n| < r$. При таких x ряд

$$|b_0| + \sum_{m=1}^{\infty} |b_m| \left(\sum_{n=0}^{\infty} |a_n| \cdot |x^n| \right)^m$$

сходится. Пусть $\left(\sum_{n=0}^{\infty} |a_n| \cdot |x^n|\right)^m = \sum_{k=0}^{\infty} \gamma_k^{(m)} |x|^k$. Здесь числа $\gamma_k^{(m)}$ получаются из чисел $|a_n|$ по таким же формулам, как и числа $c_k^{(m)}$ из чисел a_n. В эти формулы входят некоторые суммы произведений, поэтому $|c_k^{(m)}| \leq \gamma_k^{(m)}$. Следовательно, ряд

$$|b_0| + \sum_{m=1}^{\infty} |b_m| \left(\sum_{k=0}^{\infty} |c_k^{(m)}| \cdot |x|^k \right)$$

тоже сходится. Абсолютно сходящийся двойной ряд сходится, поэтому ряд

$$b_0 + \sum_{m=1}^{\infty} b_m \left(\sum_{n=0}^{\infty} a_n x^n \right)^m = b_0 + \sum_{m=1}^{\infty} b_m \left(\sum_{k=0}^{\infty} c_k^{(m)} x^k \right)$$

сходится. □

Теорема 7.26. *Пусть ряд $\sum_{n=0}^{\infty} a_n x^n = f(x)$ имеет положительный радиус сходимости и $a_0 \neq 0$. Тогда функцию $g(x) = 1/f(x)$ можно представить рядом $\sum_{n=0}^{\infty} b_n x^n$, имеющим положительный радиус сходимости.*

Доказательство. Можно считать, что $a_0 = 1$. Пусть $\varphi(x) = \sum_{n=1}^{\infty} a_n x^n = f(x) - 1$. Требуется доказать, что радиус сходимости ряда, представляющего функцию $\frac{1}{1+\varphi(x)}$, положителен. Радиус сходимости ряда $\frac{1}{1+y} = 1 - y + y^2 - y^3 + \dots$ положителен (он равен 1). В этот ряд можно подставить $y = \sum_{n=1}^{\infty} a_n x^n$. Согласно теореме 7.25 при $\sum_{n=1}^{\infty} |a_n x^n| < 1$ в результате получим ряд с положительным радиусом сходимости. □

Пример 7.5. *Ряд, представляющий функцию $x \operatorname{ctg} x$, сходится при $|x| < \pi$.*

Доказательство. Если функция $f(x)$ бесконечно дифференцируема при $|x| < R$, то она представляется рядом, сходящимся при при $|x| < R$; это следует из представления остаточного члена ряды Тейлора как в форме Лагранжа, так и в интегральной форме. Бесконечная дифференцируемость функции $x \operatorname{ctg} x$ при $|x| < \pi$ и $x \neq 0$ очевидна. Бесконечная дифференцируемость функции $x \operatorname{ctg} x$ при $x = 0$ следует из того, что ряд, представляющий эту функцию, можно представить в виде отношения рядов $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$ и $\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \dots$ □

7.13. Экспонента в комплексной области

Будем обозначать множество комплексных чисел символом $\mathbb{C}$. Для любого комплексного числа $z \in \mathbb{C}$ можно определить $e^z \in \mathbb{C}$ как сумму ряда $1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots$.

Задача 7.40. Докажите, что этот ряд сходится для любого $z \in \mathbb{C}$.

Задача 7.41. Докажите, что $e^z e^w = e^{z+w}$ для любых комплексных z и w.

Задача 7.42. Докажите, что если x — вещественное число, то $e^{ix} = \cos x + i \sin x$ (*Эйлер*, 1743).

Задача 7.43. Докажите, что $e^{\pi i} = -1$ и $e^{2\pi i} = 1$.

Замечание. В вещественной области функция $f(x) = e^x$ при разных x принимает разные значения. В комплексной области это свойство уже не выполняется. Например, $f(0) = 1 = f(2\pi i)$.

Задача 7.44. Пусть $x \neq 2m\pi$ — вещественное число.
а) Докажите, что $\sum_{k=1}^n e^{ikx} = e^{ix}\frac{1-e^{inx}}{1-e^{ix}} = \frac{\sin(nx/2)}{\sin(x/2)} e^{i(n+1)x/2}$.
б) Докажите, что $\sum_{k=1}^n \sin kx = \frac{\sin\frac{nx}{2}\sin\frac{(n+1)x}{2}}{\sin\frac{x}{2}}$.
в) Докажите, что $\sum_{k=1}^n \cos kx = \frac{\sin\frac{nx}{2}\cos\frac{(n+1)x}{2}}{\sin\frac{x}{2}} = \frac{\sin\frac{(2n+1)x}{2}}{2\sin\frac{x}{2}} - \frac{1}{2}$.

Из формулы Эйлера следует, что $\cos x = \frac{e^{ix}+e^{-ix}}{2}$ и $\sin x = \frac{e^{ix}-e^{-ix}}{2i}$ для всех вещественных x. Определим косинус и синус в комплексной области по этим же формулам: $\cos z = \frac{e^{iz}+e^{-iz}}{2}$ и $\sin z = \frac{e^{iz}-e^{-iz}}{2i}$ для всех $z \in \mathbb{C}$. Положим также $\operatorname{tg} z = \frac{\sin z}{\cos z}$ при $\cos z \neq 0$ и $\operatorname{ctg} z = \frac{\cos z}{\sin z}$ при $\sin z \neq 0$. Тогда для всех вещественных x получаем следующую связь тригонометрических и гиперболических функций: $\sin(ix) = i \operatorname{sh} x$, $\cos(ix) = \operatorname{ch} x$, $\operatorname{tg}(ix) = i \operatorname{th} x$ и $\operatorname{ctg}(ix) = -i \operatorname{cth} x$.

Аналогично можно положить $\operatorname{sh} z = \frac{e^z - e^{-z}}{2}$ и $\operatorname{ch} z = \frac{e^z + e^{-z}}{2}$ для всех $z \in \mathbb{C}$. Тогда для всех вещественных x получим: $\operatorname{sh}(ix) = i \sin x$, $\operatorname{ch}(ix) = \cos x$, $\operatorname{th}(ix) = i \operatorname{tg} x$ и $\operatorname{cth}(ix) = -i \operatorname{ctg} x$.

7.14. Степенные ряды в комплексной области

Как это часто бывает, рассмотрение комплексных чисел часто позволяет прояснить явления, происходящие в вещественной области. Например, ряд $1 - x^2 + x^4 - x^6 + \ldots$ при $|x| < 1$ сходится к функции $f(x) = \frac{1}{1+x^2}$. При $x = \pm 1$ эта функция непрерывна, в этих точках у неё нет особенностей. С чем же тогда связана расходимость ряда при $x = \pm 1$? Оказывается, что она связана с особыми точками $z = \pm i$ функции $f(z) = \frac{1}{1+z^2}$. Кроме того, рассмотрение степенных рядов к комплексной области помогает доказывать сходимость вещественных степенных рядов.

Будем говорить, что степенной ряд $\sum_{n=0}^{\infty} a_n z^n$, где $a_n \in \mathbb{C}$ и $z \in \mathbb{C}$, *сходится абсолютно*, если ряд $\sum_{n=0}^{\infty} |a_n z^n|$ сходится. В таком случае из М-теста Вейерштрасса следует, что оба ряда, члены которых — вещественные и мнимые части чисел $a_n z^n$, сходятся. Поэтому ряд $\sum_{n=0}^{\infty} a_n z^n$ сходится. Точно так же, как и в вещественном случае, доказывается, что если степенной ряд $\sum_{n=0}^{\infty} a_n z^n$ сходится при $z = z_0$, то он абсолютно сходится при $|z| < |z_0|$. Радиус R сходимости ряда $\sum_{n=0}^{\infty} a_n z^n$ тоже определяется точно так же. В комплексном случае определяется *круг сходимости*, заданный неравенством $|z| < R$. Для степенного ряда вида $\sum_{n=0}^{\infty} a_n (z-a)^n$ круг сходимости задаётся неравенством $|z - a| < R$.

Покажем теперь, как выход в комплексную область помогает при доказательстве сходимости рядов.

Пример 7.6. *Ряд, представляющий функцию $\frac{x}{e^x - 1}$, сходится при $|x| < 2\pi$.*

Доказательство. Несложно проверить, что $\frac{x}{e^x-1} + \frac{x}{2} = \frac{x}{2} \operatorname{cth} \frac{x}{2}$. Но $\operatorname{ctg}(ix) = i \operatorname{cth} x$, поэтому радиус сходимости ряда $\frac{x}{2} \operatorname{cth} \frac{x}{2}$ равен радиусу сходимости ряда $\frac{x}{2} \operatorname{ctg} \frac{x}{2}$. Согласно примеру 7.5 радиус сходимости этого ряда равен 2π. $\square$

Выход в комплексную область позволяет также доказать единственность коэффициентов ряда Лорана. *Ряд Лорана* — это ряд функций вида $\sum_{n=-\infty}^{\infty} a_n x^n$. Этот ряд можно представить в виде суммы ряда $\sum_{n=0}^{\infty} a_n x^n$, сходящегося при $|x| < (\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|})^{-1} = R$, и ряда $\sum_{n=-\infty}^{-1} a_n x^n$, сходящегося при $|x| > \overline{\lim}_{n\to\infty} \sqrt[n]{|a_{-n}|} = r$. Предположим, что ряды Лорана $\sum_{n=-\infty}^{\infty} a_n x^n$ и $\sum_{n=-\infty}^{\infty} b_n x^n$ сходятся при $r < |x| < R$ к одной и той же функции $f(x)$. Тогда $a_n = b_n$ для всех n. Аналогичное утверждение для степенных рядов следует из того, что $a_n = \frac{f^{(n)}(0)}{n!}$. Для рядов Лорана значения функции f и её производных в точке 0 не определены, поэтому нужен какой-то другой подход к доказательству. Этот подход заключается в рассмотрении интегралов по окружности $|z| = c$, где $r < c < R$.

Для $z \in \mathbb{C}$ ряд $\sum_{n=-\infty}^{\infty} a_n z^n$ сходится при $r < |z| < R$, поэтому для таких z определена функция $f(z) = \sum_{n=-\infty}^{\infty} a_n x^n$. Представим точку окружности $|z| = c$ в виде $z = ce^{i\varphi}$ и рассмотрим интеграл

$$\int_0^{2\pi} z^m f(z)dz = \int_0^{2\pi} c^m e^{im\varphi} f(ce^{i\varphi})d(ce^{i\varphi}) = ic^{m+1}\int_0^{2\pi} e^{i(m+1)\varphi} f(ce^{i\varphi})d\varphi.$$

Если $r < r_1 < R_1 < R$, то функция $f(z)$ равномерно непрерывна при $r_1 \le |z| \le R_1$, поэтому функция $f(ce^{i\varphi})$ равномерно непрерывна по φ. Следовательно,

$$\begin{aligned} ic^{m+1}\int_0^{2\pi} e^{i(m+1)\varphi} f(ce^{i\varphi})d\varphi &= ic^{m+1}\int_0^{2\pi} e^{i(m+1)\varphi} \sum_{n=-\infty}^{\infty} a_n c^n e^{in\varphi} d\varphi = \\ &= ic^{m+1} \sum_{n=-\infty}^{\infty} a_n c^n \int_0^{2\pi} e^{i(m+n+1)\varphi} d\varphi = \\ &= 2\pi i c a_{-m-1}, \end{aligned}$$

поскольку $\int_0^{2\pi} e^{ip\varphi} d\varphi = 2\pi$ при $p = 0$, а для всех остальных значений p интеграл равен 0. В итоге приходим к следующему выражению коэффициентов ряда Лорана через функцию, которую этот ряд представляет:

$$a_{-m-1} = \frac{c^m}{2\pi}\int_0^{2\pi} e^{i(m+1)\varphi} f(ce^{i\varphi})d\varphi.$$

7.15. Числа и многочлены Бернулли

Согласно примеру 7.6 ряд по степеням t, представляющий функцию $g(z,t) = \frac{te^{tz}}{e^t - 1}$, сходится при $|t| < 2\pi$. Запишем этот ряд в виде

$$g(z,t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} B_n(z).$$

Как мы сейчас увидим, $B_n(z)$ — многочлен степени n. Многочлены $B_n(z)$ называют *многочленами Бернулли*, а числа $B_n = B_n(0)$ называют *числами Бернулли*.

Ряд для функции $g(z,t)$ представляет собой произведение рядов

$$g(0,t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} B_n \quad \text{и} \quad e^{tz} = \sum_{n=0}^{\infty} \frac{t^n z^n}{n!}.$$

Поэтому

$$\frac{B_n(z)}{n!} = \sum_{k=0}^{n} \frac{B_{n-k} z^k}{k!\,(n-k)!},$$

т.е.

$$B_n(z) = \sum_{k=0}^{n} \binom{n}{k} B_{n-k} z^k.$$

Формально это равенство можно записать в виде $B_n(z) = (B+z)^n$, где под B^{n-k} будет подразумеваться B_{n-k}.

Одно из важнейших свойств многочленов Бернулли заключается в том, что

$$B_n(z+1) - B_n(z) = nz^{n-1}. \tag{7.5}$$

Для доказательства формулы (7.5) достаточно заметить, что

$$\begin{aligned}\sum_{n=0}^{\infty}\big(B_n(z+1) - B_n(z)\big)\frac{t^n}{n!} &= g(t,z+1) - g(t,z) = \\ &= \frac{te^{t(z+1)}}{e^t-1} - \frac{te^{tz}}{e^t-1} = te^{tz} = \sum_{n=0}^{\infty}\frac{t^{n+1}z^n}{n!}.\end{aligned}$$

Сложим равенства (7.5) для $z = 0, 1, \ldots, m-1$. В результате получим

$$\sum_{k=0}^{m-1} k^{n-1} = \frac{1}{n}\big(B_n(m) - B_n(0)\big). \tag{7.6}$$

Это, в частности, означает, что сумма $1 + 2^{n-1} + \cdots + (m-1)^{n-1}$ представляет собой многочлен степени n от m.

Историческое замечание. Равенство (7.6) и выражение для суммы $1+2^{n-1}+\cdots+(m-1)^{n-1}$ обнаружил Якоб Бернулли (1654-1705); его исследования опубликованы посмертно, в 1713 году. Экспоненциальную производящую функцию $\frac{te^{tz}}{e^t-1} = \sum\limits_{n=0}^{\infty}\frac{t^n}{n!}B_n(z)$ предложил Леонард Эйлер (1707-1783) в 1738 году. Название *многочлены Бернулли* предложил Йозеф Людвиг Раабе (1801-1859) в 1851 году.

При вычислении многочленов Бернулли удобно пользоваться рекуррентными соотношениями

$$\sum_{r=0}^{n-1}\binom{n}{r}B_r(z) = nz^{n-1}, \quad n \geq 2. \tag{7.7}$$

Эти соотношения можно доказать следующим образом:

$$\sum_{n=0}^{\infty}\frac{t^n}{n!}B_n(z+1) = \frac{te^{tz}}{e^t-1}e^z = \left(\sum_{r=0}^{\infty}\frac{t^r}{r!}B_r(z)\right)\left(\sum_{s=0}^{\infty}\frac{z^s}{s!}\right),$$

поэтому $B_n(z+1) = \sum\limits_{r=0}^{n}\binom{n}{r}B_r(z) = B_n(z) + \sum\limits_{r=0}^{n-1}\binom{n}{r}B_r(z)$. Остается воспользоваться соотношением (7.5).

При $z = 0$ соотношения (7.7) превращаются в рекуррентные соотношения для чисел Бернулли

$$\sum_{r=0}^{n-1}\binom{n}{r}B_r = 0, \quad n \geq 2. \tag{7.8}$$

Легко проверить, что $B_0 = 1$, Поэтому из соотношений (7.8) последовательно получаем

$$B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_3 = 0, \quad B_4 = -\frac{1}{30}, \quad B_5 = 0, \quad \ldots$$

Несложно показать, что $B_{2k+1} = 0$ при $k \geq 1$. В самом деле,

$$\frac{t}{e^t-1} = 1 - \frac{t}{2} + \sum_{n=2}^{\infty}\frac{t^n}{n!}B_n,$$

поэтому достаточно проверить, что функция

$$\frac{t}{e^t-1} + \frac{t}{2} = t\frac{e^t+1}{e^t-1}$$

четна. В четности этой функции легко убедиться.

Многочлены Бернулли связаны соотношением

$$B'_{n+1}(z) = (n+1)B_n(z).$$

Чтобы доказать это соотношение, продифференцируем по z обе части равенства

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} B_n(z) = \frac{te^{tz}}{e^t - 1}.$$

В результате получим

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} B'_n(z) = \frac{t^2 e^{tz}}{e^t - 1} = \sum_{n=0}^{\infty} \frac{t^{n+1}}{n!} B_n(z).$$

Приравнивая коэффициенты при t^{n+1}, получаем требуемое.

Подставив $t = 2ix$ в равенство

$$\frac{t}{e^t - 1} = 1 - \frac{t}{2} + \sum_{m=2}^{\infty} B_m \frac{t^m}{m!},$$

получим

$$x \operatorname{ctg} x = x \frac{\cos x}{\sin x} = ix \frac{e^{ix} + e^{-ix}}{e^{ix} - e^{-ix}} = 1 + \sum_{n=1}^{\infty} (-1)^n B_{2n} \frac{(2x)^{2n}}{(2n)!} \tag{7.9}$$

при $|x| < \pi$.

Задача 7.45. Докажите, что при $|x| < \pi/2$

$$\operatorname{tg} x = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{2^{2n}(2^{2n} - 1)B_{2n}}{(2n)!} x^{2n-1}.$$

7.16. Бесконечные произведения

Пусть $a_1, a_2, \ldots$ — действительные числа, отличные от -1. Бесконечное произведение $\prod_{k=1}^{\infty}(1 + a_k)$ называют *сходящимся*, если существует (конечный) предел $\lim_{n\to\infty} \prod_{k=1}^{n}(1 + a_k)$, причём этот предел отличен от нуля.

Теорема 7.27. *Если $a_k \geq 0$, то бесконечное произведение $\prod(1 + a_k)$ сходится тогда и только тогда, когда сходится бесконечный ряд $\sum a_k$.*

Доказательство. Последовательность $p_n = \prod_{k=1}^{n}(1 + a_k)$ неубывающая, поэтому она сходится либо к конечному положительному числу, либо к $+\infty$. Несложно доказать, что

$$a_1 + \cdots + a_n \leq (1 + a_1) \cdots (1 + a_n) \leq e^{a_1 + \cdots + a_n}. \tag{1}$$

Первое неравенство доказывается раскрытием скобок, а второе неравенство следует из того, что $1 + a \leq e^a$ при $a \geq 0$. Неравенство (1) показывает, что $\prod(1 + a_k) = +\infty$ тогда и только тогда, когда $\sum a_k = +\infty$. $\square$

Теорема 7.28. *Если $0 \leq b_k < 1$, то бесконечное произведение $\prod(1 - b_k)$ сходится тогда и только тогда, когда сходится бесконечный ряд $\sum b_k$.*

Доказательство. Последовательность $p_n = \prod_{k=1}^{n}(1 - b_k)$ невозрастающая, поэтому она сходится либо к положительному числу, любо к нулю. Легко проверить, что если $0 \leq b_k < 1$, то $1 - b \leq e^{-b}$. Поэтому

$$(1 - b_1) \cdots (1 - b_n) \leq e^{-b_1 - \cdots - b_n}.$$

Это означает, что если ряд $\sum b_k$ расходится, то $\lim_{n\to\infty} p_n = 0$.

Предположим теперь, что ряд $\sum b_k$ сходится. Тогда существует такое натуральное число N, что $\sum_{k=N}^{\infty} b_k < 1/2$. Если $\alpha_1, \alpha_2 \geq 0$, то

$$(1-\alpha_1)(1-\alpha_2) \geq 1-\alpha_1-\alpha_2.$$

Пользуясь этим неравенством, индукцией по m легко показать, что если $\alpha_1, \ldots, \alpha_m \geq 0$, то

$$(1-\alpha_1)\cdots(1-\alpha_m) \geq 1-\alpha_1-\cdots-\alpha_m.$$

Таким образом, если $n > N$, то

$$\frac{p_n}{p_{N-1}} = (1-b_N)(1-b_{N+1})\cdots(1-b_n) \geq 1-b_N-\cdots-b_n > \frac{1}{2}.$$

Эти неравенства показывают, что последовательность p_n невозрастающая и $p_n > p_{N-1}/2 > 0$. Поэтому предел $\lim_{n\to\infty} p_n$ существует и не равен нулю. □

Бесконечное произведение $\prod(1+a_k)$ называют *абсолютно сходящимся*, если бесконечное произведение $\prod(1+|a_k|)$ сходится. Теорема 7.27 показывает, что произведение $\prod(1+a_k)$ абсолютно сходится тогда и только тогда, когда сходится ряд $\sum |a_k|$.

Теорема 7.29. *Абсолютно сходящееся бесконечное произведение сходится.*

Доказательство. Пусть $p_n = \prod_{k=1}^{n}(1+a_k)$ и $P_n = \prod_{k=1}^{n}(1+|a_k|)$. Тогда

$$|p_n - p_{n-1}| = |a_n| \prod_{k=1}^{n-1} |1+a_k| \leq |a_n| \prod_{k=1}^{n-1} (1+|a_k|) = P_n - P_{n-1}.$$

По условию существует предел $\lim_{n\to\infty} P_n$, поэтому ряд $\sum(P_n - P_{n-1})$ сходится. Но в таком случае должен сходиться и ряд $\sum(p_n - p_{n-1})$, поэтому существует предел $\lim_{n\to\infty} p_n$. Нужно лишь проверить, что этот предел отличен от нуля. Рассмотрим для этого ряд $\sum \left|\dfrac{a_k}{1+a_k}\right|$. Он сходится, поскольку сходится ряд $\sum |a_k|$ и последовательность $1+a_k$ имеет предел, равный 1. Следовательно, бесконечное произведение $\prod \left(1 - \dfrac{a_k}{1+a_k}\right)$ сходится абсолютно, поэтому, как только что было доказано, последовательность $q_n = \prod\limits_{k=1}^{n} \left(1 - \dfrac{a_k}{1+a_k}\right)$ имеет конечный предел. Но $q_n = p_n^{-1}$, поэтому $\lim_{n\to\infty} p_n \neq 0$. □

Теорема 7.30. *В абсолютно сходящемся бесконечном произведении можно менять порядок множителей; произведение при этом не изменяется.*

Доказательство. Покажем сначала, что для абсолютно сходящегося бесконечного произведения $\prod_{k=1}^{\infty}(1+a_k)$ для любого $\varepsilon > 0$ можно выбрать натуральное число N так, что произведение любого конечного набора множителей $(1+a_k)$, где номера k попарно различны и больше N, отличается от 1 меньше чем на ε. В самом деле, непосредственная проверка показывает, что

$$\begin{aligned} |(1+a_{k_1})\ldots(1+a_{k_m})-1| &\leq (1+|a_{k_1}|)\ldots(1+|a_{k_m}|)-1 < \\ &< e^{|a_{k_1}|+\cdots+|a_{k_m}|}-1. \end{aligned}$$

Ряд $\sum_{k=1}^{\infty} |a_k|$ сходящийся, поэтому можно выбрать N так, что $\sum_{k=N+1}^{\infty} |a_k| < \ln(1+\varepsilon)$.

Теперь можно доказать, что если ни один множитель абсолютно сходящегося произведения не равен нулю, то такое бесконечное произведение не равно нулю. В самом деле, выберем N так, что для любого $p > 1$ выполняется неравенство $\left|\prod_{k=N+1}^{N+p}(1+a_k) - 1\right| < \varepsilon$, где $0 < \varepsilon < 1$. Тогда $\left|\prod_{k=N+1}^{\infty}(1+a_k)\right| \geq 1-\varepsilon > 0$, поэтому и $\prod_{k=1}^{\infty}(1+a_k) \neq 0$.

После этих предварительных наблюдений перейдём непосредственно к доказательству теоремы. Пусть бесконечное произведение $P' = \prod_{k=1}^{\infty}(1+a'_k)$ получено перестановкой множителей абсолютно сходящегося бесконечного произведения $P = \prod_{k=1}^{\infty}(1+a_k)$. Новое бесконечное произведение тоже абсолютно сходящееся. Рассмотрим произведение P_n первых n множителей произведения P. Выберем m так, чтобы все эти n множителей входили в произведение P'_m. Тогда

$$\frac{P'_m}{P_n} = (1+a_{k_1})\dots(1+a_{k_m}),$$

где $k_i > n$. Поэтому для любого $\varepsilon > 0$ можно выбрать N так, что $\left|\frac{P'_m}{P_n} - 1\right| < \varepsilon$ при $n > N$. Ясно, что $m \to \infty$ при $n \to \infty$, поэтому $\lim_{n\to\infty} \frac{P'_m}{P_n} = \frac{P'}{P}$ и $P' = P$. $\square$

Пусть $\{f_k(x)\}$ — последовательность функций, определённых на некотором множестве $A \subset \mathbb{R}$. Будем говорить, что бесконечное произведение $\prod_{k=1}^{\infty}\Big(1+f_k(x)\Big)$ *равномерно сходится*, на множестве A, если на множестве A последовательность функций $p_n(x) = \prod_{k=1}^{n}\Big(1+f_k(x)\Big)$ равномерно сходится к некоторой функции $p(x)$, причём $p(x) \neq 0$ для всех $x \in A$.

Теорема 7.31. *Если ряд $\sum |f_k(x)|$ равномерно сходится на множестве A к ограниченной функции, то произведение $\prod\Big(1+f_k(x)\Big)$ равномерно сходится на множестве A.*

Доказательство. По условию существует такое число M, что $\sum_{k=1}^{\infty}|f_k(x)| \leq M$ при всех $x \in A$. В таком случае

$$(1+|f_1(x)|)\cdots(1+|f_n(x)|) \leq e^{|f_1(x)|+\dots+|f_n(x)|} \leq e^M.$$

Пусть $P_n(x) = \prod\limits_{k=1}^{n}(1+|f_k(x)|)$. Тогда

$$P_{n+1}(x) - P_n(x) = |f_{n+1}(x)|(1+|f_1(x)|)\cdots(1+|f_n(x)|) \leq |f_{n+1}(x)|e^M.$$

Следовательно, ряд $\sum\Big(P_{n+1}(x) - P_n(x)\Big)$ равномерно сходится на множестве A. Доказательство завершают те же самые рассуждения, которыми мы пользовались при доказательстве теоремы 7.29. $\square$

7.17. Эйлеровы разложения тригонометрических функций

Эйлер получил следующие разложения для функций $\operatorname{ctg} x$ и $\sin x$:

$$\begin{aligned} \operatorname{ctg} x &= \frac{1}{x} + \sum_{n=1}^{\infty}\frac{2x}{x^2 - n^2\pi^2} \text{ при } x \neq n\pi, \\ \sin x &= x\prod_{n=1}^{\infty}\left(1 - \frac{x^2}{n^2\pi^2}\right). \end{aligned}$$

Эти разложения связаны друг с другом, но мы приведём независимые доказательства обоих разложений. Для этого нам потребуются некоторые тригонометрические тождества.

Теорема 7.32. *При всех $x \in \mathbb{R}$ и $n \in \mathbb{N}$ выполняется равенство*

$$\sin nx = 2^{n-1}\prod_{k=0}^{n-1}\sin\left(x + \frac{k\pi}{n}\right). \tag{7.10}$$

Доказательство. Мы воспользуемся тем, что $\sin x = \dfrac{e^{ix} - e^{-ix}}{2i}$ и $\prod_{k=0}^{n-1}\left(t - e^{-\frac{2k\pi i}{n}}\right) = t^n - 1$. Ясно, что

$$\begin{aligned} \prod_{k=0}^{n-1}\sin\left(x + \frac{k\pi}{n}\right) &= \prod_{k=0}^{n-1}\left(\frac{e^{ix+\frac{k\pi i}{n}} - e^{-ix-\frac{k\pi i}{n}}}{2i}\right) = \\ &= \frac{1}{(2i)^{n-1}}\prod_{k=0}^{n-1}e^{-ix+\frac{k\pi i}{n}}\left(e^{2ix} - e^{-\frac{2k\pi i}{n}}\right). \end{aligned}$$

Остаётся заметить, что $\prod_{k=0}^{n-1}\left(e^{2ix}-e^{-\frac{2k\pi i}{n}}\right)=e^{2inx}-1$ и $e^{\frac{\pi i}{n}(1+\dots+(n-1))}=e^{\frac{(n-1)\pi i}{2}}=i^{n-1}$. □

Следствие 1. *Для любого натурального n выполняется равенство*

$$\sin\frac{\pi}{n}\sin\frac{2\pi}{n}\cdots\sin\frac{(n-1)\pi}{n}=\frac{n}{2^{n-1}}. \tag{7.11}$$

Доказательство. Разделим обе части равенства (7.10) на $\sin x$ и рассмотрим предел при $x\to 0$. □

Следствие 2. *Пусть $x\in\mathbb{R}$, $n=2m+1$ и $nx/\pi\notin\mathbb{Z}$. Тогда*

$$\operatorname{ctg} nx=(-1)^m\operatorname{ctg} x\operatorname{ctg}\left(x+\frac{\pi}{n}\right)\cdots\operatorname{ctg}\left(x+\frac{(n-1)\pi}{n}\right). \tag{7.12}$$

Доказательство. Ясно, что

$$\sin n\left(x+\frac{\pi}{2}\right)=\sin\left(nx+\frac{\pi}{2}+m\pi\right)=(-1)^m\cos nx.$$

Поэтому, заменив в формуле (7.10) x на $x+\frac{\pi}{2}$, получим

$$\cos nx=(-1)^m 2^{n-1}\cos x\cos\left(x+\frac{\pi}{n}\right)\cdots\cos\left(x+\frac{(n-1)\pi}{n}\right). \tag{7.13}$$

Поделив (7.13) на (7.10), получаем требуемое. □

Теперь уже можно перейти непосредственно к доказательству эйлеровых разложений.

Теорема 7.33 (эйлерово разложение для котангенса). *Для любого $x\in\mathbb{R}$, $x\neq 0$, $\pm\pi$, $\pm 2\pi,\dots$, выполняется равенство*

$$\operatorname{ctg} x=\frac{1}{x}+\sum_{n=1}^{\infty}\frac{2x}{x^2-n^2\pi^2},$$

причём ряд, стоящий в правой части, равномерно сходится на любом компактном множестве, не содержащем точек 0, $\pm\pi$, $\pm 2\pi,\dots$

Доказательство. Формулу (7.12) можно записать в виде

$$\operatorname{ctg} nx=(-1)^m\operatorname{ctg} x\prod_{k=-m}^{m}\operatorname{ctg}\left(x-\frac{k\pi}{n}\right).$$

При этом

$$\operatorname{ctg}\left(x-\frac{k\pi}{n}\right)=\frac{1+\operatorname{tg} x\operatorname{tg}(k\pi/n)}{\operatorname{tg} x-\operatorname{tg}(k\pi/n)}.$$

Таким образом, $\operatorname{ctg} nx=P(\operatorname{tg} x)/Q(\operatorname{tg} x)$, где P — многочлен степени $n-1$, а Q — многочлен степени n с корнями $\operatorname{tg}(k\pi/n)$. Рациональную функцию $R=P/Q$ можно записать в виде

$$R(u)=A(u)+\sum_{k=-m}^{m}\frac{c_k}{u-\operatorname{tg}(k\pi/n)},$$

где $A(u)$ — многочлен, $c_k\in\mathbb{C}$ (см. доказательство теоремы 6.3 на с. 92). В рассматриваемой ситуации $\deg P<\deg Q$, поэтому $A=0$. Константу c_k легко вычислить, переходя к пределу $x\to k\pi/n$:

$$\begin{aligned} c_k &= \lim_{x\to k\pi/n}\operatorname{ctg} nx\big(\operatorname{tg} x-\operatorname{tg}(k\pi/n)\big)= \\ &= \lim_{x\to k\pi/n}\frac{\cos nx\sin\big(x-(k\pi/n)\big)}{\sin nx\cos x\cos(k\pi/n)}= \\ &= \lim_{y=x-k\pi/n\to 0}\frac{\sin y}{\sin ny}\frac{1}{\cos^2(k\pi/n)}=\frac{1}{n\cos^2(k\pi/n)}. \end{aligned}$$

Таким образом,

$$\begin{aligned}
\operatorname{ctg} nx &= \frac{1}{n \operatorname{tg} x} + \frac{1}{n \cos^2 \frac{k\pi}{n}} \sum_{k=1}^{m} \left(\frac{1}{\operatorname{tg} x - \operatorname{tg} \frac{k\pi}{n}} + \frac{1}{\operatorname{tg} x + \operatorname{tg} \frac{k\pi}{n}} \right) = \\
&= \frac{1}{n \operatorname{tg} x} + \sum_{k=1}^{m} \frac{2n \operatorname{tg} x}{\left(n \cos^2 \frac{k\pi}{n} \operatorname{tg} x\right)^2 - \left(n \sin \frac{k\pi}{n}\right)^2}.
\end{aligned}$$

Заменив x на x/n, эту формулу можно записать в виде

$$\operatorname{ctg} x = \frac{1}{n \operatorname{tg} \frac{x}{n}} + \sum_{k=1}^{\infty} v_k(n,\, x),$$

где $v_k(n,\, x) = 0$ при $k > m$ и

$$v_k(n,\, x) = \frac{2n \operatorname{tg} \frac{x}{n}}{\left(n \cos^2 \frac{k\pi}{n} \operatorname{tg} \frac{x}{n}\right)^2 - \left(n \sin \frac{k\pi}{n}\right)^2}$$

при $1 \le k \le m$.

Пусть $K \subset \mathbb{R}$ — компактное множество, не содержащее точек $0,\ \pm\pi,\ \pm 2\pi, \ldots$ На множестве K последовательность функций $f_n(x) = n \operatorname{tg}(x/n)$ равномерно сходится к функции x. Поэтому существует такое число M, что $|n \operatorname{tg}(x/n)| \le M$ для всех $x \in K$. Пусть $0 \le x \le \pi/2 < \sqrt{3}$. Тогда $\frac{\sin x}{x} \ge 1 - \frac{x^2}{6} \ge \frac{1}{2}$. Поэтому если $1 \le k \le m$, то $n \sin \frac{k\pi}{n} \ge \frac{k\pi}{2}$. Таким образом, если k достаточно велико, а именно $k\pi/2 > M$, то $|v_k(n,\, x)| \le \dfrac{2M}{\frac{k^2\pi^2}{4} - M^2}$ для всех $x \in K$. Поэтому ряд $\sum_{k=1}^{\infty} v_k(n,\, x)$ равномерно сходится на K. Остаётся заметить, что при фиксированном k последовательность функций $g_n(x) = v_k(n,\, x)$ равномерно сходится на K к функции $\frac{2x}{x^2 - k^2\pi^2}$. □

Теорема 7.34 (эйлерово разложение для синуса). *Для любого $x \in \mathbb{R}$ выполняется равенство*

$$\sin x = x \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2\pi^2}\right),$$

причём бесконечное произведение, стоящее в правой части, сходится равномерно и абсолютно на любом компактном множестве, не содержащем точек $\pm\pi,\ \pm 2\pi, \ldots$

Доказательство. Воспользовавшись тождеством

$$\sin\left(x + \frac{(n-k)\pi}{n}\right) = -\sin\left(x - \frac{k\pi}{n}\right),$$

равенство (7.10) при $n = 2m+1$ можно записать в виде

$$\begin{aligned}
\sin nx &= (-1)^m 2^{n-1} \prod_{k=-m}^{m} \sin\left(x - \frac{k\pi}{n}\right) = \\
&= (-1)^m 2^{n-1} \sin x \prod_{k=1}^{m} \sin\left(x - \frac{k\pi}{n}\right) \sin\left(x + \frac{k\pi}{n}\right).
\end{aligned}$$

Легко проверить, что

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

Поэтому, воспользовавшись равенством (7.11), получаем

$$\sin nx = n \sin x \prod_{k=1}^{m} \left(1 - \frac{\sin^2 x}{\sin^2(k\pi/n)}\right).$$

Заменим теперь x на x/n:

$$\sin x = n \sin \frac{x}{n} \prod_{k=1}^{m} \left(1 - \frac{\sin^2(x/n)}{\sin^2(k\pi/n)} \right).$$

Эту формулу можно записать в виде

$$\sin x = n \sin \frac{x}{n} \prod_{k=1}^{m} \bigl(1 - w_k(n, x)\bigr),$$

где $w_k(n, x) = 0$ при $k > m$ и $w_k(n, x) = \dfrac{\sin^2 x}{\sin^2(k\pi/n)}$ при $k \le m$.

Пусть $K \subset \mathbb{R}$ — компактное множество, не содержащее точек $\pm\pi$, $\pm 2\pi, \ldots$ Последовательность функций $n \sin(x/n)$ равномерно сходится на K к функции x, поэтому существует такое число M, что $|n \sin(x/n)| \le M$ для всех $x \in K$. При доказательстве теоремы 7.33 было показано, что если $1 \le k \le m$, то $n \sin(k\pi/n) \ge k\pi/2$. Поэтому если k достаточно велико, а именно $k\pi/2 > M$, то $|w_k(n, x)| \le \frac{4M^2}{k^2\pi^2}$. Из сходимости ряда $\sum \frac{4M^2}{k^2\pi^2}$ следует, что для любого $\varepsilon > 0$ существует такое натуральное число $N(\varepsilon) = N$, что

$$1 - \varepsilon < \prod_{k=N}^{\infty} \bigl(1 - w_k(n, x)\bigr) < 1 + \varepsilon.$$

А из того, что $\lim_{n\to\infty} w_k(n, x) = \frac{x^2}{k^2\pi^2}$, следует, что

$$\lim_{n\to\infty} \prod_{k=1}^{N-1} \bigl(1 - w_k(n, x)\bigr) = \prod_{k=1}^{N-1} \left(1 - \frac{x^2}{k^2\pi^2}\right).$$

□

Ряд $\zeta(s) = \sum\limits_{n=1}^{\infty} \frac{1}{n^s}$ сходится при $s > 1$. Функцию $\zeta(s)$ называют *дзета-функцией Римана.*

Теорема 7.35 (Эйлер). *Если k — натуральное число, то*

$$\zeta(2k) = \frac{(-1)^{k+1} B_{2k} 2^{2k-1} \pi^{2k}}{(2k)!},$$

где B_{2k} — число Бернулли.

Доказательство. Воспользуемся разложением функции $\sin z$ в бесконечное произведение: $\sin z = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2\pi^2}\right)$. Продифференцировав логарифмы обеих частей этого равенства, получим

$$\frac{\cos z}{\sin z} = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2\pi^2},$$

т.е.

$$z \frac{\cos z}{\sin z} = 1 - 2 \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \left(\frac{z}{n\pi}\right)^{2k}. \tag{7.14}$$

Напомним, что

$$z \frac{\cos z}{\sin z} = 1 + \sum_{k=1}^{\infty} (-1)^k B_{2k} \frac{(2z)^{2k}}{(2k)!}. \tag{7.15}$$

при $|z| < \pi$ (см. с. 141). Сравнение коэффициентов при z^{2k} в (7.14) и (7.15) даёт требуемое равенство. □

Историческое замечание. Леонард Эйлер (1707-1783) получил указанные разложения функций $\operatorname{ctg} x$ и $\sin x$ в 1734 году. В 1735 году он вычислил сумму $\sum \frac{1}{n^2}$, а в 1739 году — сумму $\sum \frac{1}{n^{2k}}$.

7.18. Решения задач

7.1. Сравним данный ряд с расходящимся положительным рядом $\sum_{n=2}^{\infty}(\ln\ln(n+1)-\ln\ln n)$. Применим формулу конечных приращений к функции $f(x)=\ln\ln x$ на отрезке $[n,n+1]$: $f(n+1)-f(n)=f'(n+\theta)$, где $0<\theta<1$. При этом $f'(x)=\frac{1}{\ln x}\cdot\frac{1}{x}$. Таким образом,

$$\ln\ln(n+1)-\ln\ln n=f'(n+\theta)=\frac{1}{(n+\theta)\ln(n+\theta)}<\frac{1}{n\ln n},$$

т.е. члены данного ряда $\sum_{n=2}^{\infty}\frac{1}{n\ln n}$ больше членов расходящегося положительного ряда $\sum_{n=2}^{\infty}(\ln\ln(n+1)-\ln\ln n)$. Следовательно, данный ряд расходится.

7.2. Ясно, что $\sqrt[n]{r^n|\sin n\alpha|}=r\sqrt[n]{|\sin n\alpha|}\le r<1$. Поэтому по признаку Коши рассматриваемый ряд сходится.

Отношение последовательных членов этого ряда равно $r\left|\frac{\sin(n+1)\alpha}{\sin n\alpha}\right|$. При $\alpha\ne 0$ эта величина может принимать значения больше 1 при сколь угодно больших n. Поэтому применить признак Даламбера нельзя.

7.3. Ясно, что $\int_1^l\frac{dx}{x^s}$ стремится к пределу при $s>1$ и неограниченно возрастает при $s\le 1$.

7.4. Пусть $S_n=\sum_{i=1}^n a_i$ и $T_k=\sum_{i=1}^k 2^i a_{2^i}$. Тогда при $n<2^k$

$$\begin{aligned} S_n &\le a_1+(a_2+a_3)+\dots+(a_{2^k}+\dots+a_{2^{k+1}-1})\le\\ &\le a_1+2a_2+\dots+2^k a_{2^k}=T_k, \end{aligned}$$

а при $n>2^k$

$$\begin{aligned} S_n &\ge a_1+a_2+(a_3+a_4)+\dots+(a_{2^{k-1}-1}+\dots+a_{2^k})\ge\\ &\ge \frac{1}{2}a_1+a_2+2a_4+\dots+2^{k-1}a_{2^k}=\frac{1}{2}T_k. \end{aligned}$$

7.5. а) Если $a_n\le 10^k-1$, то $n\le 9^k-1$, поскольку на данных k местах могут стоять любые из девяти цифр, причём все цифры не могут быть одновременно нулями. Значит, если $n\ge 9^k$, то $a_n\ge 10^k-1$ и $a_{n+1}\ge 10^k$. Поэтому $\sum_{n=9^k+1}^{9^{k+1}}\frac{1}{a_k}\le\frac{9^{k+1}-9^k}{10^k}=8\left(\frac{9}{10}\right)^k$. Итак, рассматриваемая сумма не превосходит $8\sum_{k=0}^{\infty}\left(\frac{9}{10}\right)^k<\infty$.

б) В d-ичной системе счисления, где $d=10^{100}$, мы получаем задачу о числах, в d-ичной записи которых не встречается одна цифра. Решение этой задачи аналогично решению задачи а).

7.6. Те же самые рассуждения, что и при решении задачи 7.5 б), показывают, что любая последовательность цифр встречается в одном из чисел a_n (доказательство нужно слегка изменить для последовательности из одних нулей, но мы можем обойтись и без этой последовательности). Из этого легко выводится, что рассматриваемая десятичная дробь непериодична.

7.7. Если оба ряда $\sum p_n$ и $\sum q_n$ сходятся, то ряд $\sum|a_n|=\sum p_n+\sum q_n$ сходится. Если же один из рядов $\sum p_n$ и $\sum q_n$ сходится, а другой расходится, то ряд $\sum a_n$ расходится.

7.8. Применив признак Дирихле к ряду $\sum(-1)^n$ и последовательности чисел a_n, получаем требуемое.

7.9. Можно считать, что $\alpha\ne 2k\pi$, поскольку при $\alpha=2k\pi$ все члены ряда равны нулю. По известной формуле тригонометрии (её доказательство можно найти в решении задачи 7.44) при $\alpha\ne 2k\pi$

$$\sum_{i=1}^n\sin i\alpha=\frac{\sin\frac{n\alpha}{2}\sin\frac{(n+1)\alpha}{2}}{\sin\frac{\alpha}{2}},$$

поэтому абсолютная величина такой суммы не превосходит $\frac{1}{\left|\sin\frac{\alpha}{2}\right|}$. Таким образом, можно применить признак Дирихле и получить требуемое.

7.10. По известной формуле тригонометрии (её доказательство можно найти в решении задачи 7.44) при $\alpha \neq 2k\pi$

$$\sum_{i=1}^{n} \cos i\alpha = \frac{\sin \frac{(2n+1)\alpha}{2} - \sin \frac{\alpha}{2}}{2 \sin \frac{\alpha}{2}},$$

поэтому абсолютная величина такой суммы не превосходит $\frac{1}{\left|\sin \frac{\alpha}{2}\right|}$. Таким образом, можно применить признак Дирихле и получить требуемое.

7.11. Заменим α на $\alpha + \pi$. Ясно, что $\sin(n(\alpha+\pi)) = (-1)^n \sin n\alpha$ и $\cos(n(\alpha+\pi)) = (-1)^n \cos n\alpha$, поэтому можно воспользоваться результатами задач 7.9 и 7.10.

7.12. Ясно, что

$$a_i b_i = (A_i - A_{i-1}) b_i = A_i (b_i - b_{i+1}) - A_{i-1} b_i + A_i b_{i+1}.$$

Складывая такие равенства для $i = n+1,\ n+2,\ \ldots,\ n+k$, получаем требуемое.

7.13. а) Интеграл $\int_a^b f(x)\varphi(x)dx$ — это предел сумм вида $f(a)\varphi(a)(x_1 - a) + f(x_1)\varphi(x_1)(x_2 - x_1) + \ldots$. Каждая из таких сумм заключена между $I_{max} = \sum_i M_i \varphi(x_{i-1})(x_i - x_{i-1})$ и $I_{min} = \sum_i m_i \varphi(x_{i-1})(x_i - x_{i-1})$, где M_i и m_i — точная верхняя и точная нижняя грань функции $f(x)$ в интервале (x_{i-1}, x_i). При этом

$$I_{max} - I_{min} < \varphi(a) \sum_i (M_i - m_i)(x_i - x_{i-1}) \to 0,$$

поскольку функция $f(x)$ интегрируема. Таким образом, рассматриваемый интеграл равен общему пределу сумм I_{max} и I_{min} и, тем самым, пределу сумм $I = \sum_i \mu_i \varphi(x_{i-1})(x_i - x_{i-1})$, где μ_i — произвольное число, заключённое между M_i и m_i.

По теореме 6.5 (см. с. 96) число μ_i можно выбрать так, что $\mu_i(x_i - x_{i-1}) = \int_{x_{i-1}}^{x_i} f(x)dx$. Числа $\varphi(a), \varphi(x_1), \ldots$ положительны и убывают, поэтому по лемме Абеля сумма I заключена между $A\varphi(a)$ и $B\varphi(a)$, где A и B — максимум и минимум интеграла $\int_a^c f(x)dx$, когда c изменяется от a до b. Этот интеграл непрерывно зависит от c, поэтому

$$\int_a^b f(x)\varphi(x)dx = \varphi(a) \int_a^c f(x)dx$$

для некоторого $c \in (a, b)$.

б) Сначала применим утверждение задачи а) для функций $f(b-y)$ и $\varphi(b-y)$ на отрезке $[0, b-a]$:

$$\int_0^{b-a} f(b-y)\varphi(b-y)dy = \varphi(b) \int_0^c f(b-y)dy$$

для некоторого $c \in [0, b-a]$. Затем сделаем замену переменных $x = b - y$. В результате получим

$$\int_a^b f(x)\varphi(x)dx = \varphi(b) \int_d^b f(x)dx$$

для $d = b - c$.

в) Если функция $\varphi(x)$ возрастающая, то можно воспользоваться задачей а), заменив $\varphi(x)$ на $\Phi(x) = \varphi(b) - \varphi(x)$. Если функция $\varphi(x)$ убывающая, то можно воспользоваться задачей б), заменив $\varphi(x)$ на $\Phi(x) = \varphi(x) - \varphi(b)$.

7.14. Согласно задаче 5.37 б) $\frac{1}{k+1} < \ln(k+1) - \ln k < \frac{1}{k}$ для любого натурального k. Складывая такие неравенства для $k = 1,\ 2,\ \ldots,\ n$, получаем

$$H_{n+1} - 1 < \ln(n+1) < H_n.$$

7.15. Неравенство $\frac{1}{2n+1} > \frac{1}{2n}$ показывает, что если бы сходился ряд $\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots$, то сходился бы и ряд $\frac{1}{2}\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots\right)$. Но согласно задаче 7.14 этот ряд расходится.

7.16. а) О т в е т: $\ln 2$. Согласно задаче 5.37 б) $\ln\frac{k+1}{k} < \frac{1}{k} < \ln\frac{k}{k-1}$. Сложим такие неравенства для $k = n, n+1, \ldots, 2n$. В результате получим $\ln\frac{2n+1}{n} < a_n < \ln\frac{2n}{n-1}$. Значит, $\lim_{n\to\infty} a_n = \ln 2$.

Замечание. Другое доказательство можно найти в решении задачи 6.46.

б) Ясно, что $b_{n+1} = b_n + \frac{1}{2n+1} - \frac{1}{2n+2}$ и $a_{n+1} = a_n - \frac{1}{n+1} + \frac{1}{2n+1} + \frac{1}{2n+2}$. Значит, $b_{n+1} - a_{n+1} = b_n - a_n + \frac{1}{n+1} - \frac{2}{2n+2} = b_n - a_n$. Равенство $b_1 = a_1$ очевидно, поэтому $b_n = a_n$ для всех n.

7.17. а) Заметим, что $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$. Поэтому сумма ряда из задачи 7.16 б) равна сумме рассматриваемого ряда.

б) Из равенства $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$ следует, что $\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \frac{1}{4\cdot 5} + \frac{1}{5\cdot 6} + \cdots = 1$, т.е. $\frac{1}{2\cdot 3} + \frac{1}{4\cdot 5} + \frac{1}{6\cdot 7} + \cdots + \ln 2 = 1$.

7.18. Достаточно доказать, что сумма первых $3n$ членов стремится к $\frac{1}{2}\ln 2$. Эта сумма равна

$$1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1} - \frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \cdots - \frac{1}{4n} =$$
$$= H_{2n} - \frac{1}{2}H_n - \frac{1}{2}H_{2n} = \frac{1}{2}(H_{2n-1} - H_n) + \frac{1}{4n}.$$

Согласно задаче 7.16 а) $H_{2n-1} - H_n \to \ln 2$ при $n \to \infty$.

7.19. Легко проверить, что

$$a_n = H_{4n-1} - H_{2n} - \frac{1}{2}(H_{2n-1} - H_n).$$

Согласно задаче 7.16 а) $H_{4n-1} - H_{2n} \to \ln 2$ и $H_{2n-1} - H_n \to \ln 2$.

7.20. Легко проверить, что члены ряда имеют вид

$$\frac{1}{2}\left(\frac{1}{4n-3} + \frac{1}{4n-1} - \frac{1}{2n-1}\right), \quad n = 1, 2, \ldots$$

Поэтому удвоенная сумма первых n членов ряда равна $\frac{1}{2n+1} + \frac{1}{2n+2} + \frac{1}{2n+3} + \cdots + \frac{1}{4n-1}$. Остаётся воспользоваться результатом задачи 7.19.

7.21. Заметим, что

$$\frac{1}{(n-1)n(n+1)} = \frac{1}{2}\left(\frac{1}{n-1} + \frac{1}{n+1}\right) - \frac{1}{n}.$$

Далее, сумма членов $\frac{1}{2}\left(1 + \frac{1}{3}\right) - \frac{1}{2}$, $\frac{1}{2}\left(\frac{1}{3} + \frac{1}{5}\right) - \frac{1}{4}$, $\ldots$, $\frac{1}{2}\left(\frac{1}{2n-1} + \frac{1}{2n+1}\right) - \frac{1}{2n}$ равна

$$-\frac{1}{2} + H_{2n-1} + \frac{1}{2n+1} - \frac{1}{2}H_n - \frac{1}{2}H_n.$$

Остаётся заметить, что $H_{2n-1} - H_n \to \ln 2$ (задача 7.16 а).

7.22. а) Для $x > 0$ выполняется неравенство $1 + x < e^x$, поэтому $\sqrt[n]{A_n} \le e^{c_n}$, где $c_n = \frac{a}{n}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)$. Воспользовавшись неравенствами из задачи 7.14 и тем, что $\lim_{n\to\infty}\frac{\ln n}{n} = 0$, получим неравенство $\lim_{n\to\infty}\sqrt[n]{A_n} \le 1$. Неравенство $\lim_{n\to\infty}\sqrt[n]{A_n} \ge 1$ очевидно.

б) Решение в основном такое же, как и для задачи а), но требуются некоторые изменения и дополнения. Здесь можно воспользоваться неравенством $1 - x \le e^{-x}$ для $0 \le x < 1$. Но этим неравенством можно воспользоваться только для $k > b$. Поэтому множители с номерами, не превосходящими b, нужно заменить на единицы и отбросить соответствующий им начальный участок гармонического ряда. На пределы это не окажет влияния. (Здесь, конечно, важно, что число b не целое. Если b целое, то $B_n = 0$ при $n \ge b$.)

7.23. Положим $f(x) = \sum_{k=1}^n \left(\frac{x^k}{k} - \binom{n}{k}\frac{(x-1)^k}{k}\right)$. Ясно, что $f(1) = H_n$. Поэтому достаточно доказать, что $f'(x) = 0$ для всех x. Ясно, что $f'(x) = \sum_{k=1}^n x^{k-1} - \sum_{k=1}^n \binom{n}{k}(x-1)^{k-1}$. Пусть $g(x) = \sum_{k=1}^n \binom{n}{k}(x-1)^{k-1}$. Тогда $(x-1)g(x) + 1 = (x-1+1)^n$, поэтому $g(x) = \frac{x^n-1}{x-1} = \sum_{k=1}^n x^{k-1}$, что и требовалось.

7.24. а) Ясно, что $\left(1-\frac{1}{p}\right)^{-1}=1+\frac{1}{p}+\frac{1}{p^2}+\frac{1}{p^3}+\dots$ Поэтому рассматриваемое произведения является суммой чисел $\frac{1}{n}$, где n делится по крайней мере на одно из чисел 2, 3, ..., p_m (или $n=1$). Все числа $\frac{1}{n}$, где $1\le n\le p_m$, в эту сумму входят.

б) Согласно задаче 5.37 б) $\ln\left(1+\frac{1}{n}\right)<\frac{1}{n}$. Поэтому $\sum_{n=1}^{p_m}\ln\left(1+\frac{1}{n}\right)<\sum_{n=1}^{p_m}\frac{1}{n}<\prod_{k=1}^{m}\left(1-\frac{1}{p_k}\right)^{-1}=\prod_{k=1}^{m}\frac{p_k}{p_k-1}=\prod_{k=1}^{m}\left(1+\frac{1}{p_k-1}\right)$. Ясно также, что $\sum_{n=1}^{p_m}\ln\left(1+\frac{1}{n}\right)=\sum_{n=1}^{p_m}\big(\ln(n+1)-\ln n\big)=\ln(p_m+1)$. Поэтому $\ln\ln p_m<\ln\ln(p_m+1)<\sum_{k=1}^{m}\ln\left(1+\frac{1}{p_k-1}\right)<\sum_{k=1}^{m}\frac{1}{p_k-1}$. Далее, $\frac{1}{p_k-1}=\frac{1}{p_k}+\frac{1}{(p_k-1)p_k}$. Поэтому $\sum_{k=1}^{m}\frac{1}{p_k-1}<\sum_{k=1}^{m}\frac{1}{p_k}+\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\dots$ Остаётся заметить, что $\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\dots=1$, поскольку $\frac{1}{n(n+1)}=\frac{1}{n}-\frac{1}{n+1}$.

7.25. Согласно задаче 7.24 ряд $\sum_{n=1}^{\infty}1/p_n$, где p_1, p_2, ...— последовательные простые числа, расходится. Поэтому можно воспользоваться задачей 7.6.

7.26. Сделаем замену переменной $x=1-t$. В результате получим $\int_0^1\frac{1-(1-t)^n}{t}dt=\int_0^1\frac{1-x^n}{1-x}dx$. Ясно также, что $\frac{1-x^n}{1-x}=1+x+x^2+\dots+x^{n-1}$ и $\int_0^1 x^k dx=\frac{1}{k+1}$.

7.27. Первое решение. Рассмотрим последовательность

$$u_m=\frac{1}{m}-\ln\frac{m+1}{m}=\int_0^1\left(\frac{1}{m}-\frac{1}{m+t}\right)dt.$$

Ясно, что если $0\le t\le 1$, то

$$0\le\frac{1}{m}-\frac{1}{m+t}=\frac{t}{m(m+t)}\le\frac{1}{m^2},$$

поэтому $0\le u_m\le 1/m^2$. Следовательно, ряд $\sum_{m=1}^{\infty}u_m$ сходится (теорема 7.21). Но

$$\sum_{m=1}^{\infty}u_m=1+\frac{1}{2}+\dots+\frac{1}{n}-\ln(n+1)=1+\frac{1}{2}+\dots+\frac{1}{n}-\ln n+\ln\left(\frac{n}{n+1}\right).$$

Ясно также, что $\lim_{n\to\infty}\ln\left(\frac{n}{n+1}\right)=0$.

Второе решение. Нам потребуются неравенства $\ln\frac{n}{n-1}<\frac{1}{n}<\ln\frac{n+1}{n}$ для $n>1$. Они следуют из равенств $\frac{1}{n}-\ln\frac{n+1}{n}=\int_0^1\left(\frac{1}{n}-\frac{1}{n+t}\right)dt$ и $\frac{1}{n}-\ln\frac{n}{n-1}=\int_0^1\left(\frac{1}{n}-\frac{1}{n-1+t}\right)dt$.

Из этих неравенств следует, что последовательность $a_n=1+\frac{1}{2}+\dots+\frac{1}{n}-\ln(n+1)$ возрастающая, последовательность $b_n=1+\frac{1}{2}+\dots+\frac{1}{n}-\ln n$ убывающая и $0<b_n-a_n<\frac{1}{n}$. Поэтому обе последовательности сходятся к одному и тому же пределу.

7.28. Из равенства $\sqrt{n+1}-\sqrt{n}=\frac{1}{\sqrt{n+1}+\sqrt{n}}$ следуют неравенства

$$\frac{1}{2\sqrt{n+1}}<\sqrt{n+1}-\sqrt{n}<\frac{1}{2\sqrt{n}}.$$

Положим $a_n=1+\frac{1}{\sqrt{2}}+\dots+\frac{1}{\sqrt{n}}-2\sqrt{n+1}$ и $b_n=1+\frac{1}{\sqrt{2}}+\dots+\frac{1}{\sqrt{n}}-2\sqrt{n}$. Из доказанных выше неравенств следует, что последовательность a_n возрастающая, последовательность b_n убывающая и $0<b_n-a_n<\frac{1}{\sqrt{n}}$. Поэтому обе последовательности сходятся к одному и тому же пределу.

7.29. Произведение первых n членов равно

$$\frac{e^{1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{n}}}{n+1}=\frac{n}{n+1}e^{1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{n}-\ln n}.$$

Ясно также, что $\lim_{n\to\infty}\frac{n}{n+1}=1$, $\lim_{n\to\infty}\left(1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{n}-\ln n\right)=\gamma$ и e^x — непрерывная функция.

7.30. Прежде всего заметим, что постоянная Эйлера равна также пределу $\lim_{n\to\infty}\left(1+\frac{1}{2}+\dots+\frac{1}{n}-\ln(n+1)\right)$, поскольку $\lim_{n\to\infty}\ln\frac{n+1}{n}=0$.

Пусть $x>1$. Из равенств

$$\begin{aligned}\frac{1}{x-1} &= \int_1^\infty \frac{dt}{t^x}=\sum_{n=1}^\infty\int_n^{n+1}\frac{dt}{t^x},\\ \ln(n+1) &= \int_1^{n+1}\frac{dt}{t}=\sum_{k=1}^n\int_k^{k+1}\frac{dt}{t}\end{aligned}$$

следует, что

$$\begin{aligned}\lim_{x\to1+0}\left(\sum_{n=1}^\infty\frac{1}{n^x}-\frac{1}{x-1}\right) &= \lim_{x\to1+0}\sum_{n=1}^\infty\left(\frac{1}{n^x}-\int_n^{n+1}\frac{dt}{t^x}\right),\\ \gamma &= \sum_{n=1}^\infty\left(\frac{1}{n}-\int_n^{n+1}\frac{dt}{t}\right).\end{aligned}$$

Ряд для γ получается из предыдущего ряда предельным переходом при $x\to1+0$, поэтому достаточно проверить равномерную сходимость ряда при $x\in[1,2]$. Применим для этого М-тест Вейерштрасса. Ясно, что при $x\in[1,2]$

$$\begin{aligned}0 &< \frac{1}{n^x}-\int_n^{n+1}\frac{dt}{t^x}=\int_n^{n+1}\left(\frac{1}{n^x}-\frac{1}{t^x}\right)dt=\int_n^{n+1}\left(\int_n^t xu^{-x-1}du\right)dt\le\\ &\le xn^{-x-1}\int_n^{n+1}\left(\int_n^t du\right)dt=\frac{1}{2}xn^{-x-1}\le n^{-2}.\end{aligned}$$

7.31. а) Радиус сходимости ряда

$$f_{n+1}(x)=\frac{f(x)-(a_0+a_1x+\dots+a_nx^n)}{x^{n+1}}=a_{n+1}+a_{n+2}x+\dots$$

такой же, как у исходного ряда, поэтому он положителен. Следовательно, функция $f_{n+1}(x)$ непрерывна в точке 0 и $f_{n+1}(0)=a_{n+1}$.

б) Функции f и g непрерывны в точке 0, поэтому $a_0=f(0)=\lim_{n\to\infty}f(x_n)=\lim_{n\to\infty}g(x_n)=g(0)=b_0$. Такое же рассуждение можно применить к функции $f_{n+1}(x)=\frac{f(x)-(a_0+a_1x+\dots+a_nx^n)}{x^{n+1}}$ для $n=0,1,\dots$

7.32. Ряды $f'(x)=\sum_{n=1}^\infty na_nx^{n-1}$, $f''(x)=\sum_{n=2}^\infty n(n-1)a_nx^{n-2},\dots$ имеют тот же самый радиус сходимости, что и ряд $\sum_{n=0}^\infty a_nx^n=f(x)$. Поэтому $f'(0)=1\cdot a_1$, $f''(0)=1\cdot2\cdot a_2,\dots$

7.33. Предел разности двух сходящихся последовательностей равен разности их пределов, поэтому сходящиеся ряды можно почленно вычитать. Вычитая почленно из ряда

$$x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\frac{x^5}{5}-\dots=\ln(1+x)$$

ряд

$$-x-\frac{x^2}{2}-\frac{x^3}{3}-\frac{x^4}{4}-\frac{x^5}{5}-\dots=\ln(1-x),$$

получаем требуемое.

7.34. а) Рассмотрим ряд из задачи 7.33 для $x=\frac{1}{2n+1}$. Поскольку $\frac{1+x}{1-x}=\frac{n+1}{n}$, получаем требуемое.

б) Непосредственно следует из а) при $n=1$, поскольку $\ln1=0$.

7.35. Радиус сходимости ряда $\sum_{n=1}^\infty H_nx^n$ равен 1, поскольку $\lim_{n\to\infty}\sqrt[n]{\ln n}=1$. Проверим, что $-\ln(1-x)=(1-x)\sum_{n=1}^\infty H_nx^n$ при $|x|<1$. При $|x|<1$ получаем:

$$\begin{aligned}-\ln(1-x) &= x+\frac{1}{2}x^2+\frac{1}{3}x^3+\dots=\\ &= H_1x+(H_2-H_1)x^2+(H_3-H_2)x^3+\dots=\\ &= \sum_{n=1}^\infty H_nx^n-\sum_{n=1}^\infty H_nx^{n+1}=(1-x)\sum_{n=1}^\infty H_nx^n.\end{aligned}$$

7.36. а) Формулу из задачи 7.34 а) можно переписать в виде

$$\left(n+\frac{1}{2}\right)\ln\left(1+\frac{1}{n}\right)=1+\frac{1}{3(2n+1)^2}+\frac{1}{5(2n+1)^4}+\dots$$

Ясно, что сумма этого ряда меньше

$$1+\frac{1}{3}\left(\frac{1}{(2n+1)^2}+\frac{1}{(2n+1)^4}+\dots\right)=1+\frac{1}{12n(n+1)}.$$

б) и в) Неравенство из задачи а) можно переписать в виде

$$e<\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}}<e^{1+\frac{1}{12n(n+1)}}. \qquad (1)$$

А так как

$$\frac{a_n}{a_{n+1}}=\frac{n!e^n}{n^{n+\frac{1}{2}}}\cdot\frac{(n+1)^{n+1+\frac{1}{2}}}{(n+1)!e^{n+1}}=\frac{\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}}}{e},$$

то из неравенств (1) получаем

$$1<\frac{a_n}{a_{n+1}}<e^{\frac{1}{12n(n+1)}}.$$

Таким образом, $a_ne^{-\frac{1}{12n}}<a_{n+1}e^{-\frac{1}{12(n+1)}}$ и $a_n>a_{n+1}$. Поэтому последовательность $\{a_ne^{-\frac{1}{12n}}\}$ возрастает, а последовательность $\{a_n\}$ убывает. При этом $a_1e^{-\frac{1}{12}}<a_ne^{-\frac{1}{12n}}<a_n<a_1$, поэтому обе последовательности имеют предел. Поскольку $e^{-\frac{1}{12n}}\to 1$, эти пределы равны одному и тому же числу a. Ясно, что $a_ne^{-\frac{1}{12n}}<a<a_n$, т.е. $a=a_ne^{-\frac{\theta}{12n}}$ для некоторого θ между 0 и 1. Таким образом, $a=\frac{n!e^n}{n^{n+\frac{1}{2}}}e^{-\frac{\theta}{12n}}$.

7.37. Формулу Валлиса (задача 6.19) можно переписать следующим образом: $\sqrt{\pi}=\lim_{n\to\infty}\frac{1}{\sqrt{n}}\frac{(2n)!!}{(2n-1)!!}$. Далее,

$$\frac{(2n)!!}{(2n-1)!!}=\frac{2\cdot4\cdot\ldots\cdot2n}{1\cdot3\cdot\ldots\cdot(2n-1)}=\frac{(2\cdot4\cdot\ldots\cdot2n)^2}{(2n)!}=\frac{(n!)^22^{2n}}{(2n)!}.$$

Выразим $n!$ и $(2n)!$ как $a_nn^{n+\frac{1}{2}}e^{-n}$ и $a_{2n}(2n)^{2n+\frac{1}{2}}e^{-2n}$ и подставим эти выражения в формулу Валлиса:

$$\sqrt{\pi}=\lim_{n\to\infty}\frac{1}{\sqrt{n}}\frac{a_n^2n^{2n+1}e^{-2n}2^{2n}}{a_{2n}n^{2n+\frac{1}{2}}e^{-2n}2^{2n+\frac{1}{2}}}=\lim_{n\to\infty}\frac{a_n^2}{a_{2n}\sqrt{2}}=\frac{a}{\sqrt{2}}.$$

В результате получим $a=\sqrt{2\pi}$.

7.38. Согласно неравенству Коши (задача 1.8)

$$\begin{aligned}(a_1+a_2+\dots+a_n)^2&=\left(a_1+2a_2\frac{1}{2}+\dots+na_n\frac{1}{n}\right)^2\leq\\&\leq\left(1+\frac{1}{2^2}+\dots+\frac{1}{n^2}\right)(a_1^2+2^2a_2^2+\dots+n^2a_n^2).\end{aligned}$$

Кроме того, $1+\frac{1}{2^2}+\dots+\frac{1}{n^2}<\frac{\pi^2}{6}$ (теорема 7.21).

7.39. Единственность двойного ряда с заданными суммами $S_{m,n}$ следует из тождества (7.4). Существование этого двойного ряда следует из того же самого тождества, но при этом нужно проверить, что если a_{mn} задать тождеством (7.4), то $\sum_{i\leq m,\, j\leq n}a_{ij}=S_{m,n}$. Положим $\sum_{i\leq m,\, j\leq n}S_{ij}=W_{m,n}$. Тогда

$$\begin{aligned}\sum_{i\leq m,\, j\leq n}a_{ij}&=\sum_{i\leq m,\, j\leq n}(S_{i,j}-S_{i,j-1}-S_{i-1,j}+S_{i-1,j-1})=\\&=W_{m,n}-W_{m,n-1}-W_{m-1,n}+W_{m-1,n-1}=S_{m,n}.\end{aligned}$$

7.40. Пусть $s_n = \sum_{k=0}^{n} \frac{z^k}{k!}$ и $m < n$. Ясно, что $|s_n - s_m| \le \sum_{k=m+1}^{n} \frac{|z|^k}{k!}$. Ряд $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$, где $x = |z|$, сходится. Поэтому, воспользовавшись критерием Коши, получаем, что исходный ряд тоже сходится.

7.41. Пусть $f_n(z) = \sum_{k=0}^{n} \frac{z^k}{k!}$. Равенство

$$\sum_{k+l=m} \frac{z^k}{k!} \cdot \frac{w^l}{l!} = \frac{1}{m!} \sum_{k+l=m} \binom{m}{k} z^k w^l = \frac{(z+w)^m}{m!}$$

показывает, что выражение для $f_n(z)f_n(w)$ содержит все слагаемые из выражения для $f_n(z+w)$ и ещё некоторые слагаемые из выражения для $\big(e^z - f_n(z)\big)\big(e^w - f_n(w)\big)$. Поэтому

$$|f_n(z)f_n(w) - f_n(z+w)| \le \big(e^{|z|} - f_n(|z|)\big)\big(e^{|w|} - f_n(|w|)\big).$$

Выражение в правой части стремится к нулю при $n \to \infty$, поэтому $e^z e^w = \lim_{n\to\infty} f_n(z)f_n(w) = \lim_{n\to\infty} f_n(z+w) = e^{z+w}$.

7.42. По определению

$$\begin{aligned} e^{ix} &= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} - \frac{x^6}{6!} - i\frac{x^7}{7!} + \dots = \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots\right) = \\ &= \cos x + i \sin x. \end{aligned}$$

7.43. Непосредственно следует из задачи 7.42.

7.44. а) Ясно, что $(1 - e^{ix}) \sum_{k=1}^{n} e^{ikx} = \sum_{k=1}^{n} (e^{ikx} - e^{i(k+1)x}) = e^{ix} - e^{i(n+1)x}$. Это доказывает первое равенство. Второе равенство доказывается следующим образом:

$$e^{ix} \frac{1 - e^{inx}}{1 - e^{ix}} = e^{ix} \frac{e^{-inx/2}(1 - e^{inx})}{e^{-ix/2}(1 - e^{ix})} e^{-i(n-1)x/2} = \frac{e^{inx/2} - e^{-inx/2}}{e^{ix/2} - e^{-ix/2}} e^{i(n+1)x/2}.$$

б) и в) непосредственно следуют из а).

7.45. Легко проверить, что $\operatorname{tg} x = \operatorname{ctg} x - 2 \operatorname{ctg} 2x$, поскольку $2 \operatorname{ctg} 2x = \frac{\cos^2 x - \sin^2 x}{\cos x \cdot \sin x} = \operatorname{ctg} x - \operatorname{tg} x$. Кроме того,

$$\operatorname{ctg} x - 2 \operatorname{ctg} 2x = -\sum_{n=1}^{\infty} (-1)^n B_{2n} \frac{(2x)^{2n-1}}{(2n)!} + \sum_{n=1}^{\infty} (-1)^n B_{2n} \frac{(4x)^{2n-1}}{(2n)!}.$$

Глава 8.

Мера Лебега. Интеграл Лебега

8.1. Множества меры нуль

Множество точек прямой имеет *меру нуль* (или является *множеством меры нуль*), если для любого $\varepsilon > 0$ существует не более чем счётный набор интервалов, который покрывает это множество, и сумма длин интервалов этого набора не превосходит ε.

Если некоторое свойство выполняется для всех точек, кроме точек множества меры нуль, то про такое свойство говорят, что оно выполняется *почти всюду*.

Пример 8.1. *Множество рациональных чисел имеет меру нуль.*

Доказательство. Занумеруем рациональные числа и покроем первое из них интервалом длины $\frac{\varepsilon}{2}$, второе интервалом длины $\frac{\varepsilon}{4}$, третье интервалом длины $\frac{\varepsilon}{8}$ и т.д. Сумма длин этих интервалов равна ε. $\square$

Задача 8.1. Докажите, что объединение счётного множества множеств меры нуль имеет меру нуль.

Задача 8.2. Докажите, что канторово множество — множество меры нуль.

С канторовым множеством тесно связана *функция Кантора*, которую Георг Кантор (1845-1918) описал в 1884 году. Функция Кантора $c\colon [0,1] \to [0,1]$ постоянна на всех интервалах, которые выбрасываются при построении канторова множества. Эти интервалы соответствуют числам x, которые имеют троичную запись с цифрой 1. В таких числах мы доходим до первой цифры 1 и все цифры после неё заменяем нулями (цифра 1 и предыдущие цифры при этом не изменяются); после этого каждую цифру 2 заменяем цифрой 1. Если же троичная запись числа x не содержит цифр 1, то заменяем каждую цифру 2 цифрой 1. Полученная в результате запись содержит только цифры 0 и 1; мы интерпретируем её как двоичную запись некоторого числа y, и полагаем $c(x) = y$.

Функция Кантора непрерывна и обладает многими удивительными свойствами: 1) эта функция не постоянна, хотя её производная почти всюду равна нулю; 2) эта функция имеет ограниченную вариацию, но не является абсолютно непрерывной; 3) эта функция отображает множество меры нуль (канторово множество) на весь отрезок.

Напомним, что множество первой категории — это множество, которое можно представить в виде счётного объединения нигде не плотных множеств. В некотором смысле это множество очень маленькое. Множество меры нуль тоже в каком-то смысле очень маленькое. Следующий пример показывает, что множество, маленькое в одном смысле, может быть весьма большим в другом смысле.

Пример 8.2. *Прямую можно представить в виде объединения множества первой категории и множества меры нуль.*

Доказательство. Занумеруем все рациональные числа: a_1, a_2, ... Пусть I_{ij} — это открытый интервал длины 2^{-i-j}, середина которого — точка a_i. Положим $U_j = \bigcup_{i=1}^{\infty} I_{ij}$ и $A = \bigcap_{j=1}^{\infty} U_j$. Для любого $\varepsilon > 0$ можно выбрать j так, что $2^{-j} < \varepsilon$. Тогда $A \subset \bigcup_{i=1}^{\infty} I_{ij}$ и $\sum_{i=1}^{\infty} |I_{ij}| = \sum_{i=1}^{\infty} 2^{-i-j} = 2^{-j} < \varepsilon$. Поэтому множество A имеет меру нуль.

Каждое множество U_j — открытое всюду плотное множество, поскольку оно содержит все рациональные числа и является объединением открытых множеств. Дополнение CU_j открытого всюду плотного множества U_j нигде не плотно (задача 4.36). Поэтому множество $CA = \bigcup_{j=1}^{\infty} CU_j$ является множеством первой категории. □

8.2. Критерий Лебега интегрируемости по Риману

Теорема 8.1 (Лебег). *Ограниченная функция f на отрезке $[a, b]$ интегрируема по Риману тогда и только тогда, когда множество точек разрыва имеет меру нуль.*

Доказательство. Пусть $\omega_f(x)$ — колебание функции f в точке x, X_ε — множество точек $x \in [a, b]$, для которых $\omega_f(x) \geq \varepsilon$. Точки разрыва — это те точки, в которых колебание положительно, поэтому множество точек разрыва — это объединение множеств $X_{1/n}$ для всех натуральных n.

Предположим сначала, что множество точек разрыва — это не множество меры нуль. Тогда одно из множеств $X_{1/n}$ — не множество меры нуль (см. задачу 8.1). Поэтому можно выбрать число $c > 0$ так, что для любого счётного покрытия множества $X_{1/n}$ интервалами сумма длин этих интервалов не меньше c. Для любого конечного покрытия множества $X_{1/n}$ интервалами сумма длин этих интервалов тоже не меньше c. Это верно и для множества $X_{1/n}$ без нескольких точек, поскольку несколько точек можно покрыть конечным набором интервалов сколь угодно малой длины.

Рассмотрим произвольное разбиение отрезка $[a, b]$ и возьмём только те отрезки $[x_k, x_{k+1}]$, для которых интервал (x_k, x_{k+1}) содержит хотя бы одну точку множества $X_{1/n}$. Выбранные интервалы (x_k, x_{k+1}) покрывают всё множество $X_{1/n}$, за исключением, возможно, нескольких точек — концов отрезков разбиения. Поэтому сумма длин выбранных отрезков $[x_k, x_{k+1}]$ не меньше c и на каждом из этих отрезков разность между точной верхней и точной нижней гранями функции f не меньше $1/n$. Таким образом, для любого разбиения отрезка $[a, b]$ разность между верхней и нижней интегральными суммами не меньше c/n, поэтому функция f не интегрируема по Риману.

Предположим теперь, что множество точек разрыва — множество меры нуль. Тогда любое множество $X_{1/n}$ — множество меры нуль. Покроем его счётным множеством интервалов, сумма длин которых не превосходит некоторого числа d. Множество $X_{1/n}$ компактно (оно замкнуто и ограничено), поэтому можно выбрать конечное множество интервалов, которые покрывают $X_{1/n}$ и сумма их длин не превосходит $1/n$. Пересечение отрезка $[a, b]$ и объединение этих интервалов — открытое в $[a, b]$ множество $A_{1/n}$. Дополнение множества $A_{1/n}$ в $[a, b]$ — замкнутое множество $B_{1/n}$; оно состоит из конечного числа отрезков. Для каждой точки $x \in B_{1/n}$ выполняется неравенство $\omega_f(x) < 1/n$, поэтому точку x можно покрыть интервалом (α, β) так, что на отрезке $[\alpha, \beta]$ колебание функции f (т.е. разность между точной верхней и точной нижней гранями) меньше $1/n$. Из этого покрытия компактного множества $B_{1/n}$ можно выбрать конечное подпокрытие. Рассмотрим разбиение отрезка $[a, b]$ концами интервалов покрытия, лежащих в $B_{1/n}$, и концами интервалов, которые покрывают $X_{1/n}$. Это разбиение зависит от двух параметров (n и d), и есть два вида отрезков этого разбиения. Отрезки первого типа лежат в $B_{1/n}$, а внутренности отрезков второго типа лежат в $A_{1/n}$.

Посмотрим, какой вклад в разность между верхней и нижней интегральными суммами могут дать отрезки каждого вида. Для отрезков первого типа колебание функции не превосходит $1/n$, а сумма длин отрезков не превосходит $b - a$, поэтому их вклад не превосходит $(b - a)/n$. Для отрезков второго типа колебание функции не превосходит $M - m$, где M и m — точная верхняя и точная нижняя грани функции f, а сумма длин отрезков не превосходит d, поэтому их вклад не превосходит $d(M - m)$. Поэтому для любого $\varepsilon > 0$ можно выбрать достаточно малые числа d и $1/n$ так, чтобы для соответствующего им разбиения отрезка $[a, b]$ разность между верхней и нижней интегральными суммами была меньше ε. □

Пример 8.3. *Положим $f(x) = 0$ для иррационального x и $f(x) = 1/n$ для несократимой дроби $x = m/n$. Функция $f(x)$ разрывна во всех рациональных точках, но при этом интегрируема по Риману: $\int_0^1 f(x)dx = 0$.*

Доказательство. Функция $f(x)$ разрывна в точке $x = m/n$, поскольку $f(x) = 1/n$, но в иррациональных точках, стремящихся к x, функция равна 0. Докажем теперь, что функция $f(x)$ непрерывна в любой иррациональной точке x. Пусть задано $\varepsilon > 0$. Выберем натуральное число $n > 1/\varepsilon$. Пусть $\delta > 0$ — это наименьшее из расстояний от точки x до рациональных точек со знаменателями $1, 2, \ldots, n$. Тогда если $|x - y| < \delta$, то $|f(x) - f(y)| = |f(y)| < \frac{1}{n} < \varepsilon$. Таким образом, точки разрыва функции $f(x)$ — это в точности рациональные точки. Множество точек разрыва имеет меру нуль, следовательно, функция $f(x)$ интегрируема по Риману. Поэтому достаточно найти предел некоторой последовательности интегральных сумм при измельчении разбиений. Ясно, что на любом отрезке есть иррациональная точка, поэтому для любого разбиения есть интегральная сумма, равная нулю. □

Историческое замечание. Функцию из примера 8.3 первым предложил Карл Йоханнес Тома (1840-1921) в 1875 году, модифицировав функцию Дирихле. Поэтому название *функция Тома* предпочтительнее довольно широко распространённого ошибочного названия *функция Римана*.

8.3. Мера Жордана и мера Лебега на прямой

Пусть M — ограниченное множество на прямой. *Внешняя мера Жордана* этого множества — это точная нижняя грань сумм длин конечных наборов отрезков, покрывающих M. Вместо отрезков можно брать интервалы или полуинтервалы. Действительно, каждый отрезок можно покрыть интервалом, длина которого сколь угодно мало отличается от длины интервала. Кроме того, можно считать, что отрезки покрытия не перекрываются (не имеют общих внутренних точек). Действительно, два перекрывающихся отрезка можно заменить на три неперекрывающихся отрезка с меньшей суммой длин, и т.д.

Внутренняя мера Жордана множества M — это точная верхняя грань сумм длин конечных наборов неперекрывающихся отрезков, содержащихся в M. Если множество M не содержит ни одного отрезка, то его внутренняя мера считается равной 0. Ясно, что внутренняя мера Жордана не превосходит внешней.

Задача 8.3. Докажите, что внешняя мера Жордана множества рациональных точек отрезка $[a, b]$ равна $b - a$, а внутренняя мера Жордана равна 0.

Задача 8.4. Докажите, что внутренняя мера Жордана множества M — это разность между длиной отрезка $[a, b]$, содержащего M, и внешней мерой Жордана множества $[a, b] \setminus M$.

Множество *измеримо по Жордану*, если его внешняя и внутренняя меры Жордана равны. *Мера Жордана* в этом случае равна общему значению внешней и внутренней мер Жордана. Множество рациональных точек отрезка не измеримо по Жордану (задача 8.3).

Задача 8.5. Докажите, что множество M измеримо по Жордану тогда и только тогда, когда его характеристическая функция χ_M интегрируема по Риману.

Задача 8.6. Докажите, что множество измеримо по Жордану тогда и только тогда, когда мера Лебега его границы равна 0.

Внешняя мера Лебега определяется аналогично внешней мере Жордана, только вместо конечных наборов отрезков берутся счётные наборы. А именно, *внешняя мера Лебега* $m^*(M)$ множества M — это точная нижняя грань сумм длин счётных наборов отрезков, покрывающих M. Сумма длин при этом может быть бесконечной. Если сумма длин любого набора отрезков, покрывающих M, бесконечна, то $m^*(M) = \infty$.

Вместо отрезков можно брать интервалы или полуинтервалы. Действительно, каждый отрезок можно покрыть интервалом, длина которого сколь угодно мало отличается от длины интервала; кроме того, для любого $\varepsilon > 0$ можно выбрать бесконечную последовательность положительных чисел так, чтобы их сумма была меньше ε.

Внешняя мера Лебега множества M равна 0 тогда и только тогда, когда M — множество меры нуль (см. с. 155).

Пример 8.4. *Внешняя мера Лебега отрезка* $[a, b]$ *равна* $b - a$.

Доказательство. Покажем, что отрезок $[a, b]$ нельзя покрыть интервалами, сумма длин которых меньше $b - a$. Предположим, что отрезок $[a, b]$ покрыт интервалами I_n. Пусть (a_1, b_1) — первый из этих интервалов, содержащий точку a. Если $b_1 \le b$, то пусть (a_2, b_2) — первый из интервалов, содержащий точку b_1, и т.д. Допустим, что построенная в результате последовательность интервалов бесконечна. Тогда возрастающая последовательность b_n сходится к некоторой точке $x \le b$. Эта точка принадлежит некоторому интервалу I_k. Если $b_n \in I_k$, то интервал (a_{n+1}, b_{n+1}) имеет номер не больше k. Но все точки b_n с достаточно большими номерами принадлежат I_k. Следовательно, построенная последовательность интервалов конечна. Получено противоречие.

Таким образом, отрезок $[a,b]$ покрыт интервалами (a_1,b_1), (a_2,b_2), ..., (a_N,b_N), поэтому

$$\begin{aligned} b-a &< b_N - a_1 = \\ &= (b_N - b_{N-1}) + (b_{N-1} - b_{N-2}) + \dots + (b_2 - b_1) + (b_1 - a_1) \le \\ &\le (b_N - a_N) + (b_{N-1} - a_{N-1}) + \dots + (b_2 - a_2) + (b_1 - a_1). \end{aligned}$$

□

Теорема 8.2. *Для любого счётного семейства множеств A_k выполняется неравенство*

$$m^*(\cup A_k) \le \sum m^*(A_k).$$

Доказательство. Если $m^*(A_k) = \infty$ для некоторого множества A_k, то требуемое неравенство очевидно. Поэтому будем считать, что $m^*(A_k) < \infty$ для всех A_k. Пусть задано $\varepsilon > 0$. Множество A_k можно покрыть счётным набором интервалов $I_{k,i}$, сумма длин которых меньше $m^*(A_k) + \varepsilon/2^k$. Объединение всех этих интервалов счётно, покрывает множество $\cup A_k$, и сумма длин этих интервалов меньше $\varepsilon + \sum m^*(A_k)$. □

Будем говорить, что множество A *измеримо по Лебегу*, если для любого множества S выполняется равенство

$$m^*(S) = m^*(S \cap A) + m^*(S \setminus A).$$

В таком случае *мера Лебега* $m(A)$ равна $m^*(A)$.

Историческое замечание. Такой подход к определению меры Лебега предложил Константин Каратеодори (1873-1950). Меру Лебега ввёл Анри Лебег (1875-1941) в 1901 году.

Множество $S \setminus A$ можно записать как $S \cap CA$, поэтому определение измеримости для множества A и для его дополнения CA одно и то же. Таким образом, множество измеримо тогда и только тогда, когда измеримо его дополнение.

Если множество A измеримо и множество B не пересекается с A, то

$$m^*(A \cup B) = m^*((A \cup B) \cap A) + m^*((A \cup B) \setminus A) = m^*(A) + m^*(B).$$

Для любого множества S выполняется равенство $A = (S \cap A) \cup (S \setminus A)$, поэтому согласно теореме 8.2

$$m^*(S) \le m^*(S \cap A) + m^*(S \setminus A).$$

Таким образом, множество A измеримо по Лебегу тогда и только тогда, когда для любого множества S выполняется неравенство

$$m^*(S) \ge m^*(S \cap A) + m^*(S \setminus A).$$

Теорема 8.3. *Если $m^*(A) = 0$, то множество A измеримо по Лебегу.*

Доказательство. Для произвольного множества S из включений $S \cap A \subset A$ и $S \setminus A \subset S$ следует, что $m^*(S \cap A) \le m^*(A) = 0$ и $m^*(S \setminus A) \le m^*(S)$, поэтому

$$m^*(S) \ge m^*(S \setminus A) = 0 + m^*(S \setminus A) = m^*(S \cap A) + m^*(S \setminus A).$$

□

Теорема 8.4. *Объединение двух измеримых по Лебегу множеств измеримо.*

Доказательство. Пусть S — произвольное множество, A_1 и A_2 — измеримые множества. Тогда

$$\begin{aligned} m^*(S) &= m^*(S\cap A_1)+m^*(S\setminus A_1)= \\ &= m^*(S\cap A_1)+m^*((S\setminus A_1)\cap A_2)+m^*((S\setminus A_1)\setminus A_2). \end{aligned}$$

Поэтому из равенств $(S\cap A_1)\cup(S\setminus A_1)\cap A_2)=S\cap(A_1\cup A_2)$ и $(S\setminus A_1)\setminus A_2=S\setminus(A_1\cup A_2)$ следует, что

$$m^*(S)\geq m^*(S\cap(A_1\cup A_2))+m^*(S\setminus(A_1\cup A_2)).$$

Поэтому множество $A_1\cup A_2$ измеримо. $\square$

Следствие. *Объединение конечного числа измеримых по Лебегу множеств множеств измеримо.*

Формулы $C(A\cup B)=CA\cup CB$ и $A\setminus B=A\cap CB$ показывают, что если множества A и B измеримы, то множества $A\cup B$ и $A\setminus B$ тоже измеримы.

Теорема 8.5. *Пусть S — произвольное множество, а A_1, …, A_n — попарно не пересекающиеся множества, измеримые по Лебегу. Тогда*

$$m^*\left(S\cap\left(\bigcup_{k=1}^n A_k\right)\right)=\sum_{k=1}^n m^*(S\cap A_k).$$

Доказательство. Применим индукцию по n. При $n=1$ утверждение очевидно. Предположим, что оно верно при $n-1$. Множества A_1, …, A_n попарно не пересекающиеся, поэтому

$$S\cap\left(\bigcup_{k=1}^n A_k\right)\cap A_n=S\cap A_n$$

и

$$\left(S\cap\left(\bigcup_{k=1}^n A_k\right)\right)\setminus A_n=S\cap\left(\bigcup_{k=1}^{n-1} A_k\right).$$

Поэтому, воспользовавшись измеримостью множества A_n и предположением индукции, получим

$$\begin{aligned} m^*\left(S\cap\left(\bigcup_{k=1}^n A_k\right)\right) &= m^*(S\cap A_n)+m^*\left(S\cap\left(\bigcup_{k=1}^{n} A_k\right)\right)= \\ &= m^*(S\cap A_n)+\sum_{k=1}^{n-1}m^*(S\cap A_k)= \\ &= \sum_{k=1}^{n}m^*(S\cap A_k). \end{aligned}$$

$\square$

Следствие. *Если A_1, …, A_n — попарно не пересекающиеся множества, измеримые по Лебегу, то $m^*\left(\bigcup_{k=1}^n A_k\right)=\sum_{k=1}^n m^*(A_k)$.*

Теорема 8.6. *Объединение счётного набора измеримых множеств измеримо.*

Доказательство. Прежде всего заметим, что объединение счётного набора измеримых множеств A_1, A_2, … можно представить как объединение счётного набора измеримых попарно не пересекающихся множеств $A_1'=A_1$, $A_2'=A_2\setminus A_1$, …, $A_k'=A_k\setminus(A_1\cup\dots\cup A_{k-1})$, … Поэтому будем считать, что множество A — это объединение счётного набора измеримых попарно не пересекающихся множеств A_1, A_2, … Положим $B_n=A_1\cup\dots\cup A_n$. Множество B_n измеримо и $CA\subset CB_n$, поэтому для любого множества S получаем

$$m^*(S)=m^*(S\cap B_n)+m^*(S\cap CB_n)\geq m^*(S\cap B_n)+m^*(S\cap CA).$$

Согласно теореме 8.5
$$m^*(S\cap B_n)=\sum_{k=1}^{n} m^*(S\cap A_k),$$
поэтому
$$m^*(S)\geq \sum_{k=1}^{n} m^*(S\cap A_k)+m^*(S\cap CA).$$
Следовательно,
$$m^*(S)\geq \sum_{k=\infty}^{n} m^*(S\cap A_k)+m^*(S\cap CA).$$
Воспользовавшись теоремой 8.2, получаем
$$m^*(S)\geq m^*(S\cap A)+m^*(S\cap CA),$$
поэтому множество A измеримо. □

Теорема 8.7. *Если измеримые множества A_1, A_2, ... попарно не пересекаются, то $m\left(\bigcup_{i=1}^{\infty} A_i\right)=\sum_{i=1}^{\infty} m(A_i)$.*

Доказательство. Из теоремы 8.5 и её следствия и из теоремы 8.6 следует, что
$$m\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} m(A_i)\leq \sum_{i=1}^{\infty} m(A_i)$$
для любого натурального n, поэтому $m\left(\bigcup_{i=1}^{\infty} A_i\right)\leq \sum_{i=1}^{\infty} m(A_i)$. Обратное неравенство следует из теоремы 8.2. □

Теорема 8.8. *Интервал и отрезок — измеримые по Лебегу множества.*

Доказательство. Добавление и выкалывание конечного числа точек не влияет на измеримость (и на внешнюю меру), поэтому достаточно доказать измеримость интервала. Для этого, в свою очередь, достаточно доказать измеримость полубесконечного интервала (a,∞). Действительно, интервал (a,b) можно представить в виде $(a,\infty)\setminus[b,\infty)$.

Пусть S — произвольное множество. Удалив при необходимости из S точку a, будем считать, что $a\notin S$. Пусть $S_1=S\cap(-\infty,a)$ и $S_2=S\cap(a,\infty)$. Требуется доказать, что
$$m^*(S_1)+m^*(S_2)\leq m^*(S),$$
т.е. для любого покрытия множества S интервалами I_k выполняется неравенство
$$m^*(S_1)+m^*(S_2)\leq \sum_{i=1}^{\infty}|I_k|.$$
Рассмотрим интервалы $I_{1,k}=(-\infty,a)\cap I_k$ и $I_{2,k}=(a,\infty)\cap I_k$. Интервалы $I_{1,k}$ покрывают множество S_1, интервалы $I_{2,k}$ покрывают множество S_2 и $|I_{1,k}|+|I_{2,k}|=|I_k|$, поэтому
$$\begin{aligned} m^*(S_1)+m^*(S_2) &\leq \sum_{i=1}^{\infty}|I_{1,k}|+\sum_{i=1}^{\infty}|I_{2,k}|= \\ &= \sum_{i=1}^{\infty}|I_k|. \end{aligned}$$
□

Согласно задаче 4.13 любое открытое множество на прямой является объединением не более чем счётного набора попарно не пересекающихся интервалов, поэтому любое открытое множество на прямой измеримо по Лебегу. Следовательно, любое замкнутое множество на прямой тоже измеримо по Лебегу.

Теорема 8.9. *Пусть множества A_1, A_2, … измеримы.*
а) Если $A_n \subset A_{n+1}$ для всех n и $A = \bigcup_{n=1}^{\infty} A_n$, то $m(A) = \lim_{n\to\infty} m(A_n)$.
б) Если $m(A_1) < \infty$, $A_{n+1} \subset A_n$ для всех n и $A = \bigcap_{n=1}^{\infty} A_n$, то $m(A) = \lim_{n\to\infty} m(A_n)$.

Доказательство. а) Если $m(A_n) = \infty$ для некоторого n, то утверждение очевидно. Поэтому будем считать, что $m(A_n) < \infty$ для всех n. Положим $B_1 = A_1$ и $B_n = A_n \setminus A_{n-1}$ при $n > 1$. Множества B_n измеримые, попарно не пересекающиеся и $A = \bigcup_{n=1}^{\infty} B_n$. Множество A_n представляется в виде объединения непересекающихся множеств A_{n-1} и B_n, поэтому $m(B_n) = m(A_n) - m(A_{n-1})$ при $n > 1$.

Согласно теореме 8.7

$$\begin{aligned} m(A) &= \sum_{n=1}^{\infty} m(B_n) = m(A_1) + \sum_{n=1}^{\infty}(m(A_{n+1}) - m(A_n)) = \\ &= \lim_{k\to\infty}[m(A_1) + \sum_{n=1}^{k-1}(m(A_{n+1}) - m(A_n))] = \lim_{k\to\infty} m(A_k). \end{aligned}$$

б) Положим $A'_n = A_1 \setminus A_n$ и $A' = \bigcup_{n=1}^{\infty} A'_n$. Согласно утверждению а) $m(A') = \lim_{n\to\infty} m(A'_n)$, поэтому $m(A_1) - m(A) = m(A_1) - \lim_{n\to\infty} m(A_n)$. $\square$

Задача 8.7. Приведите пример, показывающий, что без предположения $m(A_1) < \infty$ теорема 8.9 б) может быть неверна.

8.4. Интеграл Лебега на прямой

При определении интеграла Римана функции $f(x)$ по отрезку $[a, b]$ этот отрезок разбивается точками $a = x_0 < x_1 < \dots < x_n = b$ и рассматривается сумма $\sum_{k=1}^{n} f(x'_k)(x_k - x_{k-1})$, где $x'_k \in [x_{k-1}, x_k]$. Основная идея построения интеграла Лебега заключается в следующем: разбивают не отрезок $[a, b]$, а область значений функции f, т.е. выбирают точки $y_0 < y_1 < \dots < y_n$ так, что $y_0 < m = \inf f$ и $y_n > M = \sup f$, и рассматривают сумму вида $\sum_{k=1}^{n} y_{k-1} m_k$, где m_k — мера Лебега множества $\{x \mid y_{k-1} \le f(x) < y_k\}$.

Перейдём теперь к деталям определения интеграла Лебега. В этом параграфе нас интересует только мера Лебега, поэтому под мерой подразумевается мера Лебега, а под измеримостью подразумевается измеримость по Лебегу.

Функцию f на отрезке $I = [a, b]$ называют *измеримой*, если для любого c множество $\{x \in I \mid f(x) > c\}$ измеримо. В таком случае измеримы множества $\{x \in I \mid f(x) \le c\}$ и $\{x \in I \mid c < f(x) \le d\}$ как разности измеримых множеств. Множество $\{x \in I \mid f(x) = c\}$ тоже измеримо как пересечение измеримых множеств $\{x \in I \mid c - \frac{1}{n} < f(x) \le c + \frac{1}{n}\}$. Следовательно, измеримы также множества $\{x \in I \mid f(x) < c\}$, $\{x \in I \mid f(x) \ge c\}$, $\{x \in I \mid \ge c f(x) \le d\}$, $\{x \in I \mid c < f(x) < d\}$. Эти рассуждения показывают, что при определении измеримости функции можно исходить из любого из упомянутых множеств, кроме множества $\{x \in I \mid f(x) = c\}$.

Любая непрерывная на отрезке функция f измерима, поскольку множество $\{x \in I \mid f(x) > c\}$ открыто.

Задача 8.8. Докажите, что если последовательность измеримых функций f_n на отрезке $I = [a, b]$ поточечно сходится к функции f, то функция f измерима.

Теорема 8.10 (Северини–Егоров). *Если последовательность измеримых функций f_n поточечно сходится на отрезке $[a, b]$ к функции f, то для любого $\varepsilon > 0$ можно выбрать множество меры меньше ε, вне которого сходимость равномерная.*

Доказательство. Согласно задаче 8.8 функция f измерима. Для натуральных n и k рассмотрим множество

$$E_{n,k} = \left\{ x \in [a, b] \mid |f_m(x) - f(x)| \ge \frac{1}{k} \quad \text{для некоторого} \quad m \ge n \right\}.$$

Множество $E_{n,k}$ измеримо и $E_{n+1,k} \subset E_{n,k}$. Из поточечной сходимости функций следует, что для фиксированного k точка не может принадлежать всем множествам $E_{n,k}$, поэтому $\bigcap_{n=1}^{\infty} E_{n,k} = \varnothing$.

Согласно теореме 8.9 б) $\lim_{n\to\infty} m(E_{n,k}) = 0$. Для каждого k выберем n_k так, что $m(E_{n_k,k}) < \frac{\varepsilon}{2^k}$. Мера множества $E = \bigcup_{k=1}^{\infty} E_{n_k,k}$ меньше ε. Вне множества E сходимость равномерная. Действительно, если $x \notin E$, то $|f_m(x) - f(x)| < \frac{1}{k}$ для всех $m \geq n_k$. $\square$

Историческое замечание. Карло Северини (1872-1951) доказал теорему 8.10 в 1910 году, а Дмитрий Фёдорович Егоров (1869-1931) независимо доказал эту теорему в 1911 году. Эту теорему часто называют теоремой Егорова.

Интеграл Лебега ограниченной измеримой функции f на отрезке $[a,b]$ определяется следующим образом. Как уже говорилось, сначала нужно выбрать точки $y_0 < y_1 < \cdots < y_n$ так, что $y_0 < m = \inf f$ и $y_n > M = \sup f$. Рассмотрим две суммы: $s = \sum_{k=1}^{n} y_{k-1}\mu_k$ и $S = \sum_{k=1}^{n} y_k\mu_k$, где μ_k — мера Лебега множества $\{x \mid y_{k-1} \leq f(x) < y_k\}$. *Интеграл Лебега* — это общий предел этих сумм, когда наибольшая длина отрезка $[y_{k-1}, y_k]$ стремится к нулю.

Теорема 8.11. *Суммы s и S стремятся к общему пределу, когда наибольшая длина отрезка $[y_{k-1}, y_k]$ стремится к нулю.*

Доказательство. Суммы s и S ограничены: они заключены между $m\Delta$ и $M\Delta$, где $\Delta = b - a$ — длина отрезка.

Разность $S - s$ стремится к нулю, когда наибольшая длина отрезка $[y_{k-1}, y_k]$ стремится к нулю. Действительно, если $y_k - y_{k-1} < \delta$, то $S - s \leq \delta \sum \mu_k = \delta(M - m)$, поскольку отрезок $[M, m]$ (возможно, без одного или обоих концов) — это объединение попарно непересекающихся множеств, меры которых равны μ_k.

Если между точками y_k вставить новые точки деления, то сумма S не возрастает, а сумма s не убывает.

Таким образом, если измельчать разбиение, вставляя новые точки деления так, чтобы наибольшая длина отрезка деления стремилась к нулю, то суммы s и S стремятся к общему пределу Σ, причём $s \leq \Sigma \leq S$. Покажем, что при любом измельчении разбиения суммы стремятся к тому же самому пределу Σ. Действительно, пусть $S - s < \varepsilon$ и при каком-то другом измельчении получились суммы s' и S', для которых $S' - s' < \varepsilon$. Рассмотрим разбиение, которое получается при объединении точек первого разбиения и второго. Для него получим суммы s'' и S''. Числа s'' и S'' расположены как между числами s и S, так и между числами s' и S', поэтому отрезки $[s, S]$ и $[s', S']$ пересекаются. Следовательно, числа s' и S' отличаются меньше, чем на 2ε от числа Σ, расположенного на отрезке $[s, S]$. $\square$

Непосредственно из определения интеграла Лебега видно, что на него не влияют значения функции на множестве меры нуль. В частности, функция Дирихле интегрируема по Лебегу.

Аналогично можно определить измеримые функции и интеграл Лебега не только на отрезке, но и на любом ограниченном измеримом множестве.

Теорема 8.12. *Если функция интегрируема по Риману, то она интегрируема по Лебегу, причём оба интеграла равны.*

Доказательство. Пусть функция f на отрезке $[a,b]$ интегрируема по Риману. Представим множество $\{x \in [a,b] \mid f(x) \geq c\}$ в виде разности его замыкания (измеримого множества) и множества предельных точек, не входящих в замыкание. Все предельные точки, не входящие в замыкание, являются точками разрыва, поэтому согласно критерию Лебега их мера равна нулю. Таким образом, функция f измерима, поэтому она интегрируема по Лебегу.

Чтобы доказать равенство интегралов Римана и Лебега, рассмотрим некоторое разбиение отрезка $[a,b]$. Для каждого отрезка $[x_k, x_{k+1}]$ этого разбиения рассмотрим M_k и m_k — наибольшее и наименьшее значения функции $f(x)$ на этот отрезке. Интеграл Лебега функции f по отрезку $[x_k, x_{k+1}]$ заключён между $m_k(x_{k+1} - x_k)$ и $M_k(x_{k+1} - x_k)$, поэтому интеграл Лебега по отрезку $[a,b]$ заключён между верхней и нижней интегральными суммами для рассматриваемого разбиения. Из этого следует требуемое равенство. $\square$

Определим теперь интеграл Лебега для измеримой неограниченной функции. Предположим сначала что функция $f(x)$ неотрицательная. Рассмотрим функцию $f_n(x)$, которая равна $f(x)$ при $f(x) \leq n$ и равна n при $f(x) > n$. Интеграл Лебега функции $f(x)$ — это предел (конечный или бесконечный) интегралов Лебега функций $f_n(x)$ при $n \to \infty$. Для неположительной функции интеграл Лебега определяется аналогично. Произвольную функцию f можно представить в виде суммы $f(x) = f_1(x) + f_2(x)$, где функция f_1 неотрицательная (она равна $f(x)$ при $f(x) \geq 0$ и 0 при $f(x) < 0$), а функция f_2 неположительная (она равна $f(x)$ при $f(x) \leq 0$ и 0 при $f(x) > 0$). Если оба интеграла Лебега функций f_1 и f_2, конечны, то интеграл Лебега функции f равен их сумме. Если же хотя бы один из этих интегралов бесконечен, то интеграл Лебега функции f мы не определяем.

Теорема 8.13 (абсолютная непрерывность интеграла Лебега). *Если функция f интегрируема по Лебегу на отрезке $[a, b]$, то для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что сумма интегралов функции f по непересекающимся отрезкам, сумма длин которых меньше δ, меньше ε.*

Доказательство. Достаточно рассмотреть случай, когда функция f неотрицательная. Пусть f_n — функция, которая рассматривалась чуть выше при определении интеграла Лебега. Выберем n так, что $\int_a^b (f - f_n)dx < \varepsilon/2$. Тогда для любого набора I_1 непересекающихся отрезков, содержащихся в отрезке $[a, b] = I$, выполняется неравенство

$$\int_{I_1} (f - f_n)dx \leq \int_{I_1} (f - f_n)dx + \int_{I \setminus I_1} (f - f_n)dx = \int_I (f - f_n)dx < \frac{\varepsilon}{2}.$$

Положим $\delta = \frac{\varepsilon}{2n}$ и будем предполагать, что сумма длин отрезков набора I_1 меньше δ. Тогда

$$\int_{I_1} f\,dx < \int_{I_1} f_n\,dx + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Последовательность функций f_n, интегрируемых по Риману на отрезке $[a, b]$, может сходиться на этом отрезке к интегрируемой по Риману функции f, но при этом равенство $\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$ может не выполняться (задача 6.52). Интегрируемые по Риману функции интегрируемы по Лебегу, поэтому и для интеграла Лебега из сходимости последовательности функций сходимость последовательности их интегралов к интегралу предельной функции следует лишь при некоторых условиях.

Напомним, что согласно задаче 8.8 поточечно сходящаяся последовательность измеримых функций сходится к измеримой функции.

Теорема 8.14. *Если последовательность измеримых функций f_n на отрезке $[a, b]$ равномерно ограничена и поточечно сходится к функции f, то $\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx$.*

Доказательство. Выберем $\varepsilon > 0$. Воспользовавшись теоремой 8.10 выберем множество $E \subset [a, b]$, мера которого меньше ε, так, что вне E сходимость равномерная. Пусть $|f_n(x)| \leq M$ для всех n и всех $x \in [a, b]$. Тогда

$$\begin{aligned}
\left| \int_{[a,b]} (f_n - f) \right| &\leq \int_{[a,b]} |f_n - f| = \\
&= \int_{[a,b]\setminus E} |f_n - f| + \int_E |f_n - f| < \\
&< \int_{[a,b]\setminus E} |f_n - f| + 2M\varepsilon.
\end{aligned}$$

Из равномерной сходимости функций на $[a, b] \setminus E$ следует, что можно выбрать N так, что $|f_n(x) - f(x)| < \varepsilon$ при $n > N$ и $x \in [a, b] \setminus E$. Поэтому при $n > N$ получаем:

$$\left| \int_{[a,b]} (f_n - f) \right| < \varepsilon m([a, b] \setminus E) + 2M\varepsilon < \varepsilon(b - a + 2M).$$

Устремляя ε к нулю, получаем требуемое. □

Историческое замечание. Интеграл Лебега ввёл Анри Лебег (1875-1941) в 1902 году.

8.5. Интеграл Стильтеса

Интеграл Стильтеса $\int_a^b f\,d\alpha$ (используется также обозначение $\int_a^b f(x)d\alpha(x)$) от функции f относительно функции α обобщает интеграл Римана. Он определяется в ситуации, когда на отрезке $[a, b]$ задана монотонно возрастающая функция α; эта функция не обязательно непрерывна. Интеграл Римана при этом получается в случае, когда $\alpha(x) = x$. Пусть на отрезке $[a, b]$ задана также ограниченная функция $f(x)$ и задано некоторое разбиение этого отрезка. Для каждого отрезка $[x_k, x_{k+1}]$ рассмотрим M_k и m_k — наибольшее и наименьшее значения функции $f(x)$ на этот отрезке. Суммы

$$\sum_{k=0}^{n-1} M_k(\alpha(x_{k+1}) - \alpha(x_k)) \quad \text{и} \quad \sum_{k=0}^{n-1} m_k(\alpha(x_{k+1}) - \alpha(x_k))$$

назовём ***верхней*** и ***нижней интегральными суммами***. Интеграл Стильтеса от функции f относительно функции α существует, если точная нижняя грань верхних интегральных сумм равна точной верхней грани нижних интегральных сумм.

Историческое замечание. Томас Стилтьес (1856-1894) ввёл интеграл Стилтьеса в 1894 году.

8.6. Решения задач

8.1. Занумеруем данные множества и покроем первое из них интервалами, сумма длин которых не превосходит $\frac{\varepsilon}{2}$, второе интервалами, сумма длин которых не превосходит $\frac{\varepsilon}{4}$, и т.д. Сумма длин всех этих интервалов не превосходит ε.

8.2. Пусть $C_0 = [0, 1]$, множество C_1 получено из C_0 вырезанием интервала $(1/3, 2/3)$. Множество C_1 состоит из двух отрезков длины $1/3$. Вырежем аналогично из каждого из этих отрезков интервал длины $1/9$, и т.д. Множество C_n состоит из 2^n отрезков длины $1/3^n$. Канторово множество — это пересечение всех множеств C_n, поэтому оно содержится в каждом множестве C_n. Для данного $\varepsilon > 0$ выберем n так, что $(2/3)^n < \varepsilon$, и покроем C_n набором интервалов, сумма длин которых меньше ε. Этот набор интервалов покрывает и канторово множество.

8.3. Будем вычислять внешнюю меру Жордана, рассматривая покрытия отрезками. Возьмём несколько отрезков. Если некоторая точка $x \in [a, b]$ не покрыта ими, то ими не покрыта и некоторая окрестность этой точки, поэтому не покрыта и некоторая рациональная точка отрезка $[a, b]$. Таким образом, если несколько отрезков покрывают все рациональные точки отрезка $[a, b]$, то они покрывают весь отрезок $[a, b]$. Отрезки при этом можно считать неперекрывающимися. В таком случае сумма их длин не меньше $b - a$.

Множество рациональных точек отрезка $[a, b]$ не содержит ни одного отрезка, поэтому его внутренняя мера равна 0.

8.4. Непосредственно из очевидного свойства $\sup\{-x \mid x \in X\} = -\inf\{x \mid x \in X\}$ вытекает следующее утверждение. Пусть X — некоторое множество чисел, c — некоторое число, $Y = \{c - x \mid x \in X\}$. Тогда $\sup Y = c - \inf X$, т.е. $\sup Y + \inf X = c$.

Будем вычислять внешнюю меру Жордана множества $[a, b] \backslash M$, покрывая его непересекающимися замкнутыми отрезками, а внутреннюю меру Жордана множества M будем вычислять с помощью содержащихся в нём непересекающихся интервалов. Дополнение к набору покрывающих множество $[a, b] \backslash M$ отрезков — это набор содержащихся в M непересекающихся интервалов, и наоборот. Поэтому сумма внутренней меры Жордана множества M и внешней меры Жордана множества $[a, b] \setminus M$ равна $b - a$.

8.5. Каждому разбиению отрезка можно сопоставить множество тех отрезков, имеющих общие точки с M, и множество тех отрезков, которые содержатся в M. Сумма длин отрезков первого множества равна верхней интегральной сумме функции χ_M, а сумма длин отрезков второго множества равна нижней интегральной сумме функции χ_M.

8.6. Граничные точки множества — это точки разрыва его характеристической функции. Согласно критерию Лебега функция интегрируема по Риману тогда и только тогда, когда множество её точек разрыва имеет меру Лебега 0.

8.7. Если $A_n = (n, \infty)$, то $A = \varnothing$.

8.8. Неравенство $f(x) \leq c$ выполняется тогда и только тогда, когда для любого натурального k можно выбрать $N(k)$ так, что $f_n(x) < c + \frac{1}{k}$ при $n > N(k)$. Поэтому

$$\{x \in I \mid f(x) \leq c\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n>N} \left\{x \in I \mid f_n(x) < c + \frac{1}{k}\right\}.$$

Счётные объединения и пересечения измеримых множеств измеримы.

Часть II

Функции многих переменных

Глава 9.

Функции многих переменных

Функции многих переменных связаны с многомерными пространствами. Длину вектора $\mathbf{v} = (v_1, \ldots, v_n)$ пространства $\mathbb{R}^n$, равную $\sqrt{v_1^2 + \cdots + v_n^2}$, будем обозначать $\|\mathbf{v}\|$.

Функция многих переменных — это отображение $f\colon \mathbb{R}^n \to \mathbb{R}$. Можно также рассматривать отображения $\mathbb{R}^n \to \mathbb{R}^m$. С помощью открытых множеств непрерывность таких отображений определяется точно так же, как непрерывность функций. Множество $U \subset \mathbb{R}^n$ называют *открытым*, если для любой точки $\mathbf{x}_0 \in U$ можно выбрать $\varepsilon > 0$ так, что все точки $\mathbf{x}$, для которых $\|\mathbf{x} - \mathbf{x}_0\| < \varepsilon$, принадлежат U. Открытые подмножества любого множества $X \subset \mathbb{R}^n$ определяются как пересечения множества X со всеми открытыми подмножествами в $\mathbb{R}^n$. Множество *замкнуто*, если его дополнение открыто. Множество *связно*, если его нельзя представить в виде объединения двух непересекающихся непустых открытых множеств.

Задача 9.1. Докажите, что существует непрерывная функция f в $\mathbb{R}^n$, которая равна 1 на шаре радиуса R с центром $\mathbf{a}$ и равна 0 вне шара радиуса $2R$ с тем же центром.

Пример 9.1. *Положим $f(0,0) = 0$ и $f(x,y) = \frac{xy}{x^2+y^2}$ при $(x,y) \neq (0,0)$. Функция f непрерывна по каждой переменной, но имеет разрыв в точке* $(0,0)$.

Доказательство. Пусть $g(x) = f(x, y_0)$. Если $y_0 = 0$, то $g(x) = 0$, а если $y_0 \neq 0$, то $g(x) = \frac{xy_0}{x^2+y_0^2}$. Во всех случаях функция $g(x)$ непрерывна. Для функции $h(y) = f(x_0, y)$ доказательство аналогично.

Сколь угодно близко к точке $(0,0)$ есть точка вида (a,a), для которой $f(a,a) = 1/2$. $\square$

Историческое замечание. Первый пример разрывной функции многих переменных, непрерывной по каждой из переменных, предложил Карл Йоханнес Тома (1840-1921) в 1870 году.

9.1. Топология пространства $\mathbb{R}^n$

Множество в $\mathbb{R}^n$ *компактно*, если из любого его покрытия открытыми множествами можно выбрать конечное подпокрытие.

Прямоугольный параллелепипед в $\mathbb{R}^n$ состоит из точек $\mathbf{x} = (x_1, \ldots, x_n)$, координаты которых удовлетворяют неравенствам $a_i \leq x_i \leq b_i$.

Теорема 9.1. *Прямоугольный параллелепипед в $\mathbb{R}^n$ компактен.*

Доказательство. В многомерном случае доказательство почти такое же, как в одномерном (теорема 4.2). Вместо деления отрезка пополам теперь нужно разделить параллелепипед на 2^n параллелепипедов, проведя через его центр пространства размерности $n-1$, параллельные граням. Так мы получим последовательность вложенных параллелепипедов, для каждого из которых нельзя выбрать конечное подпокрытие. На i-й оси координат получаем стягивающуюся последовательность отрезков, поэтому есть точка x_i, принадлежащая всем этим отрезкам. Тогда точка с координатами $(x_1, \ldots, x_n)$ принадлежит всем выбранным параллелепипедам. Она содержится в некотором множестве открытого покрытия, и в этом множестве содержится какой-то из выбранных параллелепипедов. Приходим к противоречию. $\square$

Множество в $\mathbb{R}^n$ называют *ограниченным*, если оно содержится в некотором шаре. Множество в $\mathbb{R}^n$ компактно тогда и только тогда, когда оно замкнуто и ограничено. Это доказывается точно так же, как и в одномерном случае (теорема 4.3). Никаких изменений не требуют доказательства других свойств компактных множеств: образ компактного множества при непрерывном отображении компактен; непрерывная функция на компактном множестве достигает максимума. Равномерная непрерывность функции в многомерном случае определяется точно так же, как и в одномерном; функция, непрерывная на компактном множестве, равномерно непрерывна. Равномерная сходимость последовательности функций тоже определяется точно так же; равномерно сходящаяся последовательность функций, непрерывных в точке $\mathbf{x}_0$, сходится к функции, непрерывной в точке $\mathbf{x}_0$.

Задача 9.2. Докажите, что ограниченная последовательность точек в $\mathbb{R}^n$ имеет предельную точку.

Задача 9.3. а) Докажите, что объединение двух связных множеств, имеющих общую точку, связно.

б) Докажите, что произведение двух связных множеств связно.

Пусть $A \subset \mathbb{R}^n$ — произвольное подмножество. Для произвольной точки $\mathbf{x} \in \mathbb{R}^n$ величину $d(\mathbf{x}, A) = \inf_{\mathbf{a}\in A} \|\mathbf{x} - \mathbf{a}\|$ называют *расстоянием* от точки $\mathbf{x}$ до множества A.

Задача 9.4. Докажите, что функция $f(\mathbf{x}) = d(\mathbf{x}, A)$ непрерывна для любого подмножества $A \subset \mathbb{R}^n$

Задача 9.5. а) Докажите, что если множество A замкнуто, то функция $f(\mathbf{x}) = d(\mathbf{x}, A)$ для всех $\mathbf{x} \notin A$ принимает положительные значения.

б) Докажите, что любое замкнутое множество в $\mathbb{R}^n$ является множеством нулей некоторой непрерывной функции.

Пусть $A, B \subset \mathbb{R}^n$ — произвольные подмножества. Величину $d(A, B) = \inf\limits_{\mathbf{a}\in A,\ \mathbf{b}\in B} \|\mathbf{a} - \mathbf{b}\|$ называют *расстоянием* между множествами A и B.

Задача 9.6. Пусть $A \subset \mathbb{R}^n$ — замкнутое подмножество, $C \subset \mathbb{R}^n$ — компактное подмножество. Докажите, что тогда существует такая точка $\mathbf{c}_0 \in C$, что $d(A, C) = d(A, \mathbf{c}_0)$. Докажите также, что если множество A компактно, то существует ещё и такая точка $\mathbf{a}_0 \in A$, что $d(A, C) = d(\mathbf{a}_0, \mathbf{c}_0)$.

Пусть $C \subset \mathbb{R}^n$ — замкнутое множество. Рассмотрим в $\mathbb{R}^n$ множество X тех точек, для которых ближайшая точка множества C единственна. Сопоставляя точке $\mathbf{x} \in X$ ближайшую к ней точку множества C, получаем отображение $f\colon X \to C$.

Задача 9.7. Докажите, что отображение $f\colon X \to C$ непрерывно.

Задача 9.8. Пусть $K \subset U \subset \mathbb{R}^n$, причём множество K компактно, а множество U открыто. Докажите, что можно выбрать $\varepsilon > 0$ так, что если $\mathbf{x} \in K$ и $\|\mathbf{x} - \mathbf{y}\| < \varepsilon$, то $\mathbf{y} \in U$.

Задача 9.9. *Докажите, что функция $f(x)$ на отрезке $[a, b]$ непрерывна тогда и только тогда, когда её график $\{(x, y) \in \mathbb{R}^2 \mid x \in [a, b],\ y = f(x)\}$ — замкнутое связное множество на плоскости $\mathbb{R}^2$.

Задача 9.10. *Каждая из прямых, параллельных сторонам квадрата, пересекает некоторое множество точек квадрата не более чем в одной точке. Может ли граница этого множества совпадать со всем квадратом?

9.2. Дифференциал

Отображение $f\colon U \to \mathbb{R}^m$, где $U \subset \mathbb{R}^n$ — открытое множество, называют дифференцируемым в точке $\mathbf{x} \in U$, если существует линейное отображение $A\colon \mathbb{R}^n \to \mathbb{R}^m$, для которого

$$\lim_{\mathbf{h}\to 0} \frac{\|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - A\mathbf{h}\|}{\|\mathbf{h}\|} = 0.$$

Линейное отображение A называют при этом *производной* или *дифференциалом* отображения f в точке $\mathbf{x}$ и обозначают $df_{\mathbf{x}}$. Для функции $f(x)$ одного переменного линейное отображение df_x — это отображение $h \mapsto f'(x)h$. Таким образом, $f'(x) = df_x(1)$.

Замечание. Вектор $\mathbf{h}$, на котором принимает значение дифференциал $df_{\mathbf{x}}$, возникает из выражения $\mathbf{x}+\mathbf{h}$. Естественно считать, что точка $\mathbf{x}$ — начало вектора $\mathbf{h}$. Это обстоятельство иногда приходится учитывать даже для отображений $\mathbb{R}^n \to \mathbb{R}^m$. А при переходе к отображениям многообразий это проявляется уже вполне отчётливо: дифференциал отображения в точке x определён на касательном пространстве в точке x. Касательные пространства в разных точках пространства $\mathbb{R}^n$ можно отождествить параллельным переносом, а касательные пространства в разных точках многообразия нельзя отождествить никаким естественным способом.

Задача 9.11. Докажите, что производная дифференцируемого отображения единственна.

Задача 9.12. Докажите, что если отображение дифференцируемо в некоторой точке, то в этой точке оно непрерывно.

Задача 9.13. Докажите, что отображение $f=(f_1,\dots,f_m)$ дифференцируемо тогда и только тогда, когда каждая из функций f_1, …, f_m дифференцируема.

Задача 9.14. Докажите, что производная линейного отображения $A\colon \mathbb{R}^n \to \mathbb{R}^m$ в каждой точке равна A

Задача 9.15. Докажите, что если отображения $f, g\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируемы в точке $\mathbf{a}$, то $d(f+g)_{\mathbf{a}} = df_{\mathbf{a}} + dg_{\mathbf{a}}$.

Задача 9.16. Докажите, что если функции $f, g\colon \mathbb{R}^n \to \mathbb{R}$ дифференцируемы в точке $\mathbf{a}$, то $d(fg)_{\mathbf{a}} = g(\mathbf{a})df_{\mathbf{a}} + f(\mathbf{a})dg_{\mathbf{a}}$ и, если $g(\mathbf{a}) \neq 0$, то $d\left(\frac{f}{g}\right)_{\mathbf{a}} = \frac{g(\mathbf{a})df_{\mathbf{a}} - f(\mathbf{a})dg_{\mathbf{a}}}{(g(\mathbf{a}))^2}$.

Для задания линейного оператора можно использовать элементы его матрицы. Элементами матрицы df являются частные производные, которые определяются следующим образом. Отображение $\mathbb{R}^n \to \mathbb{R}^m$ можно задать функциями f_1, …, f_m от переменных x_1, …, x_n. *Частная производная* $\frac{\partial f_i}{\partial x_j}$ — это обычная производная функции f_i по x_j при фиксированных значениях остальных переменных. Для функции $f(x_1,\dots,x_n)$ частную производную $\frac{\partial f}{\partial x_i}$ обозначают также $D_i f$ и f_{x_i}. Для функций $f(x,y)$ и $f(x,y,z)$ частные производные по x, y и z иногда обозначают f_x, f_y и f_z.

Матрица дифференциала отображения f — это матрица $\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$ с n столбцами и m строками. Матрицу дифференциала отображения f называют также *матрицей Якоби*. Если матрица Якоби квадратная, то её определитель называют *якобианом*. Для отображения f якобиан в точке x обозначают $J_f(x)$. Если $f(x_1,\dots,x_n) = (y_1,\dots,y_n)$, то якобиан обозначают также

$$\frac{\partial(y_1,\dots,y_n)}{\partial(x_1,\dots,x_n)}.$$

Частные производные входят в уравнение касательной плоскости к графику функции от двух переменных. Пусть $z = F(x,y)$, где функция F имеет непрерывные частные производные в окрестности точки (x_0,y_0). Точка M_0 с координатами (x_0,y_0,z_0), где $z_0 = F(x_0,y_0)$, лежит на графике этой функции. Пусть кривая $x=f(t)$, $y=g(t)$, $z=h(t)$ при $t=t_0$ проходит через точку M_0, причём в этой точке производные f', g' и h' не обращаются в нуль. Тогда касательная к этой кривой в точке M_0 задаётся уравнениями

$$\frac{x-x_0}{f'(t_0)} = \frac{y-y_0}{g'(t_0)} = \frac{z-z_0}{h'(t_0)}.$$

Действительно, рассмотрим на кривой точки, соответствующие значениям параметра t_0 и t_0+h. Прямая, проходящая через эти точки, задаётся уравнениями

$$\frac{x-f(t_0)}{f(t_0+h)-f(t_0)} = \frac{y-g(t_0)}{g(t_0+h)-g(t_0)} = \frac{z-h(t_0)}{h(t_0+h)-h(t_0)}.$$

Поделим знаменатели этих отношений на h и устремим h к нулю. В результате получим требуемые уравнения.

Из равенства $z = F(x, y)$ следует, что $h(t) = F(f(t), g(t))$, поэтому $h'(t_0) = f'(t_0)F'_x(x_0, y_0) + g'(t_0)F'_y(x_0, y_0)$. Исключив $f'(t_0)$, $g'(t_0)$ и $h'(t_0)$ из этого уравнения и уравнений касательной, получим

$$z - z_0 = F'_x(x_0, y_0)(x - x_0) + F'_y(x_0, y_0)(y - y_0).$$

Это уравнение задаёт плоскость, содержащую касательные ко всем кривым на рассматриваемой поверхности, проходящим через точку M_0. Эта плоскость — касательная плоскость к поверхности.

Производная в точке $\mathbf{x}$ отображения $f\colon \mathbb{R}^n \to \mathbb{R}^m$ *в направлении вектора* $\mathbf{h} \in \mathbb{R}^n$ — это предел $\lim_{t\to 0} \frac{f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x})}{t}$. Эту производную будем обозначать $\partial_{\mathbf{h}} f$ или $D_{\mathbf{h}} f$. Частная производная — это производная в направлении базисного вектора.

Задача 9.17. Докажите, что если отображение дифференцируемо в точке, то в этой точке у него есть производные во всех направлениях.

Задача 9.18. Докажите, что производная отображения f в направлении $\mathbf{v} = (v_1, \ldots, v_n)$ равна $\sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}$.

Градиент функции f в точке $\mathbf{x}_0$ — это такой вектор $\operatorname{grad} f(\mathbf{x}_0)$, что

$$(\mathbf{v}, \operatorname{grad} f(\mathbf{x}_0)) = df(\mathbf{x}_0)(\mathbf{v})$$

для любого вектора $\mathbf{v}$. Координаты градиента — это частные производные функции f в точке $\mathbf{x}_0$.

Задача 9.19. Приведите пример разрывной функции, у которой есть производная в направлении любого вектора.

Теорема 9.2. *Если функция $f\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируема в точке $\boldsymbol{a}$ и функция $g\colon \mathbb{R}^m \to \mathbb{R}^k$ дифференцируема в точке $f(\boldsymbol{a})$, то композиция $h = g \circ f\colon \mathbb{R}^n \to \mathbb{R}^k$ дифференцируема в точке $\boldsymbol{a}$ и $dh_{\boldsymbol{a}} = dg_{f(\boldsymbol{a})} \circ df_{\boldsymbol{a}}$.*

Доказательство. Пусть $\mathbf{b} = f(\mathbf{a})$, $A = df_{\mathbf{a}}$ и $B = dg_{f(\mathbf{a})}$. Положим

$$\begin{aligned} F(\mathbf{x}) &= f(\mathbf{x}) - f(\mathbf{a}) - A(\mathbf{x} - \mathbf{a}), \\ G(\mathbf{y}) &= g(\mathbf{y}) - g(\mathbf{b}) - B(\mathbf{y} - \mathbf{b}), \\ H(\mathbf{x}) &= (g \circ f)(\mathbf{x}) - (g \circ f)(\mathbf{a}) - BA(\mathbf{x} - \mathbf{a}). \end{aligned}$$

Тогда $\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|F(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{a}\|} = 0$ и $\lim_{\mathbf{y}\to\mathbf{b}} \frac{\|G(\mathbf{y})\|}{\|\mathbf{y}-\mathbf{b}\|} = 0$, и нужно доказать, что $\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|H(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{a}\|} = 0$. Ясно, что

$$\begin{aligned} H(\mathbf{x}) &= (g \circ f)(\mathbf{x}) - (g \circ f)(\mathbf{a}) - BA(\mathbf{x} - \mathbf{a}) = \\ &= (g \circ f)(\mathbf{x}) - (g \circ f)(\mathbf{a}) - B(f(\mathbf{x}) - f(\mathbf{a}) - F(\mathbf{x})) = \\ &= [g(f(\mathbf{x})) - g(\mathbf{b}) - B(f(\mathbf{x}) - \mathbf{b})] + BF(\mathbf{x})) = \\ &= G(f(\mathbf{x})) + BF(\mathbf{x}). \end{aligned}$$

Поэтому достаточно проверить, что $\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|G(f(\mathbf{x}))\|}{\|\mathbf{x}-\mathbf{a}\|} = 0$ и $\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|BF(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{a}\|} = 0$. Второе равенство следует из того, что $\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|F(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{a}\|} = 0$. Для доказательства первого равенства воспользуемся тем, что $\lim_{\mathbf{y}\to\mathbf{b}} \frac{\|F(\mathbf{y})\|}{\|\mathbf{y}-\mathbf{b}\|} = 0$. Для любого $\varepsilon > 0$ существует $\delta > 0$, для которого $\|G(f(\mathbf{x}))\| \le \varepsilon \|f(\mathbf{x}) - \mathbf{b}\|$ при $\|f(\mathbf{x}) - \mathbf{b}\| < \delta$. В свою очередь, существует $\delta_1 > 0$, для которого $\|f(\mathbf{x}) - \mathbf{b}\| < \delta$ при $\|\mathbf{x} - \mathbf{a}\| < \delta_1$. Для такого δ_1 получаем

$$\begin{aligned} \|G(f(\mathbf{x}))\| &\le \varepsilon \|f(\mathbf{x}) - \mathbf{b}\| = \varepsilon \|F(\mathbf{x}) + A(\mathbf{x} - \mathbf{a})\| \le \\ &\le \varepsilon \|F(\mathbf{x})\| + \varepsilon M \|\mathbf{x} - \mathbf{a}\| \end{aligned}$$

для некоторого M. □

Замечание. Теорема 9.2 верна и в том случае, когда функции f и g определены на некоторых открытых множествах в $\mathbb{R}^n$ и в $\mathbb{R}^m$. Доказательство то же самое.

Пример 9.2. *Если* $h(t) = g(f_1(t), \dots, f_n(t))$, *то*

$$\frac{dh}{dt} = \frac{\partial g}{\partial x_1} \cdot \frac{df_1}{dt} + \ldots + \frac{\partial g}{\partial x_n} \cdot \frac{df_n}{dt}.$$

Доказательство. Матрица оператора dg — это строка $\left(\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_n}\right)$. Матрица оператора df — это столбец $\left(\frac{df_1}{dt}, \dots, \frac{df_n}{dt}\right)^T$. Произведение этих матриц — матрица, состоящая из одного элемента, равного $\frac{\partial g}{\partial x_1} \cdot \frac{df_1}{dt} + \ldots + \frac{\partial g}{\partial x_n} \cdot \frac{df_n}{dt}$. □

Задача 9.20. а) Докажите, что $\operatorname{grad}(\|\mathbf{x}\|^p) = p\|\mathbf{x}\|^{p-2}\mathbf{x}$ при $\mathbf{x} \neq \mathbf{0}$.

б) Докажите, что $\sum_{i=1}^n \frac{\partial}{\partial x_i}(\|\mathbf{x}\|^{-n}x_i) = 0$ при $\mathbf{x} \neq \mathbf{0}$.

Задача 9.21. а) Предположим, что $\frac{\partial^2 f(x,y)}{\partial x \partial y} = 0$ для всех x и y. Докажите, что существуют функции g и h, для которых $f(x, y) = g(x) + h(y)$.

б) Предположим, что $\frac{\partial^2}{\partial t^2}u(x, t) = c^2 \frac{\partial^2}{\partial x^2}u(x, t)$. Докажите, что существуют функции F и G, для которых $u(x, t) = F(x - ct) + G(x + ct)$.

Координаты $r \in [0, \infty)$ и $\varphi \in [0, 2\pi)$, для которых $x = r\cos\varphi$ и $y = r\sin\varphi$, где x, y — декартовы координаты на плоскости, называют *полярными*; r — это расстояние от точки (x, y) до начала координат, а φ — угол между осью абсцисс и лучом, идущим из начала координат в эту точку.

Задача 9.22. Пусть $g(r, \varphi) = f(r\cos\varphi, r\sin\varphi)$, где $f(x, y)$ — дифференцируемая функция. Докажите, что

$$\frac{\partial g}{\partial r} = \frac{\partial f}{\partial x}\cos\varphi + \frac{\partial f}{\partial y}\sin\varphi, \quad \frac{\partial g}{\partial \varphi} = -\frac{\partial f}{\partial x}r\sin\varphi + \frac{\partial f}{\partial y}r\cos\varphi.$$

Для дважды дифференцируемой функции $f(x_1, \dots, x_n)$ можно определить *оператор Лапласа*

$$\Delta f(x_1, \dots, x_n) = \frac{\partial^2 f(x_1, \dots, x_n)}{\partial x_1^2} + \cdots + \frac{\partial^2 f(x_1, \dots, x_n)}{\partial x_n^2}.$$

Задача 9.23. Пусть $x = r\cos\varphi$ и $y = r\sin\varphi$. Докажите, что при $r \neq 0$ для дважды дифференцируемой функции $u(x, y)$ выполняется равенство

$$\Delta u(x, y) = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2}.$$

Задача 9.24. Пусть $\mathbf{x} \in \mathbb{R}^n$ и $u(\mathbf{x}) = v(r)$, где $r = \|\mathbf{x}\|$. Докажите, что

$$\Delta u = v''(r) + \frac{n-1}{r}v'(r).$$

Задача 9.25. Докажите, что для оператора Лапласа в $\mathbb{R}^n$ выполняется равенство $\Delta(\|\mathbf{x}\|^p) = p(p + n - 2)\|\mathbf{x}\|^{p-2}$.

Функцию $f\colon \mathbb{R}^n \to \mathbb{R}$ называют *однородной* степени m, если $f(t\mathbf{x}) = t^m f(\mathbf{x})$ для всех $\mathbf{x} \in \mathbb{R}^n$.

Задача 9.26. Докажите, что если однородная функция f степени m дифференцируема, то $\sum_{i=1}^n x_i \frac{\partial f}{\partial x_i}(\mathbf{x}) = mf(\mathbf{x})$.

Множество линейных отображений $\mathbb{R}^n \to \mathbb{R}^m$ является линейным пространством. Будем обозначать это линейное пространство $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$. В пространствах $\mathbb{R}^n$ и $\mathbb{R}^m$ есть естественные базисы, поэтому пространство $\mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$ можно отождествить с пространством матриц размера $n \times m$. Размерность этого линейного пространства равна mn. Квадрат длины вектора в этом пространстве, соответствующего данной матрице A, — сумма квадратов всех элементов этой матрицы. Таким образом, $\|A\|^2$ — это сумма квадратов элементов матрицы A.

Задача 9.27. Докажите, что $\|A\mathbf{x}\| \leq \|A\| \cdot \|\mathbf{x}\|$.

В пространстве линейных отображений помимо евклидовой нормы $\|A\|$ часто используется *операторная* (или *спектральная*) норма $\|A\|_{\mathrm{s}} = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$.

Задача 9.28. Докажите, что $\|AB\|_{\mathrm{s}} \leq \|A\|_{\mathrm{s}}\|B\|_{\mathrm{s}}$.

Задача 9.29. а) Докажите, что $\|A\|_{\mathrm{s}} \leq \|A\|$.
б) Докажите, что если ранг матрицы A равен r, то $\|A\| \leq \sqrt{r}\|A\|_{\mathrm{s}}$.

Замечание. Из задачи 9.29, в частности, следует, что евклидова норма и спектральная норма задают на пространстве матриц одну и ту же топологию.

Теорема 9.3. *а) Пусть линейное отображение A невырожденное и $\|B - A\|_s < \|A^{-1}\|_s$. Тогда линейное отображение B тоже невырожденное.*
б) Множество невырожденных линейных отображений открыто.
в) Отображение $A \mapsto A^{-1}$ пространства невырожденных матриц в себя непрерывно.

Доказательство. а) Пусть $\|A^{-1}\|_{\mathrm{s}} = \frac{1}{\varepsilon}$ и $\|B - A\|_{\mathrm{s}} = \varepsilon_1$, где $\varepsilon_1 < \varepsilon$. Тогда $\|\mathbf{x}\| = \|A^{-1}A\mathbf{x}\| \leq \frac{1}{\varepsilon}\|A\mathbf{x}\|$, поэтому

$$(\varepsilon - \varepsilon_1)\|\mathbf{x}\| \leq \|A\mathbf{x}\| - \varepsilon_1\|\mathbf{x}\| \leq \|A\mathbf{x}\| - \|(B - A)\mathbf{x}\| \leq \|B\mathbf{x}\|. \tag{9.1}$$

Таким образом, если $\mathbf{x} \neq \mathbf{0}$, то $B\mathbf{x} \neq \mathbf{0}$, т.е. линейное отображение B невырожденное.

б) Непосредственно следует из а) и из того, что евклидова норма и спектральная норма задают на пространстве матриц одну и ту же топологию.

в) Положим в (9.1) $\mathbf{x} = B^{-1}\mathbf{y}$. В результате получим

$$(\varepsilon - \varepsilon_1)\|B^{-1}\mathbf{y}\| \leq \|BB^{-1}\mathbf{y}\| = \|\mathbf{y}\|.$$

Поэтому $\|B^{-1}\|_{\mathrm{s}} \leq (\varepsilon - \varepsilon_1)^{-1}$.

Запишем тождество $B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1}$ и применим к нему неравенство из задачи 9.28. В результате получим

$$\|B^{-1} - A^{-1}\|_{\mathrm{s}} \leq \|B^{-1}\|_{\mathrm{s}}\|A - B\|_{\mathrm{s}}\|A\|_{\mathrm{s}} \leq \frac{\varepsilon_1}{\varepsilon(\varepsilon - \varepsilon_1)}.$$

Поэтому $B^{-1} \to A^{-1}$ при $B \to A$, поскольку $\varepsilon_1 \to 0$. □

Задача 9.30. а) Докажите, что $\operatorname{grad}(\|A\|^2) = 2A$ для произвольной матрицы A.
б) Докажите, что $\operatorname{grad}(\det A) = (\det A)(A^{-1})^T$ для невырожденной матрицы A.

Задача 9.31. а) Пусть $f(A) = A^2$. Докажите, что $df_A(X) = AX + XA$.
б) Пусть $f(A) = A^T A$. Докажите, что $df_A(X) = A^T X + X^T A$.

Задача 9.32. Пусть $f(A) = A^{-1}$. Докажите, что $df_A(X) = -A^{-1}XA^{-1}$.

Задача 9.33. Пусть $f(A) = \det A$. Докажите, что: а) $df_I(X) = \operatorname{tr}(X)$; б) $df_A(X) = (\det A)\operatorname{tr}(A^{-1}X)$ для невырожденной матрицы A.

Будем говорить, что отображение $f\colon U \to \mathbb{R}^m$, где $U \subset \mathbb{R}^n$ — открытое подмножество, *непрерывно дифференцируемо* на U, если отображение $df\colon U \to \mathrm{L}(\mathbb{R}^n, \mathbb{R}^m)$, сопоставляющее точке $\mathbf{x} \in U$ линейное отображение $df_{\mathbf{x}}$, непрерывно. Непрерывно дифференцируемые отображения называют также C^1-отображениями (отображениями класса C^1). Используется также обозначение $f \in C^1(U)$.

По индукции можно определить отображения класса C^k: отображение f класса C^{k+1}, если отображение df класса C^k. Непрерывные отображения естественно называть отображениями класса C^0.

Теорема 9.4. *Отображение $f\colon U \to \mathbb{R}^m$, где U — открытое множество в $\mathbb{R}^n$, непрерывно дифференцируемо на U тогда и только тогда, когда все его частные производные непрерывны на U.*

Доказательство. Пусть $f \in C^1(U)$. Тогда согласно неравенству из задачи 9.27

$$\|df_{\mathbf{y}}(\mathbf{e}_i) - df_{\mathbf{x}}(\mathbf{e}_i)\| \leq \|df_{\mathbf{y}} - df_{\mathbf{x}}\| \cdot \|\mathbf{e}_i\| = \|df_{\mathbf{y}} - df_{\mathbf{x}}\|.$$

Из полученного неравенства следует, что каждая частная производная $D_i f_j(\mathbf{y}) = (df_j)_{\mathbf{y}}(\mathbf{e}_i)$ непрерывна на U.

Для доказательства обратного утверждения достаточно рассмотреть случай $m = 1$. Фиксируем $\mathbf{x} \in U$ и $\varepsilon > 0$. Множество U открыто, поэтому существует открытый шар D с центром $\mathbf{x}$ радиуса r, лежащий в U. Из непрерывности функций $D_j f$ следует, что радиус r можно выбрать столь малым, что для всех $\mathbf{y} \in D$ при $1 \leq j \leq n$ будут выполняться неравенства

$$|(D_j f)(\mathbf{y}) - (D_j f)(\mathbf{x})| < \frac{\varepsilon}{n}.$$

Пусть $\mathbf{h} = \sum h_j \mathbf{e}_j$. Положим $\mathbf{v}_0 = 0$ и $\mathbf{v}_k = h_1\mathbf{e}_1 + \dots + h_k\mathbf{e}_k$ при $1 \leq k \leq n$. Тогда

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \sum_{j=1}^{n} [f(\mathbf{x} + \mathbf{v}_j) - f(\mathbf{x} + \mathbf{v}_{j-1})].$$

Множество D выпуклое и $\|\mathbf{v}_k\| < r$ при $1 \leq k \leq n$, поэтому отрезок с концами $\mathbf{x} + \mathbf{v}_{j-1}$ и $\mathbf{x} + \mathbf{v}_j$ лежит в D. Поскольку $\mathbf{v}_j = \mathbf{v}_{j-1} + h_j \mathbf{e}_j$, по теореме о среднем значении (теорема 5.9 на с. 72) j-е слагаемое равно

$$h_j (D_j f)(\mathbf{x} + \mathbf{v}_{j-1} + \theta_j h_j \mathbf{e}_j)$$

для некоторого $\theta_j \in (0, 1)$. Поэтому j-е слагаемое отличается от $h_j(D_j f)(\mathbf{x})$ меньше, чем на $\frac{|h_j|\varepsilon}{n}$. Следовательно,

$$|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) - \sum_{j=1}^{n} h_j (D_j f)(\mathbf{x})| \leq \frac{1}{n} \sum_{j=1}^{n} |h_j| \varepsilon \leq \|\mathbf{h}\| \varepsilon$$

при всех $\mathbf{h}$, для которых $\|\mathbf{h}\| < r$.

Это означает, что функция f дифференцируема в точке $\mathbf{x}$ и $df_{\mathbf{x}}$ — линейная функция, принимающая на векторе $\mathbf{h} = \sum h_j \mathbf{e}_j$ значение $\sum h_j (D_j f)(\mathbf{x})$. Матрица этой линейной функции — строка $((D_1 f)(\mathbf{x}), \dots, (D_n f)(\mathbf{x}))$. Функции $D_1 f, \dots, D_n f$ непрерывны на E, поэтому дифференциал df непрерывен на E. □

9.3. Теорема о среднем значении

Докажем сначала теорему о среднем значении для функции n переменных. Для точек $\mathbf{a}$ и $\mathbf{b}$ в $\mathbb{R}^n$ будем обозначать $[\mathbf{a}, \mathbf{b}]$ отрезок с концами в этих точках

Теорема 9.5 (о среднем значении). *Пусть функция f дифференцируема на открытом множестве $U \subset \mathbb{R}^n$, содержащем отрезок $[\boldsymbol{a}, \boldsymbol{b}]$. Тогда*

$$f(\boldsymbol{b}) - f(\boldsymbol{a}) = df_{\boldsymbol{c}}(\boldsymbol{b} - \boldsymbol{a})$$

для некоторой точки $\boldsymbol{c} \in [\boldsymbol{a}, \boldsymbol{b}]$.

Доказательство. Рассмотрим отображение $\varphi \colon [0, 1] \to [\mathbf{a}, \mathbf{b}]$, заданное формулой $\varphi(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a}) = (1 - t)\mathbf{a} + t\mathbf{b}$. Это отображение дифференцируемо: его производная равна $\mathbf{b} - \mathbf{a}$. Поэтому отображение $g = f \circ \varphi$ тоже дифференцируемо. Отображение g — это функция одного переменного, поэтому можно применить обычную теорему о среднем значении (теорема 5.9 на с. 72) и получить, что $g(1) - g(0) = g'(\theta)$, для некоторого $\theta \in [0, 1]$. Положим $\mathbf{c} = \varphi(\theta)$. Тогда

$$\begin{aligned} f(\mathbf{b}) - f(\mathbf{a}) &= g(1) - g(0) = g'(\theta) = dg_\theta(1) = \\ &= df_{\varphi(\theta)} \circ d\varphi_\theta(1) = df_{\mathbf{c}}(\mathbf{b} - \mathbf{a}). \end{aligned}$$

□

Задача 9.34. а) Отображение F дифференцируемо на открытом множестве $U \subset \mathbb{R}^n$, содержащем отрезок $[\mathbf{a}, \mathbf{b}]$, и $dF_{\mathbf{x}}(\mathbf{b}-\mathbf{a}) = \mathbf{0}$ для всех $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Докажите, что $F(\mathbf{b}) = F(\mathbf{a})$.

б) Отображение F дифференцируемо на открытом выпуклом множестве $U \subset \mathbb{R}^n$ и $dF_{\mathbf{x}} = 0$ для всех $\mathbf{x} \in U$. Докажите, что отображение F постоянное.

Задача 9.35. Пусть отображение $F = (f_1, \dots, f_m)$ дифференцируемо на открытом множестве $U \subset \mathbb{R}^n$, содержащем отрезок $[\mathbf{a}, \mathbf{b}]$, и на множестве U частные производные отображения f по абсолютной величине не превосходят числа M. Докажите, что тогда $\|F(\mathbf{b}) - F(\mathbf{a})\| \le nmM\|\mathbf{b}-\mathbf{a}\|$.

Для доказательства теоремы 9.7 о том, что $D_iD_jf = D_jD_if$, нам понадобится аналог теоремы 9.5 для вторых разностей. При этом вместо первой разности $\Delta f_{\mathbf{a}}(\mathbf{h}) = f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})$ рассматривается вторая разность $\Delta^2 f_{\mathbf{a}}(\mathbf{h}, \mathbf{k}) = \Delta f_{\mathbf{a}+\mathbf{k}}(\mathbf{h}) - \Delta f_{\mathbf{a}}(\mathbf{h}) = f(\mathbf{a}+\mathbf{h}+\mathbf{k}) - f(\mathbf{a}+\mathbf{k}) - f(\mathbf{a}+\mathbf{h}) + f(\mathbf{a}) = \Delta^2 f_{\mathbf{a}}(\mathbf{k}, \mathbf{h})$.

Теорема 9.6. *Пусть открытое множество $U \subset \mathbb{R}^n$ содержит параллелограмм с вершинами $\boldsymbol{a}$, $\boldsymbol{a}+\boldsymbol{h}$, $\boldsymbol{a}+\boldsymbol{k}$ и $\boldsymbol{a}+\boldsymbol{h}+\boldsymbol{k}$. Если функции f и $D_{\boldsymbol{h}}f$ дифференцируемы на U, то существуют числа $\alpha, \beta \in (0,1)$, для которых выполняется равенство*

$$\Delta^2 f_{\boldsymbol{a}}(\boldsymbol{h}, \boldsymbol{k}) = D_{\boldsymbol{k}} D_{\boldsymbol{h}} f(\boldsymbol{a} + \alpha \boldsymbol{h} + \beta \boldsymbol{k}).$$

Доказательство. Рассмотрим на отрезке $[\mathbf{a}, \mathbf{a}+\mathbf{h}]$ функцию $g(\mathbf{x}) = f(\mathbf{x}+\mathbf{k}) - f(\mathbf{x})$. Эта функция дифференцируема: $dg_{\mathbf{x}} = df_{\mathbf{x}+\mathbf{k}} - df_{\mathbf{x}}$. Согласно теореме 9.5 $\Delta^2 f_{\mathbf{a}}(\mathbf{h}, \mathbf{k}) = g(\mathbf{a}+\mathbf{h}) - g(\mathbf{a}) = dg_{\mathbf{a}+\alpha\mathbf{h}}(\mathbf{h})$ для некоторого $\alpha \in (0,1)$. Далее, $dg_{\mathbf{a}+\alpha\mathbf{h}}(\mathbf{h}) = df_{\mathbf{a}+\alpha\mathbf{h}+\mathbf{k}}(\mathbf{h}) - df_{\mathbf{a}+\alpha\mathbf{h}}(\mathbf{h}) = D_{\mathbf{h}}f(\mathbf{a}+\alpha\mathbf{h}+\mathbf{k}) - D_{\mathbf{h}}f(\mathbf{a}+\alpha\mathbf{h})$. Применив ещё раз теорему 9.5, получим, что $D_{\mathbf{h}}f(\mathbf{a}+\alpha\mathbf{h}+\mathbf{k}) - D_{\mathbf{h}}f(\mathbf{a}+\alpha\mathbf{h}) = d(D_{\mathbf{h}}f)_{\mathbf{k}}(\mathbf{a}+\alpha\mathbf{h}+\beta\mathbf{k})$ для некоторого $\beta \in (0,1)$. Наконец, $d(D_{\mathbf{h}}f)_{\mathbf{a}+\alpha\mathbf{h}+\beta\mathbf{k}}(\mathbf{k}) = D_{\mathbf{k}}D_{\mathbf{h}}f(\mathbf{a}+\alpha\mathbf{h}+\beta\mathbf{k})$. □

Теорема 9.7. *Если f — функция класса C^2 на открытом множестве $U \subset \mathbb{R}^n$, то $D_iD_jf = D_jD_if$ для любых индексов i и j.*

Доказательство. Рассмотрим точку $\mathbf{a} \in U$. Для достаточно малых чисел h и k параллелограмм с вершинами $\mathbf{a}$, $\mathbf{a}+h\mathbf{e}_i$, $\mathbf{a}+k\mathbf{e}_j$ и $\mathbf{a}+h\mathbf{e}_i+k\mathbf{e}_j$ содержится в U. Поэтому, воспользовавшись теоремой 9.6, получим, что

$$\Delta^2 f_{\mathbf{a}}(h\mathbf{e}_i, k\mathbf{e}_j) = D_{k\mathbf{e}_j} D_{h\mathbf{e}_i} f(\mathbf{a} + \alpha_1 h\mathbf{e}_i + \beta_1 k\mathbf{e}_j)$$

для некоторых чисел $\alpha_1, \beta_1 \in (0,1)$. Теорему 9.6 можно ещё раз применить, поменяв местами $h\mathbf{e}_i$ и $k\mathbf{e}_j$:

$$\Delta^2 f_{\mathbf{a}}(k\mathbf{e}_j, h\mathbf{e}_i) = D_{h\mathbf{e}_i} D_{k\mathbf{e}_j} f(\mathbf{a} + \alpha_2 h\mathbf{e}_i + \beta_2 k\mathbf{e}_j)$$

для некоторых чисел $\alpha_2, \beta_2 \in (0,1)$.

Непосредственно из определения $\Delta^2 f$ следует, что $\Delta^2 f_{\mathbf{a}}(h\mathbf{e}_i, k\mathbf{e}_j) = \Delta^2 f_{\mathbf{a}}(k\mathbf{e}_j, h\mathbf{e}_i)$. Ясно также, что $D_{k\mathbf{e}_j} D_{h\mathbf{e}_i} = khD_jD_i$. Следовательно,

$$D_jD_if(\mathbf{a} + \alpha_1 h\mathbf{e}_i + \beta_1 k\mathbf{e}_j) = D_iD_jf(\mathbf{a} + \alpha_2 h\mathbf{e}_i + \beta_2 k\mathbf{e}_j).$$

Устремляя h и k к нулю, получаем требуемое. □

Историческое замечание. В 1721 г. Николай Бернулли (1687-1759) обнаружил, что результат дифференцирования по нескольким переменным не зависит от порядка дифференцирования. Первым это утверждение доказал Леонард Эйлер (1707-1783) в 1734 г. Независимое доказательство получил Алексис Клод Клеро (1713-1765) в 1739 г.

Задача 9.36. Приведите пример функции $f(x,y)$, у которой вторые частные производные $\frac{\partial}{\partial x}\frac{\partial}{\partial y}f$ и $\frac{\partial}{\partial y}\frac{\partial}{\partial x}f$ существуют, но не равны.

Следующий пример показывает, что для отображений в многомерное пространство теорема о среднем значении в виде равенства неверна.

Пример 9.3. *Рассмотрим отображение $f\colon [0, 2\pi] \to \mathbb{R}^3$, заданное формулой $f(t) = (\cos t, \sin t, t)$. Ни для какого $t \in (0,1)$ не может выполняться равенство $f(1) - f(0) = df_t(1)$.*

Доказательство. Матрица линейного отображения df_t равна $(-\sin t, \cos t, 1)$, поэтому вектор $df_t(1)$ равен $(-\sin t, \cos t, 1)$. Разность $f(1) - f(0)$ равна $(1,0,1)-(1,0,0)=(0,0,1)$. Но $\sin t$ и $\cos t$ не могут быть одновременно равны нулю. □

Но для отображений в многомерное пространство можно доказать теорему о среднем значении в виде неравенства. Одно из неравенств такого вида приведено в задаче 9.35.

Теорема 9.8 (о среднем значении). *Пусть $U \subset \mathbb{R}^n$ — выпуклое множество, $F\colon U \to \mathbb{R}^m$ — дифференцируемое отображение. Предположим, что $\|dF_{\boldsymbol{x}}\| \le L$ для всех $\boldsymbol{x} \in U$. Тогда $\|F(\boldsymbol{b}) - F(\boldsymbol{a})\| \le L\|\boldsymbol{b}-\boldsymbol{a}\|$ для всех $\boldsymbol{a}, \boldsymbol{b} \in U$.*

Доказательство. Для любого вектора $v \in \mathbb{R}^m$ можно рассмотреть функцию $f(\mathbf{x}) = (v, F(\mathbf{x}))$ и применить к ней теорему 9.5. В результате получим $(v, F(\mathbf{b}) - F(\mathbf{a})) = f(\mathbf{b}) - f(\mathbf{a}) = df_{\mathbf{c}}(\mathbf{b}-\mathbf{a}) = (v, dF_{\mathbf{c}}(\mathbf{b}-\mathbf{a}))$ для некоторой точки $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$.

Напомним, что $\|A\mathbf{x}\| \le \|A\| \cdot \|\mathbf{x}\|$ (задача 9.27). Поэтому, положив $v = F(\mathbf{b}) - F(\mathbf{a})$, получим

$$\begin{aligned}
\|F(\mathbf{b}) - F(\mathbf{a})\|^2 &= (F(\mathbf{b}) - F(\mathbf{a}), dF_{\mathbf{c}}(\mathbf{b}-\mathbf{a})) \le \\
&\le \|F(\mathbf{b}) - F(\mathbf{a})\| \cdot \|dF_{\mathbf{c}}\| \cdot \|\mathbf{b}-\mathbf{a}\| \le \\
&\le L\|F(\mathbf{b}) - F(\mathbf{a})\| \cdot \|\mathbf{b}-\mathbf{a}\|.
\end{aligned}$$

После сокращения на $\|F(\mathbf{b}) - F(\mathbf{a})\|$ получаем требуемое. (Если $F(\mathbf{b}) = F(\mathbf{a})$, то доказывать нечего.) □

9.4. Формула Тейлора

Пусть f — функция класса C^m на открытом множестве $U \subset \mathbb{R}^n$, содержащем отрезок $[\mathbf{a}, \mathbf{a}+\mathbf{h}]$. Положим $g(t) = f(\mathbf{a}+t\mathbf{h})$ для $t \in [0,1]$. Для вывода многомерной формулы Тейлора нам нужно вычислить $g^{(k)}(t)$. Покажем, что

$$\begin{aligned}
g^{(k)}(t) &= (h_1\partial_1 + \dots + h_n\partial_n)^k f(\mathbf{a}+t\mathbf{h}) = \\
&= \sum_{j_1+\dots+j_n=k} \binom{k}{j_1 \dots j_n} \partial_1^{j_1} \dots \partial_n^{j_n} f(\mathbf{a}+t\mathbf{h}) h_1^{j_1} \dots h_n^{j_n}
\end{aligned}$$

для $k \le m$. Применим индукцию по k. Прежде всего,

$$\begin{aligned}
g'(t) &= dg_t(1) = df_{\mathbf{a}+t\mathbf{h}} \circ d(\mathbf{a}+t\mathbf{h})_t(1) = df_{\mathbf{a}+t\mathbf{h}}(\mathbf{h}) = \\
&= D_{\mathbf{h}} f(\mathbf{a}+t\mathbf{h}) = (h_1\partial_1 + \dots + h_n\partial_n) f(\mathbf{a}+t\mathbf{h}).
\end{aligned}$$

Докажем теперь шаг индукции. Пусть $f_1(\mathbf{x}) = D_{\mathbf{h}}^k f(\mathbf{x})$. Тогда $g^{(k+1)}(t) = (df_1)_{\mathbf{a}+t\mathbf{h}}(\mathbf{h}) = D_{\mathbf{h}} f_1(\mathbf{a}+t\mathbf{h}) = D_{\mathbf{h}} D_{\mathbf{h}}^k f(\mathbf{a}+t\mathbf{h}) = D_{\mathbf{h}}^{k+1} f(\mathbf{a}+t\mathbf{h})$.

Для вектора $\mathbf{h} = (h_1, \dots, h_n)$ рассмотрим многочлен

$$\begin{aligned}
P_{m-1}(\mathbf{h}) &= \sum_{k=0}^{m-1} \frac{1}{k!} (h_1\partial_1 + \dots + h_n\partial_n)^k f(\mathbf{a}) = \\
&= f(\mathbf{a}) + \partial_1 f(\mathbf{a}) h_1 + \dots + \partial_n f(\mathbf{a}) h_n + \dots + \\
&\quad + \sum_{j_1+\dots+j_n=m-1} \binom{m-1}{j_1 \dots j_n} \partial_1^{j_1} \dots \partial_n^{j_n} f(\mathbf{a}) h_1^{j_1} \dots h_n^{j_n}.
\end{aligned}$$

Этот многочлен называют *многочленом Тейлора* степени $m-1$ функции f в точке $\mathbf{a}$. Остаточный член $R_{m-1}(\mathbf{a}; \mathbf{h})$ — это разность $f(\mathbf{a}+\mathbf{h}) - P_{m-1}(\mathbf{h})$, т.е. $f(\mathbf{a}+\mathbf{h}) = P_{m-1}(\mathbf{h}) + R_{m-1}(\mathbf{a}; \mathbf{h})$.

Остаточный член можно представить в разных формах. Получим сначала представление, аналогичное теореме 5.12 (см. с. 75): на отрезке $[\mathbf{a}, \mathbf{a}+\mathbf{h}]$ найдётся точка $\mathbf{c}$, для которой

$$R_{m-1}(\mathbf{a}; \mathbf{h}) = \frac{1}{m!} (h_1\partial_1 + \dots + h_n\partial_n)^m f(\mathbf{c}) = \frac{D_{\mathbf{h}}^m f(\mathbf{c})}{m!}.$$

Если $g(t) = f(\mathbf{a} + t\mathbf{h})$ для $t \in [0, 1]$, то $g(0) = f(\mathbf{a})$, $g(1) = f(\mathbf{a} + \mathbf{h})$ и, согласно формуле Тейлора для функции одного переменного,

$$g(1) = g(0) + \sum_{k=1}^{m-1} \frac{g^{(k)}(0)}{k!} + \frac{g^{(m)}(\theta)}{m!}$$

для некоторой точки $\theta \in (0, 1)$. Кроме того, $g^{(k)}(t) = D^k_{\mathbf{h}} f(\mathbf{a} + t\mathbf{h})$ для $k \leq m$.

Получим теперь интегральное представление

$$R_{m-1}(\mathbf{a}; \mathbf{h}) = \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} (h_1\partial_1 + \cdots + h_n\partial_n)^{m-1} f(\mathbf{a} + t\mathbf{h}) dt.$$

Оно непосредственно следует из интегрального представления остаточного члена для функции одной переменной (см. § 6.7). Положим $g(t) = f(\mathbf{a} + t\mathbf{h})$ для $t \in [0, 1]$. Это функция класса C^m на отрезке $[0, 1]$. Поэтому для каждого $\tau \in [0, 1]$ имеем равенство

$$g(\tau) = \sum_{k=0}^{m-1} \frac{g^{(k)}(0)}{k!} \tau^k + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} g^{(m)}(t\tau) \tau^m dt.$$

При $\tau = 1$ получаем, в частности,

$$g(1) = \sum_{k=0}^{m-1} \frac{g^{(k)}(0)}{k!} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} g^{(m)}(t) dt.$$

Кроме того,

$$\begin{array}{rcl} g^{(k)}(0) & = & (h_1\partial_1 + \cdots + h_n\partial_n)^k f(\mathbf{a}) \quad (k = 0, \ldots, m-1), \\ g^{(m)}(t) & = & (h_1\partial_1 + \cdots + h_n\partial_n)^k f(\mathbf{a} + t\mathbf{h}). \end{array}$$

Историческое замечание. Формулу Тейлора на функции многих переменных обобщил Жозеф-Луи Лагранж (1736-1813) в 1772 году.

Формула Тейлора позволяет выяснить, является ли невырожденная критическая точка функции f точкой локального максимума или локального минимума. Точка $\mathbf{a}$ *критическая*, если все частные производные $D_i f(\mathbf{a})$ равны нулю; критическая точка *невырожденная*, если матрица $(D_i D_j f(\mathbf{a}))$ невырожденная. Для невырожденной критической точки $\mathbf{a}$ можно рассмотреть квадратичную форму

$$q(\mathbf{h}) = \frac{1}{2} D^2_{\mathbf{h}} f(\mathbf{a}) = \frac{1}{2} \sum D_i D_j f(\mathbf{a}) h_i h_j.$$

Теорема 9.9. *Пусть f — функция класса C^3 в окрестности невырожденной критической точки $\mathbf{a}$. Если квадратичная форма q положительно определённая, то функция f имеет локальный минимум в точке a; если квадратичная форма отрицательно определённая, то функция имеет локальный минимум; если форма неопределённая, то функция не имеет в точке $\mathbf{a}$ ни локального максимума, ни локального минимума.*

Доказательство. Рассмотрим сначала случай, когда квадратичная форма $q(\mathbf{h})$ положительно определённая. Поскольку $f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) = q(\mathbf{h}) + R_2(\mathbf{h})$, достаточно выбрать $\delta > 0$ так, что если $0 < \|\mathbf{h}\| < \delta$, то $\frac{q(\mathbf{h}) + R_2(\mathbf{h})}{\|\mathbf{h}\|^2} > 0$. Ясно, что

$$\frac{q(\mathbf{h}) + R_2(\mathbf{h})}{\|\mathbf{h}\|^2} = q\left(\frac{\mathbf{h}}{\|\mathbf{h}\|}\right) + \frac{R_2(\mathbf{h})}{\|\mathbf{h}\|^2}.$$

Вектор $\frac{\mathbf{h}}{\|\mathbf{h}\|}$ лежит на единичной сфере. Непрерывная функция q на единичной сфере достигает минимума m, и этот минимум положителен. Кроме того, $\lim_{\mathbf{h} \to \mathbf{0}} \frac{R_2(\mathbf{h})}{\|\mathbf{h}\|^2} = 0$.

Остальные случаи разбираются аналогично. □

Задача 9.37. *Ограничение функции $f(x, y)$ на любую прямую, проходящую через начало координат, имеет в начале координат локальный минимум. Обязательно ли эта функция имеет локальный минимум в начале координат?

9.5. Метод множителей Лагранжа

Классический метод множителей Лагранжа предназначен для нахождения локального экстремума функции $f(\mathbf{x})$, где $\mathbf{x} \in \mathbb{R}^n$, при условии, что $g_1(\mathbf{x}) = 0, \ldots, g_k(\mathbf{x}) = 0$. Этот метод заключается в следующем. Составим функцию Лагранжа

$$\mathcal{L}(\mathbf{x}, \lambda_0, \ldots, \lambda_k) = \lambda_0 f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \cdots + \lambda_k g_k(\mathbf{x}).$$

Предположим, что $\mathbf{x}_0$ — точка локального экстремума функции $f(\mathbf{x})$, и в окрестности этой точки функция f непрерывно дифференцируема. Тогда найдутся числа $\lambda_0, \ldots, \lambda_n$, не все равные нулю, для которых выполняются равенства $\frac{\partial \mathcal{L}(\mathbf{x}_0,\lambda)}{\partial x_j} = 0$, т.е.

$$\lambda_0 \frac{\partial f(\mathbf{x}_0, \lambda)}{\partial x_j} + \lambda_1 \frac{\partial g_1(\mathbf{x}_0, \lambda)}{\partial x_j} + \cdots + \lambda_k \frac{\partial g_k(\mathbf{x}_0, \lambda)}{\partial x_j} = 0, \tag{9.2}$$

для $j = 1, \ldots, n$. Вместе с уравнениями $g_1(\mathbf{x}) = 0, \ldots, g_k(\mathbf{x}) = 0$ получаем $n+k$ уравнений с $n+k+1$ неизвестными. Но множители Лагранжа $\lambda_0, \ldots, \lambda_n$ можно умножать на любую константу, отличную от нуля, поэтому можно считать, что один из множителей Лагранжа равен 1, и фактически у нас $n+k$ уравнений с $n+k$ неизвестными.

Если $\lambda_0 = 0$, то равенства (9.2) означают лишь вырожденность ограничений (линейную зависимость градиентов функций $g_1, \ldots, g_k$ в точке минимума) и не связаны с функцией, экстремум которой ищется. Но правило множителей Лагранжа с дополнительным предположением $\lambda_0 \neq 0$ может оказаться неверным.

Рассмотрим, например, задачу о поиске минимума функции $f(x, y) = x$ при условии $g(x, y) = y^2 - x^3 = 0$. Единственная точка минимума — это точка $(0, 0)$. Попробуем составить функцию Лагранжа с $\lambda_0 = 1$:

$$\mathcal{L}(x, y, \lambda) = x + \lambda(y^2 - x^3).$$

Приравнивая нулю производные по x и по y, получаем $1 - 3\lambda x^2 = 0$ и $2\lambda y = 0$. Кроме того, $y^2 - x^3 = 0$. Полученные уравнения несовместны.

Одно из доказательств метода множителей Лагранжа основано на теореме о неявной функции, поэтому мы ещё вернёмся к этому методу в § 10.5. А сейчас обсудим технически более сложное доказательство, не использующее теорему о неявной функции. При этом будет доказана более общая теорема, когда помимо ограничений в виде равенств есть ограничения в виде неравенств.

При нахождении минимума функции $f(\mathbf{x})$ помимо ограничений в виде равенств $g_1(\mathbf{x}) = 0, \ldots, g_k(\mathbf{x}) = 0$ могут быть и ограничения в виде неравенств $h_1(\mathbf{x}) \leq 0, \ldots, h_s(\mathbf{x}) \leq 0$. В таком случае условием Куна–Таккера для точки минимума $\mathbf{x}_0$ называют следующее условие: градиенты всех функций g_i и тех функций h_j, для которых $h_j(\mathbf{x}_0) = 0$, линейно независимы в точке $\mathbf{x}_0$.

Доказательство теоремы 9.10 следует статье [McSh]. При доказательстве этой теоремы мы будем использовать следующее обозначение: $f^+(\mathbf{x}) = \max(f(\mathbf{x}), 0)$.

Теорема 9.10. *Пусть функции $f, g_1, \ldots, g_k, h_1, \ldots, h_s$ непрерывно дифференцируемы на множестве $G \subset \mathbb{R}^n$, $\boldsymbol{x}_0$ — внутренняя точка множества G, для которой выполняются равенства $g_1(\boldsymbol{x}) = 0, \ldots, g_k(\boldsymbol{x}) = 0$ и неравенства $h_1(\boldsymbol{x}) \leq 0, \ldots, h_s(\boldsymbol{x}) \leq 0$, причём $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$ для любой точки $\boldsymbol{x} \in G$, для которой выполняются эти равенства и неравенства. Тогда существуют числа $\lambda_0, \lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_s$, не все равные нулю, для которых*

$$\lambda_0 D_j f(\boldsymbol{x}_0) + \sum_{i=1}^{k} \lambda_i D_j g_i(\boldsymbol{x}_0) + \sum_{r=1}^{z} \mu_r D_j h_r(\boldsymbol{x}_0) = 0 \tag{9.3}$$

при $j = 1, \ldots, n$. Более того:

а) $\lambda_0 \geq 0$ и $\mu_1 \geq 0, \ldots, \mu_r \geq 0$;

б) $\mu_r = 0$ для всех r, для которых $h_r(\boldsymbol{x}_0) < 0$;

в) если выполняется условие Куна–Таккера, то можно считать, что $\lambda_0 = 1$.

Доказательство. Без ограничения общности можно считать, что $\mathbf{x}_0 = (0, \ldots, 0)$, $f(\mathbf{x}_0) = 0$ и $h_1(\mathbf{x}_0) = 0, \ldots, h_z(\mathbf{x}_0) = 0$, $h_{z+1}(\mathbf{x}_0) < 0, \ldots, h_s(\mathbf{x}_0) < 0$. Выберем $\varepsilon_1 > 0$ так, что замкнутый шар радиуса ε_1 с центром в начале координат содержится в G и функции $h_{z+1}, \ldots, h_s$ отрицательны на этом шаре. Прежде всего докажем следующее утверждение.

Для каждого $\varepsilon \in (0, \varepsilon_1]$ *можно выбрать положительное число* N *так, что*

$$f(\boldsymbol{x}) + \|\boldsymbol{x}\|^2 + N\left(\sum_{i=1}^{k} g_i(\boldsymbol{x})^2 + \sum_{r=1}^{z} h_r^+(\boldsymbol{x})^2\right) > 0$$

для всех $\boldsymbol{x}$*, для которых* $\|\boldsymbol{x}\| = \varepsilon$*.*

Предположим, что это утверждение неверно. Тогда существует последовательность положительных чисел $N_1, N_2, \ldots$, стремящаяся к ∞, и на сфере $\|\mathbf{x}\| = \varepsilon$ существует последовательность точек $\mathbf{x}_1, \mathbf{x}_2, \ldots$, для которых выполняются неравенства

$$f(\mathbf{x}_m) + \|\mathbf{x}_m\|^2 < -N_m\left(\sum_{i=1}^{k} g_i(\mathbf{x}_m)^2 + \sum_{r=1}^{z} h_r^+(\mathbf{x}_m)^2\right). \tag{9.4}$$

Из последовательности точек на сфере можно выбрать сходящуюся подпоследовательность. Без ограничения общности можно считать, что исходная последовательность $\mathbf{x}_m$ сходится к некоторой точке $\mathbf{x}^*$. При этом $\|\mathbf{x}^*\| = \lim_{m\to\infty}\|\mathbf{x}_m\| = \varepsilon$ и $\lim_{m\to\infty} f(\mathbf{x}_m) = f(\mathbf{x}^*)$. Поделим обе части неравенства (9.4) на $-N_m$ и устремим n к бесконечности. В результате получим

$$\sum_{i=1}^{k} g_i(\mathbf{x}^*)^2 + \sum_{r=1}^{z} h_r^+(\mathbf{x}^*)^2 \le \lim_{m\to\infty} \frac{f(\mathbf{x}_m) + \|\mathbf{x}_m\|^2}{-N_m} = 0.$$

(Мы воспользовались тем, что предел последовательности дробей, числители которых ограничены, а знаменатели стремятся к $-\infty$, равен 0.) Таким образом, для точки $\mathbf{x}^*$ выполняются все требуемые равенства и неравенства, поэтому $\lim_{m\to\infty} f(\mathbf{x}_m) = f(\mathbf{x}^*) \ge f(\mathbf{x}_0) = 0$. Но из неравенства (9.4) следует, что $f(\mathbf{x}_m) \le -\varepsilon^2$. Приходим к противоречию.

Докажем ещё одно утверждение.

Для каждого $\varepsilon \in (0, \varepsilon_1]$ *можно выбрать точку* $\bar{\boldsymbol{x}}$ *и единичный вектор* $(\lambda_0, \ldots, \lambda_k, \mu_1, \ldots, \mu_z)$ *с неотрицательными координатами* λ_0*,* $\mu_1, \ldots, \mu_z$ *так, что* $\|\bar{\boldsymbol{x}}\| < \varepsilon$ *и*

$$\lambda_0(D_j f(\bar{\boldsymbol{x}}) + 2\bar{x}^j) + \sum_{i=1}^{k} \lambda_i D_j g_i(\bar{\boldsymbol{x}}) + \sum_{r=1}^{z} \mu_r D_j h_r(\bar{\boldsymbol{x}}) = 0 \tag{9.5}$$

для $j = 1, \ldots, n$ *(здесь* $\bar{x}^j$ *—* j*-я координата точки* $\bar{\boldsymbol{x}}$*).*

Для данного ε выберем N, как в первом утверждении, и рассмотрим на множестве G функцию F, заданную формулой

$$F(\mathbf{x}) = f(\mathbf{x}) + \|\mathbf{x}\|^2 + N\left(\sum_{i=1}^{k} g_i(\mathbf{x})^2 + \sum_{r=1}^{z} h_r^+(\mathbf{x})^2\right).$$

На замкнутом шаре D_ε радиуса ε с центром в начале координат функция F принимает наименьшее значение в некоторой точке $\bar{\mathbf{x}}$. Ясно, что $F(\bar{\mathbf{x}}) \le F(\mathbf{0}) = 0$, поэтому из неравенства (9.4) следует, что $\|\bar{\mathbf{x}}\| \ne \varepsilon$. Таким образом, $\bar{\mathbf{x}}$ — внутренняя точка шара D_ε, и в этой точке все частные производные $D_j F$ обращаются в нуль. В точке $\bar{\mathbf{x}}$ производная D_j функции $h_r^+(\mathbf{x})^2$ равна

$$2h_r^+(\bar{\mathbf{x}}) D_j h_r(\bar{\mathbf{x}}).$$

Действительно, если $h_r^+(\bar{\mathbf{x}}) > 0$, то это очевидно, а если $h_r^+(\bar{\mathbf{x}}) = 0$, то в точке $\bar{\mathbf{x}}$ функция $h_r^+(\mathbf{x})^2$ имеет нуль по крайней мере второго порядка. Следовательно,

$$D_j f(\bar{\mathbf{x}}) + 2\bar{x}^j + \sum_{i=1}^{k} 2N g_i(\bar{\mathbf{x}}) D_j g_i(\bar{\mathbf{x}}) + \sum_{r=1}^{z} 2N h_r^+(\bar{\mathbf{x}}) D_j h_r(\bar{\mathbf{x}}) = 0 \tag{9.6}$$

для $j = 1, \ldots, n$. Положим

$$\begin{aligned} L &= \left(1 + \sum_{i=1}^{k}(2Ng_i(\bar{\mathbf{x}}))^2 + \sum_{r=1}^{z}(2Nh_r^+(\bar{\mathbf{x}}))^2\right)^{\frac{1}{2}}, \\ \lambda_0 &= \frac{1}{L}, \\ \lambda_i &= \frac{2Ng_i(\bar{\mathbf{x}})}{L} \quad (i = 1, \ldots, k), \\ \mu_r &= \frac{2Nh_r^+(\bar{\mathbf{x}})}{L} \quad (r = 1, \ldots, z), \\ \mu_r &= 0 \quad (r = z+1, \ldots, s). \end{aligned}$$

Тогда $(\lambda_0, \ldots, \lambda_k, \mu_1, \ldots, \mu_s)$ — единичный вектор с неотрицательными координатами λ_0, $\mu_1, \ldots, \mu_s$. Поделив обе части равенства (9.6) на L, получим требуемое равенство.

Рассмотрим теперь последовательность положительных чисел $\varepsilon_1 > \varepsilon_2 > \varepsilon_3 > \ldots$, сходящуюся к нулю, и выберем для каждого числа ε_m точку $\bar{\mathbf{x}}_m$, для которой $\|\bar{\mathbf{x}}_m\| < \varepsilon_m$, и единичный вектор $(\lambda_{0,m}, \ldots, \lambda_{k,m}, \mu_{1,m}, \ldots, \mu_{z,m}, 0, \ldots, 0)$ так, чтобы выполнялись равенства (9.5) (с естественными необходимыми изменениями обозначений). Выберем подпоследовательность, для которой единичные векторы $(\lambda_{0,m}, \ldots, \lambda_{k,m}, \mu_{1,m}, \ldots, \mu_{z,m}, 0, \ldots, 0)$ стремятся к некоторому пределу $(\lambda_0, \ldots, \lambda_k, \mu_1, \ldots, \mu_s)$. Ясно, что $\bar{\mathbf{x}}_m \to \mathbf{x}_0$, поэтому для точки $\mathbf{x}_0$ выполняются равенства (9.5) (с заменой $\bar{\mathbf{x}}$ на $\mathbf{x}_0$). Тем самым доказано всё, кроме утверждения в). Докажем теперь его.

Если выполняется условие Куна–Таккера, то $\lambda_0 \neq 0$, поскольку при $\lambda = 0$ равенства (9.3) противоречат линейной независимости градиентов функций. □

Историческое замечание. Жозеф-Луи Лагранж (1736-1813) разработал метод множителей в 1797 году. Леонард Эйлер (1707-1783) применял этот метод уже в 1732 году. Метод множителей Лагранжа обобщили на случай, когда помимо ограничений в виде равенств есть ограничения в виде неравенств, Гарольд Кун (1925-2014) и Альберт Таккер (1905-1995) в 1951 году.

9.6. Лемма Адамара

Теорема 9.11 (лемма Адамара). *Пусть U — выпуклая окрестность начала координат в $\mathbb{R}^n$, $f \in C^1(U)$ и $f(\boldsymbol{0}) = \boldsymbol{0}$. Тогда существуют такие непрерывные функции $g_1, \ldots, g_n$, что $f(\boldsymbol{x}) = \sum x_i g_i(\boldsymbol{x})$ и $g_i(\boldsymbol{0}) = \frac{\partial f}{\partial x_i}(\boldsymbol{0})$.*

Доказательство. Ясно, что $\frac{df(t\mathbf{x})}{dt}dt = \sum \frac{\partial f(t\mathbf{x})}{\partial x_i} \cdot \frac{d(tx_i)}{dt} = \sum x_i \frac{\partial f(t\mathbf{x})}{\partial x_i}$, поэтому

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}) - f(\mathbf{0}) = f(t\mathbf{x})\big|_{t=0}^{t=1} = \\ &= \int_0^1 \frac{df(t\mathbf{x})}{dt}dt = \int_0^1 \sum x_i \frac{\partial f(t\mathbf{x})}{\partial x_i} dt. \end{aligned}$$

Таким образом, можно положить $g_i(\mathbf{x}) = \int_0^1 \frac{\partial f(t\mathbf{x})}{\partial x_i} dt$. □

9.7. Решения задач

9.1. Рассмотрим непрерывную функцию $\varphi(x)$, которая равна 1 при $x \leq R$, равна 0 при $x \geq 2R$, и линейна при $x \in [R, 2R]$. Функция $f(\mathbf{x}) = \varphi(\|\mathbf{x}\|)$ обладает требуемыми свойствами.

9.2. Доказательство теоремы Больцано–Вейерштрасса (теорема 2.5) легко переносится на многомерный случай.

9.3. а) Пусть множества A и B связны и $p \in A \cap B$. Предположим, что $A \cup B = C \cup D$, где C и D — открытые непересекающиеся множества. Множество A связно, поэтому оно целиком лежит в одном из множеств C или D. Пусть для определённости $p \in C$. Тогда $A \subset C$. Аналогично $B \subset C$. Следовательно, $D = \varnothing$.

б) Для любой точки $a \times b \in X \times Y$ множества $a \times Y$ и $X \times b$ связны и имеют общую точку $a \times b$, поэтому множество $(a \times Y) \cup (X \times b)$ связно. Заменив точку a на произвольную точку $x \in X$, получим, что множество $T_x = (x \times Y) \cup (X \times b)$ связно. Ясно также, что $a \times b \in T_x$ для всех x. Объединение любого семейства связных множеств, имеющих общую точку, связно; это доказывается точно так же, как и для двух множеств. Поэтому множество $\bigcup_{x \in X} T_x = X \times Y$ связно.

9.4. Пусть $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Тогда $d(\mathbf{x}, A) = \inf_{\mathbf{a} \in A} \|\mathbf{x} - \mathbf{a}\| \leq \|\mathbf{x} - \mathbf{y}\| + \inf_{\mathbf{a} \in A} \|\mathbf{y} - \mathbf{a}\| = \|\mathbf{x} - \mathbf{y}\| + d(\mathbf{y}, A)$, т.е. $d(\mathbf{x}, A) - d(\mathbf{y}, A) \leq \|\mathbf{x} - \mathbf{y}\|$. Аналогично доказывается, что $d(\mathbf{y}, A) - d(\mathbf{x}, A) \leq \|\mathbf{x} - \mathbf{y}\|$. Следовательно, $|f(\mathbf{x}) - f(\mathbf{y})| \leq \|\mathbf{x} - \mathbf{y}\|$, поэтому функция f непрерывна.

9.5. а) Если множество A замкнуто, то множество $\mathbb{R}^n \setminus A$ открыто. Поэтому для любой точки $\mathbf{x}_0 \in \mathbb{R}^n \setminus A$ найдётся такое $\delta > 0$, что шар радиуса δ с центром $\mathbf{x}_0$ принадлежит множеству $\mathbb{R}^n \setminus A$. В таком случае $d(\mathbf{x}, A) \geq \delta > 0$.

б) Замкнутое множество A является множеством нулей непрерывной функции $f(\mathbf{x}) = d(\mathbf{x}, A)$.

9.6. Функция $f(\mathbf{x}) = d(\mathbf{x}, A)$ непрерывна на компактном множестве C, поэтому она достигает минимума в некоторой точке $\mathbf{c}_0 \in C$. Если множество A компактно, то непрерывная функция $g(\mathbf{x}) = d(\mathbf{c}_0, \mathbf{x})$ на множестве A достигает минимума в некоторой точке $\mathbf{a}_0 \in A$.

9.7. Предположим, что существует последовательность точек $\mathbf{x}_1$, $\mathbf{x}_2$, ... множества X, сходящаяся к точке $\mathbf{x} \in X$, и существует $\varepsilon > 0$, для которого $\|f(\mathbf{x}_i) - f(\mathbf{x})\| \geq \varepsilon$ для всех $i = 1, 2, \ldots$. Последовательность $f(\mathbf{x}_1)$, $f(\mathbf{x}_2)$, ... ограниченная, поскольку $\|f(\mathbf{x}_i) - \mathbf{x}\| \leq \|\mathbf{x}_i - \mathbf{x}\| + \|f(\mathbf{x}_i) - \mathbf{x}_i\| \leq \|\mathbf{x}_i - \mathbf{x}\| + \|f(\mathbf{x}) - \mathbf{x}_i\|$. Поэтому из последовательности $f(\mathbf{x}_1)$, $f(\mathbf{x}_2)$, ... можно выбрать сходящуюся подпоследовательность. Можно считать, что исходная последовательность точек выбрана так, что последовательность точек $f(\mathbf{x}_1)$, $f(\mathbf{x}_2)$, ... множества C сходится. Множество C замкнутое, поэтому эта последовательность сходится к некоторой точке $\mathbf{z} \in C$. Из непрерывности расстояния до множества следует, что $d(\mathbf{x}, C) = \lim_{i \to \infty} d(\mathbf{x}_i, C) = \lim_{i \to \infty} \|\mathbf{x}_i - f(\mathbf{x}_i)\| = \|\mathbf{x} - \mathbf{z}\|$. Поэтому из единственности ближайшей точки для точки $\mathbf{x}$ следует, что $\mathbf{z} = f(\mathbf{x})$. Но это противоречит неравенству $\|f(\mathbf{x}_i) - f(\mathbf{x})\| \geq \varepsilon$.

9.8. Множество U открыто, поэтому для каждой точки $\mathbf{x} \in K \subset U$ можно выбрать $\varepsilon(\mathbf{x}) > 0$ так, что открытый шар радиуса $2\varepsilon(\mathbf{x})$ с центром $\mathbf{x}$ содержится в U. Открытые шары с теми же центрами вдвое меньшего радиуса покрывают компактное множество K, поэтому из такого покрытия можно выбрать конечное подпокрытие шарами с центрами $\mathbf{x}_1, \ldots, \mathbf{x}_n$ и радиусами $\varepsilon(\mathbf{x}_1), \ldots, \varepsilon(\mathbf{x}_n)$. Пусть число ε равно наименьшему из этих радиусов, $\mathbf{x} \in K$ и $\|\mathbf{x} - \mathbf{y}\| < \varepsilon$. Тогда $\|\mathbf{x} - \mathbf{x}_i\| < \varepsilon(\mathbf{x}_i)$ для некоторого i, поэтому

$$\|\mathbf{x}_i - \mathbf{y}\| \leq \|\mathbf{x}_i - \mathbf{x}\| + \|\mathbf{x} - \mathbf{y}\| \leq \varepsilon(\mathbf{x}_i) + \varepsilon \leq 2\varepsilon(\mathbf{x}_i),$$

а значит, $\mathbf{y} \in U$.

9.9. Предположим сначала, что функция $f(x)$ непрерывна. Тогда её график связен как образ связного множества $[a, b]$ при непрерывном отображении. Пусть при $k \to \infty$ последовательность точек $(x_k, f(x_k))$ сходится к некоторой точке (x, y). Тогда $x_k \to x$, поэтому из непрерывности функции f следует, что $f(x_k) \to f(x)$. Таким образом, точка $(x, y) = (x, f(y))$ лежит на графике, т.е. график — замкнутое множество.

Предположим теперь, что график — замкнутое связное множество. Проверим сначала, что для любых точек $x_1, x_2 \in [a, b]$ функция f принимает все значения между $f(x_1)$ и $f(x_2)$. Для определённости будем считать, что $x_1 < x_2$ и $f(x_1) < f(x_2)$. Предположим, что $y_0 \in [f(x_1), f(x_2)]$ и $f(x) \neq y_0$ для всех $x \in [x_1, x_2]$. Пусть множество U состоит из точек, для которых $x < x_1$, и точек, для которых $x < x_2$ и $y < y_0$, а множество V состоит из точек, для которых $x > x_1$, и точек, для которых $x > x_1$ и $y > y_0$. Множества U и V открытые и непересекающиеся, их замыкания покрывают всю плоскость, на их общей границе нет точек графика, $(x_1, f(x_1)) \in U$ и $(x_2, f(x_2)) \in V$. Это противоречит связности графика.

Если функция f разрывна, то можно выбрать точку $x_0 \in [a, b]$, число $\varepsilon_0 > 0$ и сходящуюся к x_0 последовательность точек x_k так, что $|f(x_k) - f(x_0)| > \varepsilon_0$. Не ограничивая общности, можно считать, что $f(x_k) > f(x_0) + \varepsilon_0$. На отрезке $[x_0, x_k]$ (или $[x_k, x_0]$) функция f принимает промежуточное значение $f(x_0) + \varepsilon_0$ в некоторой точке c_k. Последовательность точек $(c_k, f(c_k)) = (c_k, f(x_0) + \varepsilon_0)$ сходится к точке $(x_0, f(x_0) + \varepsilon_0)$, не лежащей на графике. Это противоречит замкнутости графика.

9.10. Ответ: да, может. Требуемое множество можно построить следующим образом. Разделим квадрат на 4 равных квадрата, занумеруем эти квадраты и последовательно выберем в них по одной точке так, чтобы они не лежали на прямых, проходящих через ранее выбранные точки параллельно сторонам квадрата. Затем каждый из полученных квадратов разделим на 4 равных квадрата и выберем таким же образом точки в них, и т.д.

9.11. Предположим, что два оператора A_1 и A_2 являются дифференциалами отображения f в точке $\mathbf{x}$. Рассмотрим оператор $B = A_1 - A_2$. Из неравенства

$$\|B\mathbf{h}\| \leq \|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - A_1\mathbf{h}\| + \|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - A_2\mathbf{h}\|$$

следует, что $\|B\mathbf{h}\|/\|\mathbf{h}\| \to 0$ при $\mathbf{h} \to \mathbf{0}$. Фиксируем вектор $\mathbf{h} \neq \mathbf{0}$. Тогда $\|B(t\mathbf{h})\|/\|t\mathbf{h}\| \to 0$ при $t \to 0$. Но $\|B(t\mathbf{h})\|/\|t\mathbf{h}\| = \|B(\mathbf{h})\|/\|\mathbf{h}\|$, поэтому $B\mathbf{h} = \mathbf{0}$.

9.12. Для достаточно малого $\mathbf{h} \neq \mathbf{0}$ выполняется равенство

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \|\mathbf{h}\| \cdot \frac{f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - A\mathbf{h}}{\|\mathbf{h}\|} + A\mathbf{h}.$$

При $\mathbf{h} \to \mathbf{0}$ оба слагаемых в правой части стремятся к нулю.

9.13. Пусть отображение f дифференцируемо и A — его дифференциал. Тогда дифференциал функции f_i — линейная функция $\mathbf{h} \mapsto (A\mathbf{h})_i$, где $(A\mathbf{h})_i$ — i-я координата вектора $A\mathbf{h}$. Это следует из равенства

$$\|f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - A\mathbf{h}\|^2 = \sum_{i=1}^{m} |f_i(\mathbf{x}+\mathbf{h}) - f_i(\mathbf{x}) - (A\mathbf{h})_i|^2.$$

Пусть функции $f_1, \ldots, f_m$ дифференцируемы и $A_1, \ldots, A_m$ — их дифференциалы. Тогда матрица дифференциала отображения f — это матрица со строками $A_1, \ldots, A_m$. Это следует из того же равенства.

9.14. Для линейного отображения A для всех $\mathbf{x}, \mathbf{h} \in \mathbb{R}^n$ выполняется равенство $A(\mathbf{x}+\mathbf{h}) - A\mathbf{x} - A\mathbf{h} = \mathbf{0}$.

9.15. Утверждение непосредственно следует из неравенства $\|(f+g)(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - g(\mathbf{a}) - (A+B)\mathbf{h}\| \leq \|f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - A\mathbf{h}\| + \|g(\mathbf{a}+\mathbf{h}) - g(\mathbf{a}) - B\mathbf{h}\|$.

9.16. Пусть $f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + A\mathbf{h} + \alpha(\mathbf{h})$ и $g(\mathbf{a}+\mathbf{h}) = g(\mathbf{a}) + B\mathbf{h} + \beta(\mathbf{h})$, где $\alpha(\mathbf{h})/\|\mathbf{h}\| \to 0$ и $\beta(\mathbf{h})/\|\mathbf{h}\| \to 0$ при $\mathbf{h} \to \mathbf{0}$. Тогда

$$f(\mathbf{a}+\mathbf{h})g(\mathbf{a}+\mathbf{h}) = f(\mathbf{a})g(\mathbf{a}) + f(\mathbf{a})B\mathbf{h} + g(\mathbf{a})A\mathbf{h} + \gamma(\mathbf{h}),$$

где $\gamma(\mathbf{h}) = A\mathbf{h} \cdot \beta(\mathbf{h}) + B\mathbf{h} \cdot \alpha(\mathbf{h}) + \alpha(\mathbf{h})\beta(\mathbf{h})$. Ясно, что $\gamma(\mathbf{h})/\|\mathbf{h}\| \to 0$ при $\mathbf{h} \to \mathbf{0}$.

Несложно проверить, что

$$\frac{f(\mathbf{a}+\mathbf{h})}{g(\mathbf{a}+\mathbf{h})} - \frac{f(\mathbf{a})}{g(\mathbf{a})} - \frac{g(\mathbf{a})B - f(\mathbf{a})A}{(g(\mathbf{a}))^2}\mathbf{h} = R_1(\mathbf{h}) + R_2(\mathbf{h}),$$

где

$$\begin{aligned} R_1(\mathbf{h}) &= \left(\frac{1}{g(\mathbf{a})g(\mathbf{a}+\mathbf{h})} - \frac{1}{(g(\mathbf{a}))^2}\right)(g(\mathbf{a})B - f(\mathbf{a})A)\mathbf{h}, \\ R_2(\mathbf{h}) &= \frac{g(\mathbf{a})\alpha(\mathbf{h}) - f(\mathbf{a})\beta(\mathbf{h})}{g(\mathbf{a})g(\mathbf{a}+\mathbf{h})}. \end{aligned}$$

Ясно, что $\frac{1}{g(\mathbf{a})g(\mathbf{a}+\mathbf{h})} - \frac{1}{(g(\mathbf{a}))^2} \to 0$ при $\mathbf{h} \to \mathbf{0}$, поэтому $R_1(\mathbf{h})/\|\mathbf{h}\| \to 0$ при $\mathbf{h} \to \mathbf{0}$. Очевидно также, что $R_2(\mathbf{h})/\|\mathbf{h}\| \to 0$ при $\mathbf{h} \to \mathbf{0}$.

9.17. Если $df = A$, то $D_{\mathbf{h}}f = A\mathbf{h}$, поскольку $\lim_{t\to 0}\left\|\frac{f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x})}{t} - A\mathbf{h}\right\| = \lim_{t\to 0}\left\|\frac{f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x})}{t} - \frac{A(t\mathbf{h})}{t}\right\| = \lim_{t\to 0}\|\mathbf{h}\| \cdot \frac{\|f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x})-A(t\mathbf{h})\|}{\|t\mathbf{h}\|} = 0.$

9.18. Производная отображения f в направлении $\mathbf{v}=(v_1,\ldots,v_n)$ равна $df(\mathbf{v})=\sum_{i=1}^n v_i\frac{\partial f}{\partial x_i}$.

9.19. Положим $f(0,0)=0$ и $f(x,y)=\frac{x^2y}{x^4+y^2}$. Вычислим в точке $(0,0)$ производную в направлении вектора $h=(u,v)$, где $v\neq 0$:

$$\begin{aligned}\lim_{t\to 0}\frac{f(0+tu,0+tv)-f(0,0)}{t} &= \lim_{t\to 0}\frac{(tu)^2tv}{(tu)^4+(tv)^2}\cdot\frac{1}{t}=\\ &= \lim_{t\to 0}\frac{u^2v}{t^2u^4+v^2}=\frac{u^2v}{v^2}=\frac{u^2}{v}.\end{aligned}$$

Если же $v=0$ и $u\neq 0$, то этот предел равен 0. Производная в направлении нулевого вектора равна 0.

В точке $(0,0)$ функция разрывна, поскольку сколь угодно близко к этой точке есть точка вида (t,t^2), а в этой точке значение равно $1/2$.

9.20. а) Пусть $h(\mathbf{x})=g(f(\mathbf{x}))$, где $f(\mathbf{x})=\|\mathbf{x}\|^2$ и $g(y)=y^{p/2}$. Тогда $g'(y)=\frac{p}{2}y^{(p-2)/2}$ и $\frac{\partial f}{\partial x_i}(\mathbf{x})=2x_i$. Поэтому $\frac{\partial\|\mathbf{x}\|^p}{\partial x_i}(\mathbf{x})=p\|\mathbf{x}\|^{p-2}x_i$.

б) Воспользовавшись формулой $\frac{\partial\|\mathbf{x}\|^p}{\partial x_i}(\mathbf{x})=p\|\mathbf{x}\|^{p-2}x_i$, получим

$$\begin{aligned}\frac{\partial}{\partial x_i}(\|\mathbf{x}\|^{-n}x_i) &= \|\mathbf{x}\|^{-2n}\left(\|\mathbf{x}\|^n-x_i\frac{\partial\|\mathbf{x}\|^n}{\partial x_i}\right)=\\ &= \|\mathbf{x}\|^{-2n}\left(\|\mathbf{x}\|^n-nx_i^2\|\mathbf{x}\|^{n-2}\right).\end{aligned}$$

Ясно также, что $\sum_{i=1}^n\left(\|\mathbf{x}\|^n-nx_i^2\|\mathbf{x}\|^{n-2}\right)=n\|\mathbf{x}\|^n-n\|\mathbf{x}\|^n=0$.

9.21. а) Из равенства $\frac{\partial}{\partial x}\frac{\partial f(x,y)}{\partial y}=0$ следует, что $\frac{\partial f(x,y)}{\partial y}=\varphi(y)$. Из равенства $\frac{\partial f(x,y)}{\partial y}=\varphi(y)$ следует, что $f(x,y)=g(x)+y\varphi(y)=g(x)+h(y)$.

б) Сделаем замену переменных $\xi=x-ct$ и $\eta=x+ct$, т.е. $x=\frac{\xi+\eta}{2}$ и $t=\frac{\xi-\eta}{2c}$. Тогда

$$\frac{\partial u}{\partial\xi}=\frac{\partial u}{\partial t}\frac{\partial t}{\partial\xi}+\frac{\partial u}{\partial x}\frac{\partial x}{\partial\xi}=-\frac{1}{2c}\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial u}{\partial x}.$$

Продолжая аналогичные вычисления, получаем $\frac{\partial^2u}{\partial t\partial\xi}=0$.

9.22. По правилу дифференцирования композиции функций получаем

$$\begin{pmatrix}\frac{\partial g}{\partial r} & \frac{\partial g}{\partial\varphi}\end{pmatrix}=\begin{pmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{pmatrix}\begin{pmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial\varphi}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\varphi}\end{pmatrix}=\begin{pmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\end{pmatrix}\begin{pmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{pmatrix}.$$

9.23. Первые производные в полярных координатах вычислены в задаче 9.22. В частности, $\frac{\partial u}{\partial r}=\frac{\partial u}{\partial x}\cos\varphi+\frac{\partial u}{\partial y}\sin\varphi$ (одну и ту же функцию от переменных x,y и от переменных r,φ мы теперь обозначаем не разными буквами f и g, а одной буквой u). Продолжим эти вычисления:

$$\begin{aligned}\frac{\partial^2u}{\partial r^2} &= \frac{\partial^2u}{\partial x^2}\cos^2\varphi+2\frac{\partial^2u}{\partial x\partial y}\sin\varphi\cos\varphi+\frac{\partial^2u}{\partial y^2}\sin^2\varphi,\\ \frac{1}{r^2}\frac{\partial^2u}{\partial\varphi^2} &= \frac{\partial^2u}{\partial x^2}\sin^2\varphi-2\frac{\partial^2u}{\partial x\partial y}\sin\varphi\cos\varphi+\frac{\partial^2u}{\partial y^2}\sin^2\varphi-\\ &- \frac{1}{r}\left(\frac{\partial u}{\partial x}\cos\varphi+\frac{\partial u}{\partial y}\sin\varphi\right).\end{aligned}$$

Складывая эти два выражения и выражение для $\frac{1}{r}\frac{\partial u}{\partial r}$, получаем требуемое.

9.24. Для $r=\|\mathbf{x}\|$ получаем: $\frac{\partial r}{\partial x_i}=\frac{1}{2}\|\mathbf{x}\|^{-1/2}\cdot 2x_i=\frac{x_i}{r}$. Поэтому

$$\begin{aligned}\frac{\partial u}{\partial x_i} &= \frac{dv}{dr}\cdot\frac{\partial r}{\partial x_i}=v'(r)\frac{x_i}{r},\\ \frac{\partial^2u}{\partial x_i^2} &= v''(r)\frac{x_i^2}{r^2}+v'(r)\left(\frac{1}{r}-\frac{x_i^2}{r^3}\right).\end{aligned}$$

Складывая выражения для $\frac{\partial^2u}{\partial x_i^2}$ для $i=1,\,2,\,\ldots,\,n$, получаем требуемое.

9.25. Применим формулу из задачи 9.24 для $v(r) = r^p$. В результате получим

$$\begin{aligned}\Delta(\|\mathbf{x}\|^p) &= p(p-1)r^{p-2} + \frac{n-1}{r}pr^{p-1} = \\ &= p(p+n-2)r^{p-2} = p(p+n-2)\|\mathbf{x}\|^{p-2}.\end{aligned}$$

9.26. Рассмотрим функцию $g(t) = f(t\mathbf{x})$ и вычислим $g'(t)$ двумя способами. С одной стороны, согласно примеру 9.2 $g'(t) = \sum x_i \frac{\partial f}{\partial x_i}(t\mathbf{x})$. С другой стороны, $g(t) = t^m f(\mathbf{x})$, поэтому $g'(t) = mt^{m-1}f(\mathbf{x})$. Полагая $t = 1$, получаем требуемое равенство.

9.27. Квадрат k-й координаты вектора $A\mathbf{x}$ равен

$$(a_{k1}x_1 + \dots + a_{kn}x_n)^2 \le (a_{k1}^2 + \dots + a_{kn}^2)(x_1^2 + \dots + x_n^2).$$

Складывая такие неравенства для $k = 1, \dots, n$, получаем требуемое.

9.28. Требуемое неравенство следует из того, что $\|AB\mathbf{x}\| \le \|A\|_s\|B\mathbf{x}\| \le \|A\|_s\|B\|_s\|\mathbf{x}\|$.

9.29. а) Требуемое неравенство непосредственно следует из неравенства из задачи 9.27.

б) Любую матрицу A можно представить в виде $A = UDW$, где матрицы U и W унитарные, а матрица D диагональная (см. [П1], с. 243). Поэтому $\|A\|^2 = \sum_{i,j} a_{ij}^2 = \operatorname{tr}(A^*A) = \operatorname{tr}(W^*D^*U^*UDW) = \operatorname{tr}(W^*D^*DW) = \operatorname{tr}(D^*D) = \sum |d_i|^2$, где $d_1, \dots, d_n$ — диагональные элементы матрицы D. Кроме того,

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|UDW\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|UDW\mathbf{x}\|}{\|W\mathbf{x}\|} = \frac{\|UD\mathbf{y}\|}{\|\mathbf{y}\|} = \frac{\|D\mathbf{y}\|}{\|\mathbf{y}\|}.$$

Ясно, что $\sup_{\mathbf{y}\ne\mathbf{0}} \frac{\|D\mathbf{y}\|}{\|\mathbf{y}\|} = |d_{\max}|$ — наибольшая из абсолютных величин чисел $d_1, \dots, d_n$. Требуемое неравенство следует из того, что $\sum |d_i|^2 \le r|d_{\max}|^2$.

9.30. а) Требуемое равенство следует из того, что $\frac{\partial}{\partial a_{ij}}(\sum a_{pq}^2) = 2a_{ij}$.

б) Воспользуемся разложением определителя матрицы $A = (a_{ij})$ по i-й строке (см., например, [П1], с. 23-24):

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij},$$

где матрица A_{ij} получена из матрицы A вычёркиванием i-й строки и j-го столбца. Положим $b_{ij} = (-1)^{i+j}\det A_{ij}$. Для всех k матрица A_{ik} не зависит от переменной a_{ij}, поэтому для функции $f = \det A$ от n^2 переменных a_{pq} получаем: $\frac{\partial f}{\partial a_{ij}} = b_{ij}$, т.е. $\operatorname{grad}(\det A) = (b_{ij}) = B$. Требуется доказать, что $B = (\det A)(A^{-1})^T$, т.е $A^T B = (\det A)I$, где I — единичная матрица.

Если i-ю строку матрицы A заменить на k-ю строку, то получим матрицу с двумя одинаковыми строками. Определитель такой матрицы равен 0, поэтому $\sum_{j=1}^{n} a_{kj}b_{ij} = 0$ при $k \ne i$. Вместе с формулой разложения определителя матрицы по строке, записанной в виде $\sum_{j=1}^{n} a_{ij}b_{ij} = \det A$, это доказывает требуемое равенство.

9.31. а) Ясно, что $f(A+X) = (A+X)^2 = A^2 + AX + XA + X^2$. Проверим, что $\lim_{X\to 0} \frac{\|X^2\|}{\|X\|} = 0$. Пусть $\|X\| = \varepsilon$. Тогда каждый элемент матрицы X не превосходит ε, поэтому каждый элемент матрицы X^2 не превосходит $n\varepsilon^2$, где n — порядок матрицы. Поэтому $\|X^2\|^2 \le n^2(n\varepsilon^2)^2 = (n^2\varepsilon^2)^2$ и $\frac{\|X^2\|}{\|X\|} \le n^2\varepsilon$.

б) Ясно, что $f(A+X) = (A^T + X^T)(A+X) = A^TA + A^TX + X^TA + X^TX$. Равенство $\lim_{X\to 0} \frac{\|X^TX\|}{\|X\|} = 0$ доказывается аналогично.

9.32. Если $\|Y^n\| \to 0$ при $n \to \infty$, то $(I+Y)^{-1} = I - Y + Y^2 - Y^3 + \dots$ Поэтому для достаточно малой матрицы X получаем: $(A+X)^{-1} = (A(I + A^{-1}X))^{-1} = (I + A^{-1}X)^{-1}A^{-1} = (I - A^{-1}X + \dots)A^{-1} = A^{-1} - A^{-1}XA^{-1} + \dots$ Таким образом, $(A+X)^{-1} - A^{-1} = -A^{-1}XA^{-1} + \dots$

9.33. а) Пусть собственные значения матрицы X равны $\lambda_1, \ldots, \lambda_n$. Приведя матрицу X к жордановой форме, получаем: $\det(I+tX) = \prod(1+t\lambda_i) = 1+t(\operatorname{tr} X)+\ldots$, где многоточием обозначены члены, делящиеся на t^2.

б) Воспользовавшись задачей а), для достаточно малой матрицы X получаем: $\det(A+X) - \det A = (\det A)\det(I+A^{-1}X) - \det A = (\det A)\operatorname{tr}(A^{-1}X)+\ldots$

9.34. а) Пусть $F = (f_1, \ldots, f_m)$, где $f_1, \ldots, f_m$ — функции. Требуемое утверждение достаточно доказать для каждой из этих функций. Но для функции оно непосредственно следует из теоремы о среднем значении.

б) Непосредственно следует из а).

9.35. Прежде всего заметим, что диагональ прямоугольного параллелепипеда не превосходит суммы его рёбер, поэтому

$$\|F(\mathbf{b}) - F(\mathbf{a})\| \leq \sum_{i=1}^{m} |f_i(\mathbf{b}) - f_i(\mathbf{a})|.$$

Далее, применив теорему о среднем значении к функции f_i, получим

$$|f_i(\mathbf{b}) - f_i(\mathbf{a})| \leq M \sum_{j=1}^{n} |b_j - a_j| \leq nM\|\mathbf{b} - \mathbf{a}\|,$$

поскольку $|b_j - a_j| \leq \|\mathbf{b} - \mathbf{a}\|$.

9.36. Положим $f(0,0) = 0$ и $f(x,y) = \frac{xy(x^2-y^2)}{x^2+y^2}$ при $(x,y) \neq (0,0)$. Функция f непрерывна в точке $(0,0)$, поскольку $|f(x,y)| \leq |xy|$. Из равенств $f(x,0) = 0$ и $f(0,y) = 0$ следует, что $\frac{\partial}{\partial x} f(0,0) = 0$ и $\frac{\partial}{\partial y} f(0,0) = 0$. Непосредственные вычисления частных производных показывают, что если $x \neq 0$ и $y \neq 0$, то $\frac{\partial}{\partial x} f(0,y) = -y$ и $\frac{\partial}{\partial y} f(x,0) = x$. Поэтому первые частные производные непрерывны в точке $(0,0)$. По определению

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y} f(0,0) = \lim_{x\to 0} \frac{1}{x}\left(\frac{\partial}{\partial y} f(x,0) - \frac{\partial}{\partial y} f(0,0)\right) = 1.$$

Аналогично $\frac{\partial}{\partial y}\frac{\partial}{\partial x} f(0,0) = -1$.

9.37. Ответ: нет. Положим $f(x,y) = (y-x^2)(y-2x^2)$. Если $x^2 < y < 2x^2$, то $f(x,y) < 0$, поэтому в сколь угодно малой окрестности начала координат есть точки, в которых $f(x,y) < 0 = f(0,0)$. Проверим, что ограничение функции $f(x,y)$ на любую прямую, проходящую через начало координат, имеет в начале координат локальный минимум. Для прямых $x = 0$ и $y = 0$ это очевидно: ограничения функции f на эти прямые равны y^2 и $2x^4$ соответственно. Любую другую прямую можно задать уравнением $y = ax$, где $a \neq 0$. Ограничение функции f на эту прямую равно $x^2(a-x)(a-2x)$. При $|x| < a/2$ эта величина неотрицательна.

Глава 10.

Теорема о неявной функции

10.1. Формулировка теоремы о неявной функции

В простейшем случае теорема о неявной функции заключается в следующем.

Теорема 10.1. *Пусть функция $F(x,y)$ непрерывно дифференцируема в окрестности точки (x_0,y_0), $F(x_0,y_0)=0$ и $F'_y(x_0,y_0)\neq 0$. Тогда найдутся открытые интервалы $I_x \ni x_0$ и $I_y \ni y_0$ и непрерывно дифференцируемая функция $f\colon I_x \to I_y$, обладающие следующим свойством: для точки $(x,y)\in I_x\times I_y$ равенство $F(x,y)=0$ эквивалентно равенству $y=f(x)$. Производную функции f при этом можно вычислить по формуле*

$$f'(x) = -\frac{F'_x(x,f(x))}{F'_y(x,f(x))}.$$

Рассмотрим сначала пример, поясняющий формулировку теоремы о неявной функции. Пусть $F(x,y)=x^2+y^2-1$. Уравнение $F(x,y)=0$ задаёт окружность S. Условие $F'_y(x_0,y_0)\neq 0$ эквивалентно тому, что $y_0\neq 0$, поскольку $F'_y(x,y)=2y$. Если точка (x_0,y_0) лежит на окружности S и $y_0\neq 0$, то $f(x)=\pm\sqrt{1-x^2}$ (знак здесь совпадает со знаком числа y_0). Если $y_0=0$, то требуемой функции f не существует. Действительно, в этом случае $x_0=\pm 1$. Пусть для определённости $x_0=1$. Тогда при $x>x_0$ соответствующего значения $y=f(x)$ нет вообще, а при $x<x_0$ (и $x>-1$) соответствующих значений $y=f(x)$ два.

В общей теореме о неявной функции речь идёт о m уравнениях, которые связывают $n+m$ переменных. В итоге m переменных выражаются через n независимых переменных. Формулируется эта теорема следующим образом.

Теорема 10.2. *Пусть $\boldsymbol{x}\in\mathbb{R}^n$, $\boldsymbol{y}\in\mathbb{R}^m$, W — окрестность точки $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ в $\mathbb{R}^n\times\mathbb{R}^m$, отображение $F\colon W\to\mathbb{R}^m$ непрерывно дифференцируемо, $F(\boldsymbol{x}_0,\boldsymbol{y}_0)=\boldsymbol{0}$ и якобиан отображения $\boldsymbol{y}\mapsto F(\boldsymbol{x}_0,\boldsymbol{y})$ в точке $\boldsymbol{y}_0$ отличен от нуля. Тогда найдутся открытые окрестности U и V точек $\boldsymbol{x}_0$ и $\boldsymbol{y}_0$ в $\mathbb{R}^n$ и $\mathbb{R}^m$ и непрерывно дифференцируемое отображение $f\colon U\to V$, обладающие следующим свойством: для точки $(\boldsymbol{x},\boldsymbol{y})\in U\times V$ равенство $F(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{0}$ эквивалентно равенству $\boldsymbol{y}=f(\boldsymbol{x})$.*

Для точки $\mathbf{x}\in U$ дифференциал отображения f при этом можно вычислить по формуле

$$df_{\mathbf{x}} = -(d_2F_{(\mathbf{x},f(\mathbf{x}))})^{-1}\circ d_1F_{(\mathbf{x},f(\mathbf{x}))},$$

где d_1 — дифференциал отображения F с фиксированными переменными $\mathbf{y}$, а d_2 — дифференциал отображения F с фиксированными переменными $\mathbf{x}$. Действительно, дифференцируя равенство $F(\mathbf{x},f(\mathbf{x}))=\mathbf{0}$, получаем

$$\big(d_1F_{(\mathbf{x},f(\mathbf{x}))}, d_2F_{(\mathbf{x},f(\mathbf{x}))}\big)\begin{pmatrix} I_n \\ df_{\mathbf{x}} \end{pmatrix} = \mathbf{0},$$

т.е.

$$d_1F_{(\mathbf{x},f(\mathbf{x}))} + d_2F_{(\mathbf{x},f(\mathbf{x}))}\circ df_{\mathbf{x}} = \mathbf{0}.$$

В следующих двух параграфах мы приведём два независимых доказательства теоремы о неявной функции. В § 10.2 мы сначала докажем теорему о неявной функции, а затем выведем из неё теорему об обратной функции. В § 10.3 мы докажем теорему о неявной фунции с помощью сжимающих отображений, но при этом будет доказана только непрерывность неявной функции.

Историческое замечание. Теорему об обратной функции для комплексной функции одной переменной доказал в 1770 году Жозеф-Луи Лагранж (1736-1813); он применил её в небесной механике. Теорема о неявной комплексной функции одной переменной появилась в неопубликованных работах Огюстена Луи Коши (1789-1857). Первым теорему о неявной функции для вещественных функций многих переменных опубликовал Улисс Дини (1845-1918) в 1876-1877 годы. Он же первым сформулировал условие невырожденности отображения с помощью якобиана.

10.2. Теорема об обратной функции

Теорема о неявной функции тесно связана с теоремой об обратной функции. Один из подходов к доказательству теоремы о неявной функции заключается в том, чтобы сначала доказать теорему об обратной функции.

Теорема 10.3 (об обратной функции). *Пусть $f\colon U \to V$ — отображение класса C^1 открытых множеств в $\mathbb{R}^n$. Предположим, что отображение $df_{\boldsymbol{a}}$ обратимо для некоторой точки $\boldsymbol{a} \in U$, и пусть $\boldsymbol{b} = f(\boldsymbol{a})$. Тогда:*

а) можно выбрать открытые множества $\widetilde{U} \ni \boldsymbol{a}$ и $\widetilde{V} \ni \boldsymbol{b}$ в пространстве $\mathbb{R}^n$ так, что f — взаимно однозначное отображение $\widetilde{U}$ на $\widetilde{V}$;

б) отображение $g\colon \widetilde{V} \to \widetilde{U}$, обратное к f, имеет класс C^1.

Доказательство. Пусть B — некоторое выпуклое множество, на котором определено отображение $f = (f_1, \dots, f_n)$, и $\mathbf{y}, \mathbf{z} \in B$. Для $i = 1, \dots, n$ рассмотрим на отрезке $[0,1]$ функцию $\varphi_i(t) = f_i(t\mathbf{y} + (1-t)\mathbf{z})$. Из теоремы о среднем значении следует, что

$$f_i(\mathbf{y}) - f_i(\mathbf{z}) = \varphi_i(1) - \varphi_i(0) = \frac{d\varphi_i}{dt}(t_i) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\mathbf{w}_i)(y_i - z_i),$$

где $t_i \in (0,1)$ и $\mathbf{w}_i = t_i\mathbf{y} + (1-t_i)\mathbf{z} \in B$. Эти равенства можно записать в матричном виде:

$$f(\mathbf{y}) - f(\mathbf{z}) = J(\mathbf{y} - \mathbf{z}), \tag{10.1}$$

где $J = J(\mathbf{w}_1, \dots, \mathbf{w}_n) = \left(\frac{\partial f_i}{\partial x_j}(\mathbf{w}_i)\right)$ — матрица, зависящая от n^2 переменных.

Функция $\det J(\mathbf{w}_1, \dots, \mathbf{w}_n)$ непрерывна и $\det J(\mathbf{a}, \dots, \mathbf{a}) \neq 0$. Поэтому существует такое число ε, что если точки $\mathbf{w}_1, \dots, \mathbf{w}_n$ принадлежат замкнутому шару $D = D^n_{\mathbf{a},\varepsilon}$, то $\det J(\mathbf{w}_1, \dots, \mathbf{w}_n) \neq 0$. Это неравенство и формула (10.1) показывают, что если $\mathbf{y}, \mathbf{z} \in D$ и $\mathbf{y} \neq \mathbf{z}$, то $f(\mathbf{y}) \neq f(\mathbf{z})$. Таким образом, ограничение отображения f на множество D является взаимно однозначным.

Пусть $\widetilde{U} = \operatorname{Int} D$. Докажем, что множество $f(\widetilde{U})$ открыто в $\mathbb{R}^n$. Пусть $\mathbf{u} \in \widetilde{U}$ — произвольная точка. Требуется доказать, что открытый шар достаточно малого радиуса с центром $f(\mathbf{u})$ целиком лежит в $f(\widetilde{U})$. На компактном множестве ∂D функция $\varphi(\mathbf{y}) = \|f(\mathbf{y}) - f(\mathbf{u})\|$ достигает минимума. Этот минимум положителен, поскольку $f(\mathbf{u}) \notin f(\partial D)$. Поэтому можно выбрать положительное число δ так, что $\|f(\mathbf{y}) - f(\mathbf{u})\| > 2\delta$ для всех $\mathbf{y} \in \partial D$. Покажем, что открытый шар радиуса δ с центром $f(\mathbf{u})$ принадлежит $f(\widetilde{U})$. Действительно, пусть $\|f(\mathbf{u}) - \mathbf{z}\| < \delta$. Тогда если $\mathbf{y} \in D$, то

$$\|f(\mathbf{y}) - \mathbf{z}\| \geq \|f(\mathbf{y}) - f(\mathbf{u})\| - \|f(\mathbf{u}) - \mathbf{z}\| > 2\delta - \delta = \delta.$$

Поэтому на множестве D гладкая функция $\psi(\mathbf{w}) = \|f(\mathbf{w}) - \mathbf{z}\|^2$ достигает минимума во внутренней точке (на границе значение этой функции больше δ^2, а в точке $\mathbf{u}$ значение меньше δ^2). Пусть $\mathbf{v}$ — точка минимума функции ψ на множестве D. Тогда

$$0 = \frac{\partial \psi}{\partial x_j} = 2\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(\mathbf{v})(f_i(\mathbf{v}) - z_i).$$

Как мы уже знаем, если $\mathbf{v} \in D$, то $J(\mathbf{v}, \dots, \mathbf{v}) = \det\left(\frac{\partial f_i}{\partial x_j}(\mathbf{v})\right) \neq 0$. Поэтому $\mathbf{z} = f(\mathbf{v}) \in f(\widetilde{U})$, что и требовалось.

Докажем теперь, что отображение $g\colon \widetilde{V} \to \widetilde{U}$, обратное к f, имеет класс C^1. На компактном множестве ($\mathbf{w}_1, \dots, \mathbf{w}_n \in D$, $\|\mathbf{h}\| = 1$) функция $\frac{\|J(\mathbf{w}_1,\dots,\mathbf{w}_n)\mathbf{h}\|}{\|\mathbf{h}\|}$ принимает минимальное значение $c > 0$, поэтому

$$\|f(\mathbf{y}) - f(\mathbf{z})\| = \|J(\mathbf{y} - \mathbf{z})\| \geq c\|\mathbf{y} - \mathbf{z}\|$$

для всех $\mathbf{y}, \mathbf{z} \in D \subset \widetilde{U}$. Пусть $\mathbf{y}_1 = f(\mathbf{y})$ и $\mathbf{z}_1 = f(\mathbf{z})$, т.е. $\mathbf{y} = g(\mathbf{y}_1)$ и $\mathbf{z} = g(\mathbf{z}_1)$. Тогда

$$\|g(\mathbf{y}_1) - g(\mathbf{z}_1)\| \le \frac{1}{c}\|\mathbf{y}_1 - \mathbf{z}_1\|, \tag{10.2}$$

поэтому отображение g непрерывно.

Рассмотрим точки $\mathbf{x}_1 \in \widetilde{V}$ и $\mathbf{x}_1 + \mathbf{h}_1 \in \widetilde{V}$ и положим $\mathbf{x} = g(\mathbf{x}_1)$ и $\mathbf{h} = g(\mathbf{x}_1 + \mathbf{h}_1) - g(\mathbf{x}_1)$. Применим оператор $(df_{\mathbf{x}})^{-1}$ к равенству

$$\mathbf{h}_1 = f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = df_{\mathbf{x}}(\mathbf{h}) + r(\mathbf{h}),$$

где $\|r(\mathbf{h})\|/\|\mathbf{h}\| \to 0$ при $\mathbf{h} \to \mathbf{0}$. В результате получим

$$(df_{\mathbf{x}})^{-1}(\mathbf{h}_1) = \mathbf{h} + (df_{\mathbf{x}})^{-1}(r(\mathbf{h})),$$

т.е.

$$g(\mathbf{x}_1 + \mathbf{h}_1) - g(\mathbf{x}_1) - (df_{\mathbf{x}})^{-1}(\mathbf{h}_1) = -(df_{\mathbf{x}})^{-1}(r(\mathbf{h})).$$

Из неравенства (10.2) следует, что $\|\mathbf{h}\| \le c^{-1}\|\mathbf{h}_1\|$, поэтому $\|r(\mathbf{h})\|/\|\mathbf{h}_1\| \to 0$ при $\|\mathbf{h}_1\| \to 0$, а значит, $(df_{\mathbf{x}})^{-1}$ — дифференциал отображения g в точке $\mathbf{x}_1$. Следовательно, отображение g непрерывно дифференцируемо, поскольку отображение $g\colon V \to U$ непрерывно, отображение df множества U в невырожденные матрицы непрерывно по условию, а переход к обратной матрице — непрерывное отображение по теореме 9.3 в). □

Задача 10.1. Докажите, что в достаточно малой окрестности единичной матрицы I у каждой матрицы есть единственный квадратный корень, т.е. существуют такие окрестности U и V матрицы I, что для любой матрицы $A \in U$ существует единственная матрица $B \in V$, для которой $B^2 = A$.

Пусть U и V — открытые множества в $\mathbb{R}^n$. Отображение $f\colon U \to V$ класса C^1 называют *диффеоморфизмом* класса C^1, если отображение f биективно и обратное отображение f^{-1} имеет класс C^1. На языке диффеоморфизмов теорему об обратной функции можно сформулировать так: если отображение $df_{\mathbf{a}}$ невырожденное, то отображение f является диффеоморфизмом некоторой окрестности точки $\mathbf{a}$ на образ этой окрестности (*локальным диффеоморфизмом*).

Теорема об обратной функции непосредственно следует из теоремы о неявной функции: достаточно положить $\mathbf{x}_0 = \mathbf{b}$, $\mathbf{y}_0 = \mathbf{a}$, $F(\mathbf{x}, \mathbf{y}) = \mathbf{x} - f(\mathbf{y})$ и применить теорему о неявной функции. Покажем, как вывести теорему о неявной функции из теоремы об обратной функции. Для заданного отображения $F(\mathbf{x}, \mathbf{y})$ положим $\tilde{f}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, F(\mathbf{x}, \mathbf{y}))$. В точке $\mathbf{a} = (\mathbf{x}_0, \mathbf{y}_0)$ якобиан отображения $\tilde{f}$ отличен от нуля. Действительно, первые n строк матрицы Якоби — это единичная матрица, дополненная нулями, а на пересечении последних m строк и m столбцов стоит матрица Якоби отображения $\mathbf{y} \mapsto F(\mathbf{x}_0, \mathbf{y})$ в точке $\mathbf{y}_0$. Применив теорему об обратной функции, найдём отображение $\tilde{g}$, обратное отображению $\tilde{f}$: $\tilde{g}(\tilde{f}(\mathbf{x}, \mathbf{y})) = (\mathbf{x}, \mathbf{y})$, т.е. $\tilde{g}(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) = (\mathbf{x}, \mathbf{y})$. Искомое отображение f теперь можно построить следующим образом: вектор $f(\mathbf{x})$ получается из вектора $\tilde{g}(\mathbf{x}, \mathbf{0})$ отбрасыванием первых n координат. Ясно, что если $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, то $\mathbf{y} = f(\mathbf{x})$. И наоборот, если $\mathbf{y} = f(\mathbf{x})$, то $\tilde{g}(\mathbf{x}, F(\mathbf{x}, \mathbf{y})) = \tilde{g}(\mathbf{x}, \mathbf{0})$, поэтому $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, так как отображение $\tilde{g}$ взаимно однозначное.

10.3. Сжимающие отображения

Другой подход к доказательству теоремы о неявной функции основан на сжимающих отображениях. Но при этом мы докажем только непрерывность неявной функции.

Пусть $X \subset \mathbb{R}^n$. Отображение $f\colon X \to \mathbb{R}^n$ *сжимающее*, если $\|f(\mathbf{x}) - f(\mathbf{y})\| \le L\|\mathbf{x} - \mathbf{y}\|$ для некоторой константы $L < 1$.

Теорема 10.4 (о неподвижной точке). *Пусть $C \subset \mathbb{R}^n$ — замкнутое множество. Тогда сжимающее отображение $f\colon C \to C$ имеет единственную неподвижную точку.*

Доказательство теоремы 10.4 почти дословно повторяет доказательство теоремы о неподвижной точке для сжимающих отображений отрезка в себя (см. с. 42). Отметим, что замкнутость множества C используется только при доказательстве существования неподвижной точки; теорема о единственности неподвижной точки верна для сжимающего отображения произвольного множества.

В этом параграфе будем использовать следующие обозначения: $\overline{D}^n_{\mathbf{a},r}$ и $D^n_{\mathbf{a},r}$ — замкнутый и открытый шары в $\mathbb{R}^n$ с центром $\mathbf{a}$ радиуса r, Теорема 10.5. *Пусть* $f\colon \overline{D}^n_{\boldsymbol{a},r} \to \mathbb{R}^n$ *сжимающее отображение с положительной константой* $L < 1$. *Тогда если* $\|f(\boldsymbol{a}) - \boldsymbol{a}\| \le (1-L)r$, *то* $f(\overline{D}^n_{\boldsymbol{a},r}) \subset \overline{D}^n_{\boldsymbol{a},r}$ *и отображение* f *имеет единственную неподвижную точку.*

Доказательство. Пусть $\mathbf{x} \in \overline{D}^n_{\mathbf{a},r}$. Тогда

$$\begin{aligned} \|f(\mathbf{x}) - \mathbf{a}\| &\le \|f(\mathbf{x}) - f(\mathbf{a})\| + \|f(\mathbf{a}) - \mathbf{a}\| \le \\ &\le L\|\mathbf{x} - \mathbf{a}\| + (1-L)r \le Lr + (1-L)r = r, \end{aligned}$$

поэтому $f(\mathbf{x}) \in \overline{D}^n_{\mathbf{a},r}$. Существование и единственность неподвижной точки следует теперь из теоремы 10.4. □

Теорема 10.6. *Пусть* $f\colon D^n_{\boldsymbol{a},r} \to \mathbb{R}^n$ *сжимающее отображение с положительной константой* $L < 1$. *Тогда если* $\|f(\boldsymbol{a}) - \boldsymbol{a}\| < (1-L)r$, *то отображение* f *имеет единственную неподвижную точку.*

Доказательство. Пусть $\|f(\mathbf{a}) - \mathbf{a}\| = (1-L)r - \varepsilon$, где $\varepsilon > 0$. Тогда такие же рассуждения, как и при доказательстве 10.5, показывают, что если $\mathbf{x} \in D^n_{\mathbf{a},r}$, то $\|f(\mathbf{x}) - \mathbf{a}\| \le r - \varepsilon$. Поэтому сжимающее отображение f переводит замкнутый шар радиуса $r - \varepsilon$ с центром $\mathbf{a}$ в себя. Такое отображение замкнутого шара имеет единственную неподвижную точку. Эта точка принадлежит $D^n_{\mathbf{a},r}$, и другой неподвижной точки у сжимающего отображения $f\colon D^n_{\mathbf{a},r} \to \mathbb{R}^n$ быть не может. □

Теорема 10.7. *Пусть* $f\colon \mathbb{R}^n \to \mathbb{R}^n$ — *сжимающее отображение с положительной константой* $L < 1$, $\|f(\boldsymbol{a}) - \boldsymbol{a}\| = d$. *Тогда расстояние от точки* $\boldsymbol{a}$ *до неподвижной точки* $\boldsymbol{p}$ *сжимающего отображения* f *не превосходит* $\frac{d}{1-L}$.

Доказательство. К замкнутому шару с центром $\mathbf{a}$ радиуса $r = \frac{d}{1-L}$ можно применить теорему 10.5. Неподвижная точка $\mathbf{p}$ сжимающего отображения f лежит в этом шаре. □

Теорема 10.8. *Пусть* $U \subset \mathbb{R}^n$ *и* $V = D^m_{\boldsymbol{y}_0,s}$, $G_{\boldsymbol{x}}(\boldsymbol{y})$ — *семейство сжимающих отображений* $V \to V$ *с положительной константой* $L < 1$ *(не зависящей от* $\boldsymbol{x}$*), причём для фиксированного* $\boldsymbol{y}$ *отображение* $G_{\boldsymbol{x}}(\boldsymbol{y})$ *непрерывно зависит от* $\boldsymbol{x} \in U$. *Предположим, что для любой точки* $\boldsymbol{x} \in U$ *выполняется неравенство* $\|G(\boldsymbol{x}, \boldsymbol{y}_0) - \boldsymbol{y}_0\| < (1-L)s$. *Тогда для любой точки* $\boldsymbol{x} \in U$ *существует единственная неподвижная точка* $\boldsymbol{y}_{\boldsymbol{x}}$ *отображения* $G_{\boldsymbol{x}}$, *и точка* $\boldsymbol{y}_{\boldsymbol{x}}$ *непрерывно зависит от* $\boldsymbol{x}$.

Доказательство. Существование и единственность неподвижной точки $\mathbf{y}_{\mathbf{x}}$ отображения $G_{\mathbf{x}}$ следует из теоремы 10.6. Проверим, что отображение $\mathbf{x} \mapsto \mathbf{y}_{\mathbf{x}}$ непрерывно в точке $\mathbf{x}_0$. Пусть задано $\varepsilon > 0$. Отображение $G_{\mathbf{x}}(\mathbf{y})$ непрерывно зависит от $\mathbf{x}$, поэтому можно выбрать $\delta > 0$ так, что если $\|\mathbf{x} - \mathbf{x}_0\| < \delta$, то $\|G_{\mathbf{x}}(\mathbf{y}_{\mathbf{x}_0}) - G_{\mathbf{x}_0}(\mathbf{y}_{\mathbf{x}_0})\| < \varepsilon$. Поэтому согласно теореме 10.7 $\|\mathbf{y}_{\mathbf{x}} - \mathbf{y}_{\mathbf{x}_0}\| < \frac{\varepsilon}{1-L}$. Из этого следует непрерывность отображения $\mathbf{x} \mapsto \mathbf{y}_{\mathbf{x}}$ в точке $\mathbf{x}_0$. □

Теорема 10.9. *Пусть* $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{y} \in \mathbb{R}^m$, W — *окрестность точки* $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ *в* $\mathbb{R}^n \times \mathbb{R}^m$, *отображение* $F\colon W \to \mathbb{R}^m$ *непрерывно дифференцируемо,* $F(\boldsymbol{x}_0, \boldsymbol{y}_0) = \boldsymbol{0}$ *и якобиан отображения* $\boldsymbol{y} \mapsto F(\boldsymbol{x}_0, \boldsymbol{y})$ *в точке* $\boldsymbol{y}_0$ *отличен от нуля. Тогда найдутся открытые окрестности* U *и* V *точек* $\boldsymbol{x}_0$ *и* $\boldsymbol{y}_0$ *в* $\mathbb{R}^n$ *и* $\mathbb{R}^m$ *и непрерывное отображение* $f\colon U \to V$, *обладающие следующим свойством: для точки* $(\boldsymbol{x}, \boldsymbol{y}) \in U \times V$ *равенство* $F(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$ *эквивалентно равенству* $\boldsymbol{y} = f(\boldsymbol{x})$.

Доказательство. Положим $L_0 = d_2F_{(\mathbf{x}_0,\mathbf{y}_0)}$ — дифференциал отображения $\mathbf{y} \mapsto F(\mathbf{x}_0, \mathbf{y})$ в точке $\mathbf{y}_0$. Равенство $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ эквивалентно равенству

$$\mathbf{y} = \mathbf{y} - L_0^{-1}F(\mathbf{x}, \mathbf{y}).$$

Для $(\mathbf{x}, \mathbf{y}) \in W$ положим $G(\mathbf{x}, \mathbf{y}) = \mathbf{y} - L_0^{-1}F(\mathbf{x}, \mathbf{y})$. Тогда решение уравнения $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$ сводится к нахождению неподвижных точек отображения $\mathbf{y} \mapsto G(\mathbf{x}, \mathbf{y})$. Для этого можно применить принцип сжимающих отображений.

Ясно, что $G(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}$ и $d_2G_{(\mathbf{x}_0,\mathbf{y}_0)} = \mathrm{id} - L_0^{-1}L_0 = \mathbf{0}$. Отображение $d_2G_{(\mathbf{x},\mathbf{y})}$ непрерывно по $\mathbf{x}$ и $\mathbf{y}$, поэтому можно выбрать открытые шары $D^n_{\mathbf{x}_0,r} = U$ и $D^m_{\mathbf{y}_0,s} = V$ так, что $\|d_2G_{(\mathbf{x},\mathbf{y})}\| \le 1/2$ при $\mathbf{x} \in U$ и $\mathbf{y} \in V$. Уменьшив при необходимости радиус r, можно добиться выполнения неравенства $\|G(\mathbf{x}, \mathbf{y}_0) - \mathbf{y}_0\| < s/2$ для всех $\mathbf{x} \in U$.

Согласно теореме о среднем (теорема 9.8) для любого фиксированного $\mathbf{x} \in U$ отображение $G_{\mathbf{x}}(\mathbf{y}) = G(\mathbf{x}, \mathbf{y})$ является сжимающим отображением $V \to V$ с константой $L = 1/2$. Поэтому согласно теореме 10.8 для любого $\mathbf{x} \in U$ существует единственная точка $\mathbf{y} \in V$, для которой $G(\mathbf{x}, \mathbf{y}) = \mathbf{y}$, причём отображение $\mathbf{x} \mapsto \mathbf{y}$ непрерывно. $\square$

Историческое замечание. Идея применения сжимающих отображений для доказательства теоремы о неявной функции восходит к итерационному доказательству этой теоремы, которое предложил Эдуард Гурса (1858-1936) в 1903 году.

10.4. Уравнение касательной плоскости

Рассмотрим непрерывно дифференцируемую функцию $F(x, y, z)$. Предположим, что $F(x_0, y_0, z_0) = 0$ и $F'_z(x_0, y_0, z_0) \neq 0$. Тогда согласно теореме о неявной функции в некоторой окрестности точки (x_0, y_0, z_0) поверхность $F(x, y, z) = 0$ является графиком некоторой функции: $z = f(x, y)$. Выражение для производной неявной функции показывает, что этой окрестности выполняются равенства $\frac{\partial z}{\partial x} = -\frac{F'_x}{F'_z}$ и $\frac{\partial z}{\partial y} = -\frac{F'_y}{F'_z}$. Уравнение касательной плоскости к поверхности $z = f(x, y)$ задаётся уравнением

$$F_x(x - x_0) + F_y(y - y_0) + F_z(z - z_0) = 0.$$

Это следует из того, что уравнение касательной плоскости имеет вид

$$z - z_0 = \frac{\partial z}{\partial x}(x - x_0) + \frac{\partial z}{\partial y}(y - y_0).$$

В точке (x_0, y_0, z_0) касательная плоскость к поверхности $F(x, y, z) = 0$ перпендикулярна градиенту $\operatorname{grad} F(x_0, y_0, z_0)$.

10.5. Метод множителей Лагранжа-2

Пусть функция f и функции $g_1, \ldots, g_m$ непрерывно дифференцируемы на открытом множестве $S \subset \mathbb{R}^{n+m}$. Нас интересует локальный экстремум $\mathbf{x}_0 \in S$ функции f при условии, что выполняются соотношения $g_1 = 0, \ldots, g_m = 0$. В случае локального минимума это означает, что у точки $\mathbf{x}_0$, для которой выполняются соотношения $g_1 = 0, \ldots, g_m = 0$, есть такая окрестность U, что $f(\mathbf{x}_0) \leq f(\mathbf{x})$ для всех точек $\mathbf{x} \in U$, для которых выполняются эти соотношения.

Теорема 10.10. *Пусть $\boldsymbol{x}_0$ — точка локального экстремума функции f при условии $g_1 = 0, \ldots, g_m = 0$, причём в этой точке матрица Якоби отображения $g = (g_1, \ldots, g_m)$ имеет ранг m. Тогда*

$$\operatorname{grad} f(\boldsymbol{x}_0) = \lambda_1 \operatorname{grad} g_1(\boldsymbol{x}_0) + \cdots + \lambda_m \operatorname{grad} g_m(\boldsymbol{x}_0) \tag{10.3}$$

для некоторых чисел $\lambda_1, \ldots, \lambda_m$.

Доказательство. Можно считать, что координаты в $\mathbb{R}^{n+m}$ занумерованы так, что матрица $\left(\frac{\partial g_i}{\partial x_j}(\mathbf{x}_0)\right)$, где $1 \leq i \leq m$ и $n + 1 \leq j \leq n + m$, невырожденная. Рассмотрим векторное равенство (10.3) для последних m координат как систему линейных уравнений относительно $\lambda_1, \ldots, \lambda_m$. Определитель этой системы линейных уравнений не равен нулю, поэтому она имеет единственное решение $\lambda_1, \ldots, \lambda_m$. Требуется проверить, что при таких $\lambda_1, \ldots, \lambda_m$ равенство (10.3) выполняется и для координат с номерами $1, \ldots, n$.

Введём следующее обозначение: $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, где $\mathbf{y} = (x_1, \ldots, x_n)$ и $\mathbf{z} = (x_{n+1}, \ldots, x_{n+m})$. Для точки $\mathbf{x}_0 = (\mathbf{y}_0, \mathbf{z}_0)$ выполняется равенство $g(\mathbf{y}_0, \mathbf{z}_0) = \mathbf{0}$. Поэтому согласно теореме о неявной функции можно выбрать открытые окрестности U и V точек $\mathbf{y}_0$ и $\mathbf{z}_0$ и непрерывное отображение $h\colon U \to V$ так, что для точки $(\mathbf{y}, \mathbf{z}) \in U \times V$ равенство $g(\mathbf{y}, \mathbf{z}) = \mathbf{0}$ эквивалентно равенству $\mathbf{z} = h(\mathbf{y})$. Положим $H(\mathbf{y}) = f(\mathbf{y}, h(\mathbf{y}))$. В точке $\mathbf{y}_0$ функция H имеет локальный экстремум, поэтому все частные производные функции H в этой точке обращаются в нуль:

$$\frac{\partial f}{\partial x_i}(\mathbf{x}_0) + \sum_{k=n+1}^{n+m} \frac{\partial f}{\partial x_k}(\mathbf{x}_0) \frac{\partial h_k}{\partial x_i}(\mathbf{z}_0) = 0$$

для $i = 1, \ldots, n$ и $h = (h_{n+1}, \ldots, h_{n+m})$. Эти равенства можно записать в виде

$$d_1 f_{\mathbf{x}_0} + d_2 f_{\mathbf{x}_0} \circ dh_{\mathbf{y}_0} = \mathbf{0},$$

где $d_1 f$ и $d_2 f$ — дифференциалы отображения f по первым n переменным и по последним m переменным. Матрицы отображений $d_1 f$ и $d_2 f$ — это первые n координат вектора $\operatorname{grad} f(\mathbf{x}_0)$ и последние m его координат.

Дифференциал $dh_{\mathbf{y}_0}$ неявной функции находится из соотношения

$$d_1 g_{\mathbf{x}_0} + d_2 g_{\mathbf{x}_0} \circ dh_{\mathbf{y}_0} = \mathbf{0},$$

где $d_1 g$ и $d_2 g$ — дифференциалы отображения g по первым n переменным и по последним m переменным. Матрицы отображений $d_1 g$ и $d_2 g$ составлены из первых n координат векторов $\operatorname{grad} g_1(\mathbf{x}_0)$, $\ldots$, $\operatorname{grad} g_m(\mathbf{x}_0)$ и последних m координат этих векторов.

Числа $\lambda_1, \ldots, \lambda_m$ выбраны так, что $d_2 f_{\mathbf{x}_0} = \Lambda d_2 g_{\mathbf{x}_0}$, где Λ — диагональная матрица с диагональю $(\lambda_1, \ldots, \lambda_m)$. Поэтому

$$d_1 f_{\mathbf{x}_0} = -d_2 f_{\mathbf{x}_0} \circ dh_{\mathbf{y}_0} = d_2 f_{\mathbf{x}_0} \circ (d_2 g_{\mathbf{x}_0})^{-1} \circ d_1 g_{\mathbf{x}_0} = \Lambda d_1 g_{\mathbf{x}_0},$$

что и требовалось. □

Применим метод множителей Лагранжа для нахождения локального экстремума функции $f(\mathbf{x}) = \|\mathbf{x}\|^2$ при условии $\mathbf{x}A\mathbf{x}^T = 1$, где $A = (a_{ij})$ — симметрическая положительно определённая матрица. Согласно этому методу в точке локального экстремума выполняется равенство $\operatorname{grad} f(\mathbf{x}) = \lambda \operatorname{grad} g(\mathbf{x})$, где $g(\mathbf{x}) = \mathbf{x}A\mathbf{x}^T - 1$. Из симметричности матрицы A следует, что $\operatorname{grad}(\mathbf{x}A\mathbf{x}^T) = \mathbf{x}A + A\mathbf{x}^T = 2\mathbf{x}A$, поэтому приходим к равенству $\mathbf{x} = \lambda\mathbf{x}A$. Таким образом, $\lambda^{-1} = \mu$ — собственное значение матрицы A. Все собственные значения симметрической положительно определённой матрицы положительные. Если точка $\mathbf{x}$ удовлетворяет соотношению $\mathbf{x} = \lambda\mathbf{x}A$, то

$$f(\mathbf{x}) = \|\mathbf{x}\|^2 = \mathbf{x}\mathbf{x}^T = \lambda\mathbf{x}A\mathbf{x}^T = \lambda = \mu^{-1},$$

где μ — собственное значение матрицы A. Отметим, что локальным экстремумам соответствуют не все собственные значения матрицы A, а только наибольшее и наименьшее. Это проще всего понять для эллипсоида $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Пусть $0 < a < b < c$. Рассмотрим сферы с центром в начале координат. Сфера радиуса a содержится в эллипсоиде, сфера радиуса c содержит эллипсоид, а сфера радиуса b пересекает эллипсоид: часть эллипсоида расположена внутри сферы, а часть вне.

Пример 10.1. *Наибольшее значение определителя матрицы, суммы квадратов элементов строк которой равны $h_1^2, \ldots, h_n^2$, равно $|h_1 \ldots h_n|$.*

Доказательство. Пусть наибольшее значение определителя достигается для матрицы A. Ясно, что это значение отлично от нуля. Согласно задаче 9.30 б) $\operatorname{grad}(\det A) = (\det A)(A^{-1})^T$ для невырожденной матрицы A. Для функции $g_i = a_{i1}^2 + \cdots + a_{in}^2$ получаем: $\frac{\partial g_i}{\partial a_{ij}} = 2a_{ij}$ и $\frac{\partial g_k}{\partial a_{ij}} = 0$ при $k \neq i$. Поэтому матрица $G = \lambda_1 \operatorname{grad} g_1 + \cdots + \lambda_n \operatorname{grad} g_n$ имеет следующий вид: её строка с номером k равна строке с номером k матрицы A, умноженной на $2\lambda_k$. Из соотношения $\operatorname{grad}(\det A) = G$ следует, что все числа $\lambda_1, \ldots, \lambda_n$ отличны от нуля и строки матрицы A попарно ортогональны. Поэтому матрица AA^T диагональная, и на диагонали стоят числа $h_1^2, \ldots, h_n^2$. Таким образом, $(\det A)^2 = \det(AA^T) = (h_1 \ldots h_n)^2$. □

Задача 10.2. а) Числа $a_1, \ldots, a_n$ положительны, $p > 1$. Найдите максимум функции $\sum a_i x_i$ при условии $\sum |x_i|^p = 1$, $x_i > 0$.

б) Выведите неравенство Гёльдера из задачи а)

10.6. Отображения постоянного ранга

Пусть $U \subset \mathbb{R}^n$ — открытое множество, $\mathbf{a} \in U$. Будем называть *локальной системой координат с началом* $\mathbf{a}$ любой диффеоморфизм φ класса C^1 некоторой окрестности U' точки $\mathbf{a}$ в U на окрестность начала координат в $\mathbb{R}^n$, переводящий точку $\mathbf{a}$ в начало координат.

Теорема 10.11. *Пусть $U \subset \mathbb{R}^n$ — открытое множество, $f\colon U \to \mathbb{R}^m$ — отображение класса C^1, причём отображение $df_{\boldsymbol{x}}$ имеет постоянный ранг r. Тогда для любой точки $\boldsymbol{a} \in U$ можно выбрать локальную систему координат $\varphi\colon U' \to \mathbb{R}^n$ с началом $\boldsymbol{a}$ и локальную систему координат $\psi\colon V \to \mathbb{R}^m$ с началом $f(\boldsymbol{a}) \in \mathbb{R}^m$ так, что $f(U') \subset V$ и*

$$\psi \circ f \circ \varphi^{-1}(x_1, \dots, x_n) = (x_1, \dots, x_r, 0, \dots, 0).$$

Доказательство. Сделав при необходимости параллельные переносы, можно считать, что $\mathbf{a} = \mathbf{0}$ и $f(\mathbf{a}) = \mathbf{0}$. Изменив нумерацию координат, можно считать, что матрица, образованная первыми r строками и первыми r столбцами матрицы $df_{\mathbf{0}}$, невырожденная. Выделив в $\mathbb{R}^n$ и в $\mathbb{R}^m$ первые r координат, представим векторы в этих пространствах в виде $(\mathbf{x}, \mathbf{y})$ и $(\mathbf{v}, \mathbf{w})$; векторы, у которых первые r координат ненулевые, соответствуют векторам $\mathbf{x}$ и $\mathbf{v}$. Запишем отображение f с выделением первых r координат:

$$f(\mathbf{x}, \mathbf{y}) = (\mathbf{v}(\mathbf{x}, \mathbf{y}), \mathbf{w}(\mathbf{x}, \mathbf{y})).$$

При таких обозначениях матрица $\frac{\partial \mathbf{v}}{\partial \mathbf{x}}(\mathbf{0})$ — это как раз и есть невырожденная матрица, образованная первыми r строками и первыми r столбцами матрицы $df_{\mathbf{0}}$.

Рассмотрим отображение $\varphi\colon U \to \mathbb{R}^n$, заданное формулой

$$\varphi(\mathbf{x}, \mathbf{y}) = (\mathbf{v}(\mathbf{x}, \mathbf{y}), \mathbf{y}).$$

Матрица $d\varphi_{\mathbf{0}} = \begin{pmatrix} \frac{\partial \mathbf{v}}{\partial \mathbf{x}}(\mathbf{0}) & * \\ 0 & I_{n-r} \end{pmatrix}$ невырожденная, поэтому по теореме об обратной функции в достаточно малой окрестности начала координат отображение φ является диффеоморфизмом, и у него есть обратное отображение

$$\varphi^{-1}(\mathbf{x}, \mathbf{y}) = (a(\mathbf{x}, \mathbf{y}), b(\mathbf{x}, \mathbf{y})).$$

Будем считать, что отображение φ^{-1} определено на открытом кубе $\widetilde{U}$ (с рёбрами, параллельными осям координат), содержащем точку $\varphi(\mathbf{0}) = \mathbf{0}$.

Ясно, что

$$\begin{aligned} (\mathbf{x}, \mathbf{y}) &= \varphi \circ \varphi^{-1}(\mathbf{x}, \mathbf{y}) = \varphi(a(\mathbf{x}, \mathbf{y}), b(\mathbf{x}, \mathbf{y})) = \\ &= (\mathbf{v}(a(\mathbf{x}, \mathbf{y}), b(\mathbf{x}, \mathbf{y}))), b(\mathbf{x}, \mathbf{y})), \end{aligned}$$

поэтому $b(\mathbf{x}, \mathbf{y}) = \mathbf{y}$ и $\mathbf{v}(a(\mathbf{x}, \mathbf{y}), \mathbf{y}) = \mathbf{v}(a(\mathbf{x}, \mathbf{y}), b(\mathbf{x}, \mathbf{y})) = \mathbf{x}$. Следовательно,

$$f \circ \varphi^{-1}(\mathbf{x}, \mathbf{y}) = f(a(\mathbf{x}, \mathbf{y}), \mathbf{y}) = (\mathbf{x}, \mathbf{w}(a(\mathbf{x}, \mathbf{y}), \mathbf{y})) = (\mathbf{x}, R(\mathbf{x}, \mathbf{y})),$$

где $R(\mathbf{x}, \mathbf{y}) = \mathbf{w}(a(\mathbf{x}, \mathbf{y}), \mathbf{y})$.

При композиции с диффеоморфизмом ранг матрицы Якоби отображения не изменяется, поэтому в окрестности начала координат ранг отображения $d(f \circ \varphi^{-1})$ равен r. Матрица этого отображения имеет вид $\begin{pmatrix} I_r & 0 \\ \frac{\partial R}{\partial \mathbf{x}} & \frac{\partial R}{\partial \mathbf{y}} \end{pmatrix}$. Ранг матрицы такого вида равен r тогда и только тогда, когда $\frac{\partial R}{\partial \mathbf{y}} = 0$. Для функции $R(\mathbf{x}, \mathbf{y})$, определённой на открытом выпуклом множестве $\widetilde{U}$, это означает, что эта функция не зависит от $\mathbf{y}$ (задача 9.34 б), поэтому можно считать, что $R(\mathbf{x}, \mathbf{y}) = S(\mathbf{x})$.

Построим теперь локальную систему координат $\psi\colon V \to \mathbb{R}^m$ с началом $\mathbf{0}$. Можно считать, что $f \circ \varphi^{-1}(\widetilde{U})$ содержится в некотором открытом множестве W. Рассмотрим открытое множество

$$V = \{(\mathbf{v}, \mathbf{w}) \subset W \mid (\mathbf{v}, \mathbf{0}) \subset \widetilde{U}\}.$$

Ясно, что $f \circ \varphi^{-1}(\widetilde{U}) \subset V$, поскольку $\widetilde{U}$ — куб и $f \circ \varphi^{-1}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, S(\mathbf{x}))$. Положим $\psi(\mathbf{v}, \mathbf{w}) = (\mathbf{v}, \mathbf{w} - S(\mathbf{v}))$. Это отображение является диффеоморфизмом множества V на свой образ $\psi(V)$, поскольку обратное отображение задаётся формулой $\psi^{-1}(\mathbf{p}, \mathbf{q}) = (\mathbf{p}, \mathbf{q} + S(\mathbf{p}))$. В итоге получаем

$$\psi \circ f \circ \varphi^{-1}(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{x}, S(\mathbf{x})) = (\mathbf{x}, \mathbf{0}),$$

что и требовалось. □

Замечание. Несложно проверить, что $\mathbf{x} = \mathbf{v}(a(\mathbf{x},\mathbf{y}),\mathbf{y})$. Поэтому из равенств, полученных при доказательстве теоремы 10.11, следует, что

$$\mathbf{w}(a(\mathbf{x},\mathbf{y}),\mathbf{y}) = R(\mathbf{x},\mathbf{y}) = S(\mathbf{x}) = S(\mathbf{v}(a(\mathbf{x},\mathbf{y}),\mathbf{y})).$$

Отображение $(\mathbf{x},\mathbf{y}) \mapsto (a(\mathbf{x},\mathbf{y}),\mathbf{y})$ — диффеоморфизм, поэтому приходим к равенству $\mathbf{w}(\mathbf{x},\mathbf{y}) = S(\mathbf{v}(\mathbf{x},\mathbf{y}))$, которое означает, что последние $m-r$ координат отображения f функционально выражаются через первые r координат. Напомним, что предварительно координаты были занумерованы так, что ранг матрицы Якоби отображения, составленного из первых r координат отображения f, равен r.

10.7. Якобиан и функциональная зависимость

Пусть $U \subset \mathbb{R}^n$ — открытое множество, $f_1, \ldots, f_m$ — функции класса C^1 на U. Эти функции называют *функционально зависимыми* в точке $\mathbf{a} \in U$, если на некоторой открытой окрестности V точки $\mathbf{b} = (f_1(\mathbf{a}),\ldots,f_m(\mathbf{a}))$ существует функция F класса C^1, для которой $\operatorname{grad} F(\mathbf{b}) \neq \mathbf{0}$ и $F(f_1(\mathbf{x}),\ldots,f_m(\mathbf{x})) = 0$ для всех $\mathbf{x}$ из некоторой открытой окрестности точки $\mathbf{a} \in U$. Функции $f_1, \ldots, f_m$ функционально зависимы на множестве U, если они функционально зависимы в каждой точке $\mathbf{a} \in U$.

Если функции $f_1, \ldots, f_m$ функционально зависимы в точке $\mathbf{a}$, то $\operatorname{grad} F(\mathbf{b}) \neq \mathbf{0}$, поэтому по теореме о неявной функции в достаточно малой окрестности точки $\mathbf{b}$ уравнение $F(\mathbf{y}_1,\ldots,\mathbf{y}_m) = 0$ позволяет выразить одну переменную через остальные. Пусть для определённости можно выразить переменную y_m. Тогда

$$f_m(\mathbf{x}) = \Phi(f_1(\mathbf{x}),\ldots,f_{m-1}(\mathbf{x}))$$

в некоторой окрестности точки $\mathbf{a}$.

Теорема 10.12. *Если функции $f_1, \ldots, f_m$ функционально зависимы на открытом множестве $U \subset \mathbb{R}^n$ и $m \leq n$, то в каждой точке $\boldsymbol{a} \in U$ ранг матрицы Якоби отображения $f = (f_1,\ldots,f_m)$ строго меньше m.*

Доказательство. Пусть $\mathbf{b} = f(\mathbf{a})$. Требуется доказать, что ранг матрицы $df_{\mathbf{b}}$, состоящей из m строк и n столбцов, меньше m, т.е. любая матрица A, составленная из m столбцов этой матрицы, вырожденная.

Рассмотрим функцию F, для которой $\operatorname{grad} F(\mathbf{b}) \neq \mathbf{0}$ и $F(f(\mathbf{x})) = 0$ для всех $\mathbf{x}$ из некоторой открытой окрестности точки $\mathbf{a} \in U$. Тогда $dF_{\mathbf{b}} \circ df_{\mathbf{a}} = \mathbf{0}$. По условию матрица-строка $dF_{\mathbf{b}}$ ненулевая (эта матрица совпадает с вектором $\operatorname{grad} F(\mathbf{b})$). Произведение этой ненулевой строки на матрицу A равно нулю, поэтому матрица A вырожденная. □

Теорема 10.13. *Пусть функции $f_1, \ldots, f_m$ заданы на открытом множестве $U \subset \mathbb{R}^n$ и $m \leq n$. Предположим, что в каждой точке $\boldsymbol{a} \in U$ ранг матрицы Якоби отображения $f = (f_1,\ldots,f_m)$ равен одному и тому же числу $r < m$. Тогда можно выбрать r функций (пусть для определённости это будут функции $f_1, \ldots, f_r$) так, что остальные $m-r$ функций в некоторой окрестности точки $\boldsymbol{a}$ представляются в виде*

$$f_i(\boldsymbol{x}) = g_i(f_1(\boldsymbol{x}),\ldots,f_r(\boldsymbol{x})).$$

Здесь функции $g_{r+1}, \ldots, g_m$ определены в некоторой окрестности точки $(f_1(\boldsymbol{a}),\ldots,f_r(\boldsymbol{a}))$.

Доказательство. Требуемое утверждение доказано в замечании после теоремы 10.11. □

10.8. Локальное разложение диффеоморфизма

Пусть $U \subset \mathbb{R}^n$ — открытое множество. Назовём отображение $F\colon U \to \mathbb{R}^n$ *простейшим*, если оно имеет вид

$$\begin{aligned} F(\mathbf{x}) &= F(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) = \\ &= (x_1,\ldots,x_{i-1},f(\mathbf{x}),x_{i+1},\ldots,x_n), \end{aligned}$$

т.е. при простейшем отображении может изменяться только одна координата (номер этой координаты фиксирован).

Теорема 10.14. *Пусть F — диффеоморфизм некоторой окрестности начала координат в $\mathbb{R}^n$, причём $F(\boldsymbol{0}) = \boldsymbol{0}$. Тогда в некоторой окрестности начала координат диффеоморфизм F можно представить в виде композиции*

$$F = G_n \circ B_n \circ G_{n-1} \circ B_{n-1} \circ \dots \circ G_1 \circ B_1,$$

где G_k — простейший диффеоморфизм, переводящий $\boldsymbol{0}$ в $\boldsymbol{0}$, а B_k — линейное преобразование, которое либо тождественно, либо заключается в перестановке двух координат.

Доказательство. Построим по индукции отображения F_m, обладающие следующими свойствами. Прежде всего, F_m — диффеоморфизм некоторой окрестности начала координат в $\mathbb{R}^n$, причём $F(\mathbf{0}) = \mathbf{0}$; кроме того, выполняются следующие дополнительные ограничения: $F_m(\mathbf{x}) = (x_1, \dots, x_{m-1}, \dots)$ (т.е. отображение F_m не изменяет первые $m-1$ координат вектора $\mathbf{x}$) и $F_{m+1} = F_m \circ B_m \circ G_m^{-1}$ при $m > 1$. На отображение F_1 никаких дополнительных ограничений нет, поэтому можно положить $F_1 = F$. Отображение F_{n+1} тождественное, поэтому последовательно получим равенства

$$\begin{aligned} F &= F_1 = F_2 \circ G_1 \circ B_1 = F_3 \circ G_2 \circ B_2 \circ G_1 \circ B_1 = \\ &= F_{n+1} \circ G_n \circ B_n \circ G_{n-1} \circ B_{n-1} \circ \dots \circ G_1 \circ B_1 = \\ &= G_n \circ B_n \circ G_{n-1} \circ B_{n-1} \circ \dots \circ G_1 \circ B_1, \end{aligned}$$

что и требуется.

Предположим, что отображение F_m, $m \geq 1$, построено (напомним, что $F_1 = F$), и займёмся построением отображения F_{m+1}. Рассмотрим матрицу невырожденного отображения $(dF_m)_{\mathbf{0}}$ с элементами $a_{ij} = (\mathbf{e}_i, (dF_m)_{\mathbf{0}}\mathbf{e}_j)$. Эта матрица имеет вид $\begin{pmatrix} I_{m-1} & 0 \\ * & * \end{pmatrix}$. У невырожденной матрицы не может быть нулевого столбца, поэтому столбец с номером m ненулевой, т.е. $a_{mj} \neq 0$ для некоторого $j \geq m$. Рассмотрим линейное преобразование B_m, переставляющее координаты m и j (если $j = m$, то это преобразование тождественное), и проектор P_m, отображающий $\mathbf{e}_m$ в $\mathbf{0}$ и $\mathbf{e}_i$ в $\mathbf{e}_i$ при $i \neq m$. Положим

$$G_m(\mathbf{x}) = P_m\mathbf{x} + (\mathbf{e}_m, F_m(B_m\mathbf{x}))\mathbf{e}_m. \tag{10.4}$$

Ясно, что $G_m(\mathbf{0}) = \mathbf{0}$. Проверим, что отображение $(dG_m)_{\mathbf{0}}$ невырожденное (тогда по теореме об обратной функции отображении G_m является диффеоморфизмом некоторой окрестности начала координат в $\mathbb{R}^n$). Из равенства

$$(dG_m)_{\mathbf{0}}\mathbf{h} = P_m\mathbf{h} + (\mathbf{e}_m, A_m(B_m\mathbf{h}))\mathbf{e}_m,$$

где $A_m = (a_{ij})$, следует, что если $(dG_m)_{\mathbf{0}}\mathbf{h} = \mathbf{0}$, то $P_m\mathbf{h} = \mathbf{0}$, поэтому $\mathbf{h} = \lambda\mathbf{e}_m$. В таком случае

$$(\mathbf{e}_m, A_m(B_m\mathbf{h})) = (\mathbf{e}_m, A_m(B_m\lambda\mathbf{e}_m)) = (\mathbf{e}_m, A_m(\lambda\mathbf{e}_j)) = \lambda a_{mj},$$

поэтому $\mathbf{0} = (dG_m)_{\mathbf{0}}\mathbf{h} = \lambda a_{mj}\mathbf{e}_m$, а значит, $\lambda = 0$, поскольку $a_{mj} \neq 0$. Таким образом, $\mathbf{h} = \mathbf{0}$, т.е. отображение $(dG_m)_{\mathbf{0}}$ невырожденное.

Пусть $\mathbf{y} = G_m(\mathbf{x})$. Тогда $F_{m+1}(\mathbf{y}) = F_m \circ B_m \circ G_m^{-1} \circ G_m(\mathbf{x}) = F_m \circ B_m(\mathbf{x})$. Непосредственно из равенства (10.4) видно, что первые $m-1$ координат у точек $\mathbf{x}$ и $\mathbf{y} = G_m(\mathbf{x})$, одинаковые. Поэтому первые $m-1$ координат у точек $\mathbf{y}$ и $F_{m+1}(\mathbf{y})$ тоже одинаковые, поскольку F_m и B_m их не изменяют. Легко проверить, что координаты с номером m у точек $\mathbf{y}$ и $F_{m+1}(\mathbf{y})$ тоже одинаковые. Действительно, из равенства (10.4) следует, что $(\mathbf{e}_m, \mathbf{y}) = (\mathbf{e}_m, F_m \circ B_m(\mathbf{x})) = (\mathbf{e}_m, F_{m+1}(\mathbf{y}))$. □

10.9. Решения задач

10.1. Пусть $f(A) = A^2$. Тогда $df_A(X) = AX + XA$ (задача 9.31). В частности, $df_I(X) = 2X$. Это отображение невырожденное, поэтому можно применить теорему об обратной функции.

10.2. а) Пусть $f(\mathbf{x}) = \sum a_i x_i$ и $g(\mathbf{x}) = \sum |x_i|^p$. Равенство $\operatorname{grad} f = \lambda \operatorname{grad} g$ означает, что $a_i = p\lambda x_i^{p-1}$, $i = 1, \dots, n$. Тогда из равенства $g(\mathbf{x}) = 1$ следует, что $p\lambda = (\sum a_i^q)^{1/q}$, где $q = \frac{p}{p-1}$, т.е. $\frac{1}{p} + \frac{1}{q} = 1$. Таким образом, в точке экстремума значение функции f равно $\frac{\sum a_i^q}{(\sum a_i^q)^{1/p}} = (\sum a_i^q)^{1/q}$.

Это именно максимальное значение. Действительно, если $x_i = 1$ и $x_j = 0$ при $j \neq i$, то значение функции f равно $a_i = (a_i^q)^{1/q} < (\sum a_i^q)^{1/q}$.

б) Пусть числа $t_1, \ldots, t_n$ положительны. Положим $x_i = \frac{t_i}{(\sum t_i^p)^{1/p}}$. Тогда $\sum |x_i|^p = 1$, поэтому $\sum a_i x_i = f(\mathbf{x}) \leq (\sum a_i^q)^{1/q}$, т.е.

$$\frac{\sum a_i t_i}{(\sum t_i^p)^{1/p}} \leq (\sum a_i^q)^{1/q}.$$

Следовательно, $\sum a_i t_i \leq (\sum a_i^q)^{1/q} (\sum t_i^p)^{1/p}$, что и требовалось.

Глава 11.

Кратные интегралы

11.1. Повторный интеграл

В интеграле $\int_a^b f(x)dx$ подынтегральная функция $f(x)$ сама может представляться интегралом:

$$f(x) = \int_{\alpha(x)}^{\beta(x)} g(x,y)dy.$$

В таком случае выражение $\int_a^b \left(\int_{\alpha(x)}^{\beta(x)} g(x,y)dy\right) dx$ называют *повторным интегралом*; интеграл от функции, представленной повторным интегралом, тоже называют повторным интегралом. Большие скобки при записи повторных интегралов обычно опускают.

Задача 11.1. а) Докажите, что если хотя бы одно из чисел $b-a$ и $d-c$ кратно 2π, то $\int_a^b \int_c^d \sin x \sin y\, dy\, dx = 0$.

б) Докажите, что если $\int_0^b \int_0^d \sin x \sin y\, dy\, dx = 0$, то хотя бы одно из чисел b и d кратно 2π.

в)* Прямоугольник разрезан на несколько прямоугольников, у каждого из которых есть целочисленная сторона. Докажите, что у исходного прямоугольника тоже есть целочисленная сторона.

11.2. Изменение порядка интегрирования

Теорема 11.1 (об изменении порядка интегрирования). *Пусть функция $f(x,y)$ непрерывна при $x \in [a,\, b]$ и $y \in [c,\, d]$. Тогда*

$$\int_a^b \int_c^d f(x,y)\, dy\, dx = \int_c^d \int_a^b f(x,y)\, dx\, dy.$$

Доказательство. Пусть $F_1(t) = \int_a^b \int_c^t f(x,y)\, dy\, dx$ и $F_2(t) = \int_c^t \int_a^b f(x,y)\, dx\, dy$. Функции $F_1(t)$ и $F_2(t)$ непрерывны при $t \in [c,d]$ и $F_1(c) = F_2(c) = 0$. Ясно также, что $F_2'(t) = \int_a^b f(x,y)\, dx$ при $t \in (c,d)$. С другой стороны, по теореме 11.2 (с учётом замечания)

$$F_1'(t) = \int_a^b \frac{d}{dt}\left(\int_c^t f(x,y)\, dy\right) dx = \int_a^b f(x,y)\, dx$$

при $t \in (c,\, d)$. Поэтому $F_1(t) = F_2(t)$ при $t \in [c,\, d]$. При $t = d$ получаем требуемое равенство. □

Пример 11.1 (интеграл Фруллани). *Пусть $0 < a < b$. Тогда*

$$\int_0^\infty \frac{e^{-at} - e^{-bt}}{t} = \ln \frac{b}{a}.$$

Доказательство. Применим теорему 11.1 к функции $f(x,y) = x^y$, где $x \in [0,1]$ и $y \in [a,b]$. С одной стороны,

$$\int_0^1 \int_a^b f(x,y)\, dx\, dy = \int_0^1 \int_a^b e^{y \ln x}\, dx\, dy = \int_0^1 \frac{x^b - x^a}{\ln x} dx.$$

С другой стороны,

$$\int_a^b \int_0^1 f(x,y)\,dy\,dx = \int_a^b \frac{1}{y+1}dy = \ln\frac{b+1}{a+1}.$$

Поэтому

$$\int_0^1 \frac{x^b - x^a}{\ln x}dx = \ln\frac{b+1}{a+1}.$$

Применив замену переменных $x = e^{-t}$, получаем требуемое. □

Следующий пример показывает, что теорема 11.1 неверна для разрывных функций.

Пример 11.2. *Пусть* $f(x,y) = \frac{x-y}{(x+y)^3}$ *при* $xy \neq 0$ *и* $f(x,y) = 0$ *при* $xy = 0$. *Тогда*

$$\int_0^1 \int_0^1 f(x,y)dy\,dx = \frac{1}{2}, \quad \int_0^1 \int_0^1 f(x,y)dx\,dy = -\frac{1}{2}.$$

Доказательство. При $x > 0$ получаем:

$$\begin{aligned}
\int_0^1 \frac{x-y}{(x+y)^3}dy &= \int_0^1 \frac{2x-(x+y)}{(x+y)^3}dy = \\
&= \int_0^1 \left(\frac{2x}{(x+y)^3} - \frac{1}{(x+y)^2}\right)dy = \\
&= \left[-\frac{x}{(x+y)^2} + \frac{1}{x+y}\right]_{y=0}^{y=1} = \\
&= -\frac{x}{(x+1)^2} + \frac{1}{x+1} + \frac{x}{x^2} - \frac{1}{x} = \frac{1}{(x+1)^2}.
\end{aligned}$$

Следовательно,

$$\int_0^1 \int_0^1 \frac{x-y}{(x+y)^3}dy\,dx = \int_0^1 \frac{1}{(x+1)^2}dx = \left[\frac{-1}{x+1}\right]_0^1 = -\frac{1}{2} + 1 = \frac{1}{2}.$$

Меняя местами x и y, получаем

$$\int_0^1 \int_0^1 \frac{y-x}{(x+y)^3}dx\,dy = \frac{1}{2} \quad \text{т.е.} \quad \int_0^1 \int_0^1 \frac{x-y}{(x+y)^3}dx\,dy = -\frac{1}{2}.$$

□

Следующий пример показывает, что теорема 11.1 неверна для бесконечных интервалов. (Для бесконечных интервалов нужна равномерная сходимость интеграла, см. теорему 11.5).

Пример 11.3.

$$\int_0^\infty \int_0^1 (e^{-xy} - xye^{-xy})dy\,dx \neq \int_0^1 \int_0^\infty (e^{-xy} - xye^{-xy})dx\,dy.$$

Доказательство. Из равенства $\int_0^1 (e^{-xy} - xye^{-xy})dy = ye^{-xy}|_{y=0}^{y=1} = e^{-x}$ следует, что интеграл в левой части равен $\int_0^\infty dx = -e^{-x}|_0^\infty = 1$. Из равенства $\int_0^\infty (e^{-xy} - xye^{-xy})dx = xe^{-xy}|_{x=0}^{x=\infty} = 0$ следует, что интеграл в правой части равен 0. □

Задача 11.2. *Интеграл непрерывной на отрезке $[c,d]$ функции f по любому отрезку длины a положителен, а по любому отрезку длины b отрицателен. Докажите, что $d < a + b + c$.

11.3. Дифференцирование под знаком интеграла

Теорема 11.2. *Пусть $f(x) = \int_{\alpha(x)}^{\beta(x)} g(x,y)dy$, причём функции $\alpha(x)$ и $\beta(x)$ дифференцируемы в точке x_0 и при всех $x \in [x_0-\eta, x_0+\eta]$ $(\eta > 0)$ и $y \in [\alpha(x_0)-\eta, \beta(x_0)+\eta]$ существует и непрерывна частная производная $\frac{\partial g}{\partial x}(x,y)$. Тогда*

$$f'(x_0) = \int_{\alpha(x)}^{\beta(x)} \frac{\partial g}{\partial x}(x_0,y)dy + \beta'(x_0)g(x_0,\beta(x_0)) - \alpha'(x_0)g(x_0,\alpha(x_0)).$$

Доказательство. Ясно, что

$$\begin{aligned} f(x) - f(x_0) &= \int_{\alpha(x)}^{\beta(x)} [g(x,y) - g(x_0,y)]dy + \int_{\beta(x_0)}^{\beta(x)} g(x,y)dy - \\ &- \int_{\alpha(x_0)}^{\alpha(x)} g(x,y)dy. \end{aligned}$$

Поэтому достаточно проверить, что

$$\lim_{x\to x_0} \int_{\alpha(x_0)}^{\beta(x_0)} \frac{g(x,y) - g(x_0,y)}{x - x_0}dy = \int_{\alpha(x_0)}^{\beta(x_0)} \frac{\partial g}{\partial x}(x_0,y)dy \qquad (11.1)$$

и

$$\lim_{x\to x_0} \int_{\beta(x_0)}^{\beta(x)} g(x,y)dy = \beta'(x_0)g(x_0,\beta(x_0)). \qquad (11.2)$$

Пусть $x \in [x_0-\eta, x_0+\eta]$. По теореме Ролля

$$\frac{g(x,y) - g(x_0,y)}{x - x_0} = \frac{\partial g}{\partial x}(\xi,y),$$

где $\xi \in [x, x_0]$ (или $\xi \in [x_0, x]$). Функция $\frac{\partial g}{\partial x}$ непрерывна на компактном множестве, поэтому для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что если $|x - x_0| \le \delta$, то

$$\left|\frac{\partial g}{\partial x}(\xi,y) - \frac{\partial g}{\partial x}(x_0,y)\right| \le \varepsilon$$

при всех $y \in [\alpha(x_0), \beta(x_0)]$. Из этого следует равенство (11.1).

Для любого $\varepsilon > 0$ можно выбрать $\delta > 0$ так, что если $|x - x_0| \le \delta$ и $y \in [\alpha(x_0), \beta(x_0)]$, то

$$|g(x,y) - g(x_0,\beta(x_0))| \le \varepsilon \qquad (11.3)$$

и

$$\left|\frac{\beta(x) - \beta(x_0)}{x - x_0} - \beta'(x_0)\right| \le \varepsilon. \qquad (11.4)$$

Из неравенства (11.3) следует, что

$$\left|\int_{\beta(x_0)}^{\beta(x)} g(x,y)dy - (\beta(x) - \beta(x_0))g(x_0,\beta(x_0))\right| \le \varepsilon|\beta(x) - \beta(x_0)|.$$

Для доказательства равенства (11.2) остаётся воспользоваться неравенством (11.4) и тем, что $|\beta(x) - \beta(x_0)| \to 0$ при $x \to x_0$. □

Замечание. Если $\alpha(x) = \mathrm{const}$, то нет необходимости выходить за пределы отрезка $[\alpha(x_0), \beta(x_0)+\eta]$, поэтому в таком случае в формулировке теоремы 11.2 отрезок $[\alpha(x_0)-\eta, \beta(x_0)+\eta]$ можно заменить на отрезок $[\alpha(x_0), \beta(x_0)+\eta]$. В случае $\beta = \mathrm{const}$ ситуация аналогична.

С помощью теоремы 11.2 можно доказать следующую формулу Коши для повторных интегралов.

Теорема 11.3 (формула Коши). *Пусть функция f непрерывна на отрезке $[0, M]$ и $x \in (0, M)$. Тогда*

$$\int_0^x \int_0^{x_{n-1}} \cdots \int_0^{x_1} f(t)\,dt\,dx_1 \dots dx_{n-1} = \frac{1}{(n-1)!}\int_0^x (x-t)^{n-1} f(t)\,dt.$$

Доказательство. Применим индукцию по n. При $n = 1$ обе части равенства совпадают. Шаг индукции заключается в доказательстве равенства

$$\int_0^y \int_0^x (x-t)^{n-1} f(t)\,dt\,dx = \frac{1}{n}\int_0^y (y-t)^n f(t)\,dt$$

для функции $f(t)$, непрерывной на отрезке $[0,\, y+\eta]$, $\eta > 0$. Согласно теореме 11.2 производная по y правой части требуемого равенства равна

$$\frac{1}{n}\int_0^y n(y-t)^{n-1} f(t)\,dt + \frac{1}{n}(y-y)^n = \int_0^y (y-t)^{n-1} f(t)\,dt,$$

т.е. она совпадает с производной по y левой части. Кроме того, при $y = 0$ обе части равенства обращаются в нуль. □

11.4. Равномерно сходящиеся интегралы

Для несобственных интегралов, зависящих от параметра, можно ввести понятие равномерной сходимости. Для несобственного интеграла на конечном интервале это понятие вводится следующим образом. Пусть функция $f(x,t)$ непрерывна для $x \in [a,b)$ и $t \in [c,d]$. Несобственный интеграл $F(t) = \int_a^b f(x,t)dx$ *сходится равномерно* для $t \in [c,d]$, если для любого $\varepsilon > 0$ можно выбрать $x_\varepsilon \in [a,b)$ так, что для всех $y \in (x_\varepsilon, b)$ и всех $t \in [c,d]$ выполняется неравенство

$$\left|\int_y^b f(x,t)dx\right| < \varepsilon.$$

Для несобственного интеграла на бесконечном интервале это понятие вводится аналогично. Пусть функция $f(x,t)$ непрерывна для $x \in [a,\infty)$ и $t \in [c,d]$. Несобственный интеграл $F(t) = \int_a^\infty f(x,t)dx$ *сходится равномерно* для $t \in [c,d]$, если для любого $\varepsilon > 0$ можно выбрать $x_\varepsilon \in [a,\infty)$ так, что для всех $y \in (x_\varepsilon,\infty)$ и всех $t \in [c,d]$ выполняется неравенство

$$\left|\int_y^\infty f(x,t)dx\right| < \varepsilon.$$

Пример 11.4. *Интеграл $\int_0^\infty \frac{e^{-tx}\sin x}{x}dx$ сходится равномерно для $t \geq 0$.*

Доказательство. Воспользуемся теоремой Бонне о среднем (задача 7.13 а): если $f(x)$ и $\varphi(x)$ — непрерывные функции на отрезке $[a,b]$, причём функция $\varphi(x)$ неотрицательная убывающая, то

$$\int_a^b f(x)\varphi(x)dx = \varphi(a)\int_a^c f(x)dx$$

для некоторого $c \in (a,b)$. Фиксируем $t \geq 0$ и положим $f(x) = \frac{\sin x}{x}$ и $\varphi(x) = e^{-tx}$:

$$\int_a^b \frac{e^{-tx}\sin x}{x}dx = e^{-ta}\int_a^c \frac{\sin x}{x}dx. \tag{11.5}$$

Согласно примеру 6.3 на с. 102 интеграл $\int_{\pi/2}^\infty \frac{\sin x}{x}dx$ сходится, поэтому интеграл $\int_0^\infty \frac{\sin x}{x}dx$ тоже сходится, поскольку $\frac{\sin x}{x} \to 1$ при $x \to 0$. Выберем $x_\varepsilon > 0$ так, что $\left|\int_a^b \frac{\sin x}{x}dx\right| < \varepsilon$ при $b > a > x_\varepsilon$. Тогда из равенства (11.5) следует, что $\left|\int_a^b \frac{e^{-tx}\sin x}{x}dx\right| < \varepsilon$ при $b > a > x_\varepsilon$. □

Теорема 11.4. *Если интеграл* $F(t) = \int_a^\infty f(x,t)dx$ *сходится равномерно для* $t \in [c,d]$ *и функция* $f(x,t)$ *непрерывна при* $x \geq a$ *и* $t \in [c,d]$, *то функция* $F(t)$ *непрерывна.*

Доказательство. Для данного $\varepsilon > 0$ выберем $x_\varepsilon \in (a,\infty)$ так, что для всех $y \in [x_\varepsilon,\infty)$ и всех $t \in [c,d]$ выполняется неравенство

$$\left|\int_y^\infty f(x,t)dx\right| < \varepsilon.$$

Затем, воспользовавшись равномерной непрерывностью функции $f(x,t)$ на компактном множестве $[a,x_\varepsilon] \times [c,d]$, выберем $\delta > 0$ так, что

$$|f(x,t) - f(x,t')| < \frac{\varepsilon}{x_\varepsilon - a}$$

при $a \leq x \leq x_\varepsilon$ и $|t - t'| < \delta$.

В результате для данного $\varepsilon > 0$ мы выбрали $\delta > 0$ так, что если $|t - t'| < \delta$, то

$$\begin{aligned} |F(t) - F(t')| &\leq \left|\int_a^{x_\varepsilon} (f(x,t) - f(x,t'))dx\right| + \\ &+ \left|\int_{x_\varepsilon}^\infty f(x,t)dx\right| + \left|\int_{x_\varepsilon}^\infty f(x,t')dx\right| = 3\varepsilon. \end{aligned}$$

□

Напомним, что пример 11.3 показывает, что теорема 11.1 неверна для бесконечных интервалов. Но она становится верной, если дополнительно потребовать равномерную сходимость интеграла.

Теорема 11.5. *Пусть функция* $f(x,y)$ *непрерывна при* $x \in [a,b]$, $y \in [c,\infty]$ *и интеграл* $\int_c^\infty f(x,y)dy$ *равномерно сходится для* $x \in [a,b]$. *Тогда*

$$\int_a^b \int_c^\infty f(x,y)\,dy\,dx = \int_c^\infty \int_a^b f(x,y)\,dx\,dy.$$

Доказательство. Из равномерной сходимости интеграла $\int_c^\infty f(x,y)dy$ следует, что для любого $\varepsilon > 0$ можно выбрать $d_0 > c$ так, что для всех $d > d_0$ и всех $x \in [a,b]$ выполняется неравенство

$$\left|\int_d^\infty f(x,y)dy\right| < \frac{\varepsilon}{b-a}.$$

Согласно теореме 11.1

$$\int_a^b \int_c^d f(x,y)\,dy\,dx = \int_c^d \int_a^b f(x,y)\,dx\,dy.$$

Поэтому

$$\begin{aligned} &\left|\int_a^b \int_c^\infty f(x,y)\,dy\,dx - \int_c^d \int_a^b f(x,y)\,dx\,dy\right| = \\ = &\left|\int_a^b \int_d^\infty f(x,y)\,dy\,dx\right| \leq \int_a^b \frac{\varepsilon}{b-a}dx = \varepsilon. \end{aligned}$$

□

Теорема 11.6. *Если функции* $f(x,t)$ *и* $\frac{\partial f}{\partial t}(x,t)$ *непрерывны при* $x \geq a$ *и* $t \in [c,d]$, *интеграл* $G(t) = \int_a^\infty \frac{\partial f}{\partial t}(x,t)dx$ *сходится равномерно для* $t \in [c,d]$ *и интеграл* $F(t) = \int_a^\infty f(x,t)dx$ *сходится для некоторого* $t_0 \in [c,d]$, *то интеграл* $F(t)$ *сходится равномерно для* $t \in [c,d]$ *и имеет место равенство* $F'(t) = G(t)$.

Доказательство. Для любого $b \geq a$ рассмотрим функцию $F_b(t) = \int_a^b f(x,t)dx$, где $t \in [c,d]$. Функции $f(x,t)$ и $\frac{\partial f}{\partial t}(x,t)$ непрерывны на множестве $[a,b]\times[c,d]$, поэтому согласно теореме 11.2

$$\frac{dF_b}{dt}(t) = \int_a^b \frac{\partial f}{\partial t}(x,t)dx.$$

Следовательно, для любого $t \in [c,d]$ получаем равенство

$$\begin{aligned} F_b(t) &= F_b(t_0) + \int_{t_0}^t \left(\int_a^b \frac{\partial f}{\partial t}(x,t)dx\right)dt = \\ &= F(t_0) + \int_{t_0}^t G(t)dt + (F_b(t_0) - F(t_0)) - \\ &- \int_{t_0}^t \left(\int_b^\infty \frac{\partial f}{\partial t}(x,t)dx\right)dt. \end{aligned}$$

Поэтому

$$\begin{aligned} \left|F_b(t) - F_b(t_0) - \int_{t_0}^t G(t)dt\right| &\leq |F_b(t_0) - F(t_0)| + \\ &+ \left|\int_{t_0}^t \left(\int_b^\infty \frac{\partial f}{\partial t}(x,t)dx\right)dt\right|. \end{aligned}$$

Пусть задано $\varepsilon > 0$. По условию $\lim_{b\to\infty} F_b(t_0) = F(t_0)$, поэтому можно выбрать $b_1 > a$ так, что $|F_b(t_0) - F(t_0)| < \varepsilon$ при $b > b_1$. Кроме того, по условию интеграл $G(t)$ сходится равномерно для $t \in [c,d]$, поэтому можно выбрать $b_2 > a$ так, что для всех $t \in [c,d]$

$$\left|\int_b^\infty \frac{\partial f}{\partial t}(x,t)dx\right| < \varepsilon$$

при $b > b_2$. Поэтому если $b > \max(b_1, b_2)$, то

$$\left|F_b(t) - F_b(t_0) - \int_{t_0}^t G(t)dt\right| < \varepsilon + \varepsilon|t - t_0| \leq \varepsilon(1 + d - c).$$

Таким образом, интеграл $F(t)$ сходится равномерно для $t \in [c,d]$ и

$$F(t) = F(t_0) + \int_{t_0}^t G(t)dt.$$

Согласно теореме 11.4 функция $G(t)$ непрерывна, поэтому $F'(t) = G(t)$. □

Приведём пример применения теоремы 11.6 для вычисления несобственных интегралов.

Пример 11.5. *При $t > 0$ интеграл $\int_0^\infty \frac{e^{-tx}\sin x}{x}dx$ сходится к $\operatorname{arctg}\frac{1}{t}$; интеграл $\int_0^\infty \frac{\sin x}{x}dx$ сходится к $\frac{\pi}{2}$.*

Доказательство. Функции $f(x,t) = \frac{e^{-tx}\sin x}{x}$ и $\frac{\partial f}{\partial t}(x,t) = e^{-tx}\sin x$ непрерывны при $x \geq 0$ и $t \in [0,d]$ для любого $d > 0$. Рассмотрим интегралы $G(t)$ и $F(t)$ из формулировки теоремы 11.6. Согласно примеру 11.4 интеграл $G(t) = \int_0^\infty e^{-tx}\sin x\,dx$ сходится равномерно для $t \in [0,d]$. При $t = 0$ интеграл $F(t)$ равен $\int_0^\infty \frac{\sin x}{x}dx$. Этот интеграл сходится (это обсуждалось при разборе примера 11.4). Следовательно, $F'(t) = G(t)$.

Интеграл $G(t) = \int_0^\infty e^{-tx}\sin x\,dx$ можно вычислить интегрированием по частям:

$$\begin{aligned} \int_0^\infty e^{-tx}\sin x\,dx &= \frac{1}{t}\int_0^\infty e^{-tx}\cos x\,dx = \\ &= \frac{1}{t} - \frac{1}{t^2}\int_0^\infty e^{-tx}\sin x\,dx, \end{aligned}$$

поэтому $G(t) = \int_0^\infty e^{-tx} \sin x \, dx = \frac{1}{1+t^2}$ и $F(t) = \operatorname{arctg} \frac{1}{t} + C$.

Покажем, что $F(t) \to 0$ при $t \to \infty$, т.е. для любого $\varepsilon > 0$ при достаточно больших t выполняется неравенство $|F(t)| < 2\varepsilon$. Выберем сначала $a > 0$ так, что $\left|\int_0^a \frac{e^{-tx} \sin x}{x} dx\right| < \varepsilon$ для всех $t \geq 0$. Затем, как и в примере 11.4, воспользуемся теоремой Бонне о среднем:

$$\int_a^\infty \frac{e^{-tx} \sin x}{x} dx = e^{-tc} \int_c^\infty \frac{\sin x}{x} dx$$

для некоторого $c > a$. Функция $H(c) = \int_c^\infty \frac{\sin x}{x} dx$ ограниченная: $|H(c)| \leq M$. Для достаточно больших t выполняется неравенство $e^{-ta} M < \varepsilon$, поэтому $|F(t)| < 2\varepsilon$, что и требовалось.

В итоге получаем $F(t) = \operatorname{arctg} \frac{1}{t}$ и $F(0) = \frac{\pi}{2}$. $\square$

11.5. Кратный интеграл

Определение интеграла Римана ограниченной функции f, заданной на прямоугольном параллелепипеде $P = [a_1, b_1] \times \cdots \times [a_n, b_n]$, во многом похоже на определение интеграла функции по отрезку. Если задано разбиение каждого ребра параллелепипеда на отрезки, то можно построить *разбиение* параллелепипеда P на параллелепипеды P_k — произведения отрезков разбиения. В каждом параллелепипеде P_k разбиения выберем произвольную точку $\mathbf{x}_k$ и рассмотрим *интегральную сумму* $\sum f(\mathbf{x}_k) v(P_k)$, где $v(P_k)$ — объём (т.е. произведение длин рёбер) параллелепипеда P_k. Если при наибольшей длине диагонали параллелепипеда разбиения стремящейся к нулю интегральные суммы имеют конечный предел, то этот предел называют *кратным интегралом* функции f по параллелепипеду P и обозначают $\int_P f$. Функцию f при этом называют *интегрируемой по Риману* на параллелепипеде P.

Интегрируемость по Риману непрерывных функций для параллелепипеда доказывается так же, как для отрезка. Вводятся верхняя и нижняя интегральные суммы для данного разбиения, доказывается, что при измельчении разбиения нижняя интегральная сумма не уменьшается, а верхняя не увеличивается, и что любая нижняя интегральная сумма не больше любой верхней интегральной суммы. Затем доказывается, что если I — точная верхняя грань нижних интегральных сумм, σ — произвольная интегральная сумма, s и S — нижняя и верхняя интегральные суммы, соответствующие тому же разбиению, что и σ, то $|\sigma - I| \leq S - s$. Наконец, для непрерывной функции доказывается, что если наименьшая длина диагонали параллелепипеда разбиения стремится к нулю, то разность $S - s$ стремится к нулю. Для этого нужно лишь воспользоваться равномерной непрерывностью функции многих переменных на компакте.

В случае кратных интегралов нам придётся иметь дело с разрывными функциями и даже с неинтегрируемыми функциями. Поэтому введём обозначения $\underline{\int}_P f$ и $\overline{\int}_P f$ для точной верхней грани нижних интегральных сумм и точной нижней грани верхних интегральных сумм.

Множество $A \subset \mathbb{R}^n$ имеет *меру нуль* (или является *множеством меры нуль*), если для любого $\varepsilon > 0$ существует не более чем счётный набор n-мерных прямоугольных параллелепипедов (с рёбрами, параллельными осям координат), которые покрывают множество A и сумма их объёмов не превосходит ε.

Объединение счётного набора множеств меры нуль имеет меру нуль.

Для функций многих переменных критерий Лебега интегрируемости по Риману формулируется так же, как для функций одной переменной (теорема 8.1); доказательство не требует существенных изменений.

Теорема 11.7 (Лебег). *Ограниченная функция f на прямоугольном параллелепипеде в $\mathbb{R}^n$ интегрируема по Риману тогда и только тогда, когда множество точек разрыва имеет меру нуль.*

Для разных целей нужно интегрировать функции по произвольным областям в $\mathbb{R}^n$, не обязательно прямоугольным. Интеграл функции f, заданной на ограниченном множестве $S \subset \mathbb{R}^n$, определяется как интеграл функции $f\chi_S \colon \mathbb{R}^n \to \mathbb{R}$, где χ_S — характеристическая функция, равная 1 на S и 0 вне S. Этот интеграл берётся по любому прямоугольному параллелепипеду, содержащему S. Этот интеграл не зависит от выбора параллелепипеда, содержащего S.

Носителем функции f называют множество точек, в которых она отлична от нуля. Носитель функции f обозначают $\operatorname{supp} f$. Носитель непрерывной функции — замкнутое множество. Носитель функции $f\chi_S$ содержится в множестве S. Интеграл функции по всем параллелепипедам, содержащим её носитель, один и тот же.

Точки разрыва характеристической функции множества S — это граничные точки множества S. Поэтому функция χ_S интегрируема по Риману тогда и только тогда, когда граница множества S имеет меру нуль. Множество S называют *измеримым по Жордану*, если его характеристическая функция интегрируема, т.е. его граница имеет меру нуль. Интеграл характеристической функции измеримого множества S называют его *мерой Жордана* или *объёмом*.

При вычислении кратного интеграла непрерывной функции по ограниченной измеримой по Жордану области можно разбивать эту область на измеримые по Жордану множества D_k и рассматривать пределы интегральных суммы вида $\sum f(\mathbf{x}_k)v(D_k)$, где $\mathbf{x}_k \in D_k$ и $v(D_k)$ — объём множества D_k, когда наибольший из диаметров множеств D_k стремится к нулю. Доказательство то же самое, что и на с. 203.

11.6. Выражение кратного интеграла через повторный

Пусть функция $f(x, y)$ непрерывна при $x \in [a, b]$ и $y \in [c, d]$. Тогда согласно теореме об изменении порядка интегрирования (теорема 11.1)

$$\int_a^b \int_c^d f(x, y)\,dy\,dx = \int_c^d \int_a^b f(x, y)\,dx\,dy.$$

Оба эти повторных интеграла равны кратному интегралу $\int_I f$ по прямоугольнику $I = [a, b] \times [c, d]$. Действительно, рассмотрим разбиения отрезков $[a, b]$ и $[c, d]$, и пусть $[\alpha, \beta]$ и $[\gamma, \delta]$ — некоторые отрезки этих разбиений. Пусть M — максимум функции f на прямоугольнике $[\alpha, \beta] \times [\gamma, \delta]$. Тогда $\int_\gamma^\delta f(x, y)\,dy \le M(\delta - \gamma)$ и $\int_\alpha^\beta \int_\gamma^\delta f(x, y)\,dy\,dx \le M(\delta - \gamma)(\beta - \alpha)$. Последнее неравенство означает, что повторный интеграл, соответствующий прямоугольнику разбиения, не превосходит вклада этого прямоугольника в верхнюю интегральную сумму для кратного интеграла. Сложив все такие неравенства, получим, что интеграл $\int_a^b \int_c^d f(x, y)\,dy\,dx$ не превосходит верхней интегральной суммы для кратного интеграла. Аналогично доказывается, что этот интеграл не меньше нижней интегральной суммы. При измельчении разбиения верхние и нижние интегральные суммы стремятся к кратному интегралу, поэтому кратный интеграл равен повторному. Для кратных интегралов в n-мерном пространстве соответствующее утверждение формулируется и доказывается аналогично.

Выражение кратного интеграла непрерывной функции через повторный интеграл позволяет вывести формулу объёма шара в n-мерном пространстве.

Пример 11.6. *Объём шара радиуса R в n-мерном пространстве при $n = 2k + 1$ равен $2\frac{(2\pi)^k}{(2k+1)!!}R^{2k+1}$, а $n = 2k$ равен $\frac{(2\pi)^k}{(2k)!!}R^{2k} = \frac{\pi^k}{k!}R^{2k}$. Здесь $n!!$ обозначает произведение всех натуральных чисел, не превосходящих n и имеющих с n одинаковую чётность.*

Доказательство. Пусть объём шара радиуса R в n-мерном пространстве равен $V_n(R)$. Докажем индукцией по n, что $V_n(R) = c_n R^n$, где c_n — некоторая константа, и выразим c_n через c_{n-1}. База индукции $n = 1$. Предположим, что $V_{n-1}(R) = c_{n-1}R^{n-1}$. Радиус сечения шара плоскостью, заданной уравнением $x_1 = t$, является шаром радиуса $\sqrt{R^2 - t^2}$ (при $t < R$). Поэтому

$$V_n(R) = \int_{-R}^{R} c_{n-1}(R^2 - t^2)^{(n-1)/2} dt = c_{n-1}R^n \int_{-\pi/2}^{\pi/2} \cos^n x\,dx$$

(второе равенство получено заменой переменных $t = R\sin x$). Таким образом, $c_n = 2c_{n-1}U_n$, где $U_n = \int_0^{\pi/2} \cos^n x\,dx$. Согласно задаче 6.19 а), если n нечётно, то $U_n = \frac{(n-1)!!}{n!!}$, а если n чётно, то $U_n = \frac{(n-1)!!}{n!!}\frac{\pi}{2}$. □

Для формулировки и доказательства теоремы о выражении кратного интеграла через повторный интеграл для разрывных функций нам понадобятся точная верхняя грань нижних интегральных сумм $\underline{\int}_P f$ и точная нижняя грань верхних интегральных сумм $\overline{\int}_P f$.

Теорема 11.8 (Фубини). *Пусть $A \subset \mathbb{R}^n$ и $B \subset \mathbb{R}^m$ — замкнутые параллелепипеды, $f\colon A \times B \to \mathbb{R}$ — интегрируемая функция, функция $g_x\colon B \to \mathbb{R}$ задана для каждого $x \in A$ равенством*

$g_x(y) = f(x,y)$, $L(x) = \underline{\int}_B g_x$ и $U(x) = \overline{\int}_B g_x$. Тогда функции L и U интегрируемы и $\int_A L = \int_A U = \int_{A\times B} f$.

Доказательство. Рассмотрим разбиения параллелепипедов A и B, и пусть α и β — некоторые параллелепипеды разбиения, $\sigma = \alpha \times \beta$. Будем использовать обозначения $m_S(f)$ и $M_S(f)$ для минимума и максимума функции f на множестве S. Составим нижнюю интегральную сумму для кратного интеграла

$$\begin{aligned}\sum_\sigma m_\sigma(f)v(\sigma) &= \sum_{\alpha\times\beta} m_{\alpha\times\beta}(f)v(\alpha\times\beta) = \\ &= \sum_\alpha \left(\sum_\beta m_{\alpha\times\beta}(f)v(\beta)\right) v(\alpha).\end{aligned}$$

Для $x \in \alpha$ выполняется неравенство $m_{\alpha\times\beta}(f) \le m_\beta(g_x)$, поэтому для $x\in\alpha$ получаем

$$\begin{aligned}\sum_\beta m_{\alpha\times\beta}(f)v(\beta) &\le \sum_\beta m_\beta(g_x)v(\beta) \le \\ &\le \underline{\int}_B g_x = L(x).\end{aligned}$$

Следовательно,

$$\begin{aligned}\sum_\sigma m_\sigma(f)v(\sigma) &\le \sum_\alpha m_\alpha(L(x))v(\alpha) \le \\ &\le \sum_\alpha M_\alpha(L(x))v(\alpha) \le \sum_\sigma M_\sigma(f)v(\sigma);\end{aligned}$$

здесь последнее неравенство доказывается аналогично первому. При измельчении разбиения крайние члены стремятся к кратному интегралу, поэтому $\int_A L = \int_{A\times B} f$. Равенство $\int_A U = \int_{A\times B} f$ доказывается аналогично. □

Задача 11.3. Компактное множество $C \subset \mathbb{R}^n$ таково, что любое его сечение гиперплоскостью $x_1 = a$ имеет меру нуль. Докажите, что C имеет меру нуль.

Следующий пример показывает, что если функция f произвольная, то из равенства

$$\int_a^b \int_c^d f(x,y)\,dx\,dy = \int_c^d \int_a^b f(x,y)\,dy\,dx$$

не следует существование кратного интеграла $\int_I f$ по прямоугольнику $I = [a,b]\times[c,d]$.

Пример 11.7. *Пусть A — множество точек вида $(2^{-m}p, 2^{-m}q)$, где m — целое неотрицательное числа, а числа p и q нечётные. Положим $f(x,y) = 1$ для $(x,y) \notin A$ и $f(x,y) = 0$ для $(x,y)\in A$. Тогда*

$$\int_a^b \int_c^d f(x,y)\,dx\,dy = \int_c^d \int_a^b f(x,y)\,dy\,dx = (b-a)(c-d),$$

но функция f на прямоугольнике $[a,b]\times[c,d]$ неинтегрируема.

Доказательство. Пусть $x_0 = 2^{-m}p$, где m — целое неотрицательное числа, а число p нечётное. Тогда на отрезке $[c,d]$ есть лишь конечное число точек $y = 2^{-m}q$, где число q нечётное. Поэтому $\int_c^d f(x_0,y)\,dy = d-c$. Если же число x_0 не имеет вида $2^{-m}p$, то $f(x_0,y) = 1$ для всех y.

Функция f разрывна в каждой точке прямоугольника, поэтому она неинтегрируема. □

Пример 11.8. *Зададим функцию f на квадрате $I = [0,1]\times[0,1]$ следующим образом: $f(x,y) = 0$, если хотя бы одно из чисел x, y иррационально, и $f(x,y) = 1/n$, если y рационально и x — несократимая дробь со знаменателем n. Тогда*

$$\int_I f = \int_0^1 \int_0^1 f(x,y)\,dx\,dy = 0,$$

но для рациональных x интеграл $\int_0^1 f(x,y)\,dy$ не существует.

Доказательство. При фиксированном рациональном $x = m/n$ функция $f(x, y)$ равна $1/n$ при рациональном y и 0 при иррациональном y. Такая функция от y неинтегрируема.

При фиксированном иррациональном y функция $f(x, y)$ равна 0. При фиксированном рациональном y функция $f(x, y)$ равна $1/n$ в рациональной точке $x = m/n$ и равна 0 в иррациональной точке x. Это функция Тома из примера 8.3. При разборе этого примера на с. 156 доказано, что эта функция интегрируема по Риману и её интеграл равен 0.

Так же, как и в примере 8.3, для функции $f(x, y)$ доказывается, что множество её точек разрыва, расположенных внутри квадрата, — это точки, обе координаты которых рациональны. Равенство $\int_I f = 0$ тоже доказывается точно так же. □

11.7. Криволинейные и поверхностные интегралы

Пусть на плоскости или в пространстве задана параметризованная гладкая кривая γ. Это означает, что задано непрерывно дифференцируемое отображение $\gamma\colon [a, b] \to \mathbb{R}^n$, где $n = 2$ или 3, причём $d\gamma_t \neq \mathbf{0}$ для всех $t \in [a, b]$. *Криволинейные интегралы* бывают двух типов: интеграл $\int_\gamma f(\mathbf{x})ds$ функции $f(\mathbf{x})$, заданной на кривой γ, и интеграл $\int_\gamma \mathbf{F}(\mathbf{x})d\mathbf{x}$ векторного поля $\mathbf{F}(\mathbf{x})$, заданного на кривой γ.

Рассмотрим разбиение отрезка $[a, b]$ на части точками $t_0 = a < t_1 < t_2 < \cdots < t_{n-1} < b = t_n$ и сопоставим ему разбиение кривой γ точками $\mathbf{x}_k = \gamma(t_k)$. На каждом отрезке $[t_k, t_{k+1}]$ выберем произвольную точку z_k. Ей соответствует точка $\mathbf{z}_k = \gamma(z_k)$ на кривой γ.

Интеграл $\int_\gamma f(\mathbf{x})ds$ — это предел интегральных сумм $\sum_{k=0}^{n-1} f(\mathbf{z}_k)\sigma_k$, где σ_k — длина дуги кривой γ, заключённой между точками $\mathbf{x}_k$ и $\mathbf{x}_{k+1}$, когда длина наибольшего из отрезков $[t_k, t_{k+1}]$ стремится к нулю. Если в качестве параметра на кривой γ взять длину дуги кривой, то интеграл $\int_\gamma f(\mathbf{x})ds$ равен интегралу Римана $\int_0^L f(\gamma(t))dt$. Здесь имеется в виду, что длина кривой γ равна L, γ — это отображение отрезка $[0, L]$, и длина дуги кривой γ, заключённой между точками $\gamma(0)$ и $\gamma(t)$, равна t.

Интеграл $\int_\gamma \mathbf{F}(\mathbf{x})d\mathbf{x}$ — это предел интегральных сумм

$$\sum_{k=0}^{n-1} (\mathbf{F}(\mathbf{z}_k), \mathbf{x}_{k+1} - \mathbf{x}_k),$$

составленных из скалярных произведений. Если $\mathbf{F} = (P, Q, R)$, то интеграл $\int_\gamma \mathbf{F}(\mathbf{x})d\mathbf{x}$ обычно записывают в виде $\int_\gamma (P\,dx + Q\,dy + R\,dz) = \int_\gamma P\,dx + \int_\gamma Q\,dy + \int_\gamma R\,dz$.

Криволинейный интеграл второго типа $\int_\gamma P\,dx$ тоже можно записать как обычный интеграл Римана, а именно:

$$\int_\gamma P\,dx = \int_a^b P(\mathbf{x}(t))x'(t)dt.$$

Действительно, согласно формуле конечных приращений

$$P(\mathbf{z}_k)(x(t_{k+1}) - x(t_k)) = P(\mathbf{z}_k)x'(\theta_k)(t_{k+1} - t_k)$$

для некоторого $\theta_k \in [t_k, t_{k+1}]$.

Отметим, что при изменении ориентации кривой γ, т.е. при замене её на кривую $\tilde{\gamma}(x) = \gamma(a + b - x)$, интеграл $\int_\gamma f(\mathbf{x})ds$ не изменяется, а интеграл $\int_\gamma \mathbf{F}(\mathbf{x})d\mathbf{x}$ меняет знак.

Выражение $\mathbf{F}(\mathbf{x})d\mathbf{x} = P\,dx + Q\,dy + R\,dz$, которое мы интегрируем по кривой γ — это дифференциальная 1-форма. На векторе (v_1, v_2, v_3) она принимает значение $Pv_1 + Qv_2 + Rv_3$. Мы уже встречались с дифференциальной 1-формой — дифференциалом df функции f. Несложно проверить, что $\int_\gamma df = f(\gamma(b)) - f(\gamma(a))$. Действительно, пусть функция $g\colon [a, b] \to \mathbb{R}$ задана формулой $g = f \circ \gamma$. Тогда

$$\begin{aligned} \int_\gamma df &= \int_a^b df_{\gamma(t)}(\gamma'(t))dt = \int_a^b g'(t) = \\ &= g(b) - g(a) = f(\gamma(b)) - f(\gamma(a)). \end{aligned}$$

Таким образом, интеграл дифференциала зависит только от начала и конца пути, но не от самого пути.

С помощью криволинейного интеграла вида $\int_\gamma \mathbf{F}(\mathbf{x})d\mathbf{x}$ можно вычислить ориентированную площадь фигуры, ограниченной замкнутой кривой $\gamma\colon [a,b] \to \mathbb{R}^2$ (кривая γ замкнутая, если $\gamma(a) = \gamma(b)$). Напомним, что ориентированная площадь фигуры, ограниченной несамопересекающейся параметризованной замкнутой кривой, по абсолютной величине равна площади этой фигуры. Для кривых, которые обходятся против часовой стрелки, ориентированная площадь положительна, а для кривых, которые обходятся по часовой стрелке, ориентированная площадь отрицательна.

Формулу ориентированной площади фигуры, ограниченной параметризованной кривой (в том числе и самопересекающейся), можно получить из формулы ориентированной площади A треугольника с вершинами в точках $(0,0)$, (x_1, y_1), (x_2, y_2):

$$A = \frac{1}{2}(x_1y_2 - x_2y_1).$$

С помощью этой формулы ориентированную площадь многоугольника с последовательными вершинами $P_i = (x_i, y_i)$, $i = 0,\ 1,\ \ldots,\ n$, можно определить как сумму ориентированных площадей треугольников OP_0P_1, OP_1P_2, $\ldots$, OP_nP_0, где $O = (0,0)$, т.е. как сумму

$$\frac{1}{2}\sum_{i=0}^{n}(x_iy_{i+1} - x_{i+1}y_i);$$

здесь $(x_{n+1}, y_{n+1}) = (x_0, y_0)$. Многоугольник при этом может быть и самопересекающимся.

Формулу ориентированной площади фигуры, ограниченной замкнутой кривой $\gamma\colon [a,b] \to \mathbb{R}^2$, можно теперь получить следующим образом. Выберем некоторое разбиение отрезка $[a,b]$: $a = t_0 < t_1 < \cdots < t_n = b$, рассмотрим многоугольник $P_0P_1\ldots P_n$, где $P_i = (x(t_i), y(t_i))$, и найдём предел ориентированных площадей таких многоугольников, когда наибольшее из чисел $\Delta t_i = t_i - t_{i-1}$ стремится к нулю.

Положим $x_i = x(t_i)$ и $y_i = y(t_i)$ и воспользуемся формулой конечных приращений. В результате получим, что ориентированная площадь многоугольника $P_0P_1\ldots P_n$ равна

$$\frac{1}{2}\sum_{i=0}^{n-1}(x(t_i)y'(\xi_i) - y(t_i)x'(\eta_i))\Delta t_i,$$

где $\Delta t_i = t_{i+1} - t_i$, $\xi_i, \eta_i \in [t_i, t_{i+1}]$. Если числа Δ_i положительны и наибольшее из них стремится к нулю, то мы получаем выражение ориентированной площади фигуры, ограниченной кривой γ, в виде интеграла

$$\frac{1}{2}\int_a^b(xy' - yx')dt = \frac{1}{2}\int_\gamma(x\,dy - y\,dx).$$

Используя интегрирование по частям, несложно убедиться, что

$$-\int_a^b yx'dt = \int_a^b xy'dt.$$

Действительно,

$$\int_a^b yx'dt + \int_a^b xy'dt = (xy)|_b^a = 0,$$

поскольку кривая γ замкнутая. Для криволинейных интегралов равенство $\int_\gamma(x\,dy + y\,dx) = 0$ для замкнутой кривой γ следует из того, что $x\,dy + y\,dx = df$, где $f(x,y) = xy$. Действительно, $\int_\gamma(x\,dy + y\,dx) = \int_\gamma df = 0$ для любой замкнутой кривой γ.

Таким образом, ориентированную площадь A фигуры, ограниченной гладкой замкнутой кривой γ, можно вычислить по любой из следующих трёх эквивалентных формул:

$$\frac{1}{2}\int_\gamma(x\,dy - y\,dx) = \int_\gamma x\,dy = -\int_\gamma y\,dx. \tag{11.6}$$

Перейдём теперь к поверхностным интегралам. Рассмотрим гладкую параметризованную поверхность S, состоящую из точек $\mathbf{x}(u,v) \in \mathbb{R}^3$. Параметры (u,v) принадлежат некоторой области

$D \subset \mathbb{R}^2$ и в каждой точке этой области векторы $\frac{\partial \mathbf{x}}{\partial u}$ и $\frac{\partial \mathbf{x}}{\partial v}$ линейно независимы. Поверхностный интеграл первого типа функции f по поверхности S равен

$$\int_S f\,d\sigma = \int_D f(\mathbf{x}(u,v)) \left\| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right\| du\,dv.$$

Здесь знак $\times$ обозначает векторное произведение, и $\left\| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right\|$ — это площадь параллелограмма, натянутого на векторы $\frac{\partial \mathbf{x}}{\partial u}$ и $\frac{\partial \mathbf{x}}{\partial v}$.

Задача 11.4. Пусть S — это график функции $z = f(x,y)$ над областью D, 1 — это функция, тождественно равная 1. Докажите, что

$$\int_S 1 = \int_D \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\,dx\,dy.$$

Поверхностный интеграл второго типа векторного поля $\mathbf{F}(\mathbf{x})$ по поверхности S равен

$$\int_D \left(\mathbf{F}(\mathbf{x}(u,v)), \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right) du\,dv.$$

Криволинейные и поверхностные интегралы являются интегралами дифференциальных форм. К их обсуждению мы вернёмся в параграфе 12.4. Когда мы познакомимся с дифференциальными формами, станет понятно, почему естественно рассматривать криволинейные и поверхностные интегралы именно двух типов. Первый тип интегралов связан с формой объёма.

11.8. Замена переменных в кратном интеграле

Рассмотрим сначала переход к полярным координатам: $x(r,\varphi) = r\cos\varphi$ и $y(r,\varphi) = r\sin\varphi$. Матрица Якоби отображения $\Psi(r,\varphi) = (r\cos\varphi, r\sin\varphi)$ равна $\begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$; её определитель равен r. Пусть D — область на плоскости с координатами (r,φ). Покажем, что для непрерывной функции $f(x,y)$ на области $\Psi(D)$, расположенной на плоскости с координатами (x,y), справедливо равенство

$$\int_{\Psi(D)} f = \int_D (f \circ \Psi)|\det d\Psi| = \int_D r(f \circ \Psi).$$

Можно считать, что D — прямоугольник со сторонами, параллельными осям координат. Разобьём его прямыми $r = r_i$ и $\varphi = \varphi_i$ на прямоугольники D_k. Пусть прямоугольник D_k ограничен соседними параллельными прямыми $r = r_i$ и $r = r_{i-1}$, $\varphi = \varphi_i$ и $\varphi = \varphi_{i-1}$. Его площадь $v(D_k)$ равна $|r_i - r_{i-1}| \cdot |\varphi_i - \varphi_{i-1}|$. Образ $\Psi(D_k)$ этого прямоугольника на плоскости с координатами (x,y) — криволинейный четырёхугольник, ограниченный окружностями $r = r_i$ и $r = r_{i-1}$ и лучами $\varphi = \varphi_i$ и $\varphi = \varphi_{i-1}$. Из формулы площади сектора окружности следует, что площадь $v(\Psi(D_k))$ криволинейного четырёхугольника $\Psi(D_k)$ равна $\frac{1}{2}|r_i^2 - r_{i-1}^2| \cdot |\varphi_i - \varphi_{i-1}|$. Таким образом, $v(\Psi(D_k)) = \frac{r_i + r_{i-1}}{2} v(D_k)$.

В качестве точки $\mathbf{u}_k \in D_k$ выберем точку с координатами (r_k, φ_k), где $r_k = \frac{r_i + r_{i-1}}{2}$ и $\varphi_k \in [\varphi_{i-1}, \varphi_i]$. Пусть $\mathbf{x}_k = \Psi(\mathbf{u}_k) \in \Psi(D_k)$. Тогда интеграл $\int_D r(f \circ \Psi)$ равен пределу интегральных сумм $\sum r_k f(\Psi(\mathbf{u}_k)) v(D_k)$, а интеграл $\int_{\Psi(D)} f$ равен пределу интегральных сумм $\sum f(\mathbf{x}_k) v(\Psi(D_k))$. Эти интегральные суммы почленно равны, поскольку $v(\Psi(D_k)) = \frac{r_i + r_{i-1}}{2} v(D_k) = r_k v(D_k)$.

В частности, площадь фигуры, заключённой между линиями $\varphi = \varphi_1$ и $\varphi = \varphi_2$ и кривой, заданной в полярных координатах уравнением $r = r(\varphi)$, равна

$$\int_{\varphi_1}^{\varphi_2} \int_0^{r(\varphi)} r\,dr\,d\varphi = \frac{1}{2} \int_{\varphi_1}^{\varphi_2} r^2(\varphi)\,d\varphi.$$

В частности, площадь, ограниченная лемнискатой $r^2 = \cos 2\varphi$, равна

$$2\frac{1}{2} \int_{-\pi/4}^{\pi/4} \cos 2\varphi\,d\varphi = \left. \frac{\sin 2\varphi}{2} \right|_{-\pi/4}^{\pi/4} = 1.$$

С помощью полярных координат можно по-другому вычислить интеграл Эйлера–Пуассона, который мы уже вычисляли в примере 6.4.

Пример 11.9. $\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$ *(интеграл Эйлера–Пуассона).*

Доказательство. Квадрат интеграла $I_a = \int_{-a}^{a} e^{-x^2}dx = \int_{-a}^{a} e^{-y^2}dy$ равен $\int_{-a}^{a}\int_{-a}^{a} e^{-x^2-y^2}dx\,dy$. Интеграл положительной функции $e^{-x^2-y^2}$ по квадрату $|x| \le a$, $|y| \le a$ заключён между интегралами этой функции по кругам, заданных в полярных координатах неравенствами $r \le 1$ и $r \le \sqrt{2}$ (меньший круг содержится в квадрате, а больший содержит квадрат). Интеграл по кругу $r \le b$ равен

$$\int_0^{2\pi}\left(\int_0^b e^{-r^2} r\,dr\right)d\varphi = \int_0^{2\pi}\left(\frac{1}{2} - \frac{1}{2}e^{-b^2}\right)d\varphi = \pi(1-e^{-b^2}).$$

Поэтому

$$\pi(1-e^{-a^2}) < I_a^2 < \pi(1-e^{-4a^2}).$$

Устремляя a к бесконечности, получаем требуемое. □

Перейдём к многомерному случаю. Начнём с простейшего, но важного примера, когда замена переменных линейная. Напомним, что невырожденное линейное отображение $A\colon \mathbb{R}^n \to \mathbb{R}^n$ переводит параллелепипед, объём которого равен V, в параллелепипед, объём которого равен $|\det A|V$. Пусть $S \subset \mathbb{R}^n$ — измеримое по Жордану множество, $f(\mathbf{x})$ — интегрируемая функция на S, $A(S)$ — образ множества S при отображении A, $\mathbf{x} = A\mathbf{y}$ — образ точки $\mathbf{y}$ при отображении A, $g = f \circ A$, т.е. $g(\mathbf{y}) = f(A\mathbf{y}) = f(\mathbf{x})$. Выясним, какой должна быть функция $\varphi(\mathbf{y})$, чтобы заведомо выполнялось равенство $\int_{A(D)} f = \int_A \varphi$. Для выполнения этого равенства достаточно, чтобы в интегральных суммах были равны соответственные слагаемые $\varphi(\mathbf{y}_k)v(P_k)$ и $f(\mathbf{x}_k)v(AP_k) = f(\mathbf{x}_k)|\det A|v(P_k) = g(\mathbf{y}_k)|\det A|v(P_k)$. Таким образом, если $\varphi(\mathbf{y}) = g(\mathbf{y})|\det A| = (f\circ A)(\mathbf{y})|\det A|$, то рассматриваемые интегралы равны, и мы приходим к равенству $\int_{A(D)} f = \int_A (f\circ A)|\det A|$.

Напомним формулу замены переменной в интеграле Римана (см. с. 97). Если функция φ имеет непрерывную производную на отрезке $[\alpha,\beta]$, $\varphi(\alpha) = a$, $\varphi(\beta) = b$ и функция f непрерывна на области значений функции φ, то

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

Если $a < b$, то $\int_{[a,b]} f = \int_a^b f(x)dx = -\int_b^a f(x)dx$. Поэтому в том случае, когда φ' не обращается в нуль, для кратного интеграла по отрезку формула замены переменной следующая:

$$\int_{\varphi([\alpha,\beta])} f(x)dx = \int_{[\alpha,\beta]} f(\varphi(t))|\varphi'(t)|dt.$$

Теорема 11.9. *Пусть U и V — открытые подмножества в $\mathbb{R}^n$, $\Psi\colon V \to U$ — диффеоморфизм класса C^1, $f\colon U \to \mathbb{R}$ — непрерывная функция с компактным носителем, содержащимся в V. Тогда*

$$\int_{\Psi(V)} f = \int_U f = \int_V (f\circ\Psi)|\det d\Psi|. \tag{11.7}$$

Доказательство. Рассмотрим сначала случай $n = 2$. Пусть $\Psi(u,v) = (x(u,v), y(u,v))$. Воспользуемся теоремой 10.14 о представлении диффеоморфизма в виде композиции простейших диффеоморфизмов. Если Ψ — перестановка двух координат, то формула (11.7) верна; это следует из того, что кратный интеграл выражается через повторный и для повторного интеграла верна теорема об изменении порядка интегрирования. Проверим теперь, что формула (11.7) верна для простейшего диффеоморфизма, например, когда $u = x$ и $\Psi(x,v) = (x, \psi(x,v))$; для диффеоморфизма функция $\psi(x,v)$ монотонна по v. В этом случае согласно формуле замены переменной для кратного интеграла по отрезку при фиксированном $x = x_0$

$$\int_{[\psi(x_0,\alpha),\psi(x_0,\beta)]} f(x_0,y) = \int_{[\alpha,\beta]} f(x_0,\psi(x_0,v))\left|\det d\Psi_{(x_0,v)}\right|,$$

поскольку $\det d\Psi_{(x,v)} = \frac{\partial\psi}{\partial v}(x,v)$. Проинтегрировав полученное равенство по $x \in [\delta,\gamma]$, получим формулу (11.7) кратного интеграла по прямоугольнику $[\delta,\gamma]\times[\alpha,\beta]$ на плоскости $(u,v) = (x,v)$.

Если формула (11.7) верна для двух диффеоморфизмов, то она верна и для их композиции, поскольку матрица Якоби композиции двух отображений равна произведению их матриц Якоби.

Таким образом, если носитель функции f содержится в открытом множестве V, для которого диффеоморфизм $\Psi\colon V \to U$ представлен в виде композиции простейших, то формула (11.7) верна. Для каждой точки $\mathbf{x}$ компактного носителя K функции f можно выбрать открытый шар $V_{\mathbf{x}}$ с центром $\mathbf{x}$, ограничение диффеоморфизма Ψ на который представляется в виде композиции простейших. Пусть $W_{\mathbf{x}}$ — шар вдвое меньшего радиуса с тем же центром $\mathbf{x}$. Конечный набор шаров $W_1, \dots, W_m$ покрывает компактное множество K. Пусть функция β_i равна 0 вне шара V_i и равна 1 на шаре W_i вдвое меньшего радиуса с тем же центром (см. задачу 9.1). Положим $\alpha_1 = \beta_1$, $\alpha_2 = (1-\beta_1)\beta_2, \dots, \alpha_m = (1-\beta_1)\dots(1-\beta_{m-1})\beta_m$. Ясно, что $\alpha_m + (1-\beta_1)\dots(1-\beta_{m-1})(1-\beta_m) = (1-\beta_1)\dots(1-\beta_{m-1})$, поэтому

$$\alpha_1 + \dots + \alpha_m = 1 - (1-\beta_1)\dots(1-\beta_{m-1})(1-\beta_m).$$

В каждой точке обращается в нуль хотя бы один из множителей $(1-\beta_i)$, поэтому $\alpha_1 + \dots + \alpha_m = 1$. Носитель непрерывной функции $\alpha_i f$ содержится в V_i, поэтому для неё верна формула (11.7). Следовательно, эта формула верна и для функции $f = \sum \alpha_i f$.

Случай произвольного n не требует существенных изменений. □

В частности, для объёмов (когда $f = 1$) получаем соотношение $\int_{\Psi(V)} 1 = \int_V |\det d\Psi|$. Если $\det d\Psi = \pm 1$, то диффеоморфизм Ψ сохраняет объём.

Историческое замечание. Формулу замены переменных для двойного интеграла получил Эйлер в 1769 году, тройного — Лагранж в 1773 году. Для произвольной размерности формулу замены переменных получил Якоби в 1841 году.

11.9. Сферические координаты

В трёхмерном пространстве аналогом полярных координат являются *сферические координаты* (r, θ, φ). Декартовы координаты (x, y, z) выражаются через сферические координаты следующим образом:

$$\begin{aligned} x &= r\sin\theta\cos\varphi, \\ y &= r\sin\theta\sin\varphi, \\ z &= r\cos\theta. \end{aligned}$$

Полярные координаты изменяются в следующих пределах: $0 \le \theta \le \pi$, $0 \le \varphi < 2\pi$, $0 \le r < \infty$.

Задача 11.5. Докажите, что якобиан отображения $\Psi(r, \theta, \varphi) = (x, y, z)$ равен $r^2 \sin\theta$.

Согласно задаче 11.5 формула замены переменных для сферических координат выглядит следующим образом:

$$\int\int_{\Psi(V)}\int f(\mathbf{x})dx\,dy\,dz = \int\int_V\int f(\mathbf{x})r^2\sin\theta\,dr\,d\theta\,d\varphi.$$

где $\mathbf{x}$ — точка с декартовыми координатами (x, y, z) и сферическими координатами (r, θ, φ).

При $r = 1$ сферические координаты (θ, φ) являются координатами на единичной сфере. В этих координатах поверхностный интеграл первого типа по единичной сфере записывается в виде

$$\int\int f(\mathbf{x})\sin\theta\,d\theta\,d\varphi, \tag{11.8}$$

поскольку в сферических координатах элемент площади как раз и равен $\sin\theta\,d\theta\,d\varphi$.

Пусть единичная сфера разбита на области D_i, в области D_i выбрана точка $\mathbf{x}_i$, площадь области D_i равна Δ_i. Поверхностный интеграл (11.8) равен пределу интегральных сумм $\sum f(\mathbf{x}_i)\Delta_i$, когда диаметр области D_i стремится у нулю.

Поверхностный интеграл (11.8) инвариантен относительно движений пространства, сохраняющих единичную сферу, в следующем смысле. Если A — ортогональное линейное преобразование, то

$$\int\int f(A\mathbf{x})\sin\theta\,d\theta\,d\varphi = \int\int f(\mathbf{x})\sin\theta\,d\theta\,d\varphi.$$

Это следует из того, что у этих интегралов есть почленно совпадающие интегральные суммы: они получаются друг из друга преобразованиями A и A^{-1}.

Для вычисления среднего значения длины проекции вектора в трёхмерном пространстве на прямую или на плоскость нужен именно такой инвариантный интеграл, чтобы среднее значение зависело только от длины вектора, и не зависело от его направления. В § 6.18 при вычислении на плоскости среднего значения проекции вектора на прямую мы использовали интеграл $\int_0^{2\pi} f(\varphi)d\varphi$, инвариантность которого относительно поворотов $\varphi \mapsto \varphi + C$ слишком очевидна, поэтому там мы о ней специально не говорили. (Строго говоря, мы использовали интеграл $\int_0^{\pi} f(\varphi)d\varphi$ для функций, имеющих период π. Но это фактически то же самое.)

Среднее значение функции $f(\mathbf{x})$ на единичной сфере равно отношению интеграла этой функции к интегралу функции, тождественно равной 1. Поэтому вычислим сначала интеграл по сфере функции, тождественно равной 1:

$$\int_0^{2\pi} \left(\int_0^{\pi} \sin\theta \, d\theta \right) d\varphi = \int_0^{2\pi} 2 \, d\varphi = 4\pi.$$

Ничего удивительного: поверхностный интеграл первого типа функции, тождественно равной 1, равен площади поверхности.

Пример 11.10. *Среднее значение длины проекции на прямую вектора длины a в $\mathbb{R}^3$ равно $\frac{a}{2}$.*

Доказательство. Пусть $\mathbf{x}$ — точка на единичной сфере, $f(\mathbf{x})$ — длина проекции вектора $(0, 0, a)$ на прямую, проходящую через точку $\mathbf{x}$ и начало координат. Угол между векторами $(0, 0, a)$ и $\mathbf{x}$ равен θ, поэтому $f(\mathbf{x}) = a|\cos\theta|$ и интеграл функции f по единичной сфере равен

$$\int_0^{2\pi} \int_0^{\pi} a|\sin\theta\cos\theta| d\theta \, d\varphi = 2\pi a.$$

□

Задача 11.6. В $\mathbb{R}^3$ даны два набора векторов, причём сумма длин проекций векторов первого набора на любую прямую не больше суммы длин проекций векторов второго набора на ту же прямую. Докажите, что сумма длин векторов первого набора не больше суммы длин векторов второго набора.

Задача 11.7. Сумма длин векторов в $\mathbb{R}^3$ равна a. Докажите, что среди них можно выбрать несколько векторов, длина суммы которых не меньше $a/4$.

Пример 11.11. *Среднее значение длины проекции на плоскость вектора длины a в $\mathbb{R}^3$ равно $\frac{\pi a}{4}$.*

Доказательство. Пусть $\mathbf{x}$ — точка на единичной сфере, $f(\mathbf{x})$ — длина проекции вектора $(0, 0, a)$ на плоскость, перпендикулярную единичному вектору $\mathbf{x}$. Угол между векторами $(0, 0, a)$ и $\mathbf{x}$ равен θ, поэтому $f(\mathbf{x}) = a\sin\theta$ и интеграл функции f по единичной сфере равен

$$\int_0^{2\pi} \int_0^{\pi} a\sin^2\theta \, d\theta \, d\varphi = \pi^2 a.$$

□

11.10. Инвариантное интегрирование на пространстве прямых

В предыдущем параграфе мы определили инвариантный интеграл на сфере. Его геометрический смысл был вполне ясен. В этом параграфе мы определим инвариантный интеграл на множестве прямых на плоскости. Геометрия множества, точки которого — прямые на плоскости, не столь проста; это множество мы будем описывать с помощью координат.

Каждой прямой на плоскости, не проходящей через начало координат, можно сопоставить координаты (p, θ) следующим образом. Проведём из начала координат O перпендикуляр OP к этой прямой; тогда p — длина отрезка OP, θ — угол поворота от луча Ox к лучу OP. Прямая с координатами (p, θ) задаётся уравнением $x\cos\theta + y\sin\theta = p$.

Таким образом, нам нужно определить интеграл на множестве точек плоскости с координатами (p, θ), где $p > 0$ и $0 \le \theta < 2\pi$. Этот интеграл должен быть инвариантным относительно движений плоскости. Обсудим сначала, как устроены инвариантный интеграл на плоскости с декартовыми координатами (x, y). Будем вычислять площадь $A(S)$ множества S по формуле

$$A(S) = \iint_S f(x,y)dx\,dy,$$

где $f(x, y)$ — некоторая функция. Для каких функций f определённая таким образом площадь сохраняется при движениях? Убедимся, что только для констант. Рассмотрим движение $(\bar{x}, \bar{y}) \to (x, y)$, заданное формулами

$$x = a + \bar{x}\cos\varphi - \bar{y}\sin\varphi, \quad y = b + \bar{x}\sin\varphi + \bar{y}\cos\varphi \tag{11.9}$$

(так можно задать любое движение плоскости, сохраняющее ориентацию). Пусть при этом движении фигура $\bar{S}$ переходит в фигуру S. Нас интересуют функции f, для которых

$$\iint_{\bar{S}} f(\bar{x},\bar{y})d\bar{x}\,d\bar{y} = \iint_S f(x,y)dx\,dy$$

для любого движения и для любой фигуры. По формуле замены переменных в кратном интеграле

$$\iint_S f(x,y)dx\,dy = \iint_{\bar{S}} f(x(\bar{x},\bar{y}), y(\bar{x},\bar{y}))\frac{\partial(x,y)}{\partial(\bar{x},\bar{y})}d\bar{x}\,d\bar{y},$$

где

$$\frac{\partial(x,y)}{\partial(\bar{x},\bar{y})} = \begin{vmatrix} \frac{\partial x}{\partial \bar{x}} & \frac{\partial x}{\partial \bar{y}} \\ \frac{\partial y}{\partial \bar{x}} & \frac{\partial y}{\partial \bar{y}} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{vmatrix} = 1.$$

Таким образом,

$$\iint_{\bar{S}} f(x(\bar{x},\bar{y}), y(\bar{x},\bar{y}))\frac{\partial(x,y)}{\partial(\bar{x},\bar{y})}d\bar{x}\,d\bar{y} = \iint_{\bar{S}} f(x(\bar{x},\bar{y}), y(\bar{x},\bar{y}))d\bar{x}\,d\bar{y},$$

т.е. нас интересуют функции, для которых

$$\iint_{\bar{S}} f(\bar{x},\bar{y})d\bar{x}\,d\bar{y} = \iint_{\bar{S}} f(x(\bar{x},\bar{y}), y(\bar{x},\bar{y}))d\bar{x}\,d\bar{y}$$

для любой фигуры S и любого движения. Из того, что это равенство выполняется для любой фигуры, следует, что для любого движения выполняется равенство

$$f(x(\bar{x},\bar{y}), y(\bar{x},\bar{y})) = f(\bar{x},\bar{y}),$$

но тогда $f(x, y) = \mathrm{const}$, потому что движением можно перевести любую точку плоскости в любую другую точку.

Теперь мы можем ввести инвариантный интеграл на множестве прямых. Движение (11.9) переводит прямую $x\cos\theta + y\sin\theta = p$ в прямую

$$\bar{x}\cos(\theta-\varphi) + \bar{y}\sin(\theta-\varphi) = p - a\cos\varphi - b\sin\varphi.$$

Таким образом, если мы сопоставляем прямой точку с координатами (p, θ), то движение (11.9) переводит точку (p, θ) в точку $(p - a\cos\varphi - b\sin\varphi, \theta - \varphi)$. Несложно проверить, что якобиан этого отображения тоже равен 1 и таким отображением можно перевести любую точку, соответствующую прямой, в любую другую точку, соответствующую прямой. Поэтому площадь (меру) множества S прямых тоже естественно определить как $\iint_S dp\,d\theta$. При этом прямые, проходящие через начало координат, можно не учитывать, потому что эти прямые образуют множество меры нуль.

Теорема 11.10. *Пусть γ — гладкая кривая длины l. Тогда мера множества прямых (с учётом кратности), пересекающих кривую γ, равна $2l$.*

Доказательство. Сначала рассмотрим случай, когда кривая γ — отрезок длины l. Мера инвариантна относительно движений, поэтому можно считать, что этот отрезок расположен так, что его середина — начало координат, а сам он лежит на оси Ox. Тогда мера множества прямых, пересекающих γ, равна

$$\iint dp\,d\theta = \int_0^{2\pi}\left(\int_0^{(l/2)|\cos\theta|} dp\right) d\theta = \int_0^{2\pi} \frac{l}{2}|\cos\theta| d\theta = 2l.$$

Затем рассмотрим случай, когда кривая γ — ломаная длины l, состоящая из i звеньев длины l_i каждое. В этом случае прямая может пересекать несколько звеньев, и тогда её нужно учитывать с весом $n = n(p,\theta)$, равным количеству точек пересечения. (В вершине ломаной прямая пересекает два отрезка, но такие прямые образуют множество меры нуль.) Применив полученную выше формулу ко всем звеньям ломаной и просуммировав, получаем

$$\iint n\,dp\,d\theta = 2\sum_i l_i = 2l.$$

Общий случай получается предельным переходом. $\square$

Историческое замечание. Первые формулы интегральной геометрии получил Морган Крофтон (1826-1915).

11.11. Решения задач

11.1. а) Ясно, что $\int_a^b \int_c^d \sin x \sin y\,dy\,dx = (\cos b - \cos a)(\cos d - \cos c)$. Если хотя бы одно из чисел $b-a$ и $d-c$ кратно 2π, то один из этих множителей равен 0.

б) Ясно, что $\int_0^b \int_0^d \sin x \sin y\,dy\,dx = (\cos b - 1)(\cos d - 1)$. Если $\cos\varphi = 1$, то $\varphi = 2n\pi$, где число n целое.

в) Увеличим данный прямоугольник в 2π раз и расположим его так, чтобы одна вершина была началом координат, а выходящие из неё стороны располагались на осях координат. Согласно задаче а) интеграл функции $\sin x \sin y$ по каждому из прямоугольников, на которые разрезан данный прямоугольник, равен нулю. Интеграл по всему прямоугольнику равен сумме интегралов по прямоугольникам, на которые он разрезан. Это очевидно в случае, когда прямоугольник разрезан на два прямоугольника, а к этому случаю всё сводится, если продолжить все линии разрезов. Следовательно, интеграл по всему прямоугольнику равен 0. Поэтому согласно задаче б) одна из сторон этого прямоугольника кратна 2π.

11.2. Предположим, что $d \geq a+b+c$. Тогда функция $g(y) = \int_c^{c+a} f(x+y)dx$ определена при $0 \leq y \leq b$, поэтому можно рассмотреть интеграл

$$\int_0^b \left(\int_c^{c+a} f(x+y)dx\right) dy = \int_c^{c+a}\left(\int_0^b f(x+y)dy\right) dx.$$

При этом $\int_c^{c+a} f(x+y)dx > 0$, а $\int_0^b f(x+y)dy < 0$.

11.3. Из любого покрытия компактного множества открытыми множествами можно выбрать конечное подпокрытие, поэтому компактное множество имеет меру нуль тогда и только тогда, когда его мера Жордана равна нулю. Заключим множество C в прямоугольный параллелепипед $A\times B$, где A — одномерный параллелепипед (отрезок), а B — параллелепипед размерности $(n-1)$. Применим теорему Фубини к характеристической функции $f = \chi_C$ множества C. Для $a \in A$ положим $g_a(\mathbf{y}) = f(a, \mathbf{y})$. Интеграл $\int_B g_a(\mathbf{y})$ равен мере Жордана сечения множества C гиперплоскостью $x_1 = a$, т.е. равен нулю. Поэтому $\int_{A\times B} f = 0$, т.е. мера Жордана множества C равна нулю.

11.4. Если $\mathbf{x} = (x, y, f(x,y))$, то $\frac{\partial \mathbf{x}}{\partial x} = (1, 0, f_x)$ и $\frac{\partial \mathbf{x}}{\partial y} = (0, 1, f_y)$, поэтому

$$\left\|\frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v}\right\| = \|(-f_x, -f_y, 1)\| = \sqrt{f_x^2 + f_y^2 + 1}.$$

11.5. Матрица Якоби имеет вид

$$\begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}.$$

Раскладывая определитель этой матрицы по последней строке, получаем, что он равен $r^2\cos^2\theta\sin\theta\Delta + r^2\sin^3\theta\Delta = r^2\sin\theta\Delta$, где $\Delta = \begin{vmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{vmatrix} = 1$.

11.6. Из условия следует, что сумма средних значений длин проекций для первого набора не больше, чем для второго, т.е. половина суммы длин векторов первого набора не больше половины сумм длин векторов второго набора.

11.7. Сумма длин проекций данных векторов на некоторую прямую не меньше среднего значения, т.е. не меньше $a/2$. Выберем на этой прямой направление и разделим векторы на две группы: с неотрицательными проекциями на эту прямую и с неположительными. Для одной из этих групп сумма длин проекций не меньше $a/4$. Длина проекции суммы этих векторов не меньше $a/4$.

Глава 12.

Анализ на многообразиях

12.1. Определение и основные свойства

Понятие топологического пространства введено на с. 54. Для определения многообразия нам понадобятся некоторые понятия, относящиеся к топологическим пространствам.

Гомеоморфизм топологических пространств X и Y — это такое непрерывное отображение $f\colon X \to Y$, что у него есть обратное отображение f^{-1} и оно тоже непрерывно. Гомеоморфизм устанавливает взаимно однозначное соответствие не только между множествами X и Y, но и между открытыми множествами в X и открытыми множествами в Y.

База топологии — это такой набор открытых множеств, что любое открытое множество можно представить в виде объединения множеств из базы. Например, базой пространства $\mathbb{R}^n$ является множество открытых шаров рационального радиуса с центрами в точках с рациональными координатами. Эта база счётная, т.е. $\mathbb{R}^n$ — топологическое пространство со счётной базой.

Топологическое пространство X называют *хаусдорфовым*, если у любых двух его точек есть непересекающиеся окрестности, т.е. если $x, y \in X$ и $x \neq y$, то существуют непересекающиеся открытые множества $U \ni x$ и $V \ni y$. Пространство $\mathbb{R}^n$ хаусдорфово: в качестве непересекающихся окрестностей точек $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ можно взять открытые шары с центрами $\mathbf{x}$ и $\mathbf{y}$ радиуса $d/3$, где $d = \|\mathbf{x} - \mathbf{y}\|$.

Топологическое пространство X со счётной базой называют *топологическим многообразием*, если оно хаусдорфово и любая точка $x \in X$ обладает открытой окрестностью, гомеоморфной открытому множеству в $\mathbb{R}^n$. Пусть M^n — топологическое многообразие размерности n. Пару (U, φ), где $U \subset M^n$ — связное открытое множество и $\varphi\colon U \to \varphi(U)$ — гомеоморфизм на открытое подмножество в $\mathbb{R}^n$, будем называть *картой*, или *локальной системой координат*. Если при этом $\varphi(x) = 0$, то будем говорить, что (U, φ) — *локальная система координат с началом в точке* x.

Гладкая структура класса C^k на топологическом многообразии M — это семейство $\mathcal{A}$ локальных систем координат $\{(U_\alpha, \varphi_\alpha), \alpha \in A\}$, обладающее следующими свойствами:

1) множества U_α покрывают M;

2) если $U_\alpha \cap U_\beta \neq \varnothing$, то $\varphi_\alpha \varphi_\beta^{-1}$ — отображение класса C^k;

3) семейство $\mathcal{A}$ максимально в том смысле, что если (U, φ) — локальная система координат и все отображения $\varphi_\alpha \varphi^{-1}$ и $\varphi \varphi_\alpha^{-1}$ $(U_\alpha \cap U \neq \varnothing)$ гладкие, то $(U, \varphi) \in \mathcal{A}$.

Набор карт, покрывающих M и обладающих свойством 2), называют *атласом*.

Топологическое многообразие M с заданной гладкой структурой называют *гладким многообразием*. В анализе обычно рассматривают гладкие многообразия класса C^∞. В дальнейшем *многообразием* мы будем называть гладкое многообразия класса C^∞.

Чтобы получить определение *многообразия с краем*, нужно считать картами также и гомеоморфизмы $\varphi\colon U \to \varphi(U)$, где $\varphi(U)$ — открытое подмножество в топологическом пространстве $\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon x_1 \geq 0\}$.

При определении гладкой структуры на многообразии с краем мы будем предполагать, что свойство 2) таково:

2) если $U_\alpha \cap U_\beta \neq \varnothing$, то существуют взаимно обратные гладкие отображения $f_{\alpha\beta}$ и $f_{\beta\alpha}$ открытых множеств в $\mathbb{R}^n$, ограничениями которых являются отображения $\varphi_\alpha\varphi_\beta^{-1}$ и $\varphi_\beta\varphi_\alpha^{-1}$.

Пусть M^n — многообразие с краем. Будем говорить, что $x \in M^n$ — *точка края*, если у неё есть такая карта $\varphi\colon U \to \varphi(U) \subset \mathbb{R}^n_+$, что

$$\varphi(x) \in \partial\mathbb{R}^n_+ = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \colon x_1 = 0\};$$

здесь имеется в виду гладкая карта, т. е. карта из гладкой структуры. *Краем* многообразия M^n будем называть множество всех точек края. Будем говорить, что $x \in M^n$ — *внутренняя точка*, если у неё есть такая карта $\varphi\colon U \to \varphi(U)$, что $\varphi(U)$ — открытое множество в $\mathbb{R}^n$; здесь снова имеется в виду гладкая карта.

Задача 12.1. Докажите, что внутренняя точка многообразия не может быть точкой края.

Компактное многообразие без края называют *замкнутым*.

Подмножество $N \subset M^n$ называют k-мерным *подмногообразием*, если для любой точки $x \in N$ найдётся такая карта (U, φ), что $U \cap N = \varphi^{-1}\bigl(\mathbb{R}^k \cap \varphi(U)\bigr)$, где $\mathbb{R}^k$ стандартно вложено в $\mathbb{R}^n$ (т. е. рассматриваются точки, последние $n-k$ координат которых равны 0).

Пусть M^m и N^n — многообразия с гладкими структурами $\mathcal{A} = \{(U_\alpha, \varphi_\alpha)\}$ и $\mathcal{B} = \{(V_\beta, \psi_\beta)\}$. Отображение $f\colon M^m \to N^n$ называют *гладким*, если все отображения $\psi_\beta f \varphi_\alpha^{-1}$ являются гладкими (т.е. бесконечно дифференцируемыми). Отображение $\psi_\beta f \varphi_\alpha^{-1}$ определено на открытом множестве $\varphi_\alpha^{-1}(U_\alpha \cap f^{-1}(V_\beta)) \subset \mathbb{R}^m$; оно является отображением в $\mathbb{R}^n$.

Если $f\colon M^m \to N^n$ и $g\colon N^n \to M^m$ — гладкие взаимно обратные отображения, то отображение f называют *диффеоморфизмом*, а многообразия M^m и N^n называют *диффеоморфными*. Из теоремы об обратной функции следует, что если многообразия M^m и N^n диффеоморфны, то $m = n$.

Пусть $f\colon M^m \to N^n$ — гладкое отображение многообразий без края, $x \in M^m$. Выбрав локальные системы координат в точках x и $f(x)$, можно рассмотреть матрицу Якоби $\left(\frac{\partial f_i}{\partial x_j}(x)\right)$. Ранг этой матрицы не зависит от выбора локальных систем координат. Этот ранг называют *рангом* отображения f в точке x; мы будем обозначать его $\operatorname{rank} f(x)$.

Теорема 12.1. *Если гладкое отображение $f\colon M^m \to N^n$ имеет постоянный ранг r, для любой точки $a \in M^m$ можно выбрать локальную систему координат $\varphi\colon U \to \mathbb{R}^m$ с началом a и локальную систему координат $\psi\colon V \to \mathbb{R}^n$ с началом $f(a) \in N^n$ так, что $f(U) \subset V$ и*

$$\psi \circ f \circ \varphi^{-1}(x_1, \ldots, x_m) = (x_1, \ldots, x_r, 0, \ldots, 0).$$

Доказательство. Требуемое утверждение локальное, поэтому для его доказательства достаточно ввести локальные координаты и применить теорему 10.11. □

Пусть $f\colon M^m \to N^n$ — гладкое отображение. Если $\operatorname{rank} f(x) = m$ в точке $x \in M^m$, то отображение f называют *иммерсией*, или *погружением*, в точке x, а если $\operatorname{rank} f(x) = n$ в точке $x \in M^m$, то отображение f называют *субмерсией* в точке x. Если отображение является иммерсией (субмерсией) в каждой точке $x \in M^m$, то это отображение называют просто иммерсией (субмерсией). Отображение f, которое является погружением и гомеоморфно отображает M^m на $f(M^m) \subset N^n$ называют *вложением*.

В случае иммерсии и в случае субмерсии доказательство теоремы 10.11 о строении отображения постоянного существенно проще, чем в общем случае. Кроме того, в этих случаях нет необходимости менять сразу две системы координат (в образе и в прообразе). Подробности можно найти в формулировке и в решении следующей задачи.

Задача 12.2. Выведите непосредственно из теоремы об обратной функции, что при приведении гладкого отображение $f\colon M^m \to N^n$ к виду, указанному в теореме 12.1, в случае субмерсии систему координат в окрестности точки $f(a) \in N^n$ можно выбрать произвольно, а в случае иммерсии систему координат в окрестности точки $a \in M^m$ можно выбрать произвольно.

С помощью теоремы 12.1 легко доказывается следующее утверждение.

Теорема 12.2. *Пусть $f\colon M^m \to N^n$ — гладкое отображение, $X^k \subset N^n$ — подмногообразие. Предположим, что отображение f является субмерсией в каждой точке множества $f^{-1}(X^k)$. Тогда $f^{-1}(X^k)$ — подмногообразие в M^m размерности $k + m - n$.*

Доказательство. Пусть $a \in f^{-1}(X^k)$. В окрестности точки $f(a)$ можно выбрать локальные координаты $x = (u, v)$, где $u \in \mathbb{R}^k$ и $v \in \mathbb{R}^{n-k}$, так, что в этой окрестности множество X^k задаётся уравнением $v = 0$.

В точке a отображение f является субмерсией, поэтому согласно теореме 12.1 в окрестности точки a можно выбрать локальные координаты (x, y), где $x \in \mathbb{R}^n$ и $y \in \mathbb{R}^{m-n}$, так, что в локальных координатах, выбранных в точках a и $f(a)$, отображение f запишется в виде $(x, y) \mapsto x$, т.е. $(u, v, y) \mapsto (u, v)$. При этом множество $f^{-1}(X^k)$ в выбранной координатной окрестности задаётся уравнением $v = 0$. Следовательно, множество $f^{-1}(X^k)$ является подмногообразием размерности $k + m - n$. $\square$

С помощью теоремы 12.2 можно доказывать, что некоторые подмножества многообразия являются подмногообразиями и тем самым доказывать, что они являются многообразиями.

Пример 12.1. *Сфера S^n является многообразием.*

Доказательство. Рассмотрим отображение $f\colon \mathbb{R}^{n+1} \to \mathbb{R}^1$, заданное формулой $f(x_1, \ldots, x_{n+1}) = x_1^2 + \cdots + x_{n+1}^2$. В любой точке множества $f^{-1}(1)$ ранг матрицы Якоби $J = (2x_1, \ldots, 2x_{n+1})$ равен 1, поэтому множество $f^{-1}(1) = S^n$ является подмногообразием в $\mathbb{R}^{n+1}$. $\square$

Пример 12.2. *Подмножество ортогональных матриц является подмногообразием размерности $\frac{n(n-1)}{2}$ в пространстве всех квадратных матриц порядка n.*

Доказательство. Для любой матрицы A матрица A^TA симметрична, поэтому отображение $f(A) = A^TA$ переводит пространство всех матриц в пространство симметрических матриц, размерность которого равна $\frac{n(n+1)}{2}$. Прообраз единичной матрицы I при отображении f — множество ортогональных матриц. Согласно задаче 9.31 б) $df_A(X) = A^TX + X^TA$. Если матрица A ортогональная, а матрица S симметрическая, то

$$df_A\left(\frac{AS}{2}\right) = A^T\frac{AS}{2} + \left(\frac{AS}{2}\right)^T A = \frac{A^TAS}{2} + \frac{S^TA^TA}{2} = S.$$

Поэтому в каждой точке A множества $f^{-1}(I)$ отображение df_A имеет наибольший возможный ранг. $\square$

Историческое замечание. Понятие многомерного многообразия ввёл Бернхард Риман (1826-1866) в 1854 году. Он перенёс на многомерные пространства способ измерения расстояний вдоль кривых, разработанный Карлом Фридрихом Гауссом (1777-1855). Термин «многообразие» (Mannigfaltigkeit) восходит к Гауссу; он использовал этот термин в 1851 году на лекциях по многомерной геометрии.

12.2. Касательное пространство

Касательные векторы нужны для многих целей, например, для того, чтобы определить дифференциал отображения гладких многообразий.

Касательный вектор в точке $x \in M^n$ легко определить в локальной системе координат, но при переходе к другой системе координат возникают некоторые трудности. Поэтому есть несколько определений касательного вектора, которые бывают полезны в разных ситуациях.

Одно из наиболее естественных определений таково. *Касательный вектор* в точке $x \in M^n$ — это некий объект, которому в каждой локальной системе координат (U, φ) с началом в точке x соответствует определённый вектор $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$; при этом в локальной системе координат (V, ψ) тому же самому касательному вектору соответствует вектор $w = (w_1, \ldots, w_n)$, где

$$w_i = \sum \frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0)v_j. \qquad (12.1)$$

Иными словами, w — образ вектора v под действием матрицы Якоби отображения перехода $\psi\varphi^{-1}$. Корректность этого определения следует из того, что матрица Якоби композиции двух отображений является произведением матриц Якоби этих отображений.

Основной недостаток этого определения — зависимость от выбора системы координат. Чтобы получить инвариантное определение, можно поступить разными способами.

Во-первых, касательный вектор можно определить как класс эквивалентных кривых. Вектору $\mathbf{v} \in \mathbb{R}^n$ можно сопоставить семейство всех гладких кривых $\gamma\colon (-1, 1) \to \mathbb{R}^n$, для которых $\gamma(0) = 0$ и $\frac{d\gamma}{dt}(0) = \mathbf{v}$. Если (U, φ) — локальная система координат с началом в точке $x \in M$, то кривой $\gamma(t)$ можно сопоставить кривую $\widetilde{\gamma} = \varphi^{-1}\gamma$ на многообразии M^n; при этом $\widetilde{\gamma}(0) = x$. Поэтому касательный вектор в точке $x \in M$ можно определить как класс эквивалентности гладких кривых $\widetilde{\gamma}\colon (-1, 1) \to M^n$, для которых $\widetilde{\gamma}(0) = x$. Кривые $\widetilde{\gamma}_1$ и $\widetilde{\gamma}_2$ считаются эквивалентными, если для некоторой системы координат (U, φ) с началом в точке x выполняется равенство

$$\left.\frac{d(\varphi\widetilde{\gamma}_1(t))}{dt}\right|_{t=0} = \left.\frac{d(\varphi\widetilde{\gamma}_2(t))}{dt}\right|_{t=0}.$$

Если (V, ψ) — другая система координат с началом в точке x, то

$$\begin{aligned}\left.\frac{d(\psi\widetilde{\gamma}(t))_i}{dt}\right|_{t=0} &= \left.\frac{d(\psi\varphi^{-1}\varphi\widetilde{\gamma}(t))_i}{dt}\right|_{t=0} = \\ &= \sum_j \frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0)\left.\frac{d(\varphi\widetilde{\gamma}(t))_j}{dt}\right|_{t=0}.\end{aligned}$$

Поэтому, во-первых, эквивалентность кривых не зависит от выбора локальных координат, а во-вторых, координаты касательного вектора $\left.\frac{d(\psi\widetilde{\gamma}(t))_i}{dt}\right|_{t=0}$ при переходе к другой системе координат действительно преобразуются по требуемому закону (12.1).

Во-вторых, касательный вектор как оператор дифференцирования. Пусть (U, φ) — локальная система координат с началом в точке $x \in M^n$, $\mathbf{v} \in \mathbb{R}^n$ и f — гладкая функция, определённая в некоторой окрестности точки x. Функции f можно сопоставить число $\sum_i \frac{\partial(f\varphi^{-1})}{\partial x_i}(0)v_i$, которое мы будем называть *производной функции f по направлению векторного поля v*. При переходе к другой системе координат (V, ψ) вектор $\mathbf{v}$ заменится на вектор $\mathbf{w}$ с координатами $w_i = \sum_j \frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0)v_j$, поэтому функции f в новой системе координат будет сопоставлено число

$$\sum_{i,j} \frac{\partial(f\psi^{-1})}{\partial x_i}(0)\frac{\partial(\psi\varphi^{-1})_i}{\partial x_j}(0)v_j = \sum_j \frac{\partial(f\varphi^{-1})}{\partial x_j}(0)v_j.$$

Таким образом, число, сопоставляемое функции f, не зависит от выбора системы координат.

Касательному вектору v в точке $x \in M^n$ мы сопоставили линейный оператор $v\colon C^\infty(M^n) \to \mathbb{R}$ (вместо $C^\infty(M^n)$ можно взять $C^\infty(U)$, где U — некоторая окрестность точки x; число $v(f)$ зависит только от поведения функции f в сколь угодно малой окрестности точки x). При этом выполняются следующие свойства:

1) $(\lambda v + \mu w)(f) = \lambda v(f) + \mu w(f)$;

2) $v(fg) = f(x)v(g) + g(x)v(f)$.

Второе свойство следует из того, что $\frac{\partial(fg)}{\partial x_i} = g\frac{\partial f}{\partial x_i} + f\frac{\partial g}{\partial x_i}$.

Свойства 1) и 2) вместе с линейностью оператора v можно взять за определение линейного пространства касательных векторов в точке $x \in M^n$. Но при этом нужно проверить, что не появится «лишних» операторов, т. е. если $v\colon C^\infty(\mathbb{R}^n) \to \mathbb{R}$ — линейный оператор, обладающий свойством $v(fg) = f(0)v(g) + g(0)v(f)$, то $v(f) = \sum_i \frac{\partial f}{\partial x_i}(0)v_i$ для некоторых $v_1, \ldots, v_n \in \mathbb{R}^n$. Требуемое утверждение очевидным образом следует из леммы Адамара. Действительно, $f(x) - f(0) = \sum x_i g_i(x)$, поэтому

$$v(f) = \sum 0 \cdot v(g_i) + \sum g_i(0)v(x_i) = \sum \frac{\partial f}{\partial x_i}(0)v_i,$$

где $v_i = v(x_i)$.

Касательные векторы в точке $x \in M^n$ образуют линейное пространство. Это пространство называют *касательным пространством* в точке x и обозначают T_xM^n.

Определим теперь дифференциал отображения гладких многообразий. Пусть $f\colon M^m \to N^n$ — гладкое отображение, $v \in T_xM^m$ — касательный вектор. Тогда можно определить вектор $df(v) \in T_{f(x)}N^n$. Например, если вектор v задан кривой $\gamma(t)$, то вектор $df(v)$ задаётся кривой $f(\gamma(t))$. А если вектор v задан как линейный оператор на гладких функциях, то вектор $df(v)$ задаётся как оператор $df(v)(\varphi) = v(\varphi f)$; действительно, если $\varphi \in C^\infty(U_{f(x)})$, то $\varphi f \in C^\infty(V_x)$.

Отображение $df\colon T_xM^m \to T_{f(x)}N^n$ линейно; это отображение называют *дифференциалом* отображения f в точке x.

Условие, что f — иммерсия (субмерсия) в точке x, эквивалентно тому, то дифференциал отображения f в точке x — мономорфное (эпиморфное) отображение. В такой форме иногда бывает удобнее проверять, что f — иммерсия (субмерсия).

На множестве $TM^n = \bigcup_{x\in M^n} T_xM^n$ можно ввести структуру многообразия следующим образом. Пусть (U, φ) — локальная система координат на многообразии M^n. Сопоставим касательному вектору в точке $x \in M^n$ пару $(\varphi(x), v)$, где $v = (v_1, \dots, v_n)$ — координаты этого касательного вектора в данной системе координат. В результате получим взаимно однозначное отображение

$$T\varphi\colon TU = \bigcup_{x\in U} T_xM^n \to \varphi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n}.$$

Множества TU покрывают TM^n. Потребовав, чтобы все отображения $T\varphi$ были гомеоморфизмами, мы зададим на TM^n структуру топологического пространства. Карты $(TU, T\varphi)$ задают на этом топологическом пространстве структуру многообразия. Многообразие TM^n называют *касательным расслоением* многообразия M^n.

Сопоставив касательному вектору $v \in T_xM^n$ точку x, получим проекцию $p\colon TM^n \to M^n$. При этом отображение p гладкое.

Гладкое отображение $f\colon M^m \to N^n$ индуцирует гладкое отображение $df\colon TM^m \to TN^n$ (*дифференциал* отображения f).

12.3. Метод множителей Лагранжа-3

На языке многообразий и касательных пространств к ним метод множителей Лагранжа имеет естественную интерпретацию и простое доказательство. Напомним его формулировку. Пусть функция f и функции $g_1, \dots, g_m$ бесконечно[1] дифференцируемы на открытом множестве $S \subset \mathbb{R}^{n+m}$. Нас интересует локальный экстремум $\mathbf{x}_0 \in S$ функции f при условии, что выполняются соотношения $g_1 = 0, \dots, g_m = 0$. В случае локального минимума это означает, что у точки $\mathbf{x}_0$, для которой выполняются соотношения $g_1 = 0, \dots, g_m = 0$, есть такая окрестность U, что $f(\mathbf{x}_0) \le f(\mathbf{x})$ для всех точек $\mathbf{x} \in U$, для которых выполняются эти соотношения.

Пусть $\mathbf{x}_0$ — точка локального экстремума функции f при условии $g_1 = 0, \dots, g_m = 0$, причём в этой точке матрица Якоби отображения $g = (g_1, \dots, g_m)$ имеет ранг m. Тогда

$$\operatorname{grad} f(\mathbf{x}_0) = \lambda_1 \operatorname{grad} g_1(\mathbf{x}_0) + \dots + \lambda_m \operatorname{grad} g_m(\mathbf{x}_0)$$

для некоторых чисел $\lambda_1, \dots, \lambda_m$.

Перейдём к интерпретации и доказательству. Если в точке $\mathbf{x}_0$ матрица Якоби гладкого отображения $g = (g_1, \dots, g_m)$ имеет ранг максимальный ранг m, то она имеет максимальный ранг m и в некоторой окрестности этой точки. Действительно, если некоторый минор матрицы Якоби отличен от нуля в точке, то он отличен от нуля и в некоторой окрестности этой точки. Поэтому уравнения $g_1 = 0, \dots, g_m = 0$ в некоторой окрестности точки $\mathbf{x}_0$ задают многообразие $M^n \subset \mathbb{R}^{n+m}$ размерности n.

Векторы $\operatorname{grad} g_1(\mathbf{x}_0), \dots, \operatorname{grad} g_m(\mathbf{x}_0)$ ортогональны касательному пространству к многообразию M^n в точке $\mathbf{x}_0$. Действительно, рассмотрим на многообразии M произвольную гладкую кривую $\gamma(t)$, проходящую через точку $\mathbf{x}_0$: $\gamma(0) = \mathbf{x}_0$. На многообразии M функция g_i тождественно равна нулю, поэтому $g_i(\gamma(t)) = 0$, следовательно,

$$0 = \frac{dg_i(\gamma(0))}{dt} = \sum_{i=1}^{m+n} \frac{\partial g_i}{\partial x_i}(\mathbf{x}_0) \cdot \frac{d\gamma}{dt}(0).$$

Это как раз и означает, что вектор $\operatorname{grad} g_i(\mathbf{x}_0)$ ортогонален касательному вектору, заданному кривой γ.

[1] Для удобства мы здесь рассматриваем бесконечно дифференцируемые функции вместо непрерывно дифференцируемых. Но можно рассмотреть и многообразия класса C^1. Тогда можно будет рассматривать и функции класса C^1.

Вектор $\operatorname{grad} f(\mathbf{x}_0)$ тоже ортогонален касательному пространству к многообразию M в точке $\mathbf{x}_0$. Для доказательства этого снова рассмотрим на многообразии M произвольную гладкую кривую $\gamma(t)$, проходящую через точку $\mathbf{x}_0$: $\gamma(0) = \mathbf{x}_0$. На этот раз функция $f(\gamma(t))$ не обязательно равна нулю, но она имеет локальный экстремум в точке $t = 0$, поэтому мы снова получаем

$$0 = \frac{df(\gamma(0))}{dt} = \sum_{i=1}^{m+n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \cdot \frac{d\gamma}{dt}(0).$$

Векторы $\operatorname{grad} g_1(\mathbf{x}_0)$, ..., $\operatorname{grad} g_m(\mathbf{x}_0)$ лежат в пространстве, ортогональном касательному пространству к многообразию M^n в точке $\mathbf{x}_0$. Размерность ортогонального пространства равна m. Данные векторы линейно независимы, поскольку в точке $\mathbf{x}_0$ матрица Якоби отображения $g = (g_1, \dots, g_m)$ имеет ранг m. Следовательно, эти векторы порождают всё ортогональное пространство. В частности, вектор $\operatorname{grad} f(\mathbf{x}_0)$, лежащий в ортогональном пространстве, можно представить в виде линейной комбинации векторов $\operatorname{grad} g_1(\mathbf{x}_0)$, ..., $\operatorname{grad} g_m(\mathbf{x}_0)$, что и требовалось.

12.4. Дифференциальные формы в $\mathbb{R}^n$

Гладкая дифференциальная 1-форма ω в $\mathbb{R}^n$ — это выражение вида $f_1(\mathbf{x})dx_1 + \dots + f_n(\mathbf{x})dx_n$, где dx_i — линейная функция, принимающая на векторе $\mathbf{v} = (v_1, \dots, v_n)$ значение v_i, а f_1, ..., f_n — гладкие функции. В каждой точке $\mathbf{x} \in \mathbb{R}^n$ дифференциальная 1-форма — это линейная функция от векторов в $\mathbb{R}^n$. Иногда рассматривают дифференциальные формы класса C^k, для которых функции f_i имеют класс C^k. Примером дифференциальной 1-формы класса C^k является дифференциал df функции класса C^{k+1}.

Дифференциальная k-форма — это кососимметрическая полилинейная функция от k векторов. Например, по 1-формам ω_1, ..., ω_k можно построить k-форму $\omega_1 \wedge \dots \wedge \omega_k$, принимающую на наборе векторов v_1, ..., v_k значение $\det \begin{pmatrix} \omega_1(v_1) & \dots & \omega_1(v_k) \\ \dots & \dots & \dots \\ \omega_k(v_1) & \dots & \omega_k(v_k) \end{pmatrix}$. Ясно, что если в выражении $\omega_1 \wedge \dots \wedge \omega_k$ встречаются две одинаковые 1-формы, то в результате получается нулевая k-форма (определитель матрицы с двумя одинаковыми строками равен нулю). Кроме того, если поменять местами две 1-формы, то k-форма изменит знак (при перестановке двух строк определитель матрицы меняет знак). Построим по базисным 1-формам dx_1, ..., dx_n формы вида $dx_{i_1} \wedge \dots \wedge dx_{i_k}$. Среди них много нулевых и линейно зависимых k-форм. Чтобы получить линейно независимые формы, можно взять те из них, индексы которых расположены в порядке возрастания: $i_1 < i_2 < \dots < i_k$. Произвольная дифференциальная k-форма имеет вид

$$\sum_{i_1 < i_2 < \dots < i_k} f_{i_1 \dots i_k}(\mathbf{x}) dx_{i_1} \wedge \dots \wedge dx_{i_k}.$$

Линейное пространство всех дифференциальных k-форм в $\mathbb{R}^n$ будем обозначать $\Omega^k(\mathbb{R}^n)$. Можно рассматривать также пространство $\Omega^k(U)$, где $U \subset \mathbb{R}^n$ — открытое множество.

По p-форме $\alpha = \sum a_{i_1 \dots i_p} dx_{i_1} \wedge \dots \wedge dx_{i_p}$ и q-форме $\beta = \sum b_{j_1 \dots j_q} dx_{j_1} \wedge \dots \wedge dx_{j_q}$ можно построить $(p+q)$-форму

$$\alpha \wedge \beta = \sum a_{i_1 \dots i_p} b_{j_1 \dots j_q} dx_{i_1} \wedge \dots \wedge dx_{i_p} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_q}.$$

Задача 12.3. Докажите, что для p-формы ω^p и q-формы ω^q выполняется равенство $\omega^p \wedge \omega^q = (-1)^{pq} \omega^q \wedge \omega^p$.

Дифференциальные формы являются гораздо более подходящим объектом для определения кратного интеграла, чем функции. В интегральной сумме слагаемое $f(\mathbf{x}_k)v(P_k)$ зависит от выбора системы координат: при изменении системы координат объём параллелепипеда P_k изменяется. Но если мы рассмотрим n-форму $\omega = f(\mathbf{x}_k)dx_1 \wedge \dots \wedge dx_n$, то слагаемое $f(\mathbf{x}_k)v(P_k)$ можно представить как значение формы ω на векторах — рёбрах параллелепипеда. И это выражение уже почти не зависит от системы координат: оно не меняется при заменах переменной, сохраняющих ориентацию (т.е. при заменах переменной с положительным якобианом).

Дадим теперь подробное определение интеграла n-формы ω по области в $\mathbb{R}^n$ и докажем, что этот интеграл не изменяется при заменах переменной, сохраняющих ориентацию. Для удобства вместо интеграла по ограниченной области в $\mathbb{R}^n$ будем рассматривать интеграл по $\mathbb{R}^n$ формы, обращающейся в нуль вне некоторой ограниченной области.

Сначала выясним, как изменяется запись n-формы при заменах координат (сама n-форма от координат не зависит). Для форм от двух переменных $dx = \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv$ и $dy = \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv$, поэтому

$$dx \wedge dy = \left(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right) du \wedge dv.$$

Рассмотрим теперь n-мерный случай. Для краткости будем записывать выражение $\omega = \varphi(x_1, \ldots, x_n)dx_1 \wedge \cdots \wedge dx_n$ в виде $\omega = \varphi(\mathbf{x})\,d\mathbf{x}$. Перейдя к другой системе координат $\mathbf{y}(\mathbf{x})$, получим $dy_1 \wedge \cdots \wedge dy_n = \det\left(\frac{\partial y_i}{\partial x_j}\right)dx_1 \wedge \cdots \wedge dx_n$. Кратко это будем записывать так: $d\mathbf{y} = \det\left(\frac{\partial y_i}{\partial x_j}\right)d\mathbf{x}$. В новых координатах $\omega = \psi(\mathbf{y})\,d\mathbf{y}$. Выясним, как связаны функции $\psi(\mathbf{y})$ и $\varphi(\mathbf{x})$. Ясно, что $\varphi(\mathbf{x})\,d\mathbf{x} = \omega = \psi(\mathbf{y})\,d\mathbf{y} = \psi(\mathbf{y})\det\left(\frac{\partial y_i}{\partial x_j}\right)d\mathbf{x}$, поэтому $\varphi(\mathbf{x}) = \psi(\mathbf{y}(\mathbf{x}))\det\left(\frac{\partial y_i}{\partial x_j}\right)$.

Формула замены переменных в интеграле показывает, что

$$\int_{\mathbb{R}^n} \psi(\mathbf{y})\,d\mathbf{y} = \int_{\mathbb{R}^n} \psi(\mathbf{y}(\mathbf{x}))\left|\det\left(\frac{\partial y_i}{\partial x_j}\right)\right| d\mathbf{x} = \pm \int_{\mathbb{R}^n} \varphi(\mathbf{x})\,d\mathbf{x},$$

где знак $\pm$ соответствует знаку определителя $\det\left(\frac{\partial y_i}{\partial x_j}\right)$. Это позволяет определить *интеграл* дифференциальной формы следующим образом. Предположим, что в $\mathbb{R}^n$ фиксирована ориентация. Пусть ω — n-форма в $\mathbb{R}^n$, равная нулю вне некоторой ограниченной области. Положим $\int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} \varphi(\mathbf{x})\,d\mathbf{x}$, где $\mathbf{x}$ — система координат, ориентация которой согласована с ориентацией $\mathbb{R}^n$.

Пусть ω — непрерывная 1-форма на открытом множестве $U \subset \mathbb{R}^n$, $\gamma\colon [a,b] \to U$ — отображение класса C^1. Тогда по определению

$$\int_\gamma \omega = \int_a^b \omega_{\gamma(t)}(\gamma'(t))dt.$$

Пример 12.3. *Пусть* $\omega = \frac{-y\,dx + x\,dy}{x^2+y^2}$ *и кривая* $\gamma\colon [0,2] \to \mathbb{R}^2$ *задана формулой* $\gamma(t) = (\cos \pi t, \sin \pi t)$. *Тогда* $\int_\gamma \omega = 2\pi$.

Доказательство.

$$\int_\gamma \omega = \int_0^2 \frac{-(\sin \pi t)(-\pi \sin \pi t) + (\cos \pi t)(\pi \cos \pi t)}{\cos^2 \pi t + \sin^2 \pi t} dt = 2\pi.$$

В координате t эта форма имеет вид dt. □

Определим теперь *дифференциал* дифференциальной формы. Функцию f в $\mathbb{R}^n$ удобно считать 0-формой. Дифференциал 0-формы — это обычный дифференциал функции: в точке $\mathbf{x}$ значение этой формы на векторе $\mathbf{v}$ равно $df_{\mathbf{x}}(\mathbf{v})$. Дифференциал 1-формы $\omega = f_1 dx^1 + \cdots + f_n dx^n$ равен $d\omega = df_1 \wedge dx^1 + \cdots + df_n \wedge dx^n$. Дифференциал k-формы $\omega = \sum_{i_1 < i_2 < \cdots < i_k} f_{i_1 \ldots i_k}(\mathbf{x}) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ равен

$$d\omega = \sum_{i_1 < i_2 < \cdots < i_k} (df_{i_1 \ldots i_k})_{\mathbf{x}} \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}.$$

Историческое замечание. Близко к понятию дифференциальной формы подошёл Анри Пуанкаре (1854-1912), разработав в 1895 году интегрирование на многообразиях. Определение дифференциальной формы на многообразии дал Эли Картан (1869-1951) в 1899 году, основываясь на алгебре Грассмана. Картан определил также дифференциал дифференциальной формы.

Задача 12.4. а) Пусть $\omega = f(x,y)dx + g(x,y)dy$. Докажите, что $d\omega = (-f_y + g_x)dx \wedge dy$.
б) Вычислите $d(x\,dy - y\,dx)$.
в) Вычислите $d\omega$, где форма $\omega = \frac{x\,dy - y\,dx}{x^2+y^2}$ определена при $(x, y) \neq (0, 0)$.

Напомним, что $\int_\gamma df = f(\gamma(b)) - f(\gamma(a))$ (см. с. 206). Функция f — это 0-форма. Для 1-формы на плоскости справедлива аналогичная формула

$$\int_D d\omega = \int_{\partial D} \omega, \tag{12.2}$$

где D — область, ограниченная кривой ∂D, причём эта кривая ориентирована против часовой стрелки. Нам уже встречалось много разных видов интегралов, иногда они зависели от ориентации области интегрирования, иногда не зависели. Поясним ситуацию с ориентациями в формуле (12.2). Интеграл дифференциальной n-формы по n-мерной области всегда зависит от ориентации области. При интегрировании n-формы по n-мерной области в $\mathbb{R}^n$ подразумевается, что ориентация области согласована с ориентацией $\mathbb{R}^n$. Поэтому об ориентации области D ничего не говорится, но подразумевается, что она согласована с ориентацией плоскости. А ориентацию кривой ∂D необходимо указать, без этого нельзя вычислить интеграл $\int_{\partial D}\omega$.

Вскоре мы докажем формулу (12.2) в гораздо более общем виде (см. с. 229). Сейчас мы ограничимся следующим наиболее простым случаем.

Пример 12.4. *Пусть I^2 — квадрат, заданный неравенствами $0 \le x \le 1$ и $0 \le y \le 1$, ∂I^2 — граница квадрата, ориентированная в направлении против часовой стрелки. Тогда для любой гладкой 1-формы $\omega = Pdx + Qdy$ имеет место равенство $\int_{I^2} d\omega = \int_{\partial I^2}\omega$.*

Доказательство. Ясно, что $d\omega = \left(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\right) dx \wedge dy$, поэтому

$$\begin{aligned}\int_{I^2} d\omega &= \int_0^1 \left(\int_0^1 \frac{\partial Q}{\partial x}dx\right) dy - \int_0^1 \left(\int_0^1 \frac{\partial P}{\partial y}dy\right) dx = \\ &= \int_0^1 Q(1,y)dy - \int_0^1 Q(0,y)dy - \\ &- \int_0^1 P(x,1)dx + \int_0^1 P(x,0)dx.\end{aligned}$$

Последнее выражение как раз равно $\int_{\partial I^2}\omega$, где граница квадрата обходится в направлении против часовой стрелки. □

Задача 12.5. Докажите, что $d \circ d = 0$.

Задача 12.6. Докажите, что если ω_1 — k-форма, то $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2$.

Задача 12.7. Пусть $\omega = \sum_{i=1}^n (-1)^{i-1} x_i dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n$.
а) Вычислите $d\omega$.
б) Вычислите $d(\|\mathbf{x}\|^{-n}\omega)$, где $\|\mathbf{x}\| = \sqrt{x_1^2 + \dots + x_n^2} \ne 0$.

Гладкое отображение $f\colon \mathbb{R}^m \to \mathbb{R}^n$ индуцирует линейное отображение $f^*\colon \Omega^k(\mathbb{R}^n) \to \Omega^k(\mathbb{R}^m)$ следующим образом. При $k = 0$, т.е. для функции α в $\mathbb{R}^n$, мы полагаем $f^*\alpha = \alpha \circ f$. А при $k > 0$ полагаем $(f^*\omega)(\xi_1, \dots, \xi_k) = \omega(f_*\xi_1, \dots, f_*\xi_k)$, где $f_* = df$ — дифференциал отображения f. В координатах это отображение выглядит следующим образом. Пусть $x_1, \dots, x_n$ и $y_1, \dots, y_m$ — координаты в $\mathbb{R}^n$ и $\mathbb{R}^m$. Тогда

$$f^*\Big(\sum a_{i_1\dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}\Big) = \sum a_{i_1\dots i_k} \frac{\partial f_{i_1}}{\partial y_{j_1}} \dots \frac{\partial f_{i_k}}{\partial y_{j_k}} dy_{j_1} \wedge \dots \wedge dy_{j_k}.$$

Например, $f^*(dx_i) = \sum_{j=1}^m \frac{\partial f_i}{\partial y_j} dy_j = df_i$ — дифференциал i-й координаты отображения f.

Пример 12.5. *Пусть r — фиксированное число и $f(\varphi) = (r\cos\varphi, r\sin\varphi)$. Тогда $f^*(x\,dy - y\,dx) = r^2 d\varphi$ и $f^*(\frac{x\,dy - y\,dx}{x^2+y^2}) = d\varphi$.*

Доказательство. По определению

$$\begin{aligned} f^*(x\,dy - y\,dx) &= x\frac{\partial y}{\partial\varphi}d\varphi - y\frac{\partial x}{\partial\varphi}d\varphi = \\ &= r\cos\varphi\frac{\partial(r\sin\varphi)}{\partial\varphi}d\varphi - r\sin\varphi\frac{\partial(r\cos\varphi)}{\partial\varphi}d\varphi = \\ &= r^2\cos^2\varphi\,d\varphi + r^2\sin^2\varphi\,d\varphi = r^2 d\varphi.\end{aligned}$$

Равенство $f^*(\frac{x\,dy - y\,dx}{x^2+y^2}) = d\varphi$ следует из того, что 1-форма $\frac{x\,dy - y\,dx}{x^2+y^2}$ рассматривается только на окружности $x^2 + y^2 = r^2$, поэтому $f^*(\frac{x\,dy - y\,dx}{x^2+y^2}) = f^*(\frac{x\,dy - y\,dx}{r^2}) = d\varphi$. □

Пример 12.6. *Пусть $f(r,\varphi)=(r\cos\varphi, r\sin\varphi)=(x,y)$. Тогда $f^*(dx\wedge dy)=r\,dr\wedge d\varphi$*

Доказательство. По определению

$$\begin{aligned} f^*(dx\wedge dy) &= \frac{\partial(r\cos\varphi)}{\partial r}\frac{\partial(r\sin\varphi)}{\partial\varphi}dr\wedge d\varphi+ \\ &+ \frac{\partial(r\cos\varphi)}{\partial\varphi}\frac{\partial(r\sin\varphi)}{\partial r}d\varphi\wedge dr= \\ &= r\cos^2\varphi\,dr\wedge d\varphi - r\sin^2\varphi\,d\varphi\wedge dr = r\,dr\wedge d\varphi. \end{aligned}$$

□

Пример 12.7. *Интеграл формы $\omega=\frac{x\,dy-y\,dx}{x^2+y^2}$ по окружности, содержащей внутри начало координат, равен $\pm 2\pi$, а по окружности, не содержащей начала координат, равен 0.*

Доказательство. В полярных координатах форма ω равна $d\varphi$. В полярных координатах окружность, содержащая внутри начало координат, представляет собой график функции от φ на полуинтервале $[0,2\pi)$. Интеграл формы $d\varphi$ по такой кривой равен $\pm\int_0^{2\pi}d\varphi=\pm 2\pi$. Окружность, не содержащая начала координат, касается сторон угла с вершиной в начале координат, поэтому в полярных координатах она представляет собой два графика функций на отрезке $[\alpha,\beta]$. Эти графики проходятся в противоположных направлениях, поэтому интеграл формы $d\varphi$ по такой кривой равен $\int_\alpha^\beta d\varphi+\int_\beta^\alpha d\varphi=0$. □

Теорема 12.3. *Пусть $f\colon\mathbb{R}^m\to\mathbb{R}^n$ — гладкое отображение. Тогда для дифференциальной k-формы $\omega\in\Omega^k(\mathbb{R}^n)$ выполняется равенство $d(f^*\omega)=f^*(d\omega)$.*

Доказательство. Применим индукцию по k. При $k=0$ требуется доказать, что $d(f^*\alpha)=f^*(d\alpha)$ для гладкой функции α в $\mathbb{R}^n$. Ясно, что

$$d(f^*\alpha)=d(\alpha\circ f)=\sum_{i=1}^m\frac{\partial(\alpha\circ f)}{\partial y_i}dy_i=\sum_{i=1}^m\left(\sum_{j=1}^n\frac{\partial\alpha}{\partial x_j}\circ f\right)\frac{\partial f_j}{\partial y_i}dy_i$$

и

$$\begin{aligned} f^*(d\alpha) &= f^*\left(\sum_{j=1}^n\frac{\partial\alpha}{\partial x_j}dx_j\right)=\sum_{j=1}^n\left(\frac{\partial\alpha}{\partial x_j}\circ f\right)f^*(dx_j)= \\ &= \sum_{j=1}^n\left(\frac{\partial\alpha}{\partial x_j}\circ f\right)\left(\sum_{i=1}^m\frac{\partial f_j}{\partial y_i}dy_i\right). \end{aligned}$$

Полученные выражения равны.

Предположим теперь, что равенство $d(f^*\omega)=f^*(d\omega)$ выполняется для любой $(k-1)$-формы. Представим k-форму $\omega=f dx_{i_1}\wedge\dots\wedge dx_{i_k}$ в виде $\omega=dx_{i_1}\wedge\omega'$, где $\omega'=f dx_{i_2}\wedge\dots\wedge dx_{i_k}$. При вычислениях $d(f^*\omega)$ и $f^*(d\omega)$ мы будем пользоваться результатами задач 12.5 и 12.6. Ясно, что

$$d\omega=ddx_{i_1}\wedge\omega'-dx_{i_1}\wedge d\omega'=-dx_{i_1}\wedge d\omega',$$

поэтому

$$f^*(d\omega)=f^*(-dx_{i_1}\wedge d\omega')=-f^*(dx_{i_1})\wedge f^*(d\omega')=-df_{i_1}\wedge f^*(d\omega').$$

Вычислим теперь $d(f^*\omega)$:

$$d(f^*\omega)=d(f^*(dx_{i_1})\wedge f^*\omega')=ddf_{i_1}\wedge f^*\omega'-df_{i_1}\wedge f^*(d\omega')=f^*(d\omega).$$

□

Дифференциальную форму ω называют: *замкнутой*, если $d\omega=0$; *точной*, если $\omega=d\omega'$ для некоторой формы ω'. Любая точная форма является замкнутой, поскольку $d\omega=dd\omega'=0$.

Теорема 12.4 (лемма Пуанкаре). *При $k > 0$ любая замкнутая дифференциальная k-форма, определённая на всём пространстве $\mathbb{R}^n$, точна.*

Доказательство. Докажем сначала вспомогательное утверждение: если $i_0, i_1\colon \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ — вложения, заданные формулами $i_0(\mathbf{x}) = (\mathbf{x}, 0)$ и $i_1(\mathbf{x}) = (\mathbf{x}, 1)$, то для всех $k \geq 0$ существует такое линейное отображение $D\colon \Omega^k(\mathbb{R}^{n+1}) \to \Omega^{k-1}(\mathbb{R}^n)$, что $dD + Dd = i_1^* - i_0^*$.

Это отображение определим следующим образом. Для 0-формы (функции) φ положим $D\varphi = 0$. Пусть $x_1, \dots, x_n$ — первые n координат в $\mathbb{R}^{n+1}$, t — последняя координата. Тогда при $k \geq 1$ базисные формы из $\Omega^k(\mathbb{R}^{n+1})$ разбиваются на два типа. Формы, в которые не входит dt, будем обозначать $\alpha = a\,dx_{i_1} \wedge \dots \wedge dx_{i_k}$, а формы, в которые входит dt, будем обозначать $\beta = b\,dt \wedge dx_{j_1} \wedge \dots \wedge dx_{j_{k-1}}$. Положим $D\alpha = 0$ и $D\beta = \left(\int_0^1 b\,dt\right) dx_{j_1} \wedge \dots \wedge dx_{j_{k-1}}$. Здесь $b = b(x_1, \dots, x_n, t)$, и после интегрирования по t при фиксированных $x_1, \dots, x_n$ получается функция от $x_1, \dots, x_n$.

Любая форма ω является суммой форм видов α и β, поэтому требуемое утверждение достаточно доказать только для них.

Займёмся теперь проверкой равенства $dD + Dd = i_1^* - i_0^*$ во всех трёх случаях: для функции φ, для формы α и для формы β.

1. $dD\varphi = 0$ и $Dd\varphi = \int_0^1 \frac{\partial\varphi}{\partial t} dt = \varphi(1) - \varphi(0) = (i_1^* - i_0^*)\varphi$.

2. $dD\alpha = 0$ и $Dd\alpha = \left(\int_0^1 \frac{\partial\alpha}{\partial t} dt\right) dx_{i_1} \wedge \dots \wedge dx_{i_k} = (i_1^* - i_0^*)\alpha$.

3. Прежде всего заметим, что $i_0^*\beta = i_1^*\beta = 0$, поскольку ограничение формы dt на $i_0(\mathbb{R}^n)$ и на $i_1(\mathbb{R}^n)$ тождественно равно нулю (все касательные векторы имеют нулевую t-координату). Далее,

$$
\begin{aligned}
dD\beta &= \sum_{i=1}^n \left(\int_0^1 \frac{\partial b}{\partial x_i} dt\right) dx_i \wedge dx_{j_1} \wedge \dots \wedge dx_{j_{k-1}},\\
Dd\beta &= D\left(\sum_{i=1}^n \frac{\partial b}{\partial x_i} dx_i \wedge dt \wedge dx_{j_1} \wedge \dots \wedge dx_{j_{k-1}}\right) =\\
&= -\sum_{i=1}^n \left(\int_0^1 \frac{\partial b}{\partial x_i} dt\right) dx_i \wedge dx_{j_1} \wedge \dots \wedge dx_{j_{k-1}}.
\end{aligned}
$$

Поэтому $dD\beta + Dd\beta = 0$.

Перейдём теперь непосредственно к доказательству леммы Пуанкаре. Рассмотрим отображение $H\colon \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, заданное формулой $(x, t) \mapsto tx$. Тогда $Hi_1 = \mathrm{id}$ и $Hi_0 = *$ (отображение в 0). Пусть $\omega \in \Omega^k(\mathbb{R}^n)$ и $d\omega = 0$. Тогда $\omega = (i_1^* - i_0^*)H^*\omega = dD(H^*\omega) + Dd(H^*\omega)$. Но из теоремы 12.3 следует, что $d(H^*\omega) = H^*(d\omega) = 0$, поэтому $Dd(H^*\omega) = 0$ и $\omega = dD(H^*\omega)$, т.е. форма ω точна. □

Историческое замечание. Утверждение, близкое к лемме Пуанкаре, но не на языке дифференциальных форм (которого тогда ещё не было), Анри Пуанкаре (1854-1912) доказал в 1895 году, разрабатывая интегрирование на многообразиях. В 1917 году Эдуард Гурса (1858-1936) сформулировал лемму Пуанкаре на языке дифференциальных форм. Но он не заметил, что нужны ограничения на область определения: форма должна быть определена на всём пространстве $\mathbb{R}^n$. Пример замкнутой, но не точной формы, определённой на $\mathbb{R}^n$ без начала координат, привёл Эли Картан (1869-1951) в 1922 году.

Пример 12.8. *Форма $\omega = \frac{x\,dy - y\,dx}{x^2 + y^2}$, определённая на $\mathbb{R}^2 \setminus \{0\}$, замкнутая, но не точная.*

Доказательство. Замкнутость формы ω доказана в решении задачи 12.4. Покажем, что эта форма не может быть точной. Напомним, что если $f(\varphi) = (\cos\varphi, \sin\varphi)$, то $f^*(\omega) = d\varphi$ (пример 12.5). Поэтому если $\omega = dF$ для некоторой функции F, определённой на $\mathbb{R}^2 \setminus \{0\}$, то $d\varphi = f^*(dF) = d(f^*F) = d(F \circ f)$. Функция f имеет период 2π, поэтому функция $g = F \circ f$ тоже имеет период 2π. Но равенство $d\varphi = dg = \frac{dg}{d\varphi} d\varphi$ означает, что производная функции g тождественно равна 1. Это противоречит тому, что у периодической функции должны быть точки максимума и минимума, в которых производная обращается в нуль. □

12.5. Разбиение единицы

Разбиение единицы — важный инструмент, позволяющий склеивать объекты, построенные локально, в глобальные объекты. Мы начнём с разбиений единицы в $\mathbb{R}^n$, а затем перейдём к многообразиям. Пусть $\{U_\alpha, \alpha \in \mathcal{A}\}$ — открытое покрытие некоторого множества $A \subset \mathbb{R}^n$. *Разбиением единицы*, подчинённым этому покрытию, называют множество непрерывных функций $\{\varphi_\beta, \beta \in \mathcal{B}\}$, определённых на $\mathbb{R}^n$ и обладающих следующими свойствами:

1) $0 \le \varphi_\beta(x) \le 1$ для любой точки $x \in A$;

2) у любой точки $x \in A$ есть открытая окрестность, на которой лишь конечное множество функций φ_β отлично от нуля;

3) $\sum \varphi(x) = 1$ для любой точки $x \in A$ (согласно предыдущему свойству эта сумма конечна в некоторой окрестности точки x);

4) для любой функции φ_β найдётся такое множество U_α из данного покрытия, что $\operatorname{supp} \varphi_\beta \subset U_\alpha$.

Подчинённость покрытию — это свойство 4). Свойство 2) — это локальная конечность разбиения единицы.

Напомним, что при доказательстве теоремы 11.9 о замене переменных в кратном интеграле мы уже строили и использовали непрерывное разбиение единицы, подчинённое покрытию.

Разбиение единицы называют *гладким*, если все функции φ_β бесконечно дифференцируемые.

Теорема 12.5. *а) Для любого открытого покрытия $\{U_\alpha, \alpha \in \mathcal{A}\}$ множества $A \subset \mathbb{R}^n$ существует подчинённое ему гладкое разбиение единицы $\{\varphi_\beta, \beta \in \mathcal{B}\}$.*

б) Если множество индексов $\mathcal{A}$ не более чем счётно, то можно считать, что $\mathcal{B} = \mathcal{A}$ и $\operatorname{supp} \varphi_\alpha \subset U_\alpha$ для любого $\alpha \in \mathcal{A}$.

Доказательство. а) Проверим сначала, что функция, которая равна $e^{-1/x}$ при $x > 0$ и 0 при $x \le 0$, бесконечно дифференцируема. Это нужно проверить только в точке 0. Ясно, что $(e^f)' = f'e^f$ и $(e^f)'' = (f'' + (f')^2)e^f$. Продолжая такие вычисления, получаем, что n-я производная функции e^f имеет вид $P_n e^f$, где P_n — некоторый многочлен от $f', f'', \dots, f^{(n)}$. Следовательно, n-я производная функции $e^{-1/x}$ стремится к нулю при $x \to +0$. Действительно, положим $y = 1/x$. Тогда $y \to \infty$ и $y^m/e^y \to 0$ для любого натурального m.

Функция

$$\gamma(t) = \begin{cases} e^{-1/(1-t)} & \text{при } t < 1; \\ 0 & \text{при } t \ge 1 \end{cases}$$

тоже бесконечно дифференцируемая.

Введём следующее обозначение: $D^n_r = \{\mathbf{x} \in \mathbb{R}^n \colon \|\mathbf{x}\| < r\}$. Функция в $\mathbb{R}^n$, заданная формулой $\widetilde{\gamma}(\mathbf{x}) = \gamma(\|\mathbf{x}\|^2/4)$, бесконечно дифференцируемая; она положительна во всех точках открытого шара D^n_2 и равна нулю вне его.

Для $k = 1, 2, \dots$ рассмотрим множества $X_k = D^n_k$. Построим открытые шары $V_{\beta,1} \subset V_{\beta,2} \subset V_{\beta,3}$ следующим образом. Для каждой точки $z \in \bar{X}_k \setminus X_{k-1}$ выберем открытый шар $V_{z,3}$ с центром z так, что $V_{z,3} \subset U_\alpha$ для некоторого α, $V_{z,3} \subset X_{k+1}$ и $V_{z,3} \cap X_{k-2} = \varnothing$. С каждым шаром $V_{z,3}$ свяжем аффинное отображение $\psi_z \colon V_{z,3} \to D^n_3$ на шар радиуса 3 с центром в начале координат. Пусть $V_{z,1}$ и $V_{z,2}$ — открытые шары с тем же центром z, радиусы которых составляют 1/3 и 2/3 радиуса шара $V_{z,3}$. Открытые шары $V_{z,1}$ покрывают компактное множество $\bar{X}_k \setminus X_{k-1}$, поэтому существует конечный набор открытых шаров $V_{z,1}$, покрывающий $\bar{X}_k \setminus X_{k-1}$. Рассмотрим такие шары для всех k и обозначим их $V_{\beta,1}$; рассмотрим также соответствующие им открытые шары $V_{\beta,2}$ и $V_{\beta,3}$. Отметим, что множество индексов $\{\beta\}$ не более чем счётно; кроме того, открытые шары $V_{\beta,1}$ покрывают множество A и при этом покрытие $\{V_{\beta,3}\}$ локально конечно (это следует из того, что $V_{z,3} \subset X_{k+1}$ и $V_{z,3} \cap X_{k-2} = \varnothing$) и вписано в покрытие $\{U_\alpha\}$.

Каждому индексу β соответствует некоторый шар $V_{z,3}$. Соответствующее ему отображение $\psi_z \colon V_{z,3} \to D^n_3$ обозначим ψ_β. Положим

$$g_\beta(\mathbf{x}) = \begin{cases} \widetilde{\gamma}(\psi_\beta(\mathbf{x})) & \text{для } \mathbf{x} \in V_{\beta,3}; \\ 0 & \text{для } \mathbf{x} \notin V_{\beta,3} \end{cases}$$

и рассмотрим функцию $h(\mathbf{x}) = \sum_\beta g_\beta(\mathbf{x})$. Функция h бесконечно дифференцирумая, потому что покрытие $\{V_{\beta,3}\}$ локально конечное. Множества $V_{\beta,1}$ покрывают всё множество A и $g_\beta(\mathbf{x}) > 0$, если

$\mathbf{x} \in V_{\beta,1}$. Поэтому $h(\mathbf{x}) > 0$ для любой точки $\mathbf{x} \in A$. Функции $\varphi_\beta = g_\beta/h$ образуют требуемое разбиение единицы, поскольку $\operatorname{supp}\varphi_\beta \subset V_{\beta,3} \subset U_\alpha$.

б) Рассмотрим открытое покрытие $U_1, U_2, \dots$ Мы построили разбиение единицы $\varphi_1, \varphi_2, \dots$ так, что $\operatorname{supp}\varphi_i \subset U_j$ для некоторого $j = j(i)$. Определим $\tilde{\varphi}_i$ как сумму тех φ_k, для которых $\operatorname{supp}\varphi_k \subset U_i$ и $\operatorname{supp}\varphi_k \not\subset U_j$ при $j < i$. Тогда каждая функция φ_k входит в качестве слагаемого ровно в одну функцию $\tilde{\varphi}_i$ и $\operatorname{supp}\tilde{\varphi}_i \subset U_i$. □

Перейдём теперь к многообразиям. Пусть $\{U_\alpha, \alpha \in \mathcal{A}\}$ — открытое покрытие многообразия M^n. Разбиение единицы $\{\varphi_\beta, \beta \in \mathcal{B}\}$, подчинённое этому покрытию, называют *гладким*, если все функции φ_β гладкие.

Теорема 12.6. *а) Для любого открытого покрытия $\{U_\alpha, \alpha \in \mathcal{A}\}$ многообразия M^n существует подчинённое ему гладкое разбиение единицы $\{\varphi_\beta, \beta \in \mathcal{B}\}$.*

б) Если множество индексов $\mathcal{A}$ не более чем счётно, то можно считать, что $\mathcal{B} = \mathcal{A}$ и $\operatorname{supp}\varphi_\alpha \subset U_\alpha$ для любого $\alpha \in \mathcal{A}$.

Доказательство. Мы уже доказали теорему 12.5 о разбиение единицы для пространства $\mathbb{R}^n$. Для многообразий нужны лишь некоторые дополнения. Прежде всего, нужно построить открытые множества $X_k \subset M^n$ $(k = 1, 2, \dots)$ так, что множества $\bar{X}_k$ компактны, $\bar{X}_k \subset X_{k+1}$ и $M^n = \bigcup_{k=1}^\infty X_k$. Для этого рассмотрим произвольную счётную базу пространства M^n и выберем в ней те открытые множества, замыкания которых компактны. Выбранные множества обозначим $W_1, W_2, \dots$ Эти множества покрывают всё многообразие M^n. Действительно, у любой точки $x \in M^n$ есть окрестность $U(x)$, замыкание которой компактно. Множество $U(x)$ можно представить в виде объединения множеств базы; ясно, что замыкания всех этих множеств компактны. Поэтому $x \in W_i$ для некоторого i.

Положим $X_1 = W_1$. Компактное множество $\bar{X}_1$ покрыто открытыми множествами $\{W_i\}$, поэтому $\bar{X}_1 \subset W_{i_1} \cup \dots \cup W_{i_p}$, где $i_1 < i_2 < \dots < i_p$. Положим $X_2 = W_1 \cup W_2 \cup W_{i_1} \cup \dots \cup W_{i_p}$. Множества $X_3, X_4, \dots$ строятся аналогично.

При построении открытых множеств $V_{\beta,1} \subset V_{\beta,2} \subset V_{\beta,3}$ нужно поступить следующим образом. Пусть $D^n_r = \{x \in \mathbb{R}^n : \|x\| < r\}$. Для каждой точки $z \in \bar{X}_k \setminus X_{k-1}$ выберем открытое множество $V_{z,3}$ так, что $V_{z,3} \subset U_\alpha$ для некоторого α, $V_{z,3} \subset X_{k+1}$ и $V_{z,3} \cap X_{k-2} = \varnothing$; кроме того, существует карта $\psi_z : V_{z,3} \to D^n_3$. Тогда $V_{z,1} = \psi_z^{-1}(D^n_1)$ и $V_{z,2} = \psi_z^{-1}(D^n_2)$.

Оставшаяся часть доказательства никаких изменений не требует. □

С помощью разбиения единицы можно, например, построить риманову метрику на многообразии. *Риманова метрика* на многообразии M^n — это гладкое задание в касательном пространстве T_xM^n скалярного произведения (u, v). Гладкость означает, что функция $f : TM^n \to \mathbb{R}$, заданная формулой $f(v) = (v, v)$, является гладкой. Эквивалентное определение гладкости таково: для любых гладких векторных полей X и Y на M^n функция (X, Y) является гладкой функцией на M^n.

Теорема 12.7. *На любом многообразии M^n существует риманова метрика.*

Доказательство. Покроем M^n счётным набором карт $\varphi_i : U_i \to \mathbb{R}^n$ и построим гладкое разбиение единицы $\{f_i\}$, для которого $\operatorname{supp} f_i \subset U_i$.

Для $x \in U_i$ определим скалярное произведение $(\cdot, \cdot)_i$ в T_xM^n следующим образом. Пусть векторы $v, w \in T_xM^n$ имеют в локальной системе координат (U_i, φ_i) координаты $(v_1, \dots, v_n)$ и $(w_1, \dots, w_n)$. Тогда положим $(v, w)_i = v_1w_1 + \dots + v_nw_n$.

Пусть теперь x — произвольная точка M^n и $v, w \in T_xM^n$. Положим

$$(v, w) = \sum_{i=1}^\infty f_i(x)(v, w)_i.$$

Эта сумма имеет следующий смысл: если значение $(v, w)_i$ не определено, то $x \notin U_i$, а значит, $f_i(x) = 0$; в таком случае мы полагаем $f_i(x)(v, w)_i = 0$.

При фиксированном x получается выражение вида $\lambda_1A_1 + \dots + \lambda_kA_k$, где $\lambda_i > 0$, $\sum \lambda_i = 1$ и A_i — положительно определённая симметрическая билинейная форма. Сумма форм такого вида тоже положительно определена. □

12.6. Дифференциальные формы на многообразиях

Кокасательным пространством в точке $x \in M^n$ называют пространство линейных функций на пространстве T_xM^n; кокасательное пространство обозначают $T_x^*M^n$. На множестве $T^*M^n = \bigcup_{x\in M^n} T_x^*M^n$ структура многообразия задаётся аналогично тому, как это делается для TM^n. Действительно, пусть (U,φ) — локальная система координат с началом в точке x, $v \in T_xM^n$ и $l \in T_x^*M^n$. В этой локальной системе координат вектор v имеет координаты $(v_1,\dots,v_n)$ и при этом $l(v) = l_1v_1 + \dots + l_nv_n$, где числа $l_1,\dots,l_n$ одни и те же для всех векторов. Будем считать, что $(l_1,\dots,l_n)$ — координаты ковектора l в данной системе координат. Дальше действуем точно так же, как и для TM^n.

Гладкое отображение $f\colon M^m \to N^n$ индуцирует отображение касательных расслоений $df\colon TM^m \to TN^n$, которое переносит касательные векторы в том же направлении, в котором действует отображение f. Для кокасательных расслоений индуцированное отображение δf действует в противоположном направлении, т.е. $\delta f\colon T^*N^n \to T^*M^m$. Действительно, зададим отображение δf формулой $\delta f(l)(v) = l(df(v))$. Эта формула показывает, что если $l \in T^*_{f(x)}N^n \subset T^*N^n$, то $\delta f(l) \in T_x^*M^m \subset T^*M^m$.

Пусть $\Lambda_x^kM^n$ — k-я внешняя степень пространства $T_x^*M^n$. На множестве $\Lambda^kM^n = \bigcup_{x\in M^n}\Lambda_x^kM^n$ естественным образом вводится структура многообразия. *Дифференциальной k-формой* на многообразии M^n называют гладкое сечение канонической проекции $p\colon \Lambda^kM^n \to M^n$, т.е. такое гладкое отображение $s\colon M^n \to \Lambda^kM^n$, что $ps = \mathrm{id}_{M^n}$.

В локальной системе координат форма $\omega \in \Lambda^kM^n$ имеет вид

$$\omega = \sum_{i_1<\dots<i_k} a_{i_1\dots i_k}dx_{i_1}\wedge\dots\wedge dx_{i_k},$$

где $dx_i(v) = v_i$ (i-я координата вектора v в этой системе координат).

Дифференциальная k-форма ω на многообразии M^n — это полилинейная кососимметрическая функция от k векторных полей $\xi_1,\dots,\xi_k$ на M^n. Линейное пространство k-форм на многообразии M^n мы будем обозначать $\Omega^k(M^n)$. Для двух дифференциальных форм $\omega_1 \in \Omega^p(M^n)$ и $\omega_2 \in \Omega^q(M^n)$ определено их внешнее произведение $\omega_1\wedge\omega_2 \in \Omega^{p+q}(M^n)$; при этом $\omega_1\wedge\omega_2 = (-1)^{pq}\omega_2\wedge\omega_1$.

Гладкое отображение $f\colon M^m \to N^n$ индуцирует линейное отображение $f^*\colon \Omega^k(N^n) \to \Omega^k(M^m)$, а именно $(f^*\omega)(\xi_1,\dots,\xi_k) = \omega(f_*\xi_1,\dots,f_*\xi_k)$, где $f_* = df$ — дифференциал отображения f.

Многообразие M^n называют *ориентируемым*, если существует набор карт $\{U_\alpha,\varphi_\alpha\colon \alpha\in A\}$, покрывающий M^n и обладающий тем свойством, что якобиан отображения $\varphi_\alpha\varphi_\beta^{-1}$ положителен для любых $\alpha,\beta\in A$. Такой набор карт будем называть *ориентирующим атласом*.

Теорема 12.8. *Многообразие M^n ориентируемо тогда и только тогда, когда на нём есть n-форма Ω, не обращающаяся в нуль ни в какой точке.*

Доказательство. Предположим сначала, что многообразие M^n ориентируемо. Пусть $\{U_\alpha,\varphi_\alpha\}$ — ориентирующий атлас на многообразии M^n. Пространство n-форм на n-мерном многообразии одномерно; базисная форма ω в $\mathbb{R}^n$ имеет вид $\omega = dx_1\wedge\dots\wedge dx_n$. Рассмотрим на U_α форму $\omega_\alpha = (\delta\varphi_\alpha)\omega$.

Если $f\colon \mathbb{R}^n \to \mathbb{R}^n$ — гладкое отображение, то $(\delta f)\omega = J(x)\omega$, где $J(x)$ — якобиан отображения f. Поэтому $\delta(\varphi_\beta\varphi_\alpha^{-1})\omega = \lambda\omega$, где λ — положительная функция на $\varphi_\alpha(U_\alpha\cap U_\beta)$. Следовательно, $\delta(\varphi_\alpha^{-1})\delta(\varphi_\beta)\omega = \lambda\omega$, а значит, $\omega_\beta = \delta(\varphi_\beta)\omega = \delta(\varphi_\alpha)(\lambda\omega) = \lambda_\alpha\omega_\alpha$, где $\lambda_\alpha(x) = \lambda(\varphi_\alpha(x))$ — положительная функция на $U_\alpha\cap U_\beta$.

Можно считать, что покрытие $\{U_\alpha\}$ не более чем счётно. Пусть f_α — такое разбиение единицы, подчинённое покрытию $\{U_\alpha\}$, что $\operatorname{supp} f_\alpha \subset U_\alpha$ для всех α. Положим $\Omega = \sum f_\alpha\omega_\alpha$. Ясно, что форма Ω нигде не обращается в нуль.

Предположим теперь, что n-форма Ω на многообразии M^n нигде не обращается в нуль. Пусть $\{U_\alpha,\varphi_\alpha\}$ — произвольный атлас, $\omega = dx_1\wedge\dots\wedge dx_n$ — базисная форма в $\mathbb{R}^n$. Форма $\delta(\varphi_\alpha)\omega$, определённая на U_α, пропорциональна Ω, поэтому $\delta(\varphi_\alpha)\omega = \mu_\alpha\Omega$, где μ_α — некоторая функция на U_α. Функция μ_α не обращается в нуль, поэтому $\mu_\alpha > 0$ или $\mu_\alpha < 0$ на всём множестве U_α (мы предполагаем, что множество U_α связно). Если $\mu_\alpha < 0$, то заменим отображение φ_α на отображение ψ_α, которое является композицией отображения φ_α и отображения $\mathbb{R}^n \to \mathbb{R}^n$, заданного формулой

$(x_1, x_2, \dots) \mapsto (x_2, x_1, \dots)$. Ясно, что

$$\delta(\psi_\alpha)\omega = \delta(\psi_\alpha) dx_1 \wedge dx_2 \wedge \dots \wedge dx_n = \\ = \delta(\varphi_\alpha) dx_2 \wedge dx_1 \wedge \dots \wedge dx_n = -\mu_\alpha \Omega = \nu_\alpha \Omega,$$

где $\nu_\alpha = -\mu_\alpha > 0$.

После таких замен получаем ориентирующий атлас, так как

$$\delta(\varphi_\beta \varphi_\alpha^{-1})\omega = \delta(\varphi_\alpha^{-1})\delta(\varphi_\beta)\omega = \delta(\varphi_\alpha^{-1})(\mu_\beta \Omega) = \mu_\alpha^{-1}\mu_\beta\omega.$$ □

Для ориентированного n-мерного многообразия, вложенного в $\mathbb{R}^m$, в качестве n-формы, нигде не обращающейся в нуль, можно взять *форму объёма*. Эта форма принимает на наборе касательных векторов значение, равное ориентированному объёму n-мерного параллелепипеда, натянутого на эти векторы. По абсолютной величине ориентированный объём равен обычному объёму, а его знак равен знаку ориентации набора векторов (относительно ориентации многообразия).

На кривой $\gamma(t) = (x_1(t), \dots, x_n(t))$ форма объёма (в данном случае — форма длины) имеет вид $\frac{v_1}{\|\mathbf{v}\|}dx_1 + \dots + \frac{v_n}{\|\mathbf{v}\|}dx_n$, где $\mathbf{v} = \frac{d\gamma}{dt}$. Действительно, на касательном к кривой векторе $\mathbf{v}$ эта форма принимает значение

$$\frac{v_1^2 + \dots + v_n^2}{\|\mathbf{v}\|} = \frac{\|\mathbf{v}\|^2}{\|\mathbf{v}\|} = \|\mathbf{v}\|.$$

На единичной сфере форма объёма (в данном случае — форма площади) в сферических координатах имеет вид $\sin\theta\, d\theta \wedge d\varphi$ (см. с. 210).

Задача 12.8. Докажите, что форма $\omega = x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy$ является формой объёма на единичной сфере.

Задача 12.9. Пусть $\mathbf{n} = (n_1, n_2, n_3)$ — семейство векторов единичных нормалей к поверхности в $\mathbb{R}^3$. Докажите, что 2-форма $\pm(n_1 dy \wedge dz + n_2 dz \wedge dx + n_3 dx \wedge dy)$ является формой объёма на этой поверхности.

Пример 12.9. *На единичной сфере в $\mathbb{R}^n$ дифференциальная форма $\omega = \sum_{i=1}^n (-1)^{i-1} x_i dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n$ является формой объёма.*

Доказательство. Пусть $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ — касательные векторы к единичной сфере в точке $(x_1, \dots, x_n)$. Значение формы ω на векторах $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ равно ориентированному объёму V параллелепипеда, натянутого на векторы $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, (x_1, \dots, x_n)$. При этом $(x_1, \dots, x_n)$ — единичный вектор, перпендикулярный векторам $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$. Следовательно, с точностью до знака число V равно объёму параллелепипеда, натянутого на касательные векторы $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$, что и требовалось. □

12.7. Интегрирование на многообразиях

Чтобы определить *интеграл* формы $\omega \in \Omega_c^n(M^n)$, где M^n — ориентированное многообразие, поступим следующим образом. Возьмём произвольный локально конечный атлас $\{U_\alpha\}$ с картами $f_\alpha \colon U_\alpha \to \mathbb{R}^n$, сохраняющими ориентацию. Пусть $\{\lambda_\alpha\}$ — подчинённое ему гладкое разбиение единицы. Положим

$$\int_{M^n} \omega = \sum_\alpha \int_{\mathbb{R}^n} \lambda_\alpha (f_\alpha^{-1})^* \omega = \sum_\alpha \int_{U_\alpha} \lambda_\alpha \omega.$$

Носитель формы $\lambda_\alpha \omega$ — замкнутое подмножество компактного множества $\operatorname{supp} \lambda_\alpha$, поэтому он компактен. Если $\{V_\beta\}$ — другой ориентирующий атлас и $\{\mu_\beta\}$ — подчинённое ему разбиение единицы, то из равенства $\sum_\beta \mu_\beta = 1$ следует, что $\sum_\alpha \int_{U_\alpha} \lambda_\alpha \omega = \sum_{\alpha,\beta} \int_{U_\alpha} \lambda_\alpha \mu_\beta \omega$. Носитель формы $\lambda_\alpha \mu_\beta \omega$ лежит в $U_\alpha \cap V_\beta$, поэтому $\int_{U_\alpha} \lambda_\alpha \mu_\beta \omega = \int_{V_\beta} \lambda_\alpha \mu_\beta \omega$. Следовательно, $\sum_\alpha \int_{U_\alpha} \lambda_\alpha \omega = \sum_{\alpha,\beta} \int_{V_\beta} \lambda_\alpha \mu_\beta \omega = \sum_\beta \int_{V_\beta} \mu_\beta \omega$, поскольку $\sum_\alpha \lambda_\alpha = 1$. Это означает, что определение интеграла формы ω по многообразию M^n корректно.

Историческое замечание. Интегрирование на многообразиях разработал Анри Пуанкаре (1854-1912) в 1895 году.

Если $f\colon M^n \to N^n$ — диффеоморфизм ориентированных многообразий, ω — дифференциальная n-форма на N^n, то $\int_{N^n} \omega = \pm \int_{M^n} f^*\omega$, где знак зависит от того, сохраняет или изменяет ориентацию диффеоморфизм f: для диффеоморфизма, сохраняющего ориентацию, берётся знак плюс, а для диффеоморфизма, обращающего ориентацию, берётся знак минус. Для диффеоморфизма областей в $\mathbb{R}^n$ эта формула — это просто формула замены переменных. Поэтому, чтобы получить соответствующую формулу для многообразий, достаточно воспользоваться разбиением единицы.

Теорема Стокса заключается в том, что $\int_{M^n} d\omega = \pm \int_{\partial M^n} \omega$ для любой формы $\omega \in \Omega_c^n(M^n)$. Во втором интеграле формально следовало бы вместо ω написать $i^*\omega$, где $i\colon \partial M^n \to M^n$ — естественное включение; другими словами, $i^*\omega$ — ограничение формы ω на ∂M^n. От знака $\pm$ перед вторым интегралом можно избавиться посредством соглашения об ориентации ∂M^n, индуцированной ориентацией многообразия M^n.

Историческое замечание. В 1854 году Георг Габриэль Стокс (1819-1903) предложил на экзамене теорему, которую ему сообщил в 1850 году в письме лорд Кельвин (Уильям Томсон, 1824-1907): интеграл ротора векторного поля по поверхности равен интегралу этого векторного поля по краю поверхности. Эта теорема и её многомерные обобщения получила название теорема Стокса.

Займёмся доказательством теоремы Стокса.

Случай 1. $M^n = \mathbb{R}^n$.

Пусть $\omega = \varphi\, dx_1 \wedge \dots \wedge dx_{n-1} \in \Omega_c^{n-1}(\mathbb{R}^n)$. Тогда $d\omega = \frac{\partial \varphi}{\partial x_n}\, dx_n \wedge dx_1 \wedge \dots \wedge dx_{n-1} = (-1)^{n-1} \frac{\partial \varphi}{\partial x_n} dx_1 \wedge \dots \wedge dx_n$. Поэтому
$$\int_{\mathbb{R}^n} d\omega = (-1)^{n-1} \int \left(\int_{-\infty}^{\infty} \frac{\partial \varphi}{\partial x_n} dx_n \right) dx_1 \dots dx_{n-1}.$$
Но $\int_{-\infty}^{\infty} \frac{\partial \varphi}{\partial x_n}\, dx_n = \varphi(x_1, \dots, x_{n-1}, \infty) - \varphi(x_1, \dots, x_{n-1}, -\infty) = 0$, поскольку носитель функции φ компактен.

Случай 2. $M^n = \mathbb{R}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n |\, x_n \geq 0\}$.

Пусть $\omega = \sum_i \varphi_i\, dx_1 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n \in \Omega_c^{n-1}(\mathbb{R}^n)$. Тогда $d\omega = \sum_i (-1)^{n-1} \frac{\partial \varphi_i}{\partial x_n}\, dx_1 \wedge \dots \wedge dx_n$. Поэтому
$$\int_{\mathbb{R}^n_+} d\omega = \sum_{1 \leq i < n} (-1)^{i-1} \left(\int_{-\infty}^{\infty} \frac{\partial \varphi_i}{\partial x_i} dx_i \right) dx_1 \dots \widehat{dx_i} \dots dx_n + \\ + (-1)^{n-1} \left(\int_0^{\infty} \frac{\partial \varphi_n}{\partial x_n} dx_n \right) dx_1 \dots dx_{n-1}.$$
Если $i \neq n$, то $\int_{-\infty}^{\infty} \frac{\partial \varphi_i}{\partial x_i} dx_i = 0$, как и в случае 1. При $i = n$ получаем интеграл
$$\int_0^{\infty} \frac{\partial \varphi_n}{\partial x_n}\, dx_n = \varphi_n(x_1, \dots, x_{n-1}, \infty) - \varphi_n(x_1, \dots, x_{n-1}, 0) = \\ = -\varphi_n(x_1, \dots, x_{n-1}, 0).$$
Значит,
$$\begin{aligned} \int_{\mathbb{R}^n_+} d\omega &= (-1)^n \int_{\mathbb{R}^{n-1}} \varphi_n(x_1, \dots, x_{n-1}, 0) dx_1 \dots dx_{n-1} = \\ &= (-1)^n \int_{\partial \mathbb{R}^n_+} \omega; \end{aligned}$$
при записи последнего равенства мы воспользовались тем, что при ограничении на $\partial \mathbb{R}^n_+$ форма dx_n обращается в нуль; поэтому при ограничении на $\partial \mathbb{R}^n_+$ от формы ω остаётся только одно слагаемое $\varphi_n\, dx_1 \wedge \dots \wedge dx_{n-1}$.

От знака $(-1)^n$ можно избавиться, если предполагать, что ориентация $\partial\mathbb{R}^n_+$ задаётся формой $(-1)^n dx_1 \wedge \dots \wedge dx_{n-1}$. По-другому это соглашение можно сформулировать следующим образом. Пусть базис $e_1, \dots, e_{n-1}$ задаёт положительную ориентацию края, ε — вектор внешней нормали к краю. Тогда базис $\varepsilon, e_1, \dots, e_{n-1}$ задаёт положительную ориентацию многообразия. Действительно, $\varepsilon = -e_n$, поэтому базис $\varepsilon, e_1, \dots, e_{n-1}$ ориентирован так же, как базис $(-1)^n e_1, \dots, e_{n-1}, e_n$.

В дальнейшем мы будем предполагать, что ориентация $\partial\mathbb{R}^n_+$ выбрана именно так, и в формуле Стокса знак писать не будем.

Случай 3. *M^n — произвольное ориентированное многообразие.*

Переход от $\mathbb{R}^n$ и $\mathbb{R}^n_+$ к произвольным многообразиям делается посредством разбиений единицы. Пусть $\{U_\alpha\}$ — ориентирующий атлас для M^n, в котором все множества U_α диффеоморфны $\mathbb{R}^n$ или $\mathbb{R}^n_+$; $\{\lambda_\alpha\}$ — подчинённое ему гладкое разбиение единицы. Представим форму $\omega \in \Omega^{n-1}_c(M^n)$ в виде $\omega = \sum \lambda_\alpha \omega$. Теорему Стокса достаточно доказать для каждой отдельной формы $\lambda_\alpha\omega$. Носитель этой формы — замкнутое подмножество компактного множества $\operatorname{supp} \lambda_\alpha \subset U_\alpha$, поэтому он компактен.

Задача 12.10. Докажите, что интеграл точной формы по замкнутому ориентируемому многообразию равен нулю.

Пример 12.10. *Дифференциальная форма $r^{-n}\omega$, где $\omega = \sum_{i=1}^n (-1)^{i-1} x_i dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n$ и $r = \sqrt{x_1^2 + \dots + x_n^2}$, определённая на $\mathbb{R}^n \setminus \{0\}$, замкнутая, но не точная.*

Доказательство. Замкнутость формы $r^{-n}\omega$ доказана в решении задачи 12.7. Предположим, что эта форма точная. Тогда её ограничение на единичную сферу — тоже точная форма. При $r = 1$ форма $r^{-n}\omega$ совпадает с формой ω. Но согласно примеру 12.9 эта форма является формой объёма на единичной сфере, поэтому её интеграл по единичной сфере отличен от нуля. Это противоречит результату задачи 12.10. $\square$

12.8. Степень отображения

При определении степени гладкого отображения нам понадобится теорема Сарда, поэтому начнём с формулировки и доказательства этой теоремы.

Пусть M^n — многообразие. Говорят, что множество $X \subset M^n$ имеет *меру нуль*, если существует покрытие многообразия M^n счётным множеством карт $\varphi_i \colon U_i \to \mathbb{R}^n$, для которого каждое множество $\varphi_i(X \cap U_i) \subset \mathbb{R}^n$ имеет меру нуль. Проверим, что это определение не зависит от выбора отображений $\varphi_i \colon U_i \to \mathbb{R}^n$. Прежде всего напомним, что объединение счётного набора множеств меры нуль в $\mathbb{R}^n$ является множеством меры нуль. Кроме того, верна следующая лемма.

Лемма 12.1. *Пусть $f \colon \mathbb{R}^n \to \mathbb{R}^n$ — гладкое отображение, $X \subset \mathbb{R}^n$ — множество меры нуль. Тогда $f(X) \subset \mathbb{R}^n$ — множество меры нуль.*

Доказательство. Пространство $\mathbb{R}^n$ можно представить в виде счётного объединения кубов, поэтому достаточно рассмотреть случай, когда множество X содержится в кубе I^n. Воспользовавшись результатом задачи 9.35 или теоремой 9.8 о среднем значении, легко показать, что для гладкого отображения f можно выбрать константу K так, что $\mathbf{x}, \mathbf{y} \in I^n$, то $\|f(\mathbf{x}) - f(\mathbf{y})\| \le K\|\mathbf{x} - \mathbf{y}\|$. Поэтому образ при отображении f шара радиуса r, расположенного в I^n, содержится в шаре радиуса Kr. Это означает, что если множество X покрыто шарами, сумма объёмов которых меньше ε, то множество $f(X)$ можно покрыть шарами, сумма объёмов которых меньше $K^n\varepsilon$. $\square$

Пусть $f \colon M^m \to N^n$ — гладкое отображение. Точку $x \in M^m$ называют *критической*, если $\operatorname{rank} f(x) < \min(m, n)$. В противном случае точку x называют *регулярной*. Точку $y \in N^n$ называют *критическим значением*, если $y = f(x)$, где x — некоторая критическая точка.

При доказательстве теоремы Сарда нам потребуется такое следствие теоремы Фубини: если компактное множество $C \subset \mathbb{R}^n$ таково, что любое его сечение гиперплоскостью $x_1 = a$ имеет меру нуль, то C имеет меру нуль (см. задачу 11.3).

Теорема 12.9 (Сард). *Множество критических значений любого гладкого отображения многообразий имеет меру нуль.*

Доказательство. Требуемое утверждение достаточно доказать для гладкого отображения $f\colon U \to \mathbb{R}^n$, где $U \subset \mathbb{R}^m$ — открытое множество. Применим индукцию по m. При $m = 0$ утверждение очевидно: если $n = 0$, то множество критических значений пусто, а если $n \geq 1$, то подмножество $\mathbb{R}^n$, состоящее из одной точки, имеет меру нуль. Будем предполагать, что теорема Сарда верна для гладких отображений $V \to \mathbb{R}^n$, где $V \subset \mathbb{R}^{m-1}$.

Пусть C — множество критических точек отображения $f\colon U \to \mathbb{R}^n$, C_i — множество всех точек $\mathbf{x} \in U$, в которых обращаются в нуль все частные производные отображения f порядка не выше i. Ясно, что $C \supset C_1 \supset C_2 \supset \dots$

Шаг 1. *Множество $f(C \setminus C_1)$ имеет меру нуль.*

Если $n = 1$, то $C_1 = C$. Поэтому будем считать, что $n \geq 2$. Пусть $\mathbf{c} \in C \setminus C_1$. Тогда $\frac{\partial f_i}{\partial x_j}(\mathbf{c}) \neq 0$ для некоторых i, j. Можно считать, что $i = j = 1$. Рассмотрим отображение $h\colon U \to \mathbb{R}^m$, заданное формулой
$$h(x_1, \dots, x_m) = (f_1(\mathbf{x}), x_2, \dots, x_m).$$
Ясно, что матрица Якоби отображения h в точке c равна $\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{c}) & * \\ 0 & I \end{pmatrix}$, где I — единичная матрица. Поэтому в точке $\mathbf{c}$ к отображению h применима теорема об обратной функции. Это означает, что существует окрестность V точки $\mathbf{c}$, ограничение на которую отображения h является гомеоморфизмом. Положим $g = fh^{-1}$; тогда
$$g(x_1, \dots, x_m) = (x_1, g_2(\mathbf{x}), \dots, g_n(\mathbf{x})).$$
В частности, g отображает множество $(\{t\} \times \mathbb{R}^{m-1}) \cap h(V)$ в $\{t\} \times \mathbb{R}^{n-1}$; пусть g^t — ограничение отображения g на это множество. Легко проверить, что точка множества $(\{t\} \times \mathbb{R}^{m-1}) \cap h(V)$ является критической точкой отображения g^t тогда и только тогда, когда она является критической точкой отображения g. Действительно,
$$\left(\frac{\partial g_i}{\partial x_j}\right) = \begin{pmatrix} 1 & 0 \\ * & \left(\frac{\partial g_i^t}{\partial x_j}\right) \end{pmatrix}.$$

Согласно предположению индукции множество критических точек отображения g^t имеет меру нуль в $\{t\} \times \mathbb{R}^{n-1}$. Поэтому возникает желание применить следствие теоремы Фубини и показать, что мера множества $h(V \cap C)$ равна нулю (а это означает, что мера образа множества $V \cap C$ при отображении f равна нулю). Непосредственно применить следствие теоремы Фубини нельзя, потому что множество $f(V \cap C)$ не компактно. Но множество $V \cap C$ можно представить в виде объединения счётного объединения компактных множеств, поэтому множество $f(V \cap C)$ тоже можно представить в виде объединения счётного объединения компактных множеств. К каждому из этих множеств можно применить следствие теоремы Фубини и получить желаемый результат.

Итак, у каждой точки $\mathbf{c} \in C \setminus C_1$ есть такая окрестность V, что множество $f((C \setminus C_1) \cap V) \subset f(C \cap V)$ имеет меру нуль. Множество $C \setminus C_1$ можно покрыть счётным набором таких окрестностей, поэтому мера множества $f(C \setminus C_1)$ равна нулю.

Шаг 2. *Множество $f(C_k \setminus C_{k-1})$ имеет меру нуль при любом $k \geq 1$.*

Доказательство аналогично шагу 1. Пусть $\mathbf{c} \in C_k \setminus C_{k-1}$. Тогда $\frac{\partial^{k+1} f_i}{\partial x_{j_1} \dots \partial x_{j_{k+1}}}(\mathbf{c}) \neq 0$ для некоторых i, $j_1, \dots, j_{k+1}$. Можно считать, что $i = j_1 = 1$. Пусть $w(\mathbf{x}) = \frac{\partial^k f_1}{\partial x_{j_2} \dots \partial x_{j_{k+1}}}(\mathbf{x})$. Тогда $w(\mathbf{c}) = 0$, поскольку $\mathbf{c} \in C_k$, и
$$\frac{\partial w}{\partial x_1}(\mathbf{c}) = \frac{\partial^{k+1} f_i}{\partial x_{j_1} \dots \partial x_{j_{k+1}}}(\mathbf{c}) \neq 0.$$
Рассмотрим отображение $h\colon U \to \mathbb{R}^m$, заданное формулой
$$h(x_1, \dots, x_m) = (w(\mathbf{x}), x_2, \dots, x_m).$$
Дальше действуем точно так же, как и на шаге 1. Здесь нужно будет воспользоваться тем, что любая точка множества
$$h\left((C_k \setminus C_{k+1}) \cap V\right) \cap (\{t\} \times \mathbb{R}^{m-1})$$
является критической точкой для отображения g, а значит, и для отображения g^t.

Шаг 3. *Множество $f(C_k)$ имеет меру нуль при достаточно большом k (например, при $k > (m/n) - 1$).*

Достаточно рассмотреть случай, когда $U = (0,1)^m$, причём гладкое отображение f определено в некоторой окрестности куба $[0,1]^m$. Действительно, любое открытое множество U можно покрыть счётным набором открытых кубов, обладающих таком свойством.

Пусть $\mathbf{c} \in C_k \cap U$. В таком случае в разложении Тейлора для $f(\mathbf{c}+\mathbf{h})$ отсутствуют члены порядка ниже $k+1$. Поэтому существует такая константа K, что если $\mathbf{c}+\mathbf{h} \in U$, то $\|f(\mathbf{c}+\mathbf{h}) - f(\mathbf{c})\| \leq K\|\mathbf{h}\|^{k+1}$.

Разобьём открытый куб $U = (0,1)^m$ на l^m кубов с ребром $1/l$ и рассмотрим лишь те из них, которые пересекаются с C_k. Образ при отображении f любого такого куба содержится в шаре радиуса Kd^{k+1}, где $d = \sqrt{n}/l$ — максимальное расстояние между точками куба. Поэтому множество $f(U \cap C_k)$ можно покрыть шарами, сумма объёмов которых не превосходит $K' l^m (d^{k+1})^n = K'' l^{m-(k+1)n}$. Если $m < (k+1)n$, т. е. $k > (m/n) - 1$, то $\lim_{l\to\infty} l^{m-(k+1)n} = 0$. □

Историческое замечание. Артур Сард (1909-1980) доказал теорему Сарда в 1942 году.

Теперь можно перейти непосредственно к определению степени гладкого отображения. Пусть $f\colon M^n \to N^n$ — гладкое отображение многообразий одной и той же размерности n. Мы будем предполагать, что многообразия M^n и N^n замкнутые, ориентируемые и их ориентации фиксированы. Из теоремы Сарда следует, что у отображения f есть регулярное значение $y \in N^n$. Пусть $x \in f^{-1}(y)$. Отображение $df(x)\colon T_xM^n \to T_yN^n$ является изоморфизмом, поэтому можно выбрать в точках x и y локальные координаты, ориентации которых согласованы с ориентациями многообразий M^n и N^n, и рассмотреть число $\operatorname{sgn} J_f(x)$ — знак якобиана отображения f в точке x. Назовём *степенью* отображения f относительно точки y число

$$\deg(f,y) = \sum_{x \in f^{-1}(y)} \operatorname{sgn} J_f(x).$$

Эта сумма имеет смысл, потому что множество $f^{-1}(y)$ конечно. Действительно, предположим, что множество $f^{-1}(y)$ содержит бесконечно много различных точек. Из компактности многообразия M^n следует, что существует последовательность попарно различных точек $x_i \in f^{-1}(y)$, $i \in \mathbb{N}$, сходящаяся к точке x_0. Тогда $f(x_0) = y$ и по теореме об обратной функции у точки x_0 есть окрестность U, гомеоморфно отображающаяся на окрестность точки y. В частности, $(U \setminus \{x_0\}) \cap f^{-1}(y) = \varnothing$. Приходим к противоречию.

Мы предполагаем, что если множество $f^{-1}(y)$ пусто, то $\deg(f,y) = 0$.

Историческое замечание. Понятие степени отображения ввёл Лёйтзен Эгберт Ян Брауэр (1881-1966) в 1910 году для доказательства инвариантности размерности (евклидовы пространства разной размерности не гомеоморфны).

Пример 12.11. *Пусть $S^1 = \{z \in \mathbb{C}\colon |z| = 1\}$. Рассмотрим отображение $f\colon S^1 \to S^1$, заданное формулой $f(z) = z^n$, $n \in \mathbb{Z}$. Если $n \neq 0$, то $\deg(f,w) = n$ для любой точки $w \in S^1$. (Если $n = 0$, то нужно исключить нерегулярную точку $w = 1$.)*

Если $f\colon M^n \to N^n$ — гладкое отображение замкнутых ориентированных многообразий, причём многообразие N^n связно, то степень $\deg(f,y)$ не зависит от выбора точки $y \in N^n$. Мы не будем доказывать это утверждение (доказательство можно найти, например, в [Пр2]). Отметим лишь, что в следующей теореме требование постоянства степени отображения совсем не обременительно.

Теорема 12.10. *Пусть $f\colon M^n \to N^n$ — гладкое отображение замкнутых связных ориентированных многообразий, степень которого $\deg(f,y) = \deg(f)$ не зависит от выбора точки $y \in N^n$. Тогда для любой n-формы ω на многообразии N^n имеет место равенство $\int_{M^n} f^*\omega = \deg(f) \int_{N^n} \omega$.*

Доказательство. Пусть точка $y \in N^n$ не критическая. Тогда можно выбрать такую открытую окрестность $U \subset N^n$ этой точки, что её прообраз $f^{-1}(U)$ состоит из попарно непересекающихся открытых множеств $U_1, \ldots, U_q$ в M^n, причём каждое отображение $f\colon U_i \to U$ является диффеоморфизмом. Рассмотрим сначала случай, когда носитель формы ω содержится в U. Пусть ω_i — ограничение формы $f^*\omega$ на множество U_i, ε_i — знак якобиана диффеоморфизма $f\colon U_i \to U$. Тогда

$$\begin{aligned}\int_{M^n} f^*\omega &= \sum_{i=1}^{q}\int_{U_i}\omega_i = \sum_{i=1}^{q}\varepsilon_i\int_U \omega = \\ &= \left(\sum_{i=1}^{q}\varepsilon_i\right)\int_U \omega = \deg(f)\int_{N^n}\omega\end{aligned}$$

Критические точки и их прообраз можно не учитывать, хотя и по разным причинам. Критические точки не влияют на интеграл $\int_{N^n}\omega$, потому что их мера равна нулю. Прообраз критических точек не влияет на интеграл $\int_{M^n} f^*\omega$, потому что на прообразе критических точек $f^*\omega = 0$.

Доказательство теоремы для произвольной формы ω теперь можно получить, выбрав для каждой не критической точки $y \in N^n$ открытую окрестность U указанного выше вида и рассмотрев разбиение единицы, подчинённое этому покрытию множества некритических точек. □

12.9. Функции Морса

Пусть M^n — многообразие без края и $f\colon M^n \to \mathbb{R}$ — гладкая функция. Точка $x \in M^n$ является критической тогда и только тогда, когда $\operatorname{rank} f(x) = 0$, т. е. отображение $df\colon T_xM^n \to \mathbb{R}$ нулевое. В локальных координатах $(x_1, \ldots, x_n)$ это означает, что $\frac{\partial f}{\partial x_i}(x) = 0$ при $i = 1, \ldots, n$.

Критическую точку x функции f называют *невырожденной*, если *матрица Гессе*, или *гессиан*, $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)$ невырожденная. Это определение не зависит от выбора локальных координат, поскольку при переходе к другим локальным координатам $(y_1, \ldots, y_n)$ гессиан преобразуется следующим образом:

$$\left(\frac{\partial^2 f}{\partial y_i \partial y_j}(x)\right) = J^T\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)J,$$

где $J = \left(\frac{\partial x_i}{\partial y_j}\right)$.

Гладкую функцию $f\colon M^n \to \mathbb{R}$ называют *функцией Морса*, если все её критические точки невырожденные.

Напомним, что индекс квадратичной формы $\sum a_{ij}x_ix_j$, заданной симметрический матрицей (a_{ij}), определяется следующим образом. Заменой переменных (над полем $\mathbb{R}$) квадратичную форму можно привести к виду $-y_1^2 - \cdots - y_q^2 + y_{q+1}^2 + \cdots + y_n^2$. В таком случае *индексом* квадратичной формы называют число q. Индекс квадратичной формы можно также определить как максимальную размерность подпространства, на котором форма отрицательно определена.

Индексом невырожденной критической точки x функции f называют индекс гессиана функции f в точке x.

Пример 12.12. *Пусть $f(\boldsymbol{x}) = -x_1^2 - \cdots - x_q^2 + x_{q+1}^2 + \cdots + x_n^2$. Тогда точка $\boldsymbol{x}_0 = (0, \ldots, 0)$ является критической, причём её индекс равен q.*

Теорема 12.11 (лемма Морса). *В окрестности невырожденной критической точки индекса q существуют такие локальные координаты с началом в критической точке, что в этих координатах функция f имеет вид $f(x_1, \ldots, x_n) = f(0) - x_1^2 - \cdots - x_q^2 + x_{q+1}^2 + \cdots + x_n^2$.*

Доказательство. Можно считать, что $f(0) = 0$ и локальные координаты представляют собой выпуклую окрестность в $\mathbb{R}^n$. Тогда согласно лемме Адамара существуют такие гладкие функции $g_1, \ldots, g_n$, что $f(x) = \sum x_i g_i(x)$ и $g_i(0) = \frac{\partial f}{\partial x_i}(0)$. По условию точка 0 критическая, т. е. $\frac{\partial f}{\partial x_i}(0) = 0$. Ещё раз применив ту же самую лемму, получим $f(x) = \sum x_i x_j h_{ij}(x)$, где $h_{ij}(0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(0)$, т. е. $(h_{ij}(0))$ — гессиан функции f в критической точке. После замены $h_{ij}(x)$ на $\frac{1}{2}\left(h_{ij}(x) + h_{ji}(x)\right)$ можно считать, что матрица $h_{ij}(x)$ симметрическая, а после линейной замены координат можно считать, что $h_{11}(0) \neq 0$. Уменьшив при необходимости координатную окрестность, можно считать, что $h_{11}(x)/h_{11}(0) > 0$ для всех x из координатной окрестности. Положим

$$y_1 = x_1 + \frac{h_{12}(x)}{h_{11}(x)}x_2 + \cdots + \frac{h_{1n}(x)}{h_{11}(x)}x_n, \quad y_i = x_i \text{ при } i \geq 2.$$

Согласно теореме об обратной функции отображение $(x_1, \ldots, x_n) \mapsto (y_1, \ldots, y_n)$ является диффеоморфизмом (возможно, в ещё меньшей координатной окрестности). Легко проверить, что

$$\sum x_i x_j h_{ij}(x) = h_{11}(x) y_1^2 + \sum_{i,j \geq 2} y_i y_j \tilde{h}_{ij}(x).$$

Сделаем замену $z_1 = y_1 \sqrt{|h_{11}(x)|}$, $z_i = y_i$ при $i \geq 2$, а затем аналогичные преобразования применим к квадратичной форме от $n-1$ переменной и т. д. $\square$

Следствие. *Невырожденная критическая точка является изолированной критической точкой.*

Докажем теперь, что на любом многообразии существуют функции Морса.

Теорема 12.12. *На любом многообразии M^n существует функция Морса.*

Доказательство. Пусть $g\colon M^n \to \mathbb{R}$ — произвольная гладкая функция (например, постоянная). Функцию Морса f мы будем строить, последовательно изменяя функцию g. Области $U_{i,1} \subset U_{i,2} \subset U_{i,3}$, карты $\varphi_i\colon U_{i,3} \to \mathbb{R}^n$ и функцию $\lambda\colon \mathbb{R}^n \to \mathbb{R}$ мы определим так же, как в доказательстве этой теоремы. Изменить функцию g так, чтобы у новой функции не было вырожденных критических точек в области $U_{i,1}$, можно с помощью следующего утверждения.

Лемма 1. *Пусть $U \subset \mathbb{R}^n$ — открытое множество и $f\colon U \to \mathbb{R}$ — гладкая функция. Тогда для почти всех линейных функций $A\colon \mathbb{R}^n \to \mathbb{R}$ функция $f + A$ имеет только невырожденные критические точки.*

Доказательство. Рассмотрим отображение $F\colon U \to \mathbb{R}^n$, заданное формулой

$$F(x) = \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right).$$

Точка x_0 является критической точкой отображения F тогда и только тогда, когда гессиан функции f в точке x_0 является вырожденной матрицей. Поэтому условие, что функция $f(x) - a_1 x_1 - \cdots - a_n x_n$ имеет вырожденную критическую точку x_0, эквивалентно тому, что $F(x_0) = (a_1, \ldots, a_n)$ и x_0 — критическая точка отображения F, т. е. $(a_1, \ldots, a_n)$ — образ критической точки отображения F. Остаётся воспользоваться теоремой Сарда. $\square$

Из леммы 1 следует, что если g_{i-1} — гладкая функция на многообразии M^n, то существует линейная функция $A(x) = a_1 x_1 + \cdots + a_n x_n$ со сколь угодно малыми коэффициентами a_i, для которой функция $g_i(y) = g_{i-1}(y) + \lambda(\varphi_i(y)) A(\varphi_i(y))$ не имеет вырожденных критических точек на множестве $\overline{U_{i,1}}$ (лемму 1 нужно применить к множеству $U = U_{i,2} \supset \overline{U_{i,1}}$; отметим, что $\lambda(\varphi_i(y)) = 1$ для всех точек $y \in \overline{U_{i,1}}$).

Мы научились исправлять функцию g_{i-1} на множестве $U_{i,1}$. Остаётся научиться делать это так, чтобы не портить достигнутого ранее. А именно, пусть функция g_{i-1} не имеет вырожденных критических точек на множестве $\bigcup_{j=1}^{i-1} \overline{U_{j,1}}$; мы хотим, чтобы функция g_i тоже не имела вырожденных критических точек на этом множестве. Функция g_{i-1} изменяется только на компактном множестве $\overline{U_{i,2}}$; при этом на компактном множестве $\overline{U_{i,2}} \cap \left(\bigcup_{j=1}^{i-1} \overline{U_{j,1}} \right)$ у неё нет вырожденных критических точек.

Лемма 2. *Пусть $f, g\colon U \to \mathbb{R}$ — гладкие функции на открытом множестве $U \subset \mathbb{R}^n$, причём функция f не имеет вырожденных критических точек на компактном множестве $K \subset U$. Тогда существует такое число $\varepsilon > 0$, что если все первые и вторые производные функции $f - g$ во всех точках множества K по модулю меньше ε, то функция g не имеет на K вырожденных критических точек.*

Доказательство. Функция

$$F = \sum \left(\frac{\partial f}{\partial x_i} \right)^2 + \left[\det \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right) \right]^2$$

обращается в нуль только в вырожденных критических точках функции f, поэтому на компактном множестве K функция F достигает положительного минимума δ. Если число ε достаточно мало, то

$$\sum \left(\frac{\partial f}{\partial x_i}\right)^2 - \sum \left(\frac{\partial g}{\partial x_i}\right)^2 < \delta/2$$

и

$$\left[\det\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)\right]^2 - \left[\det\left(\frac{\partial^2 g}{\partial x_i \partial x_j}\right)\right]^2 < \delta/2,$$

поэтому

$$\sum \left(\frac{\partial g}{\partial x_i}\right)^2 + \left[\det\left(\frac{\partial^2 g}{\partial x_i \partial x_j}\right)\right]^2 > 0,$$

а значит, функция g не имеет на K вырожденных критических точек. $\square$

Если числа $a_1, \ldots, a_n$ достаточно малы, то все первые и вторые производные функции $\lambda(x)(a_1x_1 + \cdots + a_nx_n)$ тоже малы. Поэтому требуемую функцию g_i можно построить, воспользовавшись леммой 2. $\square$

Историческое замечание. В 1925-1931 годы Марстон Морс (1892-1977) разработал теорию, позволяющую получать информацию о глобальном строении многообразия, исходя из информации об индексах критических точек функций Морса. Эта теория получила название теория Морса.

12.10. Решения задач

12.1. Если точка многообразия M^n одновременно является как внутренней, так и точкой края, то существуют гладкие взаимно обратные отображения $f\colon U \to V$ и $f\colon V \to U$ открытых множеств в $\mathbb{R}^n$, причём выполняются следующие свойства:

1) $0 \in U$, $0 \in V$ и $f(0) = g(0) = 0$;

2) в U есть такое открытое подмножество U', содержащее точку 0, что $f(U') \subset \mathbb{R}^n_+$.

Из того, что у отображения f в окрестности точки 0 есть гладкое обратное отображение, следует, что $\det\left(\frac{\partial f_i}{\partial x_j}(0)\right) \neq 0$. Поэтому согласно теореме об обратной функции у точки 0 есть такая окрестность $\widetilde{U} \subset U$, что множество $f(\widetilde{U})$ открыто в $\mathbb{R}^n$ и отображение $f|_{\widetilde{U}}\colon \widetilde{U} \to f(\widetilde{U})$ является гомеоморфизмом. Следовательно, $f(\widetilde{U} \cap U')$ — открытая $\mathbb{R}^n$ окрестность точки 0. Но это противоречит тому, что $f(\widetilde{U} \cap U') \subset f(\widetilde{U}') \subset \mathbb{R}^n_+$.

12.2. Пусть (U_a, φ) и $(V_{f(a)}, \psi)$ — локальные системы координат с началами в точках $a \in M^m$ и $f(a) \in N^n$; при этом будем считать, что $f(U_a) \subset V_{f(a)}$. Запишем отображение f в локальных координатах, т. е. рассмотрим отображение $\tilde{f} = \psi f \varphi^{-1}$. Для отображения $\tilde{f}$ можно рассмотреть матрицу Якоби $J = \left(\frac{\partial \tilde{f}_i}{\partial x_j}(x)\right)$. Если f — иммерсия, то ранг матрицы J в начале координат равен m, а если f — субмерсия, то ранг равен n. Поэтому для иммерсии можно выбрать $i_1, \ldots, i_m \in \{1, \ldots, n\}$, а для субмерсии можно выбрать $j_1, \ldots, j_n \in \{1, \ldots, m\}$ так, что в начале координат

$$\det\begin{pmatrix} \frac{\partial \tilde{f}_{i_1}}{\partial x_1} & \ldots & \frac{\partial \tilde{f}_{i_1}}{\partial x_m} \\ \multicolumn{3}{c}{\dotfill} \\ \frac{\partial \tilde{f}_{i_m}}{\partial x_1} & \ldots & \frac{\partial \tilde{f}_{i_m}}{\partial x_m} \end{pmatrix} \neq 0 \quad \text{или} \quad \det\begin{pmatrix} \frac{\partial \tilde{f}_1}{\partial x_{j_1}} & \ldots & \frac{\partial \tilde{f}_1}{\partial x_{j_n}} \\ \multicolumn{3}{c}{\dotfill} \\ \frac{\partial \tilde{f}_n}{\partial x_{j_1}} & \ldots & \frac{\partial \tilde{f}_n}{\partial x_{j_n}} \end{pmatrix} \neq 0.$$

Не теряя общности, будем считать, что $i_k = k$ (соответственно, $j_k = k$).

Матрицу Якоби J можно дополнить до квадратной матрицы вида $\begin{pmatrix} J_1 & J_2 \\ 0 & I \end{pmatrix}$ или $\begin{pmatrix} J_1 & 0 \\ J_2 & I \end{pmatrix}$, где I — единичная матрица порядка $|n - m|$. Существует отображение $\widetilde{F}$, матрица Якоби которого как раз и является такой матрицей. А именно, в случае иммерсии полагаем

$$\widetilde{F}(x_1, \ldots, x_m, y_1, \ldots, y_{n-m}) =$$
$$= (\tilde{f}_1(x), \ldots, \tilde{f}_m(x), \tilde{f}_{m+1}(x) + y_1, \ldots, \tilde{f}_n(x) + y_{n-m}),$$

а в случае субмерсии полагаем

$$\widetilde{F}(x_1, \dots, x_m, y_1, \dots, y_{m-n}) = (\tilde{f}(x, y), y_1, \dots, y_{m-n});$$

сокращённо эти отображения можно записать в виде $\widetilde{F}(x, y) = \tilde{f}(x) + (0, y)$ и $\widetilde{F}(x, y) = (\tilde{f}(x, y), y)$.

В начале координат якобиан отображения $\widetilde{F}$ отличен от нуля, поэтому согласно теореме об обратной функции в некоторой окрестности начала координат у отображения $\widetilde{F}$ есть гладкое обратное отображение $\widetilde{F}^{-1}$.

В случае иммерсии заменим отображение ψ на $\widetilde{F}^{-1}\psi$ (иными словами, мы изменяем локальную систему координат в образе отображения f). В новых локальных координатах отображение f устроено следующим образом:

$$x \mapsto \psi f \varphi^{-1} \tilde{f}(x) = \widetilde{F}(x, 0) \mapsto \widetilde{F}^{-1}(x, 0).$$

В случае субмерсии заменим отображение φ на $\widetilde{F}\varphi$ (иными словами, мы изменяем локальную систему координат в прообразе отображения f). В новых локальных координатах отображение f устроено следующим образом:

$$(\tilde{f}(x, y), y) \mapsto \widetilde{F}^{-1}(x, y) \mapsto \psi f \varphi^{-1} \tilde{f}(x, y),$$

т.е. $(x', y) \mapsto x'$, где $x' = \tilde{f}(x, y)$. Ясно также, что в малой окрестности начала координат при фиксированном y отображение $x \mapsto \tilde{f}(x, y)$ локально эпиморфно, поскольку якобиан этого отображения отличен от нуля.

12.3. Сначала переставим dx_{j_1} через p множителей $dx_{i_1} \wedge \dots \wedge dx_{i_p}$, затем переставим dx_{j_2} через те же p множителей, и т.д. Для этого потребуется совершить pq транспозиций.

12.4. а) По определению $d\omega = df \wedge dx + dg \wedge dy = (f_x dx + f_y dy) \wedge dx + (g_x dx + g_y dy) \wedge dy = (-f_y + g_x) dx \wedge dy$.

б) Ясно, что $d(x\,dy - y\,dx) = dx \wedge dy - dy \wedge dx = 2dx \wedge dy$.

в) Покажем, что $d\frac{x\,dy - y\,dx}{x^2+y^2} = 0$. Согласно задаче а) для этого достаточно проверить, что $f_y = g_x$ для $f(x, y) = \frac{-y}{x^2+y^2}$ и $g(x, y) = \frac{x}{x^2+y^2}$. Несложные вычисления показывают, что $f_y(x, y) = g_x(x, y) = \frac{y^2 - x^2}{(x^2+y^2)^2}$.

12.5. Ясно, что $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$. Поэтому требуемое равенство достаточно доказать для формы $\omega = f dx_{i_1} \wedge \dots \wedge dx_{i_k}$. Для такой формы

$$\begin{aligned} d(d\omega) &= d\left(\sum_{r=1}^{n} \frac{\partial f}{\partial x_r} dx_r \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}\right) = \\ &= \sum_{s=1}^{n}\sum_{r=1}^{n} \frac{\partial^2 f}{\partial x_s \partial x_r} dx_s \wedge dx_r \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}. \end{aligned}$$

В этой сумме слагаемые попарно сокращаются, поскольку $dx_r \wedge dx_s = -dx_s \wedge dx_r$.

12.6. Требуемое утверждение достаточно доказать в случае, когда $\omega_1 = f dx_{i_1} \wedge \dots \wedge dx_{i_k}$ и $\omega_2 = g dx_{j_1} \wedge \dots \wedge dx_{j_l}$. В этом случае

$$\begin{aligned} d(\omega_1 \wedge \omega_2) &= d(fg) dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l} = \\ &= (df \wedge g + f \wedge dg) dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l} = \\ &= (df \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}) \wedge (g dx_{j_1} \wedge \dots \wedge dx_{j_l}) + \\ &+ (-1)^k (f dx_{i_1} \wedge \dots \wedge dx_{i_k}) \wedge (dg \wedge dx_{j_1} \wedge \dots \wedge dx_{j_l}) = \\ &= d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2. \end{aligned}$$

Знак $(-1)^k$ появляется из-за перестановки 1-формы dg с k множителями $dx_{i_1} \wedge \dots \wedge dx_{i_k}$.

12.7. а) Дифференциал каждого из слагаемых равен $dx_1 \wedge \dots \wedge dx_n$, поэтому $d\omega = n dx_1 \wedge \dots \wedge dx_n$.

б) Покажем, что $d(\|\mathbf{x}\|^{-n}\omega) = 0$. Дифференциал i-го слагаемого равен форме $dx_1 \wedge \dots \wedge dx_n$ с коэффициентом $\frac{\partial}{\partial x_i}(\|\mathbf{x}\|^{-n} x_i)$, поэтому нужно проверить, что $\sum_{i=1}^{n} \frac{\partial}{\partial x_i}(\|\mathbf{x}\|^{-n} x_i) = 0$ при $\mathbf{x} \neq \mathbf{0}$. Это равенство доказано в решении задачи 9.20.

12.8. Пусть $x = \sin\theta\cos\varphi$, $y = \sin\theta\sin\varphi$ и $z = \cos\theta$. Тогда $dy \wedge dz = \sin^2\theta\cos\varphi\, d\theta \wedge d\varphi$, $dz \wedge dx = \sin^2\theta\sin\varphi\, d\theta \wedge d\varphi$ и $dx \wedge dy = \sin\theta\cos\theta\, d\theta \wedge d\varphi$. Поэтому $\omega = (\sin^3\theta\cos^2\varphi + \sin^3\theta\sin^2\varphi + \sin\theta\cos^2\theta) d\theta \wedge d\varphi = \sin\theta\, d\theta \wedge d\varphi$.

Другое решение этой задачи можно получить, воспользовавшись результатом задачи 12.9, поскольку вектор нормали к единичной сфере в точке (x, y, z) равен $\pm(x, y, z)$.

12.9. Если $\mathbf{v} = (v_1, v_2, v_3)$ и $\mathbf{w} = (w_1, w_2, w_3)$ — касательные векторы к данной поверхности, то площадь параллелограмма, натянутого на эти векторы равна $\pm \det \begin{pmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \\ n_1 & n_2 & n_3 \end{pmatrix}$. Значение указанной 2-формы на векторах $\mathbf{v}$ и $\mathbf{w}$ как раз и равно этому определителю.

12.10. Если многообразие M^n замкнутое, то $\partial M^n = \varnothing$, поэтому согласно теореме Стокса $\int_{M^n} d\omega = \int_{\partial M^n} \omega = \int_{\varnothing} \omega = 0$.

Литература

[Бу] Бурбаки Н., *Функции действительного переменного*, М.: Наука, 1965.

[ВШ] Верещагин Н. К., Шень А., *Начала теории множеств*, М.: МЦНМО, 2012.

[ГР] Гаспер Дж., Рахман М., *Базисные гипергеометрические ряды*, М.: Мир, 1993.

[Гу] Гурса Э., *Курс математического анализа*, Т. 1, ч. 1-2, М.-Л.: ГТТИ, 1933.

[Зо] Зорич В. А., *Математический анализ*, Ч. I-II, М.: МЦНМО, 2017.

[Кац] Кац В. Г., Чен П., *Квантовый анализ*, М.: МЦНМО, 2005.

[Ку] Курант Р., *Курс дифференциального и интегрального исчисления*, Т.1 (1967), Т.2 (1970), М.: Наука.

[Ма] Макаров Б. М., Голузина М. Г., Лодкин А. А., Подкорытов А. Н., *Избранные задачи по вещественному анализу*, СПб: Невский диалект, 2004.

[Ок] Окстоби Дж., *Мера и категория*, М.: Мир, 1974.

[П1] Прасолов В. В., *Задачи и теоремы линейной алгебры*, М.: МЦНМО, 2015.

[Пр2] Прасолов В. В., *Элементы комбинаторной и дифференциальной топологии*, М.: МЦНМО, 2014.

[Пр3] Прасолов В. В., *Элементы теории гомологий*, М.: МЦНМО, 2014.

[Ру] Рудин У., *Основы математического анализа*, М.: Мир, 1976.

[Сп] Спивак М., *Математический анализ на многообразиях*, М.: Мир, 1968.

[Ф] Фихтенгольц Г. М., *Курс дифференциального и интегрального исчисления*, т. 1-3, М.: Физматлит, 2003.

[Э] Энгелькинг Р., *Общая топология*, М.: Мир, 1986.

[Ap1] Apostol T. M., *Mathematical Analysis*, Reading MA, Addison-Wesley, 1974.

[Ap2] Apostol T. M., *An Elementary View of Euler's Summation Formula*, Amer. Math. Monthly **106** (1999), 409–418.

[Br] Bromwich T. J. I'A, *An introduction to the theory of infinite series*, London: Macmillan, 1908.

[Du1] Duistermaat J. J., Kolk J. A. C., *Multidimensional Real Analysis I: Differentiation*, Cambridge University Press, 2004.

[Du2] Duistermaat J. J., Kolk J. A. C., *Multidimensional Real Analysis II: Integration*, Cambridge University Press, 2004.

[Ed] Edwards C. H., *Advanced Calculus of Several Variables*, New York and London: Academic Press, 1973.

[Er] Ernst T., *A Comprehensive Treatment of q-Calculus*, Basel: Springer, 2012.

[Fa] Farrel O. J., Ross B., *Solved Problems in Analysis. As Applied to Gamma, Beta, Legendre and Bessel Functions*, NY: Dover Publications, 1971.

[HW] Hairer E., Wanner G., *Analysis by Its History*, NY: Springer.

[Ha] Havil J., *Gamma: exploring Euler's constant*, Princeton University Press, 2003.

[Joh1] Johnson W. P., *A q-analogue of Faà di Bruno's formula*, J. Comb. Theory, Ser. A **76** (1996), 305–314.

[Joh2] Johnson W. P., *The curious history of Faà di Bruno's formula*, Amer. Math. Monthly **109** (2002), 217–234.

[Jos] Jost J., *Postmodern Analysis*, Berlin, Heidelberg: Springer, 2005.

[Kn] Knopp K., *Theorie und Anwendlung der unendlichen Reihen*, Berlin: Springer, 1922.

[KP] Krantz S. G., Parks H. R., *The Implicit Function Theorem. Hystory, Theory, and Applications*, Birkhäuser, 2003.

[Ge] Gessel I. M., *A q-analog of the exponential formula*, Discrete Math. **40** (1982), 69–80.

[Ha] Havil J., *Gamma. Exploring Euler's Constant*, Princeton University Press, 2003.

[Ji] Jingcheng Tong, *Kummer's Test Gives Characterizations for Convergence or Divergence of all Positive Series*, American Math. Monthly, **101** (1994), 450–452.

[La] Lafontaine J., *An Introduction to Differentiable Manifolds*, Springer, 2015.

[Li1] Liouville J., *Mémoire sur le calcul différentielles à indices quelconques*, J. École Polytechn. **13** (1832), cah. 21, 71–162.

[Li2] Liouville J., *Sur la detérmination des integrales dont la valeur est algébrique*, J. École Polytech. **14** (1833), 124–193.

[McSh] McShane E. T., *The Lagrange Multiplier Rule*, American Math. Monthley, **80** (1973), 922-925.

[Mo] Monsky P., *Simpifying the Proof of Dirichlet's Theorem*, Amer. Math. Monthly **100** (1993), 861–862.

[Mu] Munkres J. R., *Analysis on Manifolds*, Addison-Wesley Publishing Company, 1990.

[Pi] Picard É., *Traité d'Analyse*, T. 1 (1891), 2 (1893), 3 (1896), Paris: Gauthier-Villars.

[Ros1] Rosenlicht M., *Liouville's theorem for functions with elementary integrals*, Pacific J. Math. **24** (1968), 153–161.

[Ros2] Rosenlicht M., *Integration in finite terms*, Amer. Math. Monthly **79** (1972), 963–972.

[Roy] Royden H. L., Fizpatrick P. M., *Real Analysis*, Pearson Education, Prentice Hall, 2010.

Предметный указатель