

НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

М. В. Коробков

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

Конспект лекций

Новосибирск
2014

Оглавление

1	Вещественные числа	1
§1	Предварительные сведения	1
§1.1	Логика высказываний	1
§1.2	Понятия теории множеств	2
§1.3	Операции над множествами	3
§1.4	Кванторы	4
§2	Числовая прямая	5
§2.1	Аксиоматика вещественных чисел	5
§2.2	Абсолютная величина числа	6
§2.3	Верхняя и нижняя грани. Аксиома непрерывности.	7
§2.4	Натуральный ряд	10
§2.5	Бином Ньютона	13
§2.6	Отображения	14
§2.7	Счетные множества	17
§2.8	Свойства счетных множеств	19
§2.9	Расширенная числовая прямая	21
2	Теория предела	23
§1	Предел последовательности	23
§2	Ряды	26
§2.1	Признаки сходимости числовых рядов	26
§2.2	Степенные ряды, радиус сходимости, асимптотика	34
§3	Предел функции	35
§3.1	Точки прикосновения множества	35
§3.2	Предел функции	37
§3.3	Асимптотические соотношения	41
§4	Непрерывные функции	43
§4.1	Определение и локальные свойства непрерывных функций	43
§4.2	Глобальные свойства непрерывных функций	44
§4.3	Основные элементарные функции	47

3	Дифференциальное исчисление	52
§1	Классические теоремы дифференциального исчисления	52
§2	Формула Тэйлора с остаточным членом в форме Лагранжа . . .	54
§3	Выпуклые функции	55
§4	Бесконечные произведения	59
§4.1	Приложения к сходимости рядов: модификации мажорируемого признака сходимости	63
§5	Приложения дифференциального исчисления	65
§5.1	Формула Циолковского	65
§5.2	Падение с парашютом	66
§5.3	Формула радиоактивного распада. Критический радиус.	66
4	Теория интеграла	68
§1	Определенный интеграл	69
§1.1	Формула Тейлора с остаточным членом в интегральной форме	70
§1.2	Иррациональность чисел e^r и π	72
§1.3	Признаки интегрируемости	73
§1.4	Теоремы о среднем	75
§2	Вектор-функции одной переменной	77
§2.1	Об аддитивных функция отрезка	79
§2.2	Площадь криволинейного сектора	80
§2.3	Объем тела вращения	81
§2.4	Площадь поверхности вращения	81
5	Метрические пространства	83
§1	Предел последовательности	83
§2	Открытые и замкнутые множества	85
§3	Внутренность, замыкание и граница множеств	89
§4	Относительная внутренность и замыкание множеств	91
§5	Пределы отображений метрических пространств	91
§6	Непрерывные отображения метрических пространств	92
§7	Гомеоморфизмы	94
§8	Равномерно непрерывные, липшицевы, сжимающие отображения	96
§9	Компактность	97
§9.1	Непрерывные отображения компактных множеств	99

6	Дифференциальное исчисление в многомерном случае	102
§1	Предел функции в точке и предел по направлению	102
§2	Частные производные и дифференцируемость	103
§3	Градиент	109
§4	Старшие производные	111
§5	Мультииндексный формализм	113
§6	Дифференциал k -го порядка	114
§7	Формула Тейлора	115
§8	Локальный экстремум	117
7	Теорема об обратном отображении	120
§1	Предварительные сведения	120
§2	Теорема о липшицевом гомеоморфизме	120
§3	Теорема об обратной функции. Диффеоморфизмы	121
§4	Теорема о неявной функции	124
§5	Правило множителей Лагранжа для поиска условного экстре- мума	126
	УЧЕБНЫЕ ПОСОБИЯ	132

Глава 1

Вещественные числа

§1 Предварительные сведения

§1.1 Логика высказываний

Красота природы издавна вызывает восхищение человека. Но глубже можно вникнуть и оценить эту красоту, если изучать законы природы. А законы природы, как сказал один из классиков, написаны на языке математики. И вот, чтобы ощутить радость от соприкосновения с красотой замысла Творца, надо постичь этот Божественный язык, язык царицы наук, математики.

Классический трактат по математике — "Начала" Евклида. Этот текст состоит из аксиом (постулатов) и теорем. Аксиомы — это положения (высказывания), принимаемые без доказательства, например, как самоочевидные основы теории. Теоремы — это логические следствия аксиом. Небольшие вспомогательные теоремы называют леммами. Итак, чтобы приступить к изучению математического языка, нужно познакомиться с логикой.

Логика оперирует (имеет дело) с высказываниями (суждениями), которые могут быть истинными или ложными, для высказывания P будем обозначать этот факт символами $P = И$ и $P = Л$ соответственно.

Число π иррациональное, число 28 совершенное — примеры истинных высказываний.

Число $2^{2^p} + 1$ простое при всяком натуральном p — это пример ложного высказывания.

Высказывания называются равносильными, $P \Leftrightarrow Q$, если они истинны или ложны одновременно, т.е. $P = И$ в том и только том случае, когда $Q = И$.

Над высказываниями можно делать логические операции.

$P \wedge Q$, P "и" Q , истинно тогда и только тогда, когда истинны и P , и Q .

$$[(P \wedge Q) = И] \Leftrightarrow [P = И \text{ и } Q = И].$$

$P \vee Q$, P "или" Q , истинно тогда и только тогда, когда истинно **хотя бы одно из высказываний** P или Q .

$\neg P$, или "не" P , истинно тогда и только тогда, когда P ложно. Отметим, что $\neg \neg P \Leftrightarrow P$.

Наибольший интерес для нас представляет импликация: $P \Rightarrow Q$, т.е. "из P следует Q " ("истинность P влечет истинность Q "), она истинна тогда и только тогда, когда Q истинно, или P ложно.

$$(P \Rightarrow Q) \Leftrightarrow (\neg P \vee Q). \quad (1.1)$$

$$\neg(P \Rightarrow Q) \Leftrightarrow \neg(\neg P \vee Q) \Leftrightarrow (\neg \neg P \wedge \neg Q) \Leftrightarrow (P \wedge \neg Q).$$

В частности, если P ложно, то высказывание $(P \Rightarrow Q)$ всегда истинно, независимо от Q . Т.е. из лжи следует что угодно. Если кошалоты летают, то кошки гавкают на луну, и т.п.

Отметим также на этом примере, что при взятии отрицания символ $\wedge$ заменяется на $\vee$ и наоборот.

Можно свести сказанное выше в таблицу.

P	Q	$\neg P$	$(P \wedge Q)$	$(P \vee Q)$	$(P \Rightarrow Q)$
И	И	Л	И	И	И
Л	И	И	Л	И	И
И	Л	Л	Л	И	Л
Л	Л	И	Л	Л	И

С помощью этой таблицы можно доказывать разные формулы логики высказываний. Например, такие

$$P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R),$$

$$P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R).$$

Чтобы доказать истинность этих формул, нужно перебрать всевозможные варианты значений P, Q, R и убедиться, что истинность формул в правой и левой частях наступает одновременно (**Упражнение**).

§1.2 Понятия теории множеств

Говоря о множестве, понимают под этим совокупность объектов произвольного рода, рассматриваемую как единое целое. Это, конечно, несколько "наивное" определение, но для наших целей достаточное, т.к. множество не является у нас основным объектом изучения, но играет вспомогательную роль в построении мат. анализа. Однако важно научиться работать с множествами и выполнять действия над ними.

Можно говорить о множестве деревьев в лесу, множестве городов в России, о множестве точек плоскости, множестве слов русского языка и т.д.

Объекты, составляющие данное множество, называются его *элементами*. Множество называется заданным, если заданы его элементы, т.е. если про каждый объект можно выяснить, является ли он элементом этого множества, или нет.

Если x является элементом множества A , то этот факт обозначают символом $x \in A$.

Как же задаются множества? Один из способов — перечисление элементов. Например, $\{a, b, c, d\}$ — это множество, состоящее из четырех элементов, a, b, c и d . Или, если устроить переключку, то можно задать множество студентов в данной аудитории. В деканате можно выяснить, каковы множества студентов в той или иной группе и т.п.

Но студентов-то конечное количество. А как описать бесконечные множества, например, множество четных чисел? Поэтому более распространенный способ задания множества таков: выделить из данного множества все его элементы, обладающие каким-нибудь свойством. Так, формула

$$\{x \in A : x \text{ обладает свойством } P(x)\}$$

задает совокупность элементов из A , обладающих свойством $P(x)$. Например, множество четных чисел определяется как $\{x \in \mathbb{Z} : x \text{ делится на } 2 \text{ без остатка}\}$. (Напомним, что символами $\mathbb{N}$ и $\mathbb{Z}$ обозначаются множества натуральных и целых чисел соответственно.)

Говорят, что множество A содержится в множестве B , если каждый элемент множества A принадлежит множеству B . Этот факт обозначают символом $A \subset B$. Сформулированное определение можно переписать на языке логики высказываний

$$A \subset B \Leftrightarrow (x \in A \Rightarrow x \in B)$$

(проверьте, что эти определения в самом равносильны!). Множества считаются равными, если они состоят из одних и тех же элементов, т.е.

$$A = B \Leftrightarrow [(A \subset B) \wedge (B \subset A)]$$

Замечание 1.1.1. Из сделанных определений можно заключить, что для всякого $x \in A$ справедливо включение $\{x\} \subset A$, однако, вообще говоря, не верна формула $\{x\} \in A$.

Важную роль играет так называемое *пустое множество*, в котором нет ни одного элемента. Пустое множество обозначается символом $\emptyset$ и играет роль нуля в числовом множестве. Это означает, в частности, что высказывание $x \in \emptyset$ ложно для всякого объекта x . Значит, в силу правила для импликации (1.1), справедливо включение $\emptyset \subset A$ для каждого множества A . Более того, про элементы пустого множества справедливы любые утверждения, например, все элементы пустого множества зеленые, но в то же время все элементы пустого множества красные, и т.д.

§1.3 Операции над множествами

Объединение $A \cup B$ множеств A и B состоит из элементов, которые принадлежат хотя бы одному из множеств A или B :

$$A \cup B := \{x : x \in A \vee x \in B\}.$$

Пересечение $A \cap B$ множеств A и B состоит из элементов, которые принадлежат одновременно обоим множествам A и B :

$$A \cap B := \{x : x \in A \wedge x \in B\}.$$

Разность $A \setminus B$ множеств A и B состоит из элементов, которые принадлежат множеству A , но не принадлежат B :

$$A \setminus B := \{x : x \in A \wedge x \notin B\}.$$

Иногда определяют еще *симметрическую разность* $A \Delta B$ множеств A и B , состоящую из элементов, принадлежащих в точности одному из множеств A либо B .

$$A \Delta B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

Упражнение 1.1.2. Докажите равенства

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C),$$

$$A \Delta (B \Delta C) = (A \Delta B) \Delta C.$$

Указание. Можно воспользоваться соответствующими формулами логики высказываний.

Пусть даны два объекта x, y . Будем говорить, что они образуют упорядоченную пару (x, y) , если x считается *первым*, а y — *вторым*. То есть, это не просто совокупность элементов $\{x, y\}$, но важен еще и их порядок. Из этого определения вытекает, что упорядоченные пары (x, y) , (u, v) совпадают в том и только том случае, когда выполнены оба равенства $x = u$ и $y = v$.

Упражнение 1.1.3. Покажите, что упорядоченную пару (x, y) можно отождествить с множеством $\{\{x\}, \{x, y\}\}$. **Указание.** Нужно проверить, что $\{\{x\}, \{x, y\}\} = \{\{u\}, \{u, v\}\}$ в том и только том случае, когда $x = u$ и $y = v$.

Декартовым произведением $A \times B$ множеств A и B называется совокупность всех упорядоченных пар (x, y) таких, что $x \in A$, $y \in B$:

$$A \times B := \{(x, y) : x \in A, y \in B\}.$$

Пример 1.1.4. Если множества $A = \{a_1, \dots, a_n\}$, $B = \{b_1, \dots, b_m\}$ конечны, то их декартово произведение удобно можно записать в виде таблицы (матрицы)

$$A \times B := \begin{Bmatrix} (a_1, b_1) & (a_1, b_2) & \dots & (a_1, b_m) \\ (a_2, b_1) & (a_2, b_2) & \dots & (a_2, b_m) \\ \dots & \dots & \dots & \dots \\ (a_n, b_1) & (a_n, b_2) & \dots & (a_n, b_m) \end{Bmatrix}.$$

Если множества A, B представляют собой отрезки $A = [a, b]$, $B = [c, d]$, то $A \times B$ представляет собой прямоугольник.

§1.4 Кванторы

Следующие обозначения весьма удобны, а потому широко применяются.

$\forall x \in A P(x)$ означает, что *для любого* элемента x множества A выполнено свойство $P(x)$.

$\exists x \in A P(x)$ означает, что *существует* элемент x множества A , для которого выполнено свойство $P(x)$.

Отметим, что при взятии отрицаний к утверждению, в котором используются кванторы, знак отрицания перемещается вправо и становится после кванторов перед высказыванием, а сами кванторы "по пути" меняются на противоположные. Это иллюстрируется следующим примером.

$$\neg(\forall x \in A \exists y \in B P(x, y)) \Leftrightarrow \exists x \in A \forall y \in B \neg P(x, y).$$

Например, утверждение: Во всех странах есть честные политики. Отрицание: существует страна, в которой нет ни одного честного политика.

Упражнение 1.1.5. Составите отрицание (в словесной форме и формулой) к следующему высказыванию: *имеется галактика, в которой вокруг каждой звезды вращаются ровно три планеты.*

$$\neg[\forall x \in \text{Gall} \exists y \in \text{Star}(x) \text{card}(\text{Planets}(y)) = 3] \Leftrightarrow ???^1$$

¹**Ответы.** Отрицание: во всякой галактике найдется звезда, вокруг которой вращается либо меньше

§2 Числовая прямая

§2.1 Аксиоматика вещественных чисел

Числовая прямая обозначается символом $\mathbb{R}$. Для любой пары чисел определены операции сложения $+$ умножения $\cdot$, удовлетворяющие следующим аксиомам.

Аксиомы сложения.

$$\forall x \in \mathbb{R} \forall y \in \mathbb{R} x + y = y + x$$

$$\exists ! 0 \in \mathbb{R} \forall x \in \mathbb{R} x + 0 = x$$

$$\forall x \in \mathbb{R} \exists ! (-x) \in \mathbb{R} x + (-x) = 0$$

$$\forall x \in \mathbb{R} \forall y \in \mathbb{R} \forall z \in \mathbb{R} x + (y + z) = (x + y) + z$$

Аксиомы умножения.

$$\forall x \in \mathbb{R} \forall y \in \mathbb{R} x \cdot y = y \cdot x$$

$$\exists 1 \in \mathbb{R} \forall x \in \mathbb{R} x \cdot 1 = x$$

$$\forall x \in \mathbb{R} (x \neq 0 \Rightarrow \exists \frac{1}{x} \in \mathbb{R} x \cdot \frac{1}{x} = 1)$$

$$\forall x \in \mathbb{R} \forall y \in \mathbb{R} \forall z \in \mathbb{R} (x \cdot y) \cdot z = x \cdot (y \cdot z)$$

$$\forall x \in \mathbb{R} \forall y \in \mathbb{R} \forall z \in \mathbb{R} x \cdot (y + z) = x \cdot y + x \cdot z.$$

Алгебраические системы, удовлетворяющие перечисленным аксиомам, называются *полями*. Таким образом, содержание предыдущих двух абзацев можно охарактеризовать одним предложением: вещественные числа образуют алгебраическое поле. Но о вещественных числах мы знаем дополнительно, что это поле *упорядоченное*, ниже идет блок аксиом, описывающих структуру порядка.

Отношение порядка.

Для каждой пары элементов x, y можно выяснить, верно ли для них отношение $x \leq y$, и, если оно верно, то будем говорить, что x не превосходит y , или x предшествует y .

Свойства этого отношения:

$$x \leq x$$

$$(x \leq y) \wedge (y \leq z) \Rightarrow x \leq z$$

$$(x \leq y) \wedge (y \leq x) \Rightarrow x = y$$

$$x \leq y \Rightarrow \forall z \in \mathbb{R} (x + z \leq y + z)$$

$$(x \leq y) \wedge (z \geq 0) \Rightarrow x \cdot z \leq y \cdot z.$$

Однако этим свойствам удовлетворяют также другие алгебраические системы, например, поле рациональных чисел. Имеется еще одна аксиома, которая однозначно (с точностью до алгебраического изоморфизма) характеризует поле вещественных чисел. Но,

трех планет, либо больше трех планет.

$$\neg [\forall x \in \text{Gall} \exists y \in \text{Star}(x) \text{card}(\text{Planets}(y)) = 3] \Leftrightarrow \forall x \in \text{Gall} \exists y \in \text{Star}(x) \text{card}(\text{Planets}(y)) = 3.$$

прежде чем перейти к ней, нужно выяснить, в чем "недостаток" системы рациональных чисел, почему нельзя ими ограничиться. Отчего возникает потребность расширения этого "с детства" знакомого числового поля?

ТЕОРЕМА 1.2.1. *Не существует рационального числа x такого, что $x^2 = 2$.*

ДОКАЗАТЕЛЬСТВО. Пусть имеется несократимая дробь $\frac{m}{n}$ такова, что $\frac{m^2}{n^2} = 2$, m, n — натуральные числа. Тогда $m^2 = 2n^2$, значит, число m — четное (если бы m было числом нечетным, то нечетным было бы и m^2). Но тогда $m = 2k$, $m^2 = 4k^2$, т.е. m^2 делится на 4. Тогда $n^2 = \frac{1}{2}m^2$ делится на 2, следовательно, n также представляет собой четное число, однако это противоречит предположению о несократимости дроби $\frac{m}{n}$. $\square$

Упражнение 1.2.2. Докажите, что если натуральное число a не является полным квадратом, то не существует рационального числа $\frac{m}{n}$ такого, что $\frac{m^2}{n^2} = a$. **Указание.** Предположите, что дробь $\frac{m}{n}$ несократима, т.е. $\text{НОД}(m, n) = 1$, и используйте теорему Евклида о том, что каждое натуральное число единственным образом (с точностью до порядка множителей) разлагается в виде произведения простых чисел $a = p_1^{\alpha_1} \dots p_k^{\alpha_k}$, причем, по условию, здесь найдется нечетный показатель α_i (выпишите аналогичные разложения в произведение простых чисел для m, n).

§2.2 Абсолютная величина числа

Модуль, или абсолютная величина числа $x \in \mathbb{R}$ определяется по следующему правилу.

$$|x| = \begin{cases} x, & \text{если } x \geq 0; \\ -x, & \text{если } x < 0. \end{cases}$$

Ясно, что модуль всегда неотрицателен. Следующие свойства модуля очень легко выводятся непосредственно из определения.

1. $\forall x \in \mathbb{R} \forall y \in \mathbb{R} |x \cdot y| = |x| \cdot |y|$.
2. $\forall x \in \mathbb{R} \forall a \geq 0 (|x| \leq a \Leftrightarrow -a \leq x \leq a)$.

Докажем другие важные свойства.

3. $\forall x \in \mathbb{R} \forall y \in \mathbb{R}$ справедливо неравенство

$$|x + y| \leq |x| + |y|. \quad (1.2)$$

ДОКАЗАТЕЛЬСТВО. В самом деле, по свойству 2 имеем

$$-|x| \leq x \leq |x|, \quad -|y| \leq y \leq |y|. \quad (1.3)$$

Складывая эти неравенства, получаем

$$-|x| - |y| \leq |x + y| \leq |x| + |y|, \quad (1.4)$$

что в силу все того же свойства 2 равносильно искомому неравенству (1.2). $\square$

§2.3 Верхняя и нижняя грани. Аксиома непрерывности.

Венчает ряд аксиом вещественная ряда следующий постулат, который играет важнейшую роль, демонстрирует принципиальное отличие от ситуации с рациональными числами, и, вместе с предыдущими аксиомами, полностью характеризует множество вещественных чисел (с точностью до изоморфизма).

Аксиома 1.2.3 (непрерывности). Пусть A, B непустые подмножества $\mathbb{R}$ такие, что

$$\forall a \in A \forall b \in B \quad a \leq b. \quad (1.5)$$

Тогда найдется число $c \in \mathbb{R}$ такое, что $A \leq c \leq B$, или, записывая это более развернуто,

$$\forall a \in A \forall b \in B \quad a \leq c \leq b. \quad (1.6)$$

Из этой аксиомы сразу же получается масса следствий. Напомним, что для пары вещественных чисел $a \leq b$ отрезком с концами a, b называется множество точек $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$.

ТЕОРЕМА 1.2.4 (о пересекающихся отрезках). Пусть дано семейство отрезков $\{I\}_{t \in T}$, причем пересечение любой пары из этих отрезков непусто. Тогда найдется точка $c \in \bigcap_{t \in T} I_t$, лежащая в пересечении всех этих отрезков.

ДОКАЗАТЕЛЬСТВО. Пусть $I_t = [a_t, b_t]$. Обозначим $A = \{a_t : t \in T\}$, $B = \{b_t : t \in T\}$. Возьмем произвольную пару точек $a_{t_1} \in A$, $b_{t_2} \in B$. Чтобы применить аксиому непрерывности, нам необходимо доказать неравенство

$$a_{t_1} \leq b_{t_2}. \quad (1.7)$$

Предположим, что это не верно, т.е. $a_{t_1} > b_{t_2}$. Но тогда пересечение отрезков $I_{t_1} \cap I_{t_2} = [a_{t_1}, b_{t_1}] \cap [a_{t_2}, b_{t_2}]$ пусто, т.к. отрезок I_{t_1} лежит строго правее отрезка I_{t_2} . Получили противоречие с условием теоремы. Тем самым неравенство (1.7) установлено для произвольной пары элементов $a_{t_1} \in A$, $b_{t_2} \in B$. Теперь по аксиоме непрерывности получаем, что найдется точка $c \in \mathbb{R}$ такая, что

$$A \leq c \leq B,$$

т.е.

$$\forall t \in T \quad a_t \leq c \leq b_t.$$

Но последнее означает, что

$$\forall t \in T \quad c \in I_t.$$

Теорема доказана. □

СЛЕДСТВИЕ 1.2.5 (о вложенных отрезках отрезках). Пусть дано семейство вложенных отрезков $\{I\}_{t \in T}$, т.е. $\forall t_1, t_2 \in T$ имеем $I_{t_1} \subset I_{t_2}$ или $I_{t_2} \subset I_{t_1}$. Тогда найдется точка $c \in \bigcap_{t \in T} I_t$, лежащая в пересечении всех этих отрезков.

Чтобы докопаться до остальных следствий аксиомы непрерывности, нам понадобятся некоторые новые понятия, которые активно будут использоваться в дальнейшем.

Пусть $M \subset \mathbb{R}$. Будем говорить, что число $v \in \mathbb{R}$ является *верхней границей* множества M (или *ограничивает M справа*), если $\forall x \in M \quad x \leq v$. Последнее условие часто

записывают в сокращенном виде $M \leq v$. Если хотя бы одно такое число v существует, то множество M называют *ограниченным сверху*.

Аналогично, говорят, что число $u \in \mathbb{R}$ является *нижней границей* множества M (или *ограничивает M слева*), если $\forall x \in M \ x \geq u$. Если такое число u существует, то множество M называют *ограниченным снизу*.

Множество называется *ограниченным*, если оно ограничено и сверху и снизу.

Множества всех верхних границ обозначается символом $\Gamma_+(M)$, аналогично, $\Gamma_-(M)$ есть множество всех нижних границ.

Число a называется *наибольшим* элементом множества M , если $a \in M$ и $M \leq a$. Наибольший элемент M обозначается символом $\max M$. Аналогично, a называется *наименьшим* элементом множества M (обозначается $a = \min M$), если $a \in M$ и $a \leq M$. Очевидно, что если максимальный элемент существует, то он является верхней границей: $\max M \in \Gamma_+(M)$, аналогично $\min M \in \Gamma_-(M)$. Однако эти максимальные и минимальные элементы существуют далеко не всегда, например, открытый отрезок $(a, b) = \{x : a < x < b\}$ не имеет минимального и максимального элемента.

Точной верхней границей множества M называется наименьшая из его верхних границ. Точную верхнюю границу иногда называют еще *верхней гранью* и обозначают символом $\sup M$ (читается "супремум M "). Таким образом, $\sup M = \min \Gamma_+(M)$. Аналогично, *точной нижней границей* (или *нижней гранью*) множества M называется наибольшая из его нижних границ, обозначается $\inf M = \max \Gamma_-(M)$ (читается "инфимум M ").

ТЕОРЕМА 1.2.6 (о гранях). *Если множество непустое $M \subset \mathbb{R}$ ограничено сверху, то у него есть верхняя грань $\sup M \in \mathbb{R}$. Если M ограничено снизу, то у него есть нижняя грань $\inf M \in \mathbb{R}$.*

СЛЕДСТВИЕ 1.2.7 (границы ограниченного множества). *Всякое ограниченное непустое множество имеет верхнюю и нижнюю грани в $\mathbb{R}$.*

ДОКАЗАТЕЛЬСТВО. Пусть множество M ограничено сверху. Тогда множество верхних границ $\Gamma_+(M)$ непусто. Обозначим $A = M$, $B = \Gamma_+(M)$. По определению верхних границ, $\forall a \in M \ \forall b \in \Gamma_+(M) \ a \leq b$. По аксиоме непрерывности (примененной к нашей ситуации, где роль A играет M , а роль B — множество $\Gamma_+(M)$) найдется число $c \in \mathbb{R}$ такое, что

$$M \leq c \leq \Gamma_+(M).$$

Первое из этих неравенств означает, что само число c является верхней границей множества M , т.е. $c \in \Gamma_+(M)$. Тогда второе неравенство говорит о том, что c является минимальным элементом множества $\Gamma_+(M)$, т.е. $c = \min \Gamma_+(M)$. Последний факт означает, по определению, что $c = \sup M$.

Рассуждение о существовании $\inf M$ в случае, когда M ограничено снизу, проводится аналогично. $\square$

Приведем удобные для дальнейшей работы переформулировки определения верхних и нижних граней. Пусть $c = \sup M$. Это означает, что выполнены два условия.

- 1) $M \leq c$;
- 2) $c \leq \Gamma_+(M)$.

Второе из этих условий равносильно тому, что всякое число $x < c$ не является верхней границей множества M . Тогда верна следующая

ЛЕММА 1.2.8. Пусть $\emptyset \neq M \subset \mathbb{R}$ и $\Gamma_+(M) \neq \emptyset$. Тогда равенство $c = \sup M$ эквивалентно одновременному выполнению следующих двух условий:

- 1) $M \leq c$;
- 2) для всякого числа $x < c$ найдется число $y \in M$ такое, что $x < y \leq c$.

Упражнение 1.2.9. Сформулируйте аналогичное свойство для $\inf$.

Ответ. Если $\emptyset \neq M \subset \mathbb{R}$ и $\Gamma_-(M) \neq \emptyset$, то равенство $c = \inf M$ эквивалентно одновременному выполнению следующих двух условий:

- 1) $c \leq M$;
- 2) для всякого числа $x > c$ найдется число $y \in M$ такое, что $c \leq y < x$.

Воспользуемся доказанной выше теоремой, чтобы продемонстрировать одно из важных преимуществ $\mathbb{R}$ перед $\mathbb{Q}$: возможность извлечения корней из положительных чисел (ср. с Теоремой 1.2.1).

ТЕОРЕМА 1.2.10. Для любого числа $y \geq 0$ найдется число $a \in \mathbb{R}$ такое, что $a^2 = y$.

ДОКАЗАТЕЛЬСТВО. Зафиксируем $y > 0$. Обозначим через M множество $\{x > 0 : x^2 \geq y\}$. Это множество непусто. В самом деле, для всякого $t > 0$ обозначим $x_t = \frac{1}{2}(t + \frac{y}{t})$. Тогда имеем

$$(x_t)^2 - y = \frac{1}{4}(t^2 - 2y + \frac{y^2}{t^2}) = \frac{1}{4}(t - \frac{y}{t})^2 \geq 0.$$

Итак, для всякого $t > 0$ число x_t принадлежит множеству M .

Очевидно, что $0 \leq M$, так что построенное множество ограничено снизу. Поэтому существует $c = \inf M \in \mathbb{R}$, причем $c \geq 0$. Докажем, что $c^2 = y$. Предположим, что это не так, тогда имеются две гипотетических возможности.

- 1) $c^2 > y$;
- 2) $c^2 < y$.

Изучим каждый из указанных вариантов. Предположим сначала, что $c^2 > y$. Тогда возьмем $x = \frac{1}{2}(c + \frac{y}{c})$. Имеем $x \geq 0$, $x^2 > y$ (см. выше), и $c - x = \frac{1}{2}(c - \frac{y}{c}) = \frac{c^2 - y}{2c} > 0$, т.е. $x \in M$ и $x < c$. Но тогда c не является нижней границей множества M , противоречие.

Рассмотрим второй случай, когда $c^2 < y$. Тогда, если $0 < \varepsilon < 1$, то $(c + \varepsilon)^2 = c^2 + 2c\varepsilon + \varepsilon^2 < c^2 + 2c\varepsilon + \varepsilon = c^2 + (2c + 1)\varepsilon$. Возьмем теперь ε такое, что $0 < \varepsilon < 1$ и $\varepsilon < \frac{y - c^2}{2c + 1}$. Тогда $(c + \varepsilon)^2 < y$. Поскольку c является точной нижней границей, найдется число $z \in M$ такое, что $c \leq z < c + \varepsilon$. Но тогда $z^2 < (c + \varepsilon)^2 < y$, и снова получили противоречие, теперь уже с определением множества M .

Итак, ни один из перечисленных случаев 1)–2) не может иметь места. Следовательно, $c^2 = y$, и теорема доказана. $\square$

Для множеств $A, B \subset \mathbb{R}$ обозначим $A + B = \{a + b : a \in A, b \in B\}$. Имеет место следующая

ТЕОРЕМА 1.2.11 (Свойства точных границ). Пусть A, B — непустые подмножества $\mathbb{R}$, тогда справедливы следующие утверждения:

- 1) $\sup(A + B) = \sup A + \sup B$, $\inf(A + B) = \inf A + \inf B$;
- 2) $\inf(-A) = -\sup A$, $\sup(-A) = \inf A$;
- 3) если $A \subset B$, то $\sup B \geq \sup A$, $\inf B \leq \inf A$.

ДОКАЗАТЕЛЬСТВО. Докажем первое утверждение про $\sup$. Обозначим $c_1 = \sup A$, $c_2 = \sup B$, $c = c_1 + c_2$. Проверим, что для c выполнены оба условия Леммы 1.2.8 (в применении к множеству $A + B$). А именно, для любой пары элементов $a \in A$, $b \in B$ имеем $a \leq c_1$, $b \leq c_2$, следовательно $a + b \leq c_1 + c_2 = c$, т.е. $A + B \leq c$. Возьмем теперь произвольно $x < c$. Обозначим $\varepsilon = \frac{1}{2}(c - x)$. Тогда по свойству (2) Леммы 1.2.8 в применении к каждому из множеств A, B найдутся числа $\bar{a} \in A$, $\bar{b} \in B$ такие, что $c_1 - \varepsilon < \bar{a} \leq c_1$, $c_2 - \varepsilon < \bar{b} \leq c_2$. Суммируя эти равенства (и учитывая, что $c = c_1 + c_2$), получаем, что

$$x = c - 2\varepsilon < \bar{a} + \bar{b} \leq c.$$

Тем самым мы убедились в справедливости обоих условий Леммы 1.2.8 (в применении ко множеству $A + B$). Утверждение $\inf(A + B) = \sup A + \sup B$ доказано. Остальные утверждения Теоремы 1.2.11 доказываются аналогично. $\square$

§2.4 Натуральный ряд

Множество $M \subset \mathbb{R}$ называют *индуктивным*, если оно удовлетворяет следующим двум условиям

- 1) $1 \in M$;
- 2) для всякого $x \in M$ справедливо включение $x + 1 \in M$.

ЛЕММА 1.2.12. Пусть $\{M_\alpha\}_{\alpha \in A}$ — произвольное семейство индуктивных множеств. Тогда их пересечение $M = \bigcap_{\alpha \in A} M_\alpha$ также представляет собой индуктивное множество.

ДОКАЗАТЕЛЬСТВО. Предоставляется учащимся как упражнение. $\square$

Обозначим через $\mathbb{N}$ пересечение всех индуктивных множеств. По доказанной Лемме 1.2.12 получаем, что само $\mathbb{N}$ представляет собой индуктивное множество. По построению, оно является наименьшим индуктивным множеством, т.е.

$$\text{для всякого индуктивного множества } M \text{ справедливо включение } M \supset \mathbb{N} \quad (1.8)$$

На самом деле, определенное таким необычным способом $\mathbb{N}$ представляет собой не что иное, как с детства знакомое множество натуральных чисел $\mathbb{N} = \{1, 2, 3, 4, \dots\}$.

Теперь наша ближайшая задача — доказать фундаментальный принцип математической индукции.

ТЕОРЕМА 1.2.13 (Математическая индукция). Пусть имеется некоторое утверждение $P(x)$ о натуральных числах $x \in \mathbb{N}$, причем выполнены следующие два условия:

- 1) $P(1)$ истинно;
- 2) если $n \in \mathbb{N}$ и $P(n)$ истинно, то $P(n + 1)$ тоже истинно.

Тогда $P(x)$ истинно для всех натуральных чисел $x \in \mathbb{N}$.

ДОКАЗАТЕЛЬСТВО. Обозначим $M = \{x \in \mathbb{N} : P(x) \text{ истинно}\}$. Тогда в силу условий теоремы $1 \in M$, и, если $x \in M$, то $x + 1 \in M$. Значит, M является индуктивным множеством. По формуле (1.2.17) имеем $\mathbb{N} \subset M$. Но по построению справедливо также обратное включение $M \subset \mathbb{N}$, следовательно $M = \mathbb{N}$. $\square$

Принцип математической индукции является мощным и красивым инструментом для доказательства множества полезных и изящных фактов теории чисел. Мы начнем со следующего элементарного неравенства, которое будет нам не раз помогать и в дальнейшем.

ТЕОРЕМА 1.2.14 (Неравенство Бернулли). Пусть $x \geq -1$ и $n \in \mathbb{N}$. Тогда

$$(1 + x)^n \geq 1 + nx. \quad (1.9)$$

ДОКАЗАТЕЛЬСТВО. Зафиксируем $x \geq -1$. Для $n = 1$ неравенство очевидно. Попробуем применить принцип математической индукции. Предположим, что неравенство верно для некоторого $n \in \mathbb{N}$, т.е. $(1 + x)^n \geq 1 + nx$. Посмотрим, что происходит на следующем шаге при $n + 1$:

$$(1 + x)^{n+1} = (1 + x) \cdot (1 + x)^n \geq^2 (1 + x)(1 + nx) = 1 + nx + x + nx^2 \geq 1 + (n + 1)x.$$

Итак, предположив, что неравенство Бернулли справедливо для n , мы получили, что оно верно и для $n + 1$. Тогда по принципу математической индукции, это неравенство верно для всех $n \in \mathbb{N}$. Теорема доказана. $\square$

Хотя следующее утверждение интуитивно очевидно, оно играет важную роль в построении основ вещественного анализа, и недаром носит свое громкое имя.

ТЕОРЕМА 1.2.15 (Принцип Архимеда). Множество $\mathbb{N}$ не ограничено сверху в $\mathbb{R}$.

ДОКАЗАТЕЛЬСТВО. Предположим, что это неверно. Тогда, по Теореме 1.2.6 у множества $\mathbb{N}$ существует точная верхняя грань $\sup \mathbb{N} \in \mathbb{R}$, обозначим ее символом L . Возьмем число $L' = L - 1$. По свойству верхней грани (см. Теорему 1.2.8 (2)) найдется натуральное число $n \in \mathbb{N}$ такое, что $L' < n \leq L$. Тогда $n + 1 > L' + 1 = L$. Но число $n + 1$ в силу индукционного принципа также принадлежит множеству $\mathbb{N}$, а значит, поскольку $\mathbb{N} \leq \sup \mathbb{N} = L$, должно выполняться неравенство $n + 1 \leq L$. Полученное противоречие завершает доказательство Теоремы. $\square$

Обозначим через $\mathbb{Z}$ множество целых чисел:

$$\mathbb{Z} = \mathbb{N} - \mathbb{N} = \{n - m : n \in \mathbb{N}, m \in \mathbb{N}\}.$$

Следующие утверждения интуитивно очевидны, однако полезно дать их строгое доказательство, чтобы ощутить, как работает принцип индукции для построения основ арифметики.

ТЕОРЕМА 1.2.16. Для множества натуральных чисел $\mathbb{N}$ справедливы следующие утверждения:

- (i) единица является его наименьшим элементом, т.е. $\mathbb{N} \ni 1 \leq \mathbb{N}$;
- (ii) если $\mathbb{N} \ni n > 1$, то $n - 1 \in \mathbb{N}$;
- (iii) если $n \in \mathbb{N}$, $m \in \mathbb{N}$ и $m > n$, то $m - n \in \mathbb{N}$;
- (iv) если $n \in \mathbb{Z}$, $m \in \mathbb{Z}$ и $m > n$, то $m - n \in \mathbb{N}$ и $m \geq n + 1$.

ДОКАЗАТЕЛЬСТВО. Первое утверждение (i) вытекает из включения 1.8 и того факта, что луч $\{x \in \mathbb{R} : x \geq 1\}$ является индуктивным множеством. Остальные утверждения предоставляются учащимся в качестве упражнения. $\square$

ТЕОРЕМА 1.2.17 (о наименьшем элементе). Произвольное непустое подмножество натуральных чисел имеет минимальный элемент.

²по индукционному предположению

ДОКАЗАТЕЛЬСТВО. Пусть $\emptyset \neq M \subset \mathbb{N}$. Поскольку по определению натуральных чисел $1 \geq M$, то M ограничена снизу и существует $\inf M \in \mathbb{R}$, обозначим его L . По свойству нижней грани (см. Теорему 1.2.8 (2)) найдется натуральное число $n \in M$ такое, что

$$L \leq n < L + 1. \quad (1.10)$$

Проверим, что на самом деле

$$\forall n \in M \quad n \geq m. \quad (1.11)$$

Если бы это было не так, т.е. если $M \ni n < m$, то по свойству (iv) Теоремы 1.2.16 $n \leq m - 1 < L$, что противоречит определению $L = \inf M$. Значит, оценка (1.11) выполнена, следовательно, число m является нижней границей множества M , и, по определению нижней грани, $L \geq m$. Учитывая первое из неравенств (1.10), получаем, что $m = L = \inf M$, что и требовалось доказать. $\square$

Точно таким же методом можно доказать следующий факт.

ТЕОРЕМА 1.2.18. Если непустое подмножество целых чисел ограничено снизу, то оно имеет минимальный элемент, а если оно ограничено сверху — то максимальный.

ТЕОРЕМА 1.2.19. Для любой пары чисел $x > 0$ и $y \in \mathbb{R}$ существует единственное число $k \in \mathbb{Z}$ со свойством

$$kx \leq y < (k + 1)x. \quad (1.12)$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим подмножество целых чисел $M = \{n \in \mathbb{Z} : \frac{y}{x} < n\}$. Множество M непусто (в силу принципа Архимеда) и ограничено снизу, поэтому вследствие (1.2.18) существует его минимальный элемент $m = \min M$. Тогда $m - 1 \notin M$, т.е. $m - 1 \leq \frac{y}{x} < m$, что равносильно искомому неравенству (1.12) с $k = m - 1$.

Докажем утверждение об единственности. Пусть имеется два таких целых числа: $kx \leq y < (k + 1)x$ и $nx \leq y < (n + 1)x$. Тогда $-(n + 1)x < -y \leq -nx$, складывая эти неравенства, получаем $(k - n - 1)x < 0 < (k + 1 - n)x$, значит,

$$k - n - 1 < 0 \quad \text{и} \quad k - n + 1 > 0.$$

В силу свойства (iv) Теоремы 1.2.16 получаем, что $k - n \leq 0$ и $k - n \geq 0$, т.е. $k = n$. $\square$

В качестве одного из приложений полученных теорем, докажем следующее полезное утверждение

ТЕОРЕМА 1.2.20. Между любыми двумя точками $a, b \in \mathbb{R}$ с $a < b$ найдутся рациональные числа, т.е. $(a, b) \cap \mathbb{Q} \neq \emptyset$.

ДОКАЗАТЕЛЬСТВО. Поскольку по условию $b - a > 0$, в силу принципа Архимеда существует натуральное число n такое, что

$$(b - a)n > 1. \quad (1.13)$$

Применяя утверждение Теоремы 1.2.19 для $y = a$, $x = \frac{1}{n}$, получаем, что найдется целое число k такое, что

$$k \leq an < k + 1. \quad (1.14)$$

Тогда, с учетом (1.13) получаем, что

$$bn > 1 + an \geq k + 1. \quad (1.15)$$

Из неравенств (1.14)–(1.15) вытекает, что

$$a < \frac{k+1}{n} < b.$$

Поскольку число $\frac{k+1}{n}$ рационально, теорема доказана. $\square$

§2.5 Бином Ньютона

Предполагаются известными следующие комбинаторные формулы:

Число перестановок n элементов: $n! = n(n-1)(n-2)\dots 2$.

Число сочетаний из n элементов по k :

$$C_n^k = \frac{n!}{k!(n-k)!}. \quad (1.16)$$

По определению полагают $0! = 1$.

ЛЕММА 1.2.21. Для всех $n, k \in \mathbb{N}$ справедлива формула

$$C_{n+1}^k = C_n^k + C_n^{k-1}. \quad (1.17)$$

ДОКАЗАТЕЛЬСТВО. Дадим комбинаторное доказательство этой леммы. Возьмем произвольное множество S из $(n+1)$ элемента и зафиксируем один из его элементов, обозначив его $*$. Тогда каждая выборка k элементов из S либо содержит элемент $*$, либо нет. Количество сочетаний (выборок) второго типа равно C_n^k (т.к. фактически нам нужно выбирать k элементов из n -элементного множества $S \setminus \{*\}$, а число сочетаний первого типа равно C_n^{k-1} (т.к. фактически нам нужно выбирать из n -элементного множества $S \setminus \{*\}$ лишь $(k-1)$ элемент, ведь k -й элемент у нас заведомо уже есть — это $*$). Суммируя указанные значения, получаем искомую формулу (1.17). $\square$

Упражнение 1.2.22. Дайте прямое доказательство формулы (1.17) через формулу с факториалами (1.16).

ТЕОРЕМА 1.2.23 (Бином Ньютона). Для произвольной пары чисел $a, b \in \mathbb{R}$ справедлива формула

$$(a+b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}. \quad (1.18)$$

ДОКАЗАТЕЛЬСТВО. Дадим комбинаторное доказательство этой леммы. Имеем

$$(a+b)^n = \overbrace{(a+b)(a+b)\dots(a+b)}^{n \text{ раз}}.$$

Чтобы получить при раскрытии скобок моном $a^k b^{n-k}$, нужно выбрать k скобок, в которых при раскрытии берется слагаемое a , и при раскрытии остальных скобок брать слагаемое b . Таким образом, при раскрытии скобок моном $a^k b^{n-k}$ будет встречаться C_n^k раз. $\square$

Упражнение 1.2.24. Докажите формулу бинома Ньютона индукцией по n , используя установленное ранее соотношение (1.17).

Формула Бинома Ньютона позволяет установить множество красивых соотношений на комбинаторные числа.

Упражнение 1.2.25. Установите справедливость следующих формул тремя различными способами: комбинаторным, с помощью бинома Ньютона, и индукцией по n с помощью (1.17).

- (i) $\sum_{k=0}^n C_n^k = 2^n;$
- (ii) $\sum_{k=0}^n (-1)^k C_n^k = 0;$
- (iii) $\sum_{i=0}^k C_n^i C_m^{k-i} = C_{n+m}^k.$

В качестве иллюстрации, приведем доказательство тремя способами первой формулы (i). По формуле бинома Ньютона для $a = b = 1$ имеем

$$2^n = (1 + 1)^n = \sum_{k=0}^n C_n^k,$$

что сразу приводит к искомому результату.

Не намного труднее комбинаторное доказательство той же формулы. Сумма $\sum_{k=0}^n C_n^k$ представляет собой число всех подмножеств n -элементного множества S . Всякое подмножество можно задать следующим образом: каждому элементу $x \in S$ припишем значение 1, если мы его включаем в рассматриваемое подмножество, или 0 — в противном случае. Таким образом, при задании подмножества для каждого элемента S нужно делать выбор из двух вариантов. Поскольку эти выборы независимы, и всего имеется n элементов в S , количество вариантов задания подмножеств S равно 2^n .

Наконец, пользуясь индукцией по n и предполагая формулу (i) справедливой для n , при переходе к $(n + 1)$ получаем

$$\sum_{k=0}^{n+1} C_{n+1}^k \stackrel{(1.17)}{=} \sum_{k=0}^{n+1} [C_n^k + C_n^{k-1}] = \sum_{k=0}^n C_n^k + \sum_{k=1}^{n+1} C_n^{k-1} = 2 \sum_{k=0}^n C_n^k \stackrel{\text{индукц. предпол.}}{=} 2 \cdot 2^n = 2^{n+1}.$$

§2.6 Отображения

Пусть X, Y — множества произвольной природы. *Отображением* $f : X \rightarrow Y$ из X в Y называется правило (закон), который каждому элементу $x \in X$ сопоставляет **единственный** элемент $y = f(x) \in Y$. Указанный элемент $y = f(x)$ называется *значением* отображения f в точке y , его иногда обозначают также символом $f|_x$.

Множество X называется *областью определения* отображения f , и обозначается символом $\text{Dom } f$, а множество Y называется *пространством значений* отображения f .

Пример 1.2.26. Если студенты в аудитории рассажены по стульям, то сопоставление стулу студента, на нем сидящего, является отображением (т.к. на стуле сидит по одному студенту). Множество определения этого отображения — занятые стулья, множество значений — совокупность студентов, собравшихся в аудитории.

Если же студенты рассажены по скамейкам, то сопоставление скамейке студентов, на ней сидящих, может и не быть отображением, т.к. на одной скамейке скамейке может находиться несколько студентов, так что ей соответствует не одно, а несколько значений.

Упражнение 1.2.27. Укажите, какие из соответствий в этом списке являются отображениями, а какие — нет: (i) студент $\rightarrow$ номер группы, где он учится; (ii) человек $\rightarrow$ дата его рождения; (iii) страна $\rightarrow$ граждане, ее населяющие; (iv) страна $\rightarrow$ численность ее населения; (v) команда $\rightarrow$ футболисты, ее составляющие.

Множество $f(X) = \{y \in Y : \exists x \in X y = f(x)\}$ называется *множеством значений* отображения f . Более общо, для $E \subset X$ множество $\{f(x) : x \in E\}$ называется *образом* множества E под действием отображения f .

Аналогично, для *прообразом* множества T при отображении f называется множество $f^{-1}(T) = \{x \in X : f(x) \in T\}$. Соответственно, прообразом точки $y \in Y$ называется множество $\{x \in X : f(x) = y\}$.

Отображение называется *взаимнооднозначным* (*инъективным*, или *отображением в*), если оно переводит разные точки в разные, т.е. для любой пары различных элементов $x_1, x_2 \in X$ выполнено неравенство $f(x_1) \neq f(x_2)$. Таким образом, для инъективного отображения f имеем

$$f(x_1) = f(x_2) \Leftrightarrow x_1 = x_2.$$

Отображение $f : X \rightarrow Y$ называется *сюръективным* (или *отображением на*), если $f(X) = Y$. Взаимнооднозначное и сюръективное отображение называется *биективным*.

Пример 1.2.28. Отображение $f : \mathbb{R} \rightarrow \mathbb{R}$, определяемое формулой $f(x) = x^2$, не является инъективным, т.к. $f(x) \equiv f(-x)$. Оно также не является сюръективным, т.к. $f(\mathbb{R}) = \mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Отображение $\mathbb{R} \ni x \mapsto x^3 \in \mathbb{R}$ является биективным. Отображение $\mathbb{R} \ni x \mapsto 2^x \in \mathbb{R}$ является инъективным, но не сюръективным. Отображение $(0, \infty) \ni x \mapsto \ln x \in \mathbb{R}$ является биективным. Вообще, одна и та же формула для f может порождать отображения $f : X \rightarrow Y$, принадлежащие различным классам, в зависимости от того, как выбираются X, Y :

$\mathbb{R} \ni x \mapsto \sin x \in \mathbb{R}$ — не инъективно и не сюръективно.

$[0, 2\pi] \ni x \mapsto \sin x \in [0, 1]$ — сюръективное, но не инъективное отображение;

$[-\pi/2, \pi/2] \ni x \mapsto \sin x \in \mathbb{R}$ — инъективное, но не сюръективное отображение;

$[-\pi/2, \pi/2] \ni x \mapsto \sin x \in [0, 1]$ — биективное отображение.

Отображение $f : X \rightarrow X$ называется *тождественным*, если $f(x) = x$. Тождественное отображение обозначается символом id_X . Ясно, что тождественное отображение любого множества в себя биективно.

Композиция $g \circ f : X \rightarrow Z$ отображений $f : X \rightarrow Y$ и $g : Y \rightarrow Z$ определяется по формуле $g \circ f|_x := g(f(x))$.

Можно определить композицию $g \circ f$ для произвольной пары отображений $f : X \rightarrow Y$ и $g : Y \rightarrow Z$ по той же формуле $g \circ f|_x := g(f(x))$, полагая при этом $\text{Dom } h = \{x \in X : f(x) \in Y\}$.

ЛЕММА 1.2.29. *Определенная выше композиция ассоциативна: $(h \circ g) \circ f = h \circ (g \circ f)$. Кроме того, для отображения $f : X \rightarrow Y$ справедливы равенства $f \circ \text{id}_X = \text{id}_Y \circ f = f$.*

ДОКАЗАТЕЛЬСТВО. Оставляется учащимся в качестве упражнения. □

ЛЕММА 1.2.30. *Если оба отображения $f : X \rightarrow Y$ и $g : Y \rightarrow Z$ инъективны (сюръективны, биективны), то и их композиция $g \circ f : X \rightarrow Z$ является инъективным (соотв., сюръективным, биективным) отображением.*

Пусть $f : X \rightarrow Y$ и $g : Y \rightarrow X$ — некоторые отображения. Тогда определены композиции $g \circ f : X \rightarrow X$, $f \circ g : Y \rightarrow Y$. Отображение g называется *левым обратным* к f , если $g \circ f = \text{id}_X$. В этом же случае f называется *правым обратным* к g . Наконец, $g : Y \rightarrow X$ называется *обратным* к $f : X \rightarrow Y$, если $f \circ g = \text{id}_Y$, $g \circ f = \text{id}_X$, т.е. f является одновременно и правым и левым обратным к g .

Пример 1.2.31. Бывают случаи, когда левое обратное отображение существует, а обратного — нет (т.е. нет биективности). Рассмотрим сначала $f : [-1, 1] \rightarrow \mathbb{R}$, $f(x) = \arcsin x$, и $g : \mathbb{R} \rightarrow [-1, 1]$, $g(y) = \sin y$. Тогда $g \circ f = \text{id}_{[-1, 1]}$, т.е. $\forall x \in [-1, 1] \ g(f(x)) = \sin(\arcsin x) = x$.

Упражнение 1.2.32. Постройте более тонкий пример, когда $X = Y = \mathbb{N}$, и, как и выше, отображение $f : X \rightarrow X$ имеет левое обратное $g : X \rightarrow X$, а правого обратного нет³.

ТЕОРЕМА 1.2.33. Если отображение $f : X \rightarrow Y$ является правым обратным к отображению $g : Y \rightarrow X$, то f инъективно, а g — сюръективно.

ДОКАЗАТЕЛЬСТВО. Докажем первое утверждение теоремы (об инъективности f). Предположим, что оно неверно, тогда найдется пара точек $x_1, x_2 \in X$ таких, что $x_1 \neq x_2$ и $f(x_1) = f(x_2)$. Но тогда в силу условия $g(f(x)) \equiv x$ имеем

$$x_1 = g(f(x_1)) = g(f(x_2)) = x_2,$$

противоречие. Инъективность f доказана.

Докажем теперь сюръективность g . Возьмем произвольную точку $x \in X$. Обозначим $y = f(x)$. Тогда $g(y) = g(f(x)) = x$. Итак, для каждой точки $y \in X$ нашлась точка $x \in X$ такая, что $f(x) = y$. Тем самым установлена сюръективность f . $\square$

ТЕОРЕМА 1.2.34. Для отображения $f : X \rightarrow Y$ существует обратное отображение $g : Y \rightarrow X$ в том и только том случае, когда f биективно.

ДОКАЗАТЕЛЬСТВО.

($\Rightarrow$) Необходимость⁴ непосредственно вытекает из Теоремы 1.2.33.

($\Leftarrow$) Докажем достаточность. Предположим, что отображение f биективно. Определим отображение $g : Y \rightarrow X$ следующим образом. Для произвольной точки $y \in Y$ возьмем $x \in X$, удовлетворяющую условию $f(x) = y$ (такая точка найдется вследствие сюръективности f), и положим $g(y) := x$. Указанная точка x единственна в силу инъективности f , так что отображение g определено корректно (однозначно). По построению, для всех $x \in X$ справедливо тождество $g(f(x)) = x$. С другой стороны, для произвольного $y \in Y$ найдется $x \in X$ со свойством $f(x) = y$ и по построению имеем $f(g(y)) = f(x) = y$. Теорема доказана. $\square$

Введем еще несколько понятий. *Графиком* отображения $f : X \rightarrow Y$ называется множество $\{(x, y) \in X \times Y : y = f(x)\}$.

Сужением $f|_A$ отображения $f : X \rightarrow Y$ на множество $A \subset X$ называется отображение $A \ni x \mapsto f(x) \in Y$ из A в Y , которое каждой точке $x \in A$ сопоставляет значение $f(x)$.

³**Ответ.** Положим $f(n) = n + 1$ для $n \in \mathbb{N}$, а $g : \mathbb{N} \rightarrow \mathbb{N}$ определяется следующим образом:

$$g(n) = \begin{cases} n - 1, & n > 1; \\ 1, & n = 1. \end{cases}$$

Тогда вновь композиция $g \circ f : \mathbb{N} \rightarrow \mathbb{N}$ является тождественным отображением $g(f(n)) = n \ \forall n \in \mathbb{N}$, но f, g не являются биекциями.

⁴Т.е. необходимость биективности для существования обратного отображения,

§2.7 Счетные множества

Обозначим $\mathbb{I}_k = \{1, 2, \dots, k\} = \{n \in \mathbb{N} : n \leq k\}$.

Множество M называется *конечным*, если существует $k \in \mathbb{N}$ такое, что найдется биективное отображение $f : M \rightarrow \mathbb{I}_k$. Если такое число k существует, то оно единственно, как вытекает из следующей теоремы.

ТЕОРЕМА 1.2.35. Пусть $k, m \in \mathbb{N}$ и $f : \mathbb{I}_k \rightarrow \mathbb{I}_m$ — инъективное отображение из $\mathbb{I}_k$ в $\mathbb{I}_m$. Тогда $k \leq m$.

ДОКАЗАТЕЛЬСТВО. Будет вести доказательство индукцией по m . При $m = 1$ — очевидно. Пусть утверждение верно для m , и пусть имеется инъективное отображение $f : \mathbb{I}_{k+1} \rightarrow \mathbb{I}_{m+1}$.

Рассмотрим два случая.

(i) Справедливо равенство

$$f(k+1) = m+1, \quad (1.19)$$

Тогда, в силу инъективности f имеем

$$\forall n \in \mathbb{I}_{k+1} \quad n < k+1 \Rightarrow f(n) < m+1, \quad (1.20)$$

другими словами

$$\forall n \in \mathbb{I}_k \quad f(n) \in \mathbb{I}_m. \quad (1.21)$$

Построим теперь отображение $\tilde{f}(x) : \mathbb{I}_k \rightarrow \mathbb{I}_m$, определяемое формулой $\tilde{f}(x) = f(x)$, $x \in \mathbb{I}_k$. Корректность определения вытекает из формулы (1.21), а инъективность $\tilde{f}$ мгновенно вытекает из инъективности самого f . Итак, $\tilde{f}(x) : \mathbb{I}_k \rightarrow \mathbb{I}_m$ есть инъекция, значит, по индукционному предположению, $k \leq m$.

(ii) Рассмотрим случай произвольного биективного отображения $f : \mathbb{I}_{k+1} \rightarrow \mathbb{I}_{m+1}$ и сведем его к предыдущему. Пусть $f(n_0) = m+1$ (такое n_0 существует в силу сюръективности f). Определим биективное отображение $g : \mathbb{I}_{k+1} \rightarrow \mathbb{I}_{k+1}$ по правилу

$$g(x) = \begin{cases} k+1, & x = n_0; \\ n_0, & x = k+1; \\ x, & x \notin \{n_0, k+1\} \end{cases} \quad (1.22)$$

(данное отображение меняет местами точки $k+1$ и n_0 , а все остальные оставляет неподвижными на своем месте). Положим $h = f \circ g$. Тогда h является инъективным отображением (как композиция инъекций) из $\mathbb{I}_{k+1}$ в $\mathbb{I}_{m+1}$, и, по построению, $h(k+1) = f(g(k+1)) = f(n_0) = m+1$. Таким образом, для инъекции $h : \mathbb{I}_{k+1} \rightarrow \mathbb{I}_{m+1}$ выполнено условие (1.19), и мы попадаем в ситуацию уже разобранный пункта (i), где, как мы уже доказали, выполнено неравенство $k \leq m$. Тем самым данное неравенство установлено и в общем случае, и теорема доказана. $\square$

СЛЕДСТВИЕ 1.2.36. Пусть $k, m \in \mathbb{N}$ и $f : \mathbb{I}_k \rightarrow \mathbb{I}_m$ — биективное отображение из $\mathbb{I}_k$ в $\mathbb{I}_m$. Тогда $k = m$.

СЛЕДСТВИЕ 1.2.37. Пусть M — конечное множество, и пусть $f : M \rightarrow \mathbb{I}_k$ и $g : M \rightarrow \mathbb{I}_m$ — биективные отображения. Тогда $k = m$.

Множество M называется *не более чем счетным*, если существует инъективное отображение $g : M \rightarrow \mathbb{N}$. Очевидно, что определенные выше конечные множества не более чем счетны. Наконец, множество называется *счетным*, если оно *бесконечно* (т.е. *не конечно*) и не более чем счетно.

ТЕОРЕМА 1.2.38. Пусть множество B не более чем счетно и $f : A \rightarrow B$ — инъективное отображение. Тогда множество A также не более чем счетно.

ДОКАЗАТЕЛЬСТВО. Поскольку множество B по условию теоремы не более чем счетно, то существует инъекция $g : B \rightarrow \mathbb{N}$. Тогда композиция $h = f \circ g$ будет инъективным отображением (см. Лемму 1.2.30) из A в $\mathbb{N}$, следовательно, множество A не более чем счетно. $\square$

ТЕОРЕМА 1.2.39. Пусть множество A не более чем счетно и $f : A \rightarrow B$ — сюръективное отображение. Тогда множество B также не более чем счетно.

ДОКАЗАТЕЛЬСТВО. Поскольку множество A по условию теоремы не более чем счетно, то существует инъекция $g : A \rightarrow \mathbb{N}$. Зададим отображение $h : B \rightarrow \mathbb{N}$ по следующему правилу. Для произвольной точки $y \in B$ рассмотрим прообраз $f^{-1}(y) = \{x \in A : f(x) = y\}$, этот прообраз будет непустым множеством в силу сюръективности f . Теперь положим

$$h(y) = \min g(f^{-1}(y)) = \{g(x) : x \in f^{-1}(y)\}.$$

Минимальный элемент в написанной выше формуле действительно существует, т.к. соответствующее непустое множество $\{g(x) : x \in f^{-1}(y)\} = \{g(x) : x \in A \wedge f(x) = y\}$ состоит из натуральных чисел, см. Теорему 1.2.17, так что отображение h определено корректно. Докажем его инъективность. Если $y_1, y_2 \in A$ и $y_1 \neq y_2$, то $f^{-1}(y_1) \cap f^{-1}(y_2) = \emptyset$, следовательно, вследствие инъективности g , имеем равенство $g(f^{-1}(y_1)) \cap g(f^{-1}(y_2)) = \emptyset$, значит, $h(y_1) = \min g(f^{-1}(y_1)) \neq \min g(f^{-1}(y_2))$. $\square$

Пусть A — некоторое множество, и пусть дано отображение $f : \mathbb{N} \rightarrow A$. Это означает, что каждому номеру $n \in \mathbb{N}$ сопоставлен некоторый элемент $f(n) \in A$, можно обозначить его символом x_n . В таком случае говорят, что задана *последовательность* $\{x_n\}_{n \in \mathbb{N}}$ элементов множества A .

ТЕОРЕМА 1.2.40. Пусть множество A — бесконечное множество и пусть $A \subset \mathbb{N}$. Тогда существует биективное отображение $f : \mathbb{N} \rightarrow A$.

ДОКАЗАТЕЛЬСТВО. Построить биекцию из $\mathbb{N}$ в A — это то же самое, что каждому номеру $n \in \mathbb{N}$ сопоставить некоторый элемент $x_n \in A$ так, чтобы разным номерам $n_1 \neq n_2$ соответствовали разные элементы $x_{n_1} \neq x_{n_2}$, и чтобы множество всех членов последовательности $\{x_n : n \in \mathbb{N}\}$ совпадало с множеством A .

Будем строить последовательность x_n по индукции. Положим $x_1 = \min A$. Допустим теперь, что мы определили элементы $x_1, \dots, x_n \in A$. Поскольку множество A бесконечно, то множество $A \setminus \{x_1, \dots, x_n\}$ не пусто. Положим $x_{n+1} = \min(A \setminus \{x_1, \dots, x_n\})$, минимальный элемент в этой формуле найдется в силу Теоремы 1.2.17.

Согласно принципу индукции, тем самым мы можем определить элемент x_n для всех $n \in \mathbb{N}$. Используя все тот же принцип индукции, легко установить, что построенная последовательность является строго возрастающей, т.е.

$$x_k < x_l \quad \text{при } k, l \in \mathbb{N} \text{ и } k < l, \quad (1.23)$$

и, более того,

$$\{x_1, \dots, x_n\} < A \setminus \{x_1, \dots, x_n\} \quad \forall n \in \mathbb{N}. \quad (1.24)$$

Инъективность нашей нумерации мгновенно вытекает из условия строго возрастания, так что осталось проверить только лишь условие сюръективности

$$\{x_n : n \in \mathbb{N}\} = A. \quad (1.25)$$

Предположим, что это неверно, тогда найдется число $m \in A \setminus \{x_n : n \in \mathbb{N}\} \neq \emptyset$. В силу свойства (1.24) имеем

$$x_n < m \quad \text{для всех } n \in \mathbb{N}.$$

Возьмем теперь произвольный элемент $k \in \mathbb{N}$ и зафиксируем его. Тогда отображение f , задаваемое формулой $f(n) = x_n$ для $n \in \mathbb{I}_k$, является инъективным отображением из $\mathbb{I}_k$ в $\mathbb{I}_m$. По теореме 1.2.35 имеем $k \leq m$. В силу произвольности k , неравенство $k \leq m$ справедливо для всех $k \in \mathbb{N}$, однако это противоречит принципу Архимеда (неограниченности $\mathbb{N}$). Тем самым соотношение (1.25), а вместе с ним и Теорема 1.2.40, доказаны. $\square$

СЛЕДСТВИЕ 1.2.41. Множество M является счетным тогда и только тогда, когда существует биективное отображение $f : M \rightarrow \mathbb{N}$.

§2.8 Свойства счетных множеств

Добавить одного постояльца в заполненную гостиницу. Добавить 10, добавить бесконечно много (две гостин. в одну), добавить бесконечно много гостиниц.

ТЕОРЕМА 1.2.42. Множество $\mathbb{N} \times \mathbb{N}$ — счетно.

ДОКАЗАТЕЛЬСТВО. Зададим отображение $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ по формуле

$$f(n, m) := (2m - 1)2^{n-1}.$$

Докажем сначала

ИНЪЕКТИВНОСТЬ. Пусть $f(n, m) = f(k, l)$. По теореме Евклида, всякое число x однозначно представляется в виде $x = 2^{\alpha_0} p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$, где p_i суть нечетные простые числа, а $\alpha_i \in \mathbb{N}$ суть неотрицательные целые числа. Применим эту теорему к $x = f(n, m)$, тогда получим

$$(2m - 1)2^{n-1} = 2^{\alpha_0} p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k} = (2l - 1)2^{k-1}.$$

Поскольку числа $(2m - 1)$, $(2l - 1)$ нечетные, двойка не входит в их разложение на простые множители, и, в силу однозначности указанного разложения, имеем $n - 1 = \alpha_0 = k - 1$. Следовательно, $n = k$, а отсюда уже легко вывести и второе требуемое равенство $m = l$.

СЮРЪЕКТИВНОСТЬ. Как уже было сказано, всякое число $x \in \mathbb{N}$ однозначно представляется в виде $x = 2^{\alpha_0} p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$, где p_i суть нечетные простые числа, а $\alpha_i \in \mathbb{N}$ суть неотрицательные целые числа. Тогда $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ представляет собой нечетное натуральное число, тем самым $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k} = 2m - 1$ для некоторого $m \in \mathbb{N}$. Тогда имеем равенство $x = 2^{\alpha_0} (2m - 1) = f(\alpha_0 + 1, m)$.

Итак, построенное отображение f является биекцией, и доказательство теоремы окончено. $\square$

Доказанная теорема дает нам хороший инструмент доказательства не более чем счетности тех или иных множеств. Из нее и из Теорем 1.2.38–1.2.39 немедленно вытекает

СЛЕДСТВИЕ 1.2.43. Пусть M — некоторое непустое множество. Тогда, если существует инъекция $f : M \rightarrow \mathbb{N} \times \mathbb{N}$ или сюръекция $f : \mathbb{N} \times \mathbb{N} \rightarrow M$, то множество M не более чем счетно.

Визуально это можно представить себе следующим образом. Рассмотрим декартову систему координат на плоскости, и поделим первый квадрант $\{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\}$ на ячейки вертикальными линиями $\{x = n : n \in \mathbb{N}\}$ и горизонтальными линиями $\{y = m : m \in \mathbb{N}\}$. Эти линии, вместе с координатными осями, разделяют первый квадрант на бесконечное множество ячеек, каждой ячейке можно взаимнооднозначно сопоставить пару натуральных чисел (n, m) — декартовы координаты ее верхнего правого угла. Полученная бесконечная таблица носит название *канторовой таблицы*. Тогда теорема 1.2.42 означает, что множество ячеек канторовой таблицы не более чем счетно. Следствие же 1.2.43 говорит нам о том, что если все элементы какого-то множества M можно вписать в канторову таблицу так, чтобы в одной ячейке помещалось не более одного элемента, то это множество не более чем счетное. При этом не требуется, чтобы все ячейки таблицы были заполнены, и один и тот же элемент может находиться в разных ячейках, это не препятствует применению Следствия 1.2.43. Главное — чтобы каждый элемент множества попал хотя бы в одну ячейку, и чтобы разные элементы не "теснились" в одной и той же ячейке.

Доказательство следующих двух утверждений основано на применении канторовой таблицы и оставляется учащимся в качестве упражнения.

СЛЕДСТВИЕ 1.2.44. Множество $\mathbb{Z}$ целых чисел и множество $\mathbb{Q}$ рациональных чисел — счетные.

ТЕОРЕМА 1.2.45. Пусть A — не более чем счетное множество, и каждому $\alpha \in A$ сопоставлено некоторое не более чем счетное множество E_α . Тогда объединение $\bigcup_{\alpha \in A} E_\alpha$ также представляет собой не более чем счетное множество.

Другими словами, объединение не более чем счетного семейства не более чем счетных множеств — не более чем счетно.

Упражнение 1.2.46. Представьте, что Вы оказались постояльцем десятизвездочной межгалактической гостиницы со счетным количеством номеров. В период новогодних каникул, как и следовало ожидать, все номера заняты. Помогите администратору гостиницы решить следующие задачи:

- (i) принять одного нового постояльца;
- (ii) принять десять новых постояльцев;
- (iii) принять счетное количество новых постояльцев (которых в срочном порядке переезжали из соседней гостиницы, где случилась авария).

(iv) принять всех постояльцев со счетного множества гостиниц, у каждой из которых было счетное количество заполненных номеров (в результате финансового кризиса, компания — владелец межгалактических отелей — решила произвести массовое сокращение, и ликвидировать все гостиницы, кроме одной — Вашей).

Упражнение 1.2.47. Докажите, что множество алгебраических чисел — т.е. корней полиномов с целочисленными коэффициентами — счетно.

Перечисленные выше факты могут создать впечатление, что элементы любого бесконечного множества можно пересчитать, однако это неверно, как показывает следующая теорема Кантора.

ТЕОРЕМА 1.2.48. Множество $\mathbb{R}$ несчетно, то есть не существует биекции $f : \mathbb{N} \rightarrow \mathbb{R}$.

ДОКАЗАТЕЛЬСТВО. Будем доказывать от противного. Пусть найдется биекция $f : \mathbb{N} \rightarrow \mathbb{R}$. Другими словами, пусть последовательность $(x_n)_{n \in \mathbb{N}}$ такова, что

$$\{x_n : n \in \mathbb{N}\} = \mathbb{R}. \quad (1.26)$$

Возьмем отрезок $[a_1, b_1] \subset \mathbb{R}$ такой, что $a_1 < b_1$ (т.е. отрезок этот невырожденный) и $x_1 \notin [a_1, b_1]$. Далее, по индукции легко построить последовательность невырожденных отрезков $[a_n, b_n]$, $n \in \mathbb{N}$, со следующими свойствами.

- (i) $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$, т.е. это последовательность *вложенных* отрезков;
- (ii) $x_n \notin [a_n, b_n]$.

По теореме о вложенных отрезках (см. Следствие 1.2.5), пересечение всех этих отрезков не пусто, т.е. существует точка $c \in \bigcap_{n \in \mathbb{N}} [a_n, b_n]$. Согласно предположению (1.26), найдется номер $m \in \mathbb{N}$ такой, что $c = x_m$. Однако по условию (ii)

$$c = x_m \notin [a_m, b_m] \supset \bigcap_{n \in \mathbb{N}} [a_n, b_n],$$

что противоречит выбору c . Полученное противоречие завершает доказательство теоремы. $\square$

Разумеется, точно таким же методом можно доказать, что каждый невырожденный отрезок $[a, b]$ содержит несчетное количество элементов. Поскольку множество всех рациональных чисел счетно, отсюда вытекает

ТЕОРЕМА 1.2.49. Между любыми двумя точками $a, b \in \mathbb{R}$, $a < b$, найдутся иррациональные числа, т.е. $(a, b) \setminus \mathbb{Q} \neq \emptyset$.

Объединяя это с утверждением Теоремы 1.2.20, получаем

СЛЕДСТВИЕ 1.2.50. Между любыми двумя точками $a, b \in \mathbb{R}$ с $a < b$ найдутся как рациональные, так и иррациональные числа, т.е. $(a, b) \cap \mathbb{Q} \neq \emptyset \neq (a, b) \setminus \mathbb{Q}$.

ДОКАЗАТЕЛЬСТВО. $\square$

§2.9 Расширенная числовая прямая

Расширенной числовой прямой называется множество $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, формально дополненное двумя элементами $\{-\infty, \infty\}$ (читается "бесконечность и минус бесконечность"), которые по определению обладают следующими свойствами:

$$\forall x \in \mathbb{R} \quad -\infty < x < \infty,$$

$$\forall x \in \mathbb{R} \quad -\infty + x = -\infty, \quad \infty + x = \infty$$

$$x > 0 \Rightarrow -\infty \cdot x = -\infty, \quad \infty \cdot x = \infty.$$

$$x < 0 \Rightarrow -\infty \cdot x = \infty, \quad \infty \cdot x = -\infty.$$

Тогда имеем, что для любого подмножества $A \subset \overline{\mathbb{R}}$ справедлива оценка $-\infty \leq A \leq \infty$. Другими словами, если естественным образом модифицировать определение верхних и

нижних границ $\Gamma_+(A) = \{v \in \overline{\mathbb{R}} : A \leq v\}$, $\Gamma_-(A) = \{u \in \overline{\mathbb{R}} : u \leq A\}$, то всегда справедливы включения, $\infty \in \Gamma_+(A)$ и $-\infty \in \Gamma_-(A)$, т.е. ∞ является верхней, а $-\infty$ — нижней границей любого множества $A \subset \overline{\mathbb{R}}$.

Как и прежде, *точной верхней границей* (ее еще называют *верхней гранью*) множества $M \subset \overline{\mathbb{R}}$ называется наименьшая из его верхних границ: $\sup M = \min \Gamma_+(M)$. Аналогично, *точной нижней границей* (или *нижней гранью*) множества M называется наибольшая из его нижних границ, т.е. $\inf M = \max \Gamma_-(M)$.

ТЕОРЕМА 1.2.51 (о гранях). *Всякое множество $A \subset \overline{\mathbb{R}}$ имеет верхнюю и нижнюю грани в $\overline{\mathbb{R}}$.*

ДОКАЗАТЕЛЬСТВО. Докажем существование верхней грани. Рассмотрим два возможных случая.

(I) Если множество A не ограничено сверху в $\mathbb{R}$, т.е. неравенство $A \leq v$ не выполнено ни для какого $v \in \mathbb{R}$, то, очевидно, $\Gamma_+(A)$ состоит из единственного элемента, т.е. $\Gamma_+(A) = \{\infty\}$, и тогда $\sup A = \min \Gamma_+(A) = \infty$.

(II) Рассмотрим теперь случай, когда множество A ограничено сверху в $\mathbb{R}$, т.е. $A \leq v$ для некоторого $v \in \mathbb{R}$. Он распадается на две ситуации.

(IIa) $A = \{-\infty\}$. Тогда, очевидно, $\Gamma_+(A) = \overline{\mathbb{R}}$, и $\sup A = \min \Gamma_+(A) = \min \overline{\mathbb{R}} = -\infty$.

(IIb) $A \cap \mathbb{R} \neq \emptyset$. Тогда существование верхней грани у A вытекает из утверждения Теоремы 1.2.6 (которое, напомним, было доказано для ограниченных сверху в $\mathbb{R}$ множеств).

Итак, существование верхней грани $\sup A$ полностью доказано. Существование $\inf A$ доказывается аналогично. $\square$

Теперь мы можем сформулировать принцип Архимеда (см. Теорему 1.2.15) следующим образом:

$$\sup \mathbb{N} = \infty. \quad (1.27)$$

Упражнение 1.2.52. Всегда ли $\sup A \geq \inf A$? Найти $\sup \emptyset$ и $\inf \emptyset$.

Глава 2

Теория предела

§1 Предел последовательности

Элементарной окрестностью точки $a \in \mathbb{R}$ называется интервал вида $(a - \delta, a + \delta)$, где $\delta \in \mathbb{R}$ — произвольное положительное число. Далее, элементарной окрестностью точки ∞ называется множество вида $(r, +\infty]$, где $r \in \mathbb{R}$ — произвольное число. Наконец, элементарной окрестностью точки $-\infty$ называется множество вида $[-\infty, r)$, где $r \in \mathbb{R}$.

Пусть $a \in \overline{\mathbb{R}}$. Множество $U \subset \overline{\mathbb{R}}$ называется *окрестностью* точки a (обозначается $U \in \mathcal{N}(a)$), если найдется элементарная окрестность V точки a такая, что $V \subset U$.

Свойства окрестностей.

1. Если $U_1, U_2 \in \mathcal{N}(a)$, то $U_1 \cap U_2 \in \mathcal{N}(a)$ и $U_1 \cup U_2 \in \mathcal{N}(a)$.
2. Если $U \in \mathcal{N}(a)$ и $U \subset T$, то $T \in \mathcal{N}(a)$.
3. Если $a, b \in \overline{\mathbb{R}}$ и $a \neq b$, то найдется пара окрестностей $U_1 \in \mathcal{N}(a)$ и $U_2 \in \mathcal{N}(b)$ таких, что $U_1 \cap U_2 = \emptyset$ (свойство отделимости).

Определение 2.1.1. Говорят, что $a \in \overline{\mathbb{R}}$ является *пределом последовательности вещественных чисел* $(x_n)_{n \in \mathbb{N}}$ (обозначают $\lim_{n \rightarrow \infty} x_n = a$), если для каждой окрестности $U \in \mathcal{N}(a)$ найдется число $m \in \mathbb{N}$ такое, что всех номеров $n \in \mathbb{N}$ при $n > m$ выполнено включение $x_n \in U$. Используя кванторы, мы можем переписать данное определение так:

$$\lim_{n \rightarrow \infty} x_n = a \Leftrightarrow \forall U \in \mathcal{N}(a) \exists m \in \mathbb{N} \forall n \in \mathbb{N} \cap (m, +\infty) \quad x_n \in U.$$

Иногда для сокращения записей мы будем использовать сокращенное обозначение $\lim x_n = a$.

Теперь "расшифруем" это определение в применении к двум возможным ситуациям: $a \in \mathbb{R}$ и $a \in \{-\infty, \infty\}$.

Случай $a \in \mathbb{R}$. Поскольку элементарные окрестности точки $a \in \mathbb{R}$ суть множества вида $(a - \varepsilon, a + \varepsilon)$, то

$$\lim_{n \rightarrow \infty} x_n = a \Leftrightarrow \forall \varepsilon > 0 \exists m \in \mathbb{N} \forall n \in \mathbb{N} \cap (m, +\infty) \quad |x_n - a| < \varepsilon. \quad (2.1)$$

Случай $a = \infty$. Поскольку элементарные окрестности точки ∞ суть множества вида $(r, \infty]$ имеем

$$\lim_{n \rightarrow \infty} x_n = a \Leftrightarrow \forall r > 0 \exists m \in \mathbb{N} \forall n \in \mathbb{N} \cap (m, +\infty) \quad x_n > r. \quad (2.2)$$

Случай $a = -\infty$. Поскольку элементарные окрестности точки $-\infty$ суть множества вида $(r, \infty]$, имеем

$$\lim_{n \rightarrow \infty} x_n = a \Leftrightarrow \forall r > 0 \exists m \in \mathbb{N} \forall n \in \mathbb{N} \cap (m, +\infty) \quad x_n > r. \quad (2.3)$$

ТЕОРЕМА 2.1.2. Пусть имеется последовательность вещественных чисел $(x_n)_{n \in \mathbb{N}}$, и пусть $a, b \in \overline{\mathbb{R}}$ являются ее пределами: $a = \lim_{n \rightarrow \infty} x_n$ и $b = \lim_{n \rightarrow \infty} x_n$. Тогда $a = b$.

ДОКАЗАТЕЛЬСТВО. Легко выводится из свойства отделимости, и оставляется читателю в качестве упражнения. $\square$

Рассмотрим следующие важные подклассы последовательностей. Последовательность $(x_n)_{n \in \mathbb{N}}$ называется *возрастающей*, если $\forall n \in \mathbb{N} \quad x_{n+1} \geq x_n$. Аналогично, последовательность называется *убывающей*, если $\forall n \in \mathbb{N} \quad x_{n+1} \leq x_n$. Возрастающие и убывающие последовательности называются *монотонными*.

Справедлива следующая красивая теорема.

ТЕОРЕМА 2.1.3. Всякая монотонная последовательность имеет предел.

ДОКАЗАТЕЛЬСТВО. Пусть $(x_n)_{n \in \mathbb{N}}$ — возрастающая последовательность. Обозначим $c = \sup\{x_n : n \in \mathbb{N}\}$. Имеются две возможности.

1. Случай $c = \infty$. Тогда множество $\sup\{x_n : n \in \mathbb{N}\}$ неограничено, т.е. для всякого числа $r \in \mathbb{R}$ найдется номер $m \in \mathbb{N}$ такой, что $x_m > r$. Тогда в силу монотонности $x_n \geq x_m > r$ для всех $n \geq m$. Ввиду произвольности $r \in \mathbb{R}$ выполнены условия критерия (2.2), следовательно, $\lim_{n \rightarrow \infty} x_n = \infty$.

2. Случай $c \in \mathbb{R}$. Возьмем произвольно $\varepsilon > 0$. Вследствие Леммы 1.2.8 (2) для найдется элемент последовательности x_m такой, что $c - \varepsilon < x_m \leq c$. Тогда в силу монотонности $c - \varepsilon < x_n \leq c$ для всех $n \geq m$. Ввиду произвольности $\varepsilon > 0$ выполнены условия критерия (2.1), следовательно, $\lim_{n \rightarrow \infty} x_n = c$. $\square$

Следующее утверждение используется очень часто и в математическом фольклоре носит наименование "принципа двух милиционеров (полицейских)": если два милиционера направляются в участок, и между ними находится нарушитель, то этому бедолаге, хочешь не хочешь, тоже приходится идти в участок.

ТЕОРЕМА 2.1.4. Пусть $a \in \overline{\mathbb{R}}$ и пусть имеются три последовательности вещественных чисел $(x_n)_{n \in \mathbb{N}}$, $(y_n)_{n \in \mathbb{N}}$, $(z_n)_{n \in \mathbb{N}}$. Предположим, что $\lim x_n = \lim z_n = a$ и $x_n \leq y_n \leq z_n$ для всех $n \in \mathbb{N}$. Тогда $\lim y_n = a$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Докажем ее утверждения для случая, когда $a \in \mathbb{R}$ (при $a \in \{-\infty, \infty\}$ делается по аналогии, даже немного проще). Возьмем произвольно $\varepsilon > 0$ и зафиксируем его. Согласно критерию (2.1) (в применении к последовательности $(x_n)_{n \in \mathbb{N}}$), найдется номер m_1 такой, что $a - \varepsilon < x_n < a + \varepsilon$ для всех $n > m_1$. По той же причине, найдется номер m_2 такой, что $a - \varepsilon < z_n < a + \varepsilon$ для всех $n > m_2$. Положим $m = \max\{m_1, m_2\}$. Тогда для всех $n > m$ имеем $n > m_1$ и $n > m_2$, следовательно,

$$a - \varepsilon < x_n \leq y_n \leq z_n < a + \varepsilon.$$

Ввиду произвольности $\varepsilon > 0$ выполнены все условия критерия (2.1) в применении к последовательности $(y_n)_{n \in \mathbb{N}}$, значит $\lim y_n = a$. $\square$

Следующее понятие, несмотря на угрожающе-необычное название, весьма полезно. Пусть имеется некоторое высказывание $P(n)$ о числах натурального ряда $n \in \mathbb{N}$. Будем говорить, что оно *асимптотически истинно*, если найдется номер m такой, что $P(n)$ истинно при всех $n > m$, обозначать этот факт будем символом $(P(n))_{n \rightarrow \infty}$ или $P(n)|_{n \rightarrow \infty}$.

Легко проверяется, что если оба высказывания $P(n)$ и $Q(n)$ асимптотически истинны, то и их конъюнкция $P(n) \wedge Q(n)$ асимптотически истинна.

ТЕОРЕМА 2.1.5. Пусть имеются последовательности вещественных чисел $(x_n)_{n \in \mathbb{N}}$ и $(y_n)_{n \in \mathbb{N}}$. Предположим, что $\lim x_n = a \in \overline{\mathbb{R}}$, $\lim y_n = b \in \overline{\mathbb{R}}$ и $a < b$. Тогда $(x_n < y_n)_{n \rightarrow \infty}$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим случай, когда оба предела конечны: $a, b \in \mathbb{R}$ (случай бесконечных пределов рассматривается по аналогии, с очевидными упрощениями). Поскольку по условию теоремы $a < b$, найдется $\varepsilon > 0$ такое, что $a + \varepsilon < b - \varepsilon$. Согласно критерию (2.1), найдется номер m_1 такой, что $a - \varepsilon < x_n < a + \varepsilon$ для всех $n > m_1$. По той же причине, найдется номер m_2 такой, что $b - \varepsilon < y_n < b + \varepsilon$ для всех $n > m_2$. Положим $m = \max\{m_1, m_2\}$. Тогда для всех $n > m$ имеем $n > m_1$ и $n > m_2$, следовательно

$$x_n < a + \varepsilon < b - \varepsilon < y_n.$$

Теорема доказана. □

ТЕОРЕМА 2.1.6. Пусть имеется последовательность вещественных чисел $(x_n)_{n \in \mathbb{N}}$, имеющая конечный предел $\lim x_n = a \in \mathbb{R}$. Тогда эта последовательность ограничена, т.е. найдется константа $M \in \mathbb{R}$ такая, что $|x_n| \leq M$ для всех $n \in \mathbb{N}$.

ДОКАЗАТЕЛЬСТВО. Возьмем произвольно $\varepsilon > 0$ и зафиксируем его. Согласно критерию (2.1), найдется номер m такой, что $a - \varepsilon < x_n < a + \varepsilon$ для всех $n > m$. Эти соотношения можно переписать в виде $-a - \varepsilon < -x_n < -a + \varepsilon$ для всех $n > m$. По свойству 2 абсолютной величины числа (см. параграф §2.2) имеем $|x_n| \leq \max\{|a + \varepsilon|, |a - \varepsilon|\}$ для всех $n > m$. Положим $M = \max\{|a + \varepsilon|, |a - \varepsilon|, |x_1|, |x_2|, \dots, |x_m|\}$. Поскольку любое конечное множество вещественных чисел имеет максимальный и минимальный элементы, то это определение корректно и $M \in \mathbb{R}$. Легко проверить, что для выбранного числа M справедлива искомая оценка $|x_n| \leq M$ для всех $n \in \mathbb{N}$ (**Указание.** Рассмотреть отдельно случай $n > m$ и $n \leq m$.) □

ТЕОРЕМА 2.1.7. Пусть имеются последовательности вещественных чисел $(x_n)_{n \in \mathbb{N}}$ и $(y_n)_{n \in \mathbb{N}}$, причем $\lim x_n = a \in \overline{\mathbb{R}}$ и $\lim y_n = b \in \overline{\mathbb{R}}$. Предположим, что $(x_n \leq y_n)_{n \rightarrow \infty}$. Тогда $a \leq b$.

ДОКАЗАТЕЛЬСТВО. Предположим противное, т.е. что $a > b$. Тогда по предыдущей Теореме 2.1.5 имеем $(x_n > y_n)_{n \rightarrow \infty}$, т.е. найдется $m \in \mathbb{N}$ такое, что $x_n > y_n$ для всех $n > m$. Но это противоречит условию доказываемой теоремы. □

Следующая теорема открывает цикл утверждений об арифметических операциях над сходящимися последовательностями.

ТЕОРЕМА 2.1.8. Пусть имеются последовательности вещественных чисел $(x_n)_{n \in \mathbb{N}}$ и $(y_n)_{n \in \mathbb{N}}$. Предположим, что $\lim x_n = a \in \overline{\mathbb{R}}$ и $\lim y_n = b \in \overline{\mathbb{R}}$, причем $\{a, b\} \neq \{-\infty, \infty\}$. Обозначим $z_n := x_n + y_n$. Тогда $\lim z_n = a + b$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим случай, когда оба предела конечны: $a, b \in \mathbb{R}$ (случай бесконечных пределов рассматривается по аналогии, с очевидными упрощениями). Возьмем произвольно $\varepsilon > 0$ и зафиксируем его. Согласно критерию (2.1) (примененным для значения параметра малости $\frac{\varepsilon}{2}$), найдется номер m_1 такой, что $a - \frac{\varepsilon}{2} < x_n < a + \frac{\varepsilon}{2}$ для всех $n > m_1$. По той же причине, найдется номер m_2 такой, что $a - \varepsilon < y_n < a + \varepsilon$ для всех $n > m_2$. Положим $m = \max\{m_1, m_2\}$. Тогда для всех $n > m$ имеем $n > m_1$ и $n > m_2$, следовательно, складывая выписанные выше неравенства, получаем

$$a + b - \varepsilon < x_n + y_n = z_n < a + b + \varepsilon.$$

Ввиду произвольности $\varepsilon > 0$ выполнены все условия критерия (2.1) в применении к последовательности $(z_n)_{n \in \mathbb{N}}$, значит $\lim z_n = a + b$. $\square$

Замечание 2.1.9. В случае, когда $\lim x_n = \infty$ и $\lim y_n = -\infty$, мы не можем сделать заключения про $\lim(x_n + y_n)$. Этот предел может в общем случае равняться любому наперед заданному числу, или вовсе не существовать (рекомендуется подобрать соответствующие примеры).

§2 Ряды

§2.1 Признаки сходимости числовых рядов

ТЕОРЕМА 2.2.1 (Необходимый признак сходимости ряда). Если ряд $\sum_{n=1}^{\infty} a_n$ сходится, то $\lim_{n \rightarrow \infty} a_n = 0$.

ДОКАЗАТЕЛЬСТВО. Обозначим через S_n частичную сумму $S_n = a_1 + \dots + a_n$. По определению, сходимость ряда $\sum_{n=1}^{\infty} a_n$ означает, что последовательность S_n имеет конечный предел. Тогда последовательность со "сдвинутой" нумерацией S_{n-1} имеет тот же предел: $\lim_{n \rightarrow \infty} S_n = \lim_{k \rightarrow \infty} S_{n-1}$. Следовательно, $0 = \lim_{n \rightarrow \infty} (S_n - S_{n-1}) = \lim_{n \rightarrow \infty} a_n$. $\square$

ТЕОРЕМА 2.2.2 (Критерий Коши). Ряд $\sum_{n=1}^{\infty} a_n$ сходится тогда и только тогда, когда для любого $\varepsilon > 0$ найдется номер $m = m(\varepsilon)$ такой, что для всякой пары натуральных чисел $k, l > m$ при $k \leq l$ имеем $|a_k + a_{k+1} + \dots + a_l| < \varepsilon$.

ДОКАЗАТЕЛЬСТВО. Обозначим через S_n частичную сумму $S_n = a_1 + \dots + a_n$. По определению, сходимость ряда $\sum_{n=1}^{\infty} a_n$ означает, что последовательность S_n имеет конечный предел. Но, согласно критерию Коши, сходимость последовательности S_n равносильна тому, что для любого $\varepsilon > 0$ найдется номер $m = m(\varepsilon)$ такой, что для всякой пары натуральных чисел $k, l > m$ при $k \leq l$ имеем $|S_k - S_{l-1}| = |a_k + a_{k+1} + \dots + a_l| < \varepsilon$. $\square$

ТЕОРЕМА 2.2.3. Если $a_n \geq 0$ для всех $n \in \mathbb{N}$ и частичные суммы ряда $\sum_{n=1}^{\infty} a_n$ ограничены (т.е. существует число $M > 0$ неотрицательны, такое, что $a_1 + \dots + a_n \leq M$ для всех $n \in \mathbb{N}$), то ряд $\sum_{n=1}^{\infty} a_n$ сходится.

ДОКАЗАТЕЛЬСТВО. Обозначим через S_n частичную сумму $S_n = a_1 + \dots + a_n$. Поскольку $a_n \geq 0$, то последовательность S_n возрастает. Далее, по условию теоремы имеем $S_n \leq M$ для всех $n \in \mathbb{N}$. Значит, S_n является монотонной ограниченной последовательностью, и по теореме Вейерштрасса эта последовательность имеет конечный предел. $\square$

ТЕОРЕМА 2.2.4 (Признак Вейерштрасса мажорируемой сходимости). Пусть ряд $\sum_{n=1}^{\infty} a_n$ сходится, а для членов ряда $\sum_{n=1}^{\infty} b_n$ справедлива оценка $|b_n| \leq a_n$ при всех $n \in \mathbb{N}$. Тогда ряд $\sum_{n=1}^{\infty} b_n$ тоже сходится.

ДОКАЗАТЕЛЬСТВО. Из условий теоремы вытекает, что $a_n \geq 0$. Согласно предыдущей теореме (критерию Коши) для любого $\varepsilon > 0$ найдется номер $m = m(\varepsilon)$ такой, что для всякой пары натуральных чисел $k, l > m$ при $k \leq l$ имеем $|a_k + a_{k+1} + \dots + a_l| = a_k + a_{k+1} + \dots + a_l < \varepsilon$. Тогда

$$|b_k + \dots + b_l| \leq |b_k| + \dots + |b_l| \leq a_k + \dots + a_l < \varepsilon.$$

Поскольку последнее неравенство выполняется для всякой пары натуральных чисел $k, l > m$ при $k \leq l$, налицо выполнение условий критерия Коши для ряда $\sum_{n=1}^{\infty} b_n$, следовательно, он тоже сходится. $\square$

Из признака мажорируемой сходимости вытекает, в частности, что ряд $\sum_{n=1}^{\infty} a_n$ заведомо сходится, если ряд из модулей $\sum_{n=1}^{\infty} |a_n|$ сходится (обратное, конечно, неверно, см. упражнение 2.2.15). В последнем случае говорят, что имеет место *абсолютная сходимость* ряда $\sum_{n=1}^{\infty} a_n$. Если же ряд $\sum_{n=1}^{\infty} a_n$ сходится, но не абсолютно (т.е. если ряд $\sum_{n=1}^{\infty} |a_n|$ расходится), то говорят, что ряд $\sum_{n=1}^{\infty} a_n$ является *условно сходящимся*.

ТЕОРЕМА 2.2.5 (Признак Коши). Ряд $\sum_{n=1}^{\infty} a_n$ сходится, если

$$\overline{\lim}_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} < 1; \quad (2.4)$$

Ряд $\sum_{n=1}^{\infty} a_n$ расходится, если

$$\overline{\lim}_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} > 1. \quad (2.5)$$

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения (2.4) теоремы. Возьмем произвольно число q из интервала $\overline{\lim}_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} < q < 1$ и зафиксируем его. По определению верхнего предела, найдется номер $m \in \mathbb{N}$ такой, что $|a_n|^{\frac{1}{n}} < q$ для всех $n \geq m$. Но тогда $|a_n| < q^n$. Напомним, что ряд $\sum_{n=1}^{\infty} q^n$ сходится (геометрическая прогрессия с множителем меньше единицы). Поэтому, в силу признака мажорируемой сходимости, ряд $\sum_{n=1}^{\infty} a_n$ также сходится.

Если же выполнено неравенство (2.5), то по определению верхнего предела найдется подпоследовательность a_{k_l} такая, что $\lim_{l \rightarrow \infty} |a_{k_l}|^{\frac{1}{k_l}} > 1$. Следовательно, $a_{k_l} > 1$ при достаточно больших l . В частности, a_{k_l} не стремится к нулю при $l \rightarrow \infty$, и по необходимому признаку (2.2.1) ряд $\sum_{n=1}^{\infty} a_n$ расходится. $\square$

ТЕОРЕМА 2.2.6 (Модифицированный признак мажорируемой сходимости). Ряд $\sum_{n=1}^{\infty} a_n$ сходится, если для него найдется сходящийся ряд с положительными слагаемыми $\sum_{n=1}^{\infty} b_n$ такой, что

$$\frac{|a_{n+1}|}{|a_n|} \leq \frac{b_{n+1}}{b_n} \Big|_{n \rightarrow \infty}. \quad (2.6)$$

ДОКАЗАТЕЛЬСТВО. Из условий теоремы вытекает, что найдется номер $m \in \mathbb{N}$ такой, что

$$\frac{|a_{n+1}|}{|a_n|} \leq \frac{b_{n+1}}{b_n} \quad \forall n \geq m. \quad (2.7)$$

Докажем по индукции, что для всех $n \geq m$ справедливо неравенство

$$|a_n| \leq C b_n, \quad (2.8)$$

где константа C , не зависящая от n , задается формулой $C = \frac{a_m}{b_m}$. При $n = m$ неравенство (2.8) очевидно (выполняется автоматически). Допустим, что оно верно для некоторого $n \geq m$ (т.е. посмотрим, что происходит при переходе от n к $n + 1$). Имеем

$$|a_{n+1}| = |a_n| \frac{|a_{n+1}|}{|a_n|} \stackrel{(2.7)}{\leq} |a_n| \frac{|b_{n+1}|}{|b_n|} \stackrel{\text{индукц. предп. (2.8)}}{\leq} C |b_n| \frac{|b_{n+1}|}{|b_n|} = C b_{n+1}.$$

Итак, если неравенство (2.8) верно для n , то оно верно и для $n + 1$. Согласно принципу математической индукции, неравенство (2.8) установлено для всех $n \geq m$, откуда в силу мажорируемого признака следует сходимость ряда $\sum_{n=1}^{\infty} a_n$. $\square$

СЛЕДСТВИЕ 2.2.7 (Признак расходимости). Ряд $\sum_{n=1}^{\infty} a_n$ расходится, если для него найдется расходящийся ряд с положительными слагаемыми $\sum_{n=1}^{\infty} b_n$ такой, что

$$\frac{a_{n+1}}{a_n} \geq \frac{b_{n+1}}{b_n} \Big|_{n \rightarrow \infty}. \quad (2.9)$$

ДОКАЗАТЕЛЬСТВО. По аналогии с предыдущим доказательством легко выводится оценка

$$a_n \geq C b_n \Big|_{n \rightarrow \infty}. \quad (2.10)$$

$\square$

СЛЕДСТВИЕ 2.2.8 (Признак Даламбера). Ряд $\sum_{n=1}^{\infty} a_n$ сходится, если

$$\overline{\lim}_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} < 1; \quad (2.11)$$

Ряд $\sum_{n=1}^{\infty} a_n$ расходится, если

$$\underline{\lim}_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} > 1. \quad (2.12)$$

ДОКАЗАТЕЛЬСТВО. Нужно воспользоваться предыдущими двумя результатами, причем сравнение проводить с геометрической прогрессией. $\square$

ТЕОРЕМА 2.2.9 (Телескопический признак). Ряд $\sum_{n=1}^{\infty} a_n$ с неотрицательными убывающими слагаемыми $a_n \geq a_{n+1} \geq 0$ сходится тогда и только тогда, когда сходится ряд $\sum_{n=1}^{\infty} 2^n a_{2^n}$.

ДОКАЗАТЕЛЬСТВО. Имеем следующую последовательность неравенств

$$2a_2 \geq a_2 + a_3,$$

$$4a_4 \geq a_4 + a_5 + a_6 + a_7,$$

$$8a_8 \geq a_8 + a_9 + \dots + a_{15},$$

...

$$2^n a_{2^n} \geq a_{2^n} + a_{2^n+1} + \dots + a_{2^{n+1}-1},$$

...

Поэтому $\sum_{n=1}^{\infty} 2^n a_{2^n} \geq \sum_{n=2}^{\infty} a_n$. Таким образом, если исходный ряд $\sum_{n=1}^{\infty} a_n$ расходится, то и "телескопический" ряд также расходится.

Докажем теперь в другую сторону: предположим, что телескопический ряд $\sum_{n=1}^{\infty} 2^n a_{2^n}$ сходится. Имеем следующую последовательность неравенств

$$a_2 \leq a_1,$$

$$2a_4 \leq a_2 + a_3$$

$$4a_8 \leq a_4 + a_5 + a_6 + a_7,$$

...

$$2^{n-1} a_{2^n} \leq a_{2^{n-1}} + a_{2^{n-1}+1} + \dots + a_{2^n-1},$$

...

Следовательно, $\frac{1}{2} \sum_{n=1}^{\infty} 2^n a_{2^n} \leq \sum_{n=1}^{\infty} a_n$. Таким образом, если исходный ряд $\sum_{n=1}^{\infty} a_n$ сходится, то и "телескопический" ряд также сходится. $\square$

СЛЕДСТВИЕ 2.2.10 (Гармонический ряд). Пусть $p \in \mathbb{R}$. Тогда ряд $\sum_{n=1}^{\infty} \frac{1}{n^p}$ сходится при $p > 1$ и расходится при $p \leq 1$.

ДОКАЗАТЕЛЬСТВО. Если $p \leq 0$, то расходимость очевидна. Если же $p \geq 0$, то слагаемые $a_n = \frac{1}{n^p}$ образуют убывающую последовательность, причем $2^n a_{2^n} = 2^n \frac{1}{2^{np}} = 2^{n(1-p)} = (2^{1-p})^n$. Очевидно, что геометрическая прогрессия $\sum_{n=1}^{\infty} q^n$ с $q = 2^{1-p}$ сходится тогда и только тогда, когда $1 - p < 0$, т.е. $p > 1$. $\square$

СЛЕДСТВИЕ 2.2.11 ("Подправленный" гармонический ряд). Пусть $p \in \mathbb{R}$. Тогда ряд $\sum_{n=1}^{\infty} \frac{1}{n \ln^p n}$ сходится при $p > 1$ и расходится при $p \leq 1$.

ДОКАЗАТЕЛЬСТВО. Обозначим $a_n = \frac{1}{n \ln^p n}$. Если $p \leq 0$, то $a_n \geq \frac{1}{n}$ и расходимость очевидна (см. предыдущее утверждение). Если же $p \geq 0$, то слагаемые $a_n = \frac{1}{n \ln^p n}$ образуют убывающую (начиная с некоторого номера) последовательность, причем $2^n a_{2^n} = 2^n \frac{1}{2^n \ln^p(2^n)} = \frac{1}{(\ln 2)^p} = (\ln 2)^p \frac{1}{n^p}$. Тем самым вопрос о сходимости искомого ряда $\sum_{n=1}^{\infty} \frac{1}{n \ln^p n}$ сведен к вопросу о сходимости ряда $\sum_{n=1}^{\infty} \frac{1}{n^p}$ из предыдущего утверждения (см. Следствие 2.2.10), а про него уже известно, что он сходится при $p > 1$. $\square$

Теперь применим Теорему 2.2.6 (модифицированный признак мажорируемой сходимости) и Следствие 2.2.7 к рядам из последних двух следствий.

ТЕОРЕМА 2.2.12 (Признак Раабе). Ряд $\sum_{n=1}^{\infty} a_n$ сходится, если найдется число $p > 1$, для которого асимптотически справедливо неравенство

$$\frac{|a_n|}{|a_{n-1}|} \leq 1 - \frac{p}{n} \Big|_{n \rightarrow \infty}. \quad (2.13)$$

Ряд $\sum_{n=1}^{\infty} a_n$ с положительными слагаемыми $a_n > 0$ расходится, если

$$\frac{|a_n|}{|a_{n-1}|} \geq 1 - \frac{1}{n} \Big|_{n \rightarrow \infty}. \quad (2.14)$$

ДОКАЗАТЕЛЬСТВО. Пусть выполнено первое предположение доказываемой теоремы с некоторым $p > 1$. Обозначим $b_n = \frac{1}{n^p}$. В силу Следствия 2.2.10 ряд $\sum_{n=1}^{\infty} b_n$ сходится. Вследствие неравенства Бернулли (см. Теорему 1.2.14) имеем

$$\frac{b_n}{b_{n-1}} = \frac{(n-1)^p}{n^p} = \left(1 - \frac{1}{n}\right)^p \geq 1 - \frac{p}{n}.$$

Отсюда по предположению (2.13) доказываемой теоремы получаем, что $\frac{|a_n|}{|a_{n-1}|} \leq \frac{|b_n|}{|b_{n-1}|} \Big|_{n \rightarrow \infty}$. Тогда по Теореме 2.2.6 ряд $\sum_{n=1}^{\infty} a_n$ тоже сходится.

Пусть теперь выполнено предположение (2.14). Обозначим $b_n = \frac{1}{n}$. В силу Следствия 2.2.10 ряд $\sum_{n=1}^{\infty} b_n$ расходится. Имеем $\frac{b_n}{b_{n-1}} = 1 - \frac{1}{n}$. Отсюда по предположению (2.14) доказываемой теоремы получаем, что $\frac{|a_n|}{|a_{n-1}|} \geq \frac{|b_n|}{|b_{n-1}|} |_{n \rightarrow \infty}$. Тогда по Следствию 2.2.10 ряд $\sum_{n=1}^{\infty} a_n$ тоже расходится. $\square$

Если теперь взять для сравнения ряд $\sum_{n=1}^{\infty} b_n$ со слагаемыми $b_n = \frac{1}{n \ln^p n}$, то получим следующий признак сходимости.

ТЕОРЕМА 2.2.13 (Признак Бертрана). Ряд $\sum_{n=1}^{\infty} a_n$ сходится, если найдется число $p > 1$, для которого асимптотически справедливо неравенство

$$\frac{|a_n|}{|a_{n-1}|} \leq 1 - \frac{1}{n} - \frac{p}{n \ln n} |_{n \rightarrow \infty}. \quad (2.15)$$

Ряд $\sum_{n=1}^{\infty} a_n$ с положительными слагаемыми $a_n > 0$ расходится, если

$$\frac{|a_n|}{|a_{n-1}|} \geq 1 - \frac{1}{n} - \frac{1}{n \ln n} |_{n \rightarrow \infty}. \quad (2.16)$$

Доказательство этого признака будет отложено до раздела §4.1 (см. там Следствие 3.4.15). Теперь коснемся вопроса об условной сходимости.

ТЕОРЕМА 2.2.14 (Признак Лейбница). Предположим, что неотрицательные числа b_n , $n \in \mathbb{N}$, образуют невозрастающую последовательность, причем $b_n \rightarrow 0$ при $n \rightarrow \infty$. Тогда ряд $\sum_{n=1}^{\infty} (-1)^n b_n$ сходится.

ДОКАЗАТЕЛЬСТВО. Обозначим $S_n = -b_1 + b_2 - b_3 + \dots + (-1)^n b_n$. Тогда, как нетрудно видеть, S_{2k+1} образуют возрастающую последовательность, а S_{2k} — убывающую. В самом деле, $S_{2(k+1)+1} = S_{2k+3} = S_{2k+1} + b_{2k+2} - b_{2k+3} \geq S_{2k+1}$, в то же время $S_{2(k+1)} = S_{2k} - b_{2k+1} + b_{2k+2} \leq S_{2k}$. Отсюда и ввиду очевидного неравенства $S_{2k+1} = S_{2k} - b_{2k+1} \leq S_{2k}$ заключаем, что отрезки $[S_{2k+1}, S_{2k}]$ образуют убывающую последовательность: $[S_{2k+1}, S_{2k}] \supset [S_{2k+3}, S_{2k+2}]$. Поскольку их длина стремится к нулю, заключаем, что существует единственная точка $s \in \bigcap_{k \in \mathbb{N}} [S_{2k+1}, S_{2k}]$. Ясно, что $S_{2k+1} \rightarrow s$ и $S_{2k} \rightarrow s$ при $k \rightarrow \infty$. Тем самым Теорема доказана. $\square$

Упражнение 2.2.15. Приведите пример условно сходящегося ряда (т.е. сам ряд $\sum_{n=1}^{\infty} a_n$ сходится, а ряд из абсолютных величин $\sum_{n=1}^{\infty} a_n$ расходится). **Ответ** см. в сноске¹.

¹Таким рядом является, например, $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$. Более общо, если $p > 0$, то ряд $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n^p}$ сходится (причем при $p > 1$ сходимость эта абсолютная, а при $0 < p \leq 1$ — условная).

Упражнение 2.2.16. Доказать, что если ряд $\sum_{n=1}^{\infty} a_n$ знакопостоянный и $a_n \neq -1 \forall n$, то его сходимость равносильна сходимости ряда $\sum_{n=1}^{\infty} \frac{a_n}{1+a_n}$. Если ряд $\sum_{n=1}^{\infty} a_n$ не знакопостоянный, то возможны ситуации, когда один из этих рядов сходится, а другой — нет (привести соответствующие примеры).

ТЕОРЕМА 2.2.17 (Теорема Мертенса). Пусть ряд $\sum_{j=0}^{\infty} b_j$ сходится, а ряд $\sum_{i=0}^{\infty} a_i$ сходится абсолютно. Положим $\gamma_k = \sum_{i+j=k} a_i b_j = \sum_{i=0}^k a_i b_{k-i}$. Тогда ряд $\sum_{k=0}^{\infty} \gamma_k$ сходится, причем

$$\left(\sum_{i=0}^{\infty} a_i \right) \cdot \left(\sum_{j=0}^{\infty} b_j \right) = \sum_{k=0}^{\infty} \gamma_k. \quad (2.17)$$

ДОКАЗАТЕЛЬСТВО. Перед основными выкладками доказательства сделаем некоторые приготовления. Пусть выполнены предположения теоремы о сходимости рядов $\sum_{j=0}^{\infty} b_j$ и $\sum_{i=0}^{\infty} |a_i|$. Тогда их частичные суммы ограничены, т.е. найдется константа рядов $M > 0$ такая, что

$$\forall k, l \in \mathbb{N} \quad \sum_{i=k}^l |a_i| \leq M, \quad \left| \sum_{j=k}^l b_j \right| \leq M. \quad (2.18)$$

Более того, по критерию Коши имеем

$$\forall \varepsilon > 0 \exists m \in \mathbb{N} \forall k, l \in \mathbb{N}_m \quad \sum_{i=k}^l |a_i| \leq \varepsilon, \quad \left| \sum_{j=k}^l b_j \right| \leq \varepsilon. \quad (2.19)$$

Перейдем теперь к доказательству теоремы. Очевидно, для проверки (2.17) достаточно будет показать, что

$$s_n = \sum_{i=0}^n a_i \sum_{j=0}^{\infty} b_j - \sum_{k=0}^n \gamma_k \rightarrow 0 \quad \text{при } n \rightarrow \infty. \quad (2.20)$$

Легко видеть, что $\sum_{k=0}^n \gamma_k = \sum_{i,j \in \mathbb{Z}_+, i+j \leq n} a_i b_j = \sum_{i=0}^n a_i \sum_{j=0}^{n-i} b_j$. Поэтому

$$|s_n| = \left| \sum_{i=0}^n a_i \sum_{j=n-i+1}^{\infty} b_j \right| \leq \sum_{i=0}^n |a_i| \cdot \left| \sum_{j=n-i+1}^{\infty} b_j \right| = \sum_{i=0}^{[n/2]} |a_i| \cdot \left| \sum_{j=n-i+1}^{\infty} b_j \right| + \sum_{i=[n/2]+1}^n |a_i| \cdot \left| \sum_{j=n-i+1}^{\infty} b_j \right| := I_1 + I_2. \quad (2.21)$$

Зафиксируем произвольно $\varepsilon > 0$, возьмем соответствующее m из формулы (2.19) и выберем число n столь большим, что $[n/2] > m$ (напомним, что $[n/2]$ означает целую часть числа $n/2$). Так как при $i \leq [n/2]$ выполнены неравенства $n - i + 1 > [n/2] > m$, для выписанных частичных сумм I_1, I_2 имеем оценки

$$I_1 \stackrel{(2.19)}{\leq} \sum_{i=0}^{[n/2]} |a_i| \varepsilon \stackrel{(2.18)}{\leq} M \varepsilon,$$

$$I_2 \stackrel{(2.18)}{\leq} \sum_{i=[n/2]+1}^n |a_i| M \stackrel{(2.19)}{\leq} \varepsilon M,$$

следовательно, $|s_n| \leq 2M\varepsilon$ для всех $n > 2m + 2$. В силу произвольности $\varepsilon > 0$ и того факта, что M есть фиксированная (не зависящая от M) константа получаем, что $s_n \rightarrow 0$ при $n \rightarrow \infty$, что и требовалось доказать. $\square$

Зададим функцию $\exp : \mathbb{R} \rightarrow \mathbb{R}$ по формуле $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Сразу же возникает вопрос о сходимости этого ряда, т.е. о корректности определения функции $\exp$. Ответом на этот вопрос служит следующая теорема.

ТЕОРЕМА 2.2.18 (Об экспоненциальной функции). Для каждого $x \in \mathbb{R}$ ряд из определения $\exp(x)$ сходится абсолютно. Более того, для произвольной пары чисел $x, y \in \mathbb{R}$ имеет место тождество

$$\exp(x) \cdot \exp(y) = \exp(x + y). \quad (2.22)$$

ДОКАЗАТЕЛЬСТВО. Возьмем и зафиксируем произвольно $x \in \mathbb{R}$ и обозначим $a_n = \frac{x^n}{n!}$. Имеем $\frac{a_n}{a_{n-1}} = \frac{x^n(n-1)!}{x^{n-1}n!} = \frac{x}{n} \rightarrow 0$ при $n \rightarrow \infty$. Следовательно, по признаку Даламбера ряд $\sum_{n=0}^{\infty} a_n$ сходится абсолютно. Первое утверждение теоремы доказано.

Возьмем теперь произвольно пару чисел $x, y \in \mathbb{R}$ и обозначим $a_n = \frac{x^n}{n!}$, $b_m = \frac{y^m}{m!}$, $\gamma_k = \sum_{n=0}^k a_n b_{k-n}$. Тогда непосредственные вычисления дают нам, что

$$\gamma_k = \sum_{n=0}^k \frac{x^n y^{k-n}}{n!(k-n)!} = \frac{1}{k!} \sum_{n=0}^k \frac{k!}{n!(k-n)!} x^n y^{k-n} \stackrel{\text{бином Ньютона (1.18)}}{=} \frac{1}{k!} (x + y)^k.$$

Тогда по Теореме 2.2.17 получаем

$$\exp(x) \cdot \exp(y) = \left(\sum_{i=0}^{\infty} a_i \right) \cdot \left(\sum_{j=0}^{\infty} b_j \right) = \sum_{k=0}^{\infty} \gamma_k = \sum_{k=0}^{\infty} \frac{1}{k!} (x + y)^k = \exp(x + y),$$

что и требовалось доказать. $\square$

СЛЕДСТВИЕ 2.2.19 (Свойства экспоненциальной функции).

- (i) $\exp(0) = 1$, $\exp(x) > 1 \quad \forall x > 0$;
- (ii) $\exp(x) > 0 \quad \forall x \in \mathbb{R}$;
- (iii) $\exp(x) < \exp(y)$, если $x < y$.
- (iv) Остаток экспоненциального ряда:

$$0 < 2|x| < n + 2 \Rightarrow 0 < \left| \exp(x) - 1 - x - \frac{x^2}{2} - \dots - \frac{x^n}{n!} \right| < \frac{|x|^{n+1}}{(n+1)!} \frac{(n+2)}{n+2-|x|}.$$

- (v) Число $e = \exp 1$ иррационально.

ДОКАЗАТЕЛЬСТВО. В доказательстве, очевидно, нуждаются только свойства (iv)–(v). Остаток экспоненциального ряда легко оценивается с помощью геометрической прогрессии. Зафиксируем точку $x \in \mathbb{R}$. Пусть число $n \in \mathbb{N}$ достаточно велико — так, чтобы $\frac{|x|}{n+2} < 1$. Тогда

$$\left| \exp(x) - 1 - x - \frac{x^2}{2} - \dots - \frac{x^n}{n!} \right| = \left| \frac{x^{n+1}}{(n+1)!} + \frac{x^{n+2}}{(n+2)!} + \frac{x^{n+3}}{(n+3)!} + \dots \right| =$$

$$\begin{aligned}
 &= \left| \frac{x^{n+1}}{(n+1)!} \left(1 + \frac{x}{(n+2)} + \frac{x^2}{(n+2)^2} + \frac{x^3}{(n+2)^3} + \dots \right) \right| \leq \frac{|x|^{n+1}}{(n+1)!} \left(1 + \frac{|x|}{(n+2)} + \frac{|x|^2}{(n+2)^2} + \frac{|x|^3}{(n+2)^3} + \dots \right) = \\
 &= \frac{|x|^{n+1}}{(n+1)!} \frac{1}{1 - \frac{|x|}{(n+2)}} = \frac{|x|^{n+1}}{(n+1)!} \frac{(n+2)}{n+2-|x|}.
 \end{aligned}$$

Из тех же оценок вытекает, что $|\exp(x) - 1 - x - \frac{x^2}{2} - \dots - \frac{x^n}{n!}| > 0$ при $0 < \frac{|x|}{n+2} < \frac{1}{2}$. В частности, при $x = 1$, обозначив $e = \exp 1$, имеем

$$0 < e - 1 - 1 - \frac{1}{2} - \dots - \frac{1}{n!} < \frac{1}{(n+1)!} \frac{(n+2)}{n+1} < \frac{1}{n \cdot n!} \quad (2.23)$$

для всех $n \in \mathbb{N}$ (последнее неравенство вытекает из того факта, что $n(n+2) < (n+1)^2$). Докажем теперь, что число e иррационально. Предположим, что это не так, тогда $e = \exp 1 = \frac{p}{q}$, где $p, q \in \mathbb{N}$, и $\frac{p}{q}$ — несократимая дробь. Тогда при $n > q$ число $e - 1 - 1 - \frac{1}{2} - \dots - \frac{1}{n!}$ будет представлять собой положительное рациональное число вида $\frac{M}{n!}$, здесь $M \in \mathbb{N}$ (т.к. общим знаменателем множества чисел $\frac{p}{q}, \frac{1}{2}, \frac{1}{3!}, \dots, \frac{1}{n!}$, очевидно, является $n!$). Поэтому $e - 1 - 1 - \frac{1}{2} - \dots - \frac{1}{n!} \geq \frac{1}{n!}$. Однако это неравенство противоречит (2.23). $\square$

§2.2 Степенные ряды, радиус сходимости, асимптотика

Пусть дана последовательность чисел $a_n \in \mathbb{R}$. Функциональный ряд $\sum_{n=0}^{\infty} a_n x^n$, x — независимая вещественная переменная, называется *степенным рядом*. *Радиус сходимости* R степенного ряда определяется по формуле $\frac{1}{R} = \overline{\lim}_{n \rightarrow \infty} |a_n|^{\frac{1}{n}}$. Возможны ситуации, когда $R = 0$ (если выписанный предел бесконечен) или $R = \infty$ (если указанный предел равен нулю). Ниже мы будем предполагать, что $R > 0$.

ТЕОРЕМА 2.2.20. Ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится абсолютно при $|x| < R$ и расходится при $|x| > R$.

ДОКАЗАТЕЛЬСТВО. Используя элементарные свойства предела, имеем

$$\overline{\lim}_{n \rightarrow \infty} |a_n x^n|^{\frac{1}{n}} = \frac{|x|}{R} < 1, \quad (2.24)$$

поэтому, согласно признаку Коши 2.2.5 ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится (при $|x| < R$).

Расходимость того же ряда при $|x| > R$ доказывается аналогично применением все того же признака Коши (см. (2.5)) и оставляется в качестве упражнения. $\square$

Для $|x| < R$ обозначим $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Далее, для $m \in \mathbb{N}$ рассмотрим частичную сумму $f_m(x) = \sum_{n=0}^m a_n x^n$. Погрешность приближения функции $f(x)$ полиномом $f_m(x)$ оценивается с помощью следующего утверждения.

ТЕОРЕМА 2.2.21. Пусть $R > 0$ и $\delta \in (0, R/2)$. Тогда найдется число $m \in \mathbb{N}$ такое, что для всякого $x \in \mathbb{R}$ при $|x| < R - 2\delta$ справедлива оценка

$$|f(x) - f_m(x)| \leq \sum_{n=m+1}^{\infty} |a_n x^n| < |x|^{m+1} \frac{1}{\delta(R - \delta)^m}. \quad (2.25)$$

ДОКАЗАТЕЛЬСТВО. Из (2.24) вытекает, что найдется $m \in \mathbb{N}$ такое, что для всякого $n \geq m$ справедливо неравенство $|a_n x^n| \leq \frac{|x|^n}{(R-\delta)^n}$. Далее в предположении $|x| < R - 2\delta$ нужно воспользоваться формулой бесконечной геометрической прогрессии (с множителем $q = \frac{|x|}{R-\delta}$). Подробности предоставляются читателю в качестве упражнения. $\square$

Замечание 2.2.22. Из (2.25) вытекают следующие асимптотические соотношения

$$|f(x) - f_m(x)| = O(x^{m+1}) = o(x^m) \quad (2.26)$$

при $x \rightarrow 0$ (об употреблении символов $O(x^{m+1})$, $o(x^m)$ см. параграф §3.3).

ТЕОРЕМА 2.2.23. Имеют место оценки

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|} \leq \frac{1}{R} \leq \overline{\lim}_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}. \quad (2.27)$$

Если же предел $\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n}$ существует, то в формуле (2.27) вместо неравенств имеют место равенства.

ДОКАЗАТЕЛЬСТВО. Обозначим $r_1 = \overline{\lim}_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}$, $r_2 = \lim_{n \rightarrow \infty} \frac{|a_{n+1}|}{|a_n|}$, $b_n = |a_n x^n|$. Легко проверить, что $\overline{\lim}_{n \rightarrow \infty} \frac{b_{n+1}}{b_n} = r_1 x$, и $\lim_{n \rightarrow \infty} \frac{b_{n+1}}{b_n} = r_2 x$. Отсюда по Теореме 2.2.8 имеем, что ряд $\sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} |a_n x^n|$ сходится при $r_1 |x| < 1$ и расходится при $r_2 |x| > 1$. Однако согласно Теореме 2.2.20 тот же самый ряд сходится при $|x| < R$ и расходится при $|x| > R$. Отсюда вытекает справедливость импликаций

$$|x| < \frac{1}{r_1} \Rightarrow |x| < R,$$

$$|x| > \frac{1}{r_2} \Rightarrow |x| > R.$$

Эти импликации могут иметь место только когда $\frac{1}{r_1} \leq R \leq \frac{1}{r_2}$. $\square$

СЛЕДСТВИЕ 2.2.24. Экспоненциальный ряд $\exp x = 1 + x + \frac{x^2}{2} + \dots + \frac{x^n}{n!} + \dots$ имеет бесконечный радиус сходимости.

ДОКАЗАТЕЛЬСТВО. Предоставляется учащимся в качестве упражнения. $\square$

§3 Предел функции

§3.1 Точки прикосновения множества

Пусть имеется непустое множество $S \subset \overline{\mathbb{R}}$. Говорят, что точка $x \in \overline{\mathbb{R}}$ является *точкой прикосновения* множества S , если $\forall U \in \mathcal{N}(x) \ U \cap S \neq \emptyset$.

Совокупность точек прикосновения множества S называется *замыканием множества S* и обозначается символом $\overline{S}$. Очевидно, что всегда выполняется включение $\overline{S} \supset S$. Если $\overline{S} = S$, то множество S называется *замкнутым*.

ТЕОРЕМА 2.3.1 (Свойства замыкания). Для всякого непустого множества $S \subset \overline{\mathbb{R}}$ справедливы следующие утверждения.

- (i) $\overline{(\overline{S})} = \overline{S}$;
- (ii) $\sup S \in \overline{S}$, $\inf S \in \overline{S}$;
- (iii) $x \in \overline{S}$ тогда и только тогда, когда найдется последовательность элементов $x_n \in S$ такая, что $x_n \rightarrow x$ при $n \rightarrow \infty$.

ДОКАЗАТЕЛЬСТВО.

На первых этапах доказательства нам понадобятся некоторые технические утверждения о свойствах системы элементарных окрестностей.

ШАГ 1. Если U является элементарной окрестностью некоторой точки x , то U является также окрестностью (уже не обязательно элементарной) каждой точки $y \in U$.

Доказательство этого факта очень просто. Проверим его для конечной точки $x \in \mathbb{R}$. Пусть U является элементарной окрестностью точки x , т.е. $U = (x - \varepsilon, x + \varepsilon)$ для некоторого $\varepsilon > 0$. Возьмем произвольно $y \in U$. Тогда $|x - y| < \varepsilon$; обозначим теперь $\varepsilon_1 = \varepsilon - |x - y|$, $V = (y - \varepsilon_1, y + \varepsilon_1)$. Ясно, что V представляет собой элементарную окрестность точки y . Возьмем произвольно $z \in V$, тогда $|z - x| \leq |z - y| + |y - x| < \varepsilon_1 + |y - x| = \varepsilon$, т.е. $z \in U$. Тем самым, ввиду произвольности z , мы доказали включение $V \subset U$. Следовательно, U является окрестностью точки y , и искомое утверждение доказано.

ШАГ 2. Для любой точки $x \in \overline{\mathbb{R}}$ существует вложенная последовательность элементарных окрестностей $U_n \in \mathcal{N}(x)$, $U_n \supset U_{n+1}$, такая, что $\bigcap_{n \in \mathbb{N}} U_n = \{x\}$, и для любой окрестности $V \in \mathcal{N}(x)$ найдется номер m такой, что $U_n \subset V$ для всех $n \geq m$.

Доказательство этого факта занимает одну строчку: если $x \in \mathbb{R}$, то достаточно взять $U_n = (x - \frac{1}{n}, x + \frac{1}{n})$, а если $x = \infty$, то $U_n = (n, \infty]$.

ШАГ 3. Перейдем теперь непосредственно к доказательству Теоремы 2.3.1.

(i) Включение $\overline{(\overline{S})} \supset \overline{S}$ очевидно. Докажем обратное включение. Пусть $x \in \overline{(\overline{S})}$. Возьмем произвольную элементарную окрестность $U \in \mathcal{N}(x)$. Тогда по определению множества $\overline{(\overline{S})}$ найдется точка $y \in U \cap \overline{S}$. Согласно утверждению Шага 1, имеем $U \in \mathcal{N}(y)$, откуда и из определения множества $\overline{S} \ni y$ вытекает, что $U \cap S \neq \emptyset$.

(ii) Обозначим $x = \sup S$. Если $x = -\infty$, то $S = \{-\infty\}$, и доказывать нечего. Пусть теперь $x \neq -\infty$. Возьмем произвольную элементарную окрестность $U \in \mathcal{N}(x)$. Ясно, что найдется точка $y \in U$ такая, что $y < x$. По определению $\sup$, y не может быть верхней границей множества S , т.е. найдется точка $z \in S$ такая, что $y < z \leq x$. Поскольку все элементарные окрестности суть выпуклые множества (т.е. вместе с любыми двумя своими точками они содержат и отрезок, их соединяющий), то $z \in U$, т.е. $U \cap S \neq \emptyset$. В силу произвольности $U \in \mathcal{N}(x)$ имеем $x \in \overline{S}$. Утверждение про $\inf S$ доказывается аналогично.

(iii) Пусть $S \ni x_n \rightarrow x$. Тогда по определению предела для каждой окрестности U точки x найдется номер m такой, что $x_n \in U$ при $n \geq m$, в частности, $x_m \in U \cap S \neq \emptyset$, следовательно $x \in \overline{S}$. Докажем в обратную сторону. Пусть $x \in \overline{S}$. Возьмем окрестности $U_n \in \mathcal{N}(x)$ из Шага 2 и точки $x_n \in U_n \cap S$ (существование таких точек вытекает из непустоты множества $U_n \cap S$). Утверждается, что имеет место сходимость $x_n \rightarrow x$. В самом деле, для произвольной окрестности $U \in \mathcal{N}(x)$ возьмем номер $m \in \mathbb{N}$ такой, что, $U_n \subset U$ для всех $n \geq m$ (см. Шаг 2). Тогда по построению справедливо включение $x_n \in U$ для всех $n \geq m$, откуда, согласно топологическому определению предела, заключаем, что $x_n \rightarrow x$ при $n \rightarrow \infty$. $\square$

§3.2 Предел функции

Всюду в данном разделе предполагается, что множество $S \subset \overline{\mathbb{R}}$ непусто и что p есть его точка прикосновения: $p \in \overline{S}$.

Говорят, что функция $f : S \rightarrow \overline{\mathbb{R}}$ имеет предел a при $x \rightarrow p$ по множеству S (обозначают $\lim_{x \rightarrow p|S} f(x) = a$ или $f(x) \rightarrow a$ при $x \rightarrow p|S$), если

$$\forall U \in \mathcal{N}(a) \exists V \in \mathcal{N}(p) \forall x \in V \cap S : f(x) \in U. \quad (2.28)$$

Данное определение можно переформулировать на арифметическом языке ε - δ для девяти возможных ситуаций (в зависимости от конечности предела p и конечности точки a). Приведем некоторые из этих ситуаций.

1) $a \in \mathbb{R}, p \in \mathbb{R}$. Тогда

$$\lim_{x \rightarrow p|S} f(x) = a \Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 \forall x \in S \cap (p - \delta, p + \delta) \quad |f(x) - a| < \varepsilon. \quad (2.29)$$

2) $a \in \mathbb{R}, p = \infty$. Тогда

$$\lim_{x \rightarrow p|S} f(x) = a \Leftrightarrow \forall \varepsilon > 0 \exists r \in \mathbb{R} \forall x \in S \cap (r, \infty) \quad |f(x) - a| < \varepsilon. \quad (2.30)$$

3) $a = \infty, p \in \mathbb{R}$. Тогда

$$\lim_{x \rightarrow p|S} f(x) = \infty \Leftrightarrow \forall r \in \mathbb{R} \exists \delta > 0 \forall x \in S \cap (p - \delta, p + \delta) \quad f(x) > r. \quad (2.31)$$

Упражнение 2.3.2. Сформулируйте определение предела на арифметическом языке ε - δ для оставшихся шести случаев (в зависимости от конечности предела p и точки a).

Обсуждаемые ниже свойства предела функции аналогичны соответствующим свойствам пределов последовательностей.

ТЕОРЕМА 2.3.3 (Единственность предела). Пусть имеется функция $f : S \rightarrow \overline{\mathbb{R}}$, точка $p \in \overline{S}$, и пусть $a, b \in \overline{\mathbb{R}}$ являются пределами f в p : $a = \lim_{x \rightarrow p|S} f(x)$ и $b = \lim_{x \rightarrow p|S} f(x)$. Тогда $a = b$.

ДОКАЗАТЕЛЬСТВО. Легко выводится из свойства отделимости, и оставляется читателю в качестве упражнения. $\square$

Функция $f : S \rightarrow \overline{\mathbb{R}}$ называется *возрастающей*, если $\forall x, y \in S \quad x < y \Rightarrow f(x) \leq f(y)$. Аналогично, f называется *убывающей*, если $\forall x, y \in S \quad x < y \Rightarrow f(x) \geq f(y)$. Возрастающие и убывающие последовательности называются *монотонными*.

Справедлива следующая красивая теорема.

ТЕОРЕМА 2.3.4. Пусть имеется множество $S \subset \overline{\mathbb{R}}$ и точка $p \in \overline{S}$, причем точка p лежит строго по одну сторону от S , т.е. либо $p < S$, либо $p > S$. Тогда всякая монотонная функция $f : S \rightarrow \overline{\mathbb{R}}$ имеет предел при $x \rightarrow p|S$.

ДОКАЗАТЕЛЬСТВО. Разберем для примера случай, когда f — возрастающая функция, а число p конечно. Обозначим $c = \sup\{f(x) : x \in S\}$. Имеются две возможности.

1. Случай $c = \infty$. Тогда множество $\sup\{f(x) : x \in S\}$ неограничено, т.е. для всякого числа $r \in \mathbb{R}$ найдется точка $y \in S$ такая, что $f(y) > r$. В силу предположения монотонности $f(x) \geq f(y) > r$ для всех $x \in S \cap (y, \infty) = S \cap (p - \delta, p + \delta)$, где через δ мы обозначили число $p - y$. Ввиду произвольности $r \in \mathbb{R}$ выполнены условия критерия (2.31), следовательно, $\lim_{x \rightarrow p|S} f(x) = \infty$.

2. Случай $c \in \mathbb{R}$. Возьмем произвольно $\varepsilon > 0$. Вследствие Леммы 1.2.8 (2) для найдется точка $y \in S$ такая, что $c - \varepsilon < f(y) \leq c$. В силу предположения монотонности $c - \varepsilon < f(x) \leq c$ для всех $x \in S \cap (y, \infty) = S \cap (p - \delta, p + \delta)$, где через δ мы обозначили число $p - y$. Ввиду произвольности $\varepsilon > 0$ выполнены условия критерия (2.29), следовательно, $\lim_{n \rightarrow \infty} x_n = c$. $\square$

Следующее утверждение используется очень часто и в математическом фольклоре носит наименование "принципа двух милиционеров (полицейских)": если два милиционера направляются в участок, и между ними находится нарушитель, то этому бедолаге, хочешь не хочешь, тоже приходится идти в участок.

ТЕОРЕМА 2.3.5. Пусть имеются три функции $f : S \rightarrow \mathbb{R}$, $g : S \rightarrow \mathbb{R}$, $h : S \rightarrow \mathbb{R}$. Предположим, что $\lim_{x \rightarrow p|S} f(x) = \lim_{x \rightarrow p|S} h(x) = a$ и $f(x) \leq g(x) \leq h(x)$ для всех $x \in S$. Тогда $\lim_{x \rightarrow p|S} g(x) = a$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Докажем ее утверждения для случая, когда $a \in \mathbb{R}$, $p \in \mathbb{R}$ (при $a \in \{-\infty, \infty\}$ делается по аналогии, даже немного проще). Возьмем произвольно $\varepsilon > 0$ и зафиксируем его. Согласно критерию (2.29) (в применении к функции $f(x)$), найдется число $\delta_1 > 0$ такое, что $a - \varepsilon < f(x) < a + \varepsilon$ для всех $x \in S \cap (p - \delta_1, p + \delta_1)$. По той же причине, найдется число $\delta_2 > 0$ такое, что $a - \varepsilon < h(x) < a + \varepsilon$ для всех $x \in S \cap (p - \delta_2, p + \delta_2)$. Положим $\delta = \min\{\delta_1, \delta_2\}$. Тогда для всех $x \in S \cap (p - \delta, p + \delta)$ имеем $x \in S \cap (p - \delta_1, p + \delta_1) \cap (p - \delta_2, p + \delta_2)$, следовательно,

$$a - \varepsilon < f(x) \leq g(x) \leq h(x) < a + \varepsilon.$$

Ввиду произвольности $\varepsilon > 0$ выполнены все условия критерия (2.29) в применении к функции $g(x)$, значит $\lim_{x \rightarrow p|S} g(x) = a$. $\square$

Замечание 2.3.6. Обычно символами $\lim_{x \rightarrow p} f(x)$ и $\lim_{S \ni x \rightarrow p} f(x)$ обозначают пределы $\lim_{x \rightarrow p|U \setminus \{p\}} f(x)$, $\lim_{x \rightarrow p|S \setminus \{p\}} f(x)$ соответственно, где U — некоторая окрестность p , т.е. $\lim_{x \rightarrow p} f(x)$ означает предел по множеству, у которого вырезана точка p (значения $f(x)$ берутся, когда x близок к p , но не совпадает с самим p).

СЛЕДСТВИЕ 2.3.7. Справедливо равенство $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$.

ДОКАЗАТЕЛЬСТВО. Для $0 < x < \frac{\pi}{2}$ сравнивая площади трех вложенных треугольников (средний из них — криволинейный, т.е. совпадает с круговым сектором угла x), получаем

$$\sin x \leq x \leq \operatorname{tg} x,$$

т.е.

$$\cos x \leq \frac{\sin x}{x} \leq 1,$$

далее применяем теорему о двух милиционерах. $\square$

Следующее понятие, несмотря на угрожающе-необычное название, весьма полезно. Пусть имеется некоторое высказывание $P(x)$ о точка множества S , и пусть $p \in \bar{S}$. Будем говорить, что высказывание P *асимптотически истинно при $x \rightarrow p|S$* , если найдется окрестность $U \in \mathcal{N}(p)$ такая, что $P(x)$ истинно при всех $x \in U \cap S$, обозначать этот факт будем символом $(P(x))_{x \rightarrow p|S}$ или $P(x)|_{x \rightarrow p|S}$.

С помощью данного понятия определение предела можно записать очень кратко. Например, для случая $a \in \mathbb{R}$

$$\lim_{x \rightarrow p|S} f(x) = a \Leftrightarrow \forall \varepsilon > 0 \quad |f(x) - a| < \varepsilon|_{x \rightarrow p|S}. \quad (2.32)$$

Отсюда легко выводится, что всякая функция, имеющая конечный предел, асимптотически ограничена, т.е.

$$\lim_{x \rightarrow p|S} f(x) = a \in \mathbb{R} \Rightarrow \exists M > 0 \quad |f(x)| \leq M|_{x \rightarrow p|S}. \quad (2.33)$$

Легко проверяется также, что если оба высказывания $P(x)$ и $Q(x)$ асимптотически истинны, то и их конъюнкция $P(x) \wedge Q(x)$ асимптотически истинна.

ТЕОРЕМА 2.3.8. Пусть имеются функции $f : S \rightarrow \mathbb{R}$ и $g : S \rightarrow \mathbb{R}$. Предположим, что $\lim_{x \rightarrow p|S} f(x) = a \in \bar{\mathbb{R}}$, $\lim_{x \rightarrow p|S} g(x) = b \in \bar{\mathbb{R}}$ и $a < b$. Тогда $(f(x) < g(x))_{x \rightarrow p|S}$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим случай, когда оба предела конечны: $a, b \in \mathbb{R}$ (случай бесконечных пределов рассматривается по аналогии, с очевидными упрощениями). Поскольку по условию теоремы $a < b$, найдется $\varepsilon > 0$ такое, что $a + \varepsilon < b - \varepsilon$. Согласно определению предела, найдется окрестность $U_1 \in \mathcal{N}(p)$ такая, что $a - \varepsilon < f(x) < a + \varepsilon$ для всех $x \in S \cap U_1$. По той же причине, найдется $U_2 \in \mathcal{N}(p)$ такое, что $b - \varepsilon < g(x) < b + \varepsilon$ для всех $x \in S \cap U_2$. Обозначим $V = U_1 \cap U_2$, по первому свойству системы окрестностей имеем $V \in \mathcal{N}(p)$. Тогда для всех $x \in S \cap V$ имеем $x \in U_1 \cap S$ и $x \in U_2 \cap S$, следовательно $f(x) < a + \varepsilon < b - \varepsilon < g(x)$. Теорема доказана. $\square$

ТЕОРЕМА 2.3.9. Пусть имеются функции $f : S \rightarrow \mathbb{R}$ и $g : S \rightarrow \mathbb{R}$, причем $\lim_{x \rightarrow p|S} f(x) = a \in \bar{\mathbb{R}}$, $\lim_{x \rightarrow p|S} g(x) = b \in \bar{\mathbb{R}}$. Предположим, что $(f(x) \leq g(x))_{x \rightarrow p|S}$. Тогда $a \leq b$.

ДОКАЗАТЕЛЬСТВО. Предположим противное, т.е. что $a > b$. Тогда по предыдущей Теореме 2.3.8 имеем $(f(x) > g(x))_{x \rightarrow p|S}$. Но это противоречит условию доказываемой теоремы. $\square$

Следующая теорема открывает цикл утверждений об арифметических операциях и предельном переходе.

ТЕОРЕМА 2.3.10. Пусть имеются функции $f : S \rightarrow \mathbb{R}$ и $g : S \rightarrow \mathbb{R}$. Предположим, что $\lim_{x \rightarrow p|S} f(x) = a \in \bar{\mathbb{R}}$, $\lim_{x \rightarrow p|S} g(x) = b \in \bar{\mathbb{R}}$, причем $\{a, b\} \neq \{-\infty, \infty\}$. Обозначим $h(x) = f(x) + g(x)$. Тогда $\lim_{x \rightarrow p|S} h(x) = a + b$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим случай, когда оба предела конечны: $a, b \in \mathbb{R}$ (случай бесконечных пределов рассматривается по аналогии, с очевидными упрощениями). Возьмем произвольно $\varepsilon > 0$ и зафиксируем его. Согласно арифметическому критерию сходимости (примененному для значения параметра малости $\frac{\varepsilon}{2}$), найдется окрестность $U_1 \in \mathcal{N}(p)$ такая, что $a - \frac{\varepsilon}{2} < f(x) < a + \frac{\varepsilon}{2}$ для всех $x \in S \cap U_1$. По той же причине, найдется окрестность $U_2 \in \mathcal{N}(p)$ такая, что $a - \varepsilon < g(x) < a + \varepsilon$ для всех $n > m_2$. Положим $V = U_1 \cap U_2$. Тогда $V \in \mathcal{N}(p)$ и для всех $x \in S \cap V$, складывая выписанные выше неравенства, получаем

$$a + b - \varepsilon < f(x) + g(x) = h(x) < a + b + \varepsilon.$$

Ввиду произвольности $\varepsilon > 0$ выполнены все условия арифметического критерия сходимости в применении к функции $h(x)$, значит $\lim_{x \rightarrow p|S} h(x) = a + b$. $\square$

Замечание 2.3.11. В случае, когда $\lim_{x \rightarrow p|S} f(x) = \infty$ и $\lim_{x \rightarrow p|S} g(x) = -\infty$, мы не можем сделать заключения про $\lim_{x \rightarrow p|S} (f(x) + g(x))$. Этот предел может в общем случае равняться любому наперед заданному числу, или вовсе не существовать (рекомендуется подобрать соответствующие примеры).

ТЕОРЕМА 2.3.12. Пусть имеются функции $f : S \rightarrow T \subset \mathbb{R}$ и $g : T \rightarrow \mathbb{R}$. Предположим, что $\lim_{x \rightarrow p|S} f(x) = q \in \overline{\mathbb{R}}$ и $\lim_{y \rightarrow q|T} g(y) = b \in \overline{\mathbb{R}}$. Обозначим $h(x) = g(f(x))$. Тогда $\lim_{x \rightarrow p|S} h(x) = b$.

ДОКАЗАТЕЛЬСТВО. Возьмем произвольную окрестность $U \in \mathcal{N}(b)$. По определению предела для $\lim_{y \rightarrow q|S} g(y) = b$, найдется окрестность $V \in \mathcal{N}(q)$ такая, что $g(y) \in U$ для всех $y \in V \cap S$. Далее, также по определению предела для $\lim_{x \rightarrow p|S} f(x) = q$, найдется окрестность $W \in \mathcal{N}(p)$ такая, что $f(x) \in V$ для всех $x \in W \cap S$. Тогда по построению $h(x) = f(g(x)) \in U$ для всех $x \in W \cap S$. $\square$

ТЕОРЕМА 2.3.13 (Критерий Гейне существования предела). Функция $f : S \rightarrow \mathbb{R}$ имеет предел в точке $p \in \overline{S}$ тогда и только тогда, когда для всякой последовательности точек $x_n \in S$ таких, что $x_n \rightarrow p$ при $n \rightarrow \infty$, последовательность значений $f(x_n)$ имеет предел при $n \rightarrow \infty$. Более того, для каждой такой последовательности справедливо равенство $\lim_{n \rightarrow \infty} f(x_n) = \lim_{x \rightarrow p|S} f(x)$.

Пример 2.3.14. Функция $f(x) = \sin \frac{1}{x}$, определенная на множестве $S = \mathbb{R} \setminus \{0\}$, не имеет предела при $x \rightarrow 0$. В самом деле, возьмем последовательность точек $x_n = \frac{1}{\frac{\pi}{2} + \pi n}$. Очевидно, что $x_n \rightarrow 0$ при $n \rightarrow \infty$, но последовательность

$$f(x_n) = \begin{cases} 1, & n \text{ четно;} \\ -1, & n \text{ нечетно,} \end{cases}$$

не имеет предела при $n \rightarrow \infty$, значит, не существует предела $\lim_{x \rightarrow 0} f(x)$.

Упражнение 2.3.15. С помощью критерия Гейне сформулировать и доказать утверждения о пределе произведения и частного двух функций (сведя дело к соответствующим утверждениям о пределах последовательностей)

§3.3 Асимптотические соотношения

Пусть имеется непустое множество $S \subset \overline{\mathbb{R}}$, точка $p \in \overline{S}$ и функции $f : S \rightarrow \mathbb{R}$, $g : S \rightarrow \mathbb{R}$. Будем говорить, что

$$f(x) = o(g(x))|_{x \rightarrow p|S}$$

(читается " f равно о малое от g при $x \rightarrow p|S$ "), если для функции $f(x)$ справедлива оценка $|f(x)| \leq \gamma(x)|g(x)|$ с некоторой функцией $\gamma : S \rightarrow \mathbb{R}$ такой, что $\lim_{x \rightarrow p|S} \gamma(x) = 0$.

Аналогично,

$$f(x) = O(g(x))|_{x \rightarrow p|S}$$

(читается " f равно о большое от g при $x \rightarrow p|S$ "), если найдется константа $C \in \mathbb{R}$ такая, что $|f(x)| \leq C|g(x)|$ при $x \rightarrow p|S$. Наконец, обозначают

$$f(x) \sim g(x)|_{x \rightarrow p|S},$$

(читается " f эквивалентна g при $x \rightarrow p|S$ "), если $f(x) - g(x) = o(g(x))|_{x \rightarrow p|S}$.

Имеют место следующие соотношения.

$$x^k = o(e^x) \text{ при } x \rightarrow \infty \text{ для всех } k \in \mathbb{Z};$$

$$e^x = o\left(\frac{1}{x^k}\right) \text{ при } x \rightarrow -\infty \text{ для всех } k \in \mathbb{Z};$$

$$\ln^k x = o(x) \text{ при } x \rightarrow \infty \text{ для всех } k \in \mathbb{Z};$$

$$\ln^k x = o\left(\frac{1}{x}\right) \text{ при } x \rightarrow 0+ \text{ для всех } k \in \mathbb{Z}.$$

В частности, для всех $\alpha \in \mathbb{R}$ имеем $\frac{e^{-\frac{1}{x}}}{x^\alpha} \rightarrow 0$ при $x \rightarrow 0+$.

ТЕОРЕМА 2.3.16. Пусть имеются функции $f : S \rightarrow \mathbb{R}$ и $g : S \rightarrow \mathbb{R}$, $p \in \overline{S}$, причем $g(x) \neq 0|_{x \rightarrow p|S}$. Тогда

$$f(x) = o(g(x))|_{x \rightarrow p|S} \Leftrightarrow \frac{f(x)}{g(x)} \rightarrow 0|_{x \rightarrow p|S}, \quad (2.34)$$

$$f(x) = O(g(x))|_{x \rightarrow p|S} \Leftrightarrow \exists C > 0 \frac{|f(x)|}{|g(x)|} \leq C|_{x \rightarrow p|S}, \quad (2.35)$$

$$f(x) \sim g(x)|_{x \rightarrow p|S} \Leftrightarrow \frac{f(x)}{g(x)} \rightarrow 1|_{x \rightarrow p|S}. \quad (2.36)$$

ДОКАЗАТЕЛЬСТВО. Легко вытекает из соответствующих определений. $\square$

ТЕОРЕМА 2.3.17. Пусть имеются функции $f : S \rightarrow \mathbb{R}$ и $g : S \rightarrow \mathbb{R}$, $p \in \overline{S}$, причем $f(x) \sim g(x)|_{x \rightarrow p|S}$. Тогда найдется функция $\gamma : S \rightarrow \mathbb{R}$ такая, что $\gamma(x) \rightarrow 1|_{x \rightarrow p|S}$ и

$$f(x) = \gamma(x)g(x)|_{x \rightarrow p|S}. \quad (2.37)$$

ДОКАЗАТЕЛЬСТВО. Легко вытекает из соответствующих определений. $\square$

СЛЕДСТВИЕ 2.3.18. Отношение $\sim$ рефлексивно, симметрично и транзитивно.

Далее предполагается, что задано непустое множество $S \subset \overline{\mathbb{R}}$, функция $f : S \rightarrow \mathbb{R}$, и точка $p \in \overline{S}$, и все асимптотические соотношения понимаются при $x \rightarrow p|S$.

ТЕОРЕМА 2.3.19. *Справедливы соотношения*

$$O(1)o(1) = o(1), \quad (2.38)$$

$$\exists \lim_{x \rightarrow p|S} f(x) = a \in \mathbb{R} \Leftrightarrow f(x) - a = o(1) \Rightarrow f(x) = O(1), \quad (2.39)$$

$$o(f) = f \cdot o(1), \quad (2.40)$$

$$O(f) = f \cdot O(1). \quad (2.41)$$

ТЕОРЕМА 2.3.20. *Справедливы соотношения*

$$o(f) + o(f) = O(f), \quad (2.42)$$

$$O(f) + o(f) = O(f), \quad (2.43)$$

$$O(f + o(f)) = O(f), \quad (2.44)$$

$$o(f + O(f)) = o(f), \quad (2.45)$$

ДОКАЗАТЕЛЬСТВО. Легко вытекает из соответствующих определений. $\square$

ТЕОРЕМА 2.3.21. *Справедливы соотношения*

$$f = o(g), \quad g = O(h) \Rightarrow f = o(h), \quad (2.46)$$

$$h \sim g \Rightarrow (f = o(g) \Leftrightarrow f = o(h)), \quad (2.47)$$

$$f_1 = o(g_1), \quad f_2 = o(g_2) \Rightarrow f_1 f_2 = o(g_1 g_2), \quad (2.48)$$

$$f_1 = o(g_1), \quad f_2 = o(g_2) \Rightarrow f_1 + f_2 = o(|g_1| + |g_2|), \quad (2.49)$$

Разумеется, можно доказать и массу других свойств.

Пример 2.3.22. (i) $x^3 = o(x^2)$ при $x \rightarrow 0$;

(ii) $x^2 = o(x^3)$ при $x \rightarrow \infty$;

(ii') Еще характерный пример: $\sqrt{x^2 + x} \sim x$ при $x \rightarrow \infty$; $\sqrt{x^2 + x} \sim \sqrt{x}$ при $x \rightarrow 0$

(iii) Для полинома $P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_{m+1} x^{m+1} + a_m x^m$, $a_n \neq 0 \neq a_m$, имеют место эквивалентности $P(x) \sim a_m x^m$ при $x \rightarrow 0$ и $P(x) \sim a_n x^n$ при $x \rightarrow \infty$;

(iv) $e^x - (1 + x + \frac{x^2}{2} + \dots + \frac{x^7}{7!}) = O(x^8) = o(x^7)$ при $x \rightarrow 0$;

(v) $\sin x - x + \frac{x^3}{6} = O(x^5) = o(x^4)$ при $x \rightarrow 0$;

(v) $\cos x - 1 + \frac{x^2}{4} = O(x^4) = o(x^3)$ при $x \rightarrow 0$;

(vi) $\ln(1 + x) - x - \frac{x^2}{2} = O(x^3) = o(x^2)$ при $x \rightarrow 0$;

(vii) Для всякого $\alpha \in \mathbb{R}$ имеем $(1 + x)^\alpha = 1 + \alpha x + \frac{1}{2}\alpha(\alpha - 1)x^2 + O(x^3)$ при $x \rightarrow 0$. В частности, $\frac{1}{1+x} = 1 - x + O(x^2)$ при $x \rightarrow 0$.

Применяя указанные соотношения и утверждения теорем, нетрудно упрощать угрожающе-громоздкие выражения, как иллюстрируют следующие примеры.

$$\begin{aligned}
o(x + \frac{\sqrt{x}}{2+\sin x+\cos x}) &= o(x + O(\sqrt{x})) = o(x) \text{ при } x \rightarrow \infty; \\
o(x + \frac{\sqrt{x}}{2+\sin x+\cos x}) &= o(x + O(\sqrt{x})) = o(\sqrt{x}) \text{ при } x \rightarrow 0. \\
\frac{x^3}{x^2+1} &\sim x \text{ при } x \rightarrow \infty; \quad \frac{x^3}{x^2+1} \sim x^3 \text{ при } x \rightarrow 0 \\
\frac{x^2-1}{x^2+1} &\sim 1 \text{ при } x \rightarrow \infty; \quad \frac{x^2-1}{x^2+1} \sim (x-1) \text{ при } x \rightarrow 1 \\
o(x + \frac{x^2}{x^4+1}) &= o(x + O(x^2)) = o(x + o(x)) = o(x) \text{ при } x \rightarrow 0; \\
o(x + \frac{x^2}{x^4+1}) &= o(x + O(\frac{1}{x^2})) = o(x + o(x)) = o(x) \text{ при } x \rightarrow \infty. \\
\frac{\sin(x+\sqrt{x})}{x} &\rightarrow \infty \text{ при } x \rightarrow 0, \\
\frac{\sin(x+\sin x)}{x} &\rightarrow 2 \text{ при } x \rightarrow 0, \\
[\exp(x^3-1)-1]/(x^2-1) &\rightarrow 3 \text{ при } x \rightarrow 0, \\
\frac{(1+\sin \sqrt{x})^2-2\sqrt{x}-1}{x} &= \frac{(1+\sqrt{x}+o(x))^2-2\sqrt{x}-1}{x} = \frac{x+o(x)}{x} = 1 + o(1). \\
\frac{1}{1+x-x^2+\sin x^3} &= \frac{1}{1+x+o(x)+x^3+o(x^3)} = \frac{1}{1+x+o(x)} = 1-x-o(x)+o(x+o(x)) = 1-x-o(x)+o(x) = \\
&1-x+o(x).
\end{aligned}$$

$$\begin{aligned}
\frac{\frac{1}{2-\exp(x+\sin x^2)} - \cos x^2}{x} &= \frac{\frac{1}{2-\exp(x+o(x))} - 1 + o(x)}{x} = \frac{\frac{1}{1-x-o(x)+o(x+o(x))} - 1 + o(x)}{x} = \\
&= \frac{\frac{1}{1-x-o(x)} - 1 + o(x)}{x} = \frac{1+x+o(x)+o(x+o(x))-1+o(x)}{x} = \frac{x+o(x)}{x} = 1.
\end{aligned}$$

§4 Непрерывные функции

§4.1 Определение и локальные свойства непрерывных функций

Пусть $\emptyset \neq S \subset \overline{\mathbb{R}}$. Говорят, что функция $f : S \rightarrow \mathbb{R}$ *непрерывна* в точке $p \in S$, если $\lim_{x \rightarrow p|S} f(x) = f(p)$, т.е. предел функции f в точке p по множеству S совпадает со значением f в этой точке.

Ясно, что понятие непрерывности в точке является *локальным*, т.е. оно зависит только от поведения f в некоторой *окрестности* точки p . Более точно, имеет место следующая

ТЕОРЕМА 2.4.1. $f : S \rightarrow \mathbb{R}$ непрерывна в точке $p \in S \Leftrightarrow \exists U \in \mathcal{N}(p)$ $f|_U$ непрерывна в p .

Сначала перечислим *локальные* свойства непрерывных функций.

ТЕОРЕМА 2.4.2. Пусть функция $f : S \rightarrow \mathbb{R}$ непрерывна в точке $p \in S$. Тогда

- (i) $f(x) = O(1)|_{x \rightarrow p|S}$;
- (ii) $f(p) = 0 \Leftrightarrow f(x) = o(1)|_{x \rightarrow p|S}$;
- (iii) $f(p) > 0 \Rightarrow f(x) > \frac{1}{2}f(p)|_{x \rightarrow p|S}$, аналогично, $f(p) < 0 \Rightarrow f(x) < \frac{1}{2}f(p)|_{x \rightarrow p|S}$.
- (iv) $f(p) > c \Rightarrow f(x) > c|_{x \rightarrow p|S}$, аналогично, $f(p) < c \Rightarrow f(x) < c|_{x \rightarrow p|S}$.

ТЕОРЕМА 2.4.3. Пусть функции $f : S \rightarrow \mathbb{R}$, $g : S \rightarrow \mathbb{R}$ непрерывны в точке $p \in S$, и пусть $\lambda \in \mathbb{R}$. Тогда λf , $f + g$, fg непрерывны в точке p . Более того, если $g(p) \neq 0$, то функция $\frac{f}{g}$ также непрерывна в точке p .

СЛЕДСТВИЕ 2.4.4. Всякая рациональная функция (т.е. отношение двух полиномов) непрерывна на своей области определения (т.е. на множестве, где знаменатель не обращается в ноль).

ТЕОРЕМА 2.4.5. Пусть функции $f : S \rightarrow \mathbb{R}$, $g : S \rightarrow \mathbb{R}$ непрерывны в точке $p \in S$, и пусть $\lambda \in \mathbb{R}$. Тогда λf , $f + g$, fg непрерывны в точке p . Более того, если $g(p) \neq 0$, то функция $\frac{f}{g}$ также непрерывна в точке p .

ТЕОРЕМА 2.4.6. Пусть имеются функции $f : S \rightarrow T \subset \mathbb{R}$ и $g : T \rightarrow \mathbb{R}$. Предположим, что f непрерывна в точке $p \in S$, и g непрерывна в точке $q = f(p)$. Тогда функция $h : S \rightarrow \mathbb{R}$, задаваемая формулой $h(x) = g(f(x))$, также непрерывна в точке p .

ТЕОРЕМА 2.4.7. Пусть имеются функции $f : S \rightarrow T \subset \mathbb{R}$ и $g : T \rightarrow \mathbb{R}$. Предположим, что f непрерывна в точке $p \in S$, и g непрерывна в точке $q = f(p)$. Тогда функция $h : S \rightarrow \mathbb{R}$, задаваемая формулой $h(x) = g(f(x))$, также непрерывна в точке p .

ТЕОРЕМА 2.4.8. Пусть функции $f : S \rightarrow \mathbb{R}$, $g : S \rightarrow \mathbb{R}$ непрерывны в точке $p \in S$, причем $f(p) = g(p)$. Тогда всякая функция $h : S \rightarrow \mathbb{R}$, удовлетворяющая условию $f(x) \leq h(x) \leq g(x)$ при $x \rightarrow p|S$, непрерывна в точке p .

ТЕОРЕМА 2.4.9 (Критерий непрерывности Гейне). Функция $f : S \rightarrow \mathbb{R}$ непрерывна в точке $p \in S$ тогда и только тогда, когда для всякой последовательности точек $x_n \in S$ таких, что $x_n \rightarrow p$ при $n \rightarrow \infty$, справедливо равенство $\lim_{n \rightarrow \infty} f(x_n) = f(p)$.

Пример 2.4.10. Функция

$$f(x) = \begin{cases} \sin \frac{1}{x}, & x \neq 0; \\ a, & x = 0 \end{cases}$$

не является непрерывной ни при каком значении $a \in \mathbb{R}$ (см. также пример 2.4.10).

ТЕОРЕМА 2.4.11 (Устойчивость строгих неравенств). Пусть функции $f : S \rightarrow \mathbb{R}$, $g : S \rightarrow \mathbb{R}$ непрерывны в точке $p \in S$, причем $f(p) < g(p)$. Тогда $f(x) < g(x)|_{x \rightarrow p|S}$.

§4.2 Глобальные свойства непрерывных функций

Всюду в дальнейшем $[a, b]$ — непустой конечный замкнутый промежуток из $\mathbb{R}$, и $S \subset [a, b]$ — непустое ограниченное замкнутое множество.

ТЕОРЕМА 2.4.12 (Теорема Вейерштрасса о максимальном значении). Если непустое множество $S \subset \mathbb{R}$ замкнуто и ограничено, и функция $f : S \rightarrow \mathbb{R}$ непрерывна, то найдутся точки $x_{\max}, x_{\min} \in S$ такие, что $f(x_{\max}) \geq f(x)$ и $f(x_{\min}) \leq f(x)$ для всех $x \in S$.

ДОКАЗАТЕЛЬСТВО. Докажем утверждение про максимальное значение. Обозначим $M = \sup\{f(x) : x \in S\}$ (пока мы не исключаем случая $M = \infty$). Ясно, что $M > -\infty$. Возьмем последовательность чисел $M_n < M$ таких, что $M_n \rightarrow M$ при $n \rightarrow \infty$. Тогда по определению $\sup$, для каждого $n \in \mathbb{N}$ найдется точка $x_n \in S$ такая, что

$$M_n < f(x_n) \leq M. \quad (2.50)$$

По теореме о двух милиционерах,

$$f(x_n) \rightarrow M \quad \text{при } n \rightarrow \infty. \quad (2.51)$$

Обозначим $p = \overline{\lim}_{n \rightarrow \infty} x_n$. Ввиду ограниченности S имеем включение $p \in \mathbb{R}$. По теореме Вейерштрасса о частичных пределах, существует подпоследовательность x_{n_k} такая, что $x_{n_k} \rightarrow p$ при $k \rightarrow \infty$. В силу замкнутости множества S имеем $p \in S$. Поскольку функция f непрерывна, то по критерию Гейне

$$f(p) = \lim_{x \rightarrow p|S} f(x) = \lim_{k \rightarrow \infty} f(x_{n_k}) = M$$

(последнее равенство выполнено в силу (2.51)). Теперь можно взять $x_{\max} = p$. □

СЛЕДСТВИЕ 2.4.13. Если непустое множество $S \subset \mathbb{R}$ замкнуто и ограничено, и функция $f : S \rightarrow \mathbb{R}$ непрерывна, то множество $f(S)$ также замкнуто и ограничено.

ДОКАЗАТЕЛЬСТВО. Ограниченность $f(S)$ напрямую вытекает из теоремы Вейерштрасса, докажем замкнутость. Пусть $y_n \in f(S)$ и $y_n \rightarrow y_0$ при $n \rightarrow \infty$. Нам требуется проверить, что $y_0 \in f(S)$. По построению, найдутся точки $x_n \in S$ такие, что $f(x_n) = y_n$. Обозначим $p = \overline{\lim}_{n \rightarrow \infty} x_n$. По теореме Вейерштрасса о частичных пределах, существует подпоследовательность x_{n_k} такая, что $x_{n_k} \rightarrow p$ при $k \rightarrow \infty$. В силу замкнутости множества S имеем $p \in S$. Поскольку функция f непрерывна, то по критерию Гейне

$$f(p) = \lim_{x \rightarrow p|S} f(x) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \lim_{k \rightarrow \infty} y_{n_k} = y_0.$$

Значит, $y_0 \in f(S)$, что и требовалось доказать. □

ТЕОРЕМА 2.4.14 (Теорема о прообразах). Пусть непустое множество $S \subset \mathbb{R}$ замкнуто. Тогда функция $f : S \rightarrow \mathbb{R}$ непрерывна в том и только том случае, если для любого замкнутого² множества $Y \subset \mathbb{R}$ прообраз $f^{-1}(y) = \{x \in S : f(x) \in Y\}$ также представляет собой замкнутое множество.

ДОКАЗАТЕЛЬСТВО. $\Rightarrow$ Пусть функция $f : S \rightarrow \mathbb{R}$ непрерывна и $Y \subset \mathbb{R}$ — замкнутое множество. Обозначим $X = f^{-1}(Y) = \{x \in S : f(x) \in Y\}$. Если $X = \emptyset$, то доказывать нечего. Предположим теперь, что множество X непусто, и возьмем произвольную сходящуюся последовательность $X \ni x_n \rightarrow x_0$. Докажем, что $x_0 \in X$. В силу замкнутости S имеем $x_0 \in S$. Вследствие непрерывности f в точке x_0 и критерия Гейне, справедливо равенство $f(x_n) \rightarrow f(x_0)$. Поскольку по построению $f(x_n) \in Y$ и Y — замкнутое множество, имеем $f(x_0) \in Y$. Отсюда немедленно вытекает искомое включение $x_0 \in X$.

$\Leftarrow$ Пусть функция $f : S \rightarrow \mathbb{R}$ такова, что $f^{-1}(Y)$ представляет собой замкнутое множество для произвольного замкнутого множества $Y \subset \mathbb{R}$. Докажем, что функция f непрерывна. Предположим, что это неверно, тогда найдется точка $p \in S$ и последовательность $x_n \in S$ такие, что $x_n \rightarrow p$, но $f(x_n) \not\rightarrow f(p)$ при $n \rightarrow \infty$. Тогда либо $f(p) < \overline{\lim}_{n \rightarrow \infty} f(x_n)$, либо $f(p) > \underline{\lim}_{n \rightarrow \infty} f(x_n)$ (т.к. если нарушены оба этих неравенства, то $\overline{\lim}_{n \rightarrow \infty} f(x_n) \leq f(p) \leq \underline{\lim}_{n \rightarrow \infty} f(x_n) \Rightarrow f(p) = \lim_{n \rightarrow \infty} f(x_n)$). Предположим, для определенности, что $f(p) < \overline{\lim}_{n \rightarrow \infty} f(x_n) = \lim_{k \rightarrow \infty} f(x_{n_k})$ (существование указанной подпоследовательности вытекает из теоремы Вейерштрасса). Возьмем $\delta > 0$ такое, что $f(p) + \delta < \overline{\lim}_{n \rightarrow \infty} f(x_n)$. Тогда

²В этом разделе и далее, если не оговорено противное, множество $Y \subset \mathbb{R}$ называется *замкнутым*, если $\bar{Y} \cap \mathbb{R} = Y$ (т.е. мы удаляем из рассмотрения бесконечно удаленные точки).

найдется номер $m \in \mathbb{N}$ такой, что $f(x_{n_k}) > f(p) + \delta$ для всех $k \geq m$. Перенумеровав подпоследовательность, можно считать, не умаляя общности, что $f(x_{n_k}) > f(p) + \delta$ для всех $k \in \mathbb{N}$. Положим $Y = [f(p) + \delta, \infty)$. Множество Y замкнуто, поэтому по нашему предположению $f^{-1}(Y) = \{x \in S : f(x) \in Y\}$ также представляет собой замкнутое множество. Однако, имеем по построению, что $x_{n_k} \in f^{-1}(Y)$, $x_{n_k} \rightarrow p$ при $k \rightarrow \infty$, но $p \notin f^{-1}(Y)$. Следовательно, множество $f^{-1}(Y)$ не замкнуто. Полученное противоречие завершает доказательство Теоремы. $\square$

ТЕОРЕМА 2.4.15 (Теорема Коши о промежуточных значениях). Если функция $f : [a, b] \rightarrow \mathbb{R}$ непрерывна, и число $y \in \mathbb{R}$ лежит между значениями $f(a)$ и $f(b)$ (т.е. y принадлежит отрезку с концами $f(a)$ и $f(b)$) то найдется точка $x \in [a, b]$ такая, что $f(x) = y$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Для определенности предположим, что $f(a) < y < f(b)$. Обозначим $S = \{x \in [a, b] : f(x) \leq y\}$. Согласно предположениям, имеем $\emptyset \neq S \neq [a, b]$, и по предыдущей теореме, S представляет собой замкнутое множество. Обозначим $p = \sup S$. Вследствие Теоремы 2.3.1 имеем $p \in S$, поэтому $f(p) \leq y$. В частности, $p < b$. Докажем, что на самом деле имеет место равенство $f(p) = y$. В самом деле, если бы не так, то $f(p) < y$, но тогда, в силу непрерывности f в точке p , для некоторого достаточно малого $\delta > 0$ справедливо неравенство $f(x) < y$ для всех точек $x \in (p - \delta, p + \delta)$. Однако это противоречит выбору $p = \sup S$. $\square$

СЛЕДСТВИЕ 2.4.16. Если функция $f : [a, b] \rightarrow \mathbb{R}$ непрерывна, то множество ее значений представляет собой замкнутый промежуток: $f([a, b]) = [c, d]$ для некоторых $c, d \in \mathbb{R}$, $c \leq d$.

ТЕОРЕМА 2.4.17 (Теорема Больцано о монотонной функции). Если функция $f : [a, b] \rightarrow \mathbb{R}$ непрерывна и инъективна, то она строго монотонна.

ДОКАЗАТЕЛЬСТВО. Тот факт, что $f([a, b]) = [c, d]$ для некоторых $c, d \in \mathbb{R}$, $c \leq d$, вытекает из Следствия 2.4.16. Предположим, далее, что $a < b$ (иначе доказывать нечего). Тогда $f(a) \neq f(b)$. Для определенности, будем считать, что $f(a) < f(b)$.

ШАГ 1. Докажем, что

$$f(a) = \min f([a, b]). \quad (2.52)$$

В самом деле, если бы не так, то найдется точка $x_0 \in (a, b)$ такая, что $f(x_0) < f(a)$. Возьмем $y = \frac{1}{2}(f(a) + f(x_0))$. Тогда элементарными вычислениями проверяется, что $f(x_0) < y < f(a)$ и $f(x_0) < y < f(b)$. По теореме Коши о промежуточных значениях найдутся точки $x_2 \in (a, x_0)$, $x_3 \in (x_0, b)$ такие, что $f(x_2) = y = f(x_3)$. Получили противоречие с предположением об инъективности f .

ШАГ 2. Аналогично доказывается, что

$$f(b) = \max f([a, b]). \quad (2.53)$$

ШАГ 3. Итак, мы установили, что для инъективных функций, заданных на промежутке, неравенство $f(a) < f(b)$ на концах промежутка влечет справедливость неравенств $f(a) < f(x) < f(b)$ для всех $x \in (a, b)$. Возьмем теперь произвольно точку $x_0 \in (a, b)$ и зафиксируем ее. Применим доказанное только что на шагах 1-2 утверждение к сужению функции f на промежуток $[a, x_0]$ (мы имеем право это сделать, т.к. сужение $f|_{[a, x_0]}$ также инъективно и по доказанному выше $f(a) < f(x_0)$). Тогда получаем, что $f(a) < f(x) < f(x_0)$ для всех

$x \in (a, x_0)$. В силу произвольности точек $x, x_0 \in \mathbb{R}$, мы установили строгое возрастание функции f . Теорема доказана. $\square$

Напомним, что промежутком с концами $a, b \in \overline{\mathbb{R}}$ называется множество одного из четырех типов: (a, b) , $(a, b]$, $[a, b)$, $[a, b]$. Промежуток обозначается символом $< a, b >$. Напомним также, что для множества $S \subset \mathbb{R}$ символом $\lim_{S \ni x \rightarrow a} f(x)$ обозначается предел $\lim_{x \rightarrow a | S \setminus \{a\}} f(x)$.

ТЕОРЕМА 2.4.18 (Теорема об обратной функции). *Если функция $f : T \rightarrow \mathbb{R}$ на промежутке $T \subset \mathbb{R}$ непрерывная и строго возрастающая, то*

- (i) *образ $f(T)$ этого промежутка является промежутком того же типа, что и T ;*
- (ii) *Отображение $f : T \rightarrow f(T)$ обратимо;*
- (iii) *Обратное к нему отображение $g : f(T) \rightarrow T$ непрерывное и строго возрастающее.*

ДОКАЗАТЕЛЬСТВО.

(i) Достаточно рассмотреть случай $T = (a, b]$. Возьмем открытый промежуток $T' = (a, b)$ и обозначим $c = \lim_{x \rightarrow a | T'} f(x) = \inf f(T')$, $d = \lim_{x \rightarrow b | T'} f(x) = \sup f(T')$ (указанные пределы существуют в силу теоремы Вейерштрасса о пределе монотонных функций). Докажем, что $f(T) = (c, d]$. Включение $f(T) \subset (c, d]$ очевидно. Пусть $y \in (c, d)$, тогда возьмем $c < c_1 < y < d_1 < d$. Вследствие определения $\sup, \inf$, найдутся точки $x_1, x_2 \in (a, b)$ такие, что $f(x_1) < c_1$, $f(x_2) > d_1$. Тогда по теореме о промежуточных значениях найдется точка $x \in (x_1, x_2)$ такая, что $f(x) = y$. Итак, мы доказали, что $(c, d) \subset f(T) \subset [c, d]$. Осталось разобраться с концевыми точками. Равенство $d = f(b)$ вытекает из непрерывности f . Если же $c \in f(T)$, то $c = f(x_0)$ для некоторого $x_0 > a$. Но тогда в силу строгого возрастания $f(x) < f(x_0) = c$ для $x \in (a, x_0)$, что противоречит равенству $c = \inf f$. Следовательно, $c \notin f(T)$, т.е. $f(T) = (c, d]$.

(ii) По построению, отображение $f : T \rightarrow f(T)$ сюръективно. В силу строгой монотонности, оно также и инъективно. Значит, оно биективно, а потому и обратимо в силу Теоремы 1.2.34.

(iii) Рассмотрим обратное отображение $g : (c, d] \rightarrow (a, b]$. Его строгое возрастание непосредственно вытекает из строгой монотонности f . Осталось проверить непрерывность. Пусть $(c, d] \ni y_n \rightarrow y \in (c, d]$. Тогда найдутся точки $x_n \in (a, b]$ и $x \in (a, b]$ такие, что $y_n = f(x_n)$ и $f(x) = y$. Докажем, что $\lim x_n = x$. Возьмем произвольную имеющую предел подпоследовательность $x_{n_k} \rightarrow x_0$. Проверим сначала, что $x_0 \in (a, b]$. Если бы не так, то $x_0 = a$, но тогда $f(x_{n_k}) \rightarrow c \neq y$, противоречие. Следовательно, $x_0 \in (a, b]$. Вследствие непрерывности имеем $f(x_0) = \lim f(x_{n_k}) = y$. Значит, $x_0 = x$. Тем самым мы показали, что множество частичных пределов последовательности x_n совпадает с одноточечным множеством $\{x\}$, поэтому $\lim_{n \rightarrow \infty} x_n = x$. Теорема полностью доказана. $\square$

§4.3 Основные элементарные функции

Функция $\sqrt{n}$

Пусть n — положительное натуральное число. Определим отображение $f_n : \mathbb{R} \ni x \mapsto x^n$. Имеем классическое тождество $(x^n - y^n) = (x - y)(x^{n-1} + x^{n-2}y + \dots + y^{n-1})$. Далее, если n нечетно, то f_n строго возрастает на $\mathbb{R}$, причем, очевидно, $f(x) \rightarrow \infty$ при $x \rightarrow \infty$

и $f(x) \rightarrow -\infty$ при $x \rightarrow -\infty$. Тогда по Теореме 2.4.18 обратное отображение $g_n : \mathbb{R} \rightarrow \mathbb{R}$ определено на всем $\mathbb{R}$ и является строго возрастающим. Это отображение и является корнем n -й степени. Если же n четно, то, опять-таки очевидно, что f_n строго возрастает на промежутке $\mathbb{R}_+ = [0, +\infty)$ и убывает на промежутке $(-\infty, 0]$. Тогда обратная функция $g_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ определена на $\mathbb{R}_+$.

Функция e^x .

Ранее была введена строго возрастающая функция $\exp : \mathbb{R} \rightarrow (0, +\infty)$ со свойствами:

(E1) Функция $\exp$ строго возрастает, $\exp(x) \rightarrow \infty$ при $x \rightarrow \infty$, $\exp(x) = \frac{1}{\exp(-x)} \rightarrow 0$ при $x \rightarrow -\infty$.

(E2) $\exp(0) = 1$;

(E3) $\exp(x + y) = \exp(x) \exp(y)$ для всех $x, y \in \mathbb{R}$.

Обозначим через e число $e = \exp(1)$. Наша ближайшая цель — показать, что функция $\exp$ совпадает с e^x . Для рациональных $x = \frac{m}{n}$ функция e^x определяется как $\sqrt[n]{e^m}$.

ТЕОРЕМА 2.4.19 (Теорема об экспоненте). Функция $\exp : \mathbb{R} \rightarrow \mathbb{R}$ непрерывна, более того, $\exp(x) = e^x$ для всех $x \in \mathbb{Q}$.

ДОКАЗАТЕЛЬСТВО. Мы докажем более общий факт: всякий степенной ряд $f(x) = \sum_{n=0}^{\infty} a_n x^n$ представляет собой непрерывную функцию на интервале $(-R, R)$, здесь R — радиус сходимости степенного ряда (см. Теорему 2.2.20). Возьмем произвольно точку $x_0 \in (-R, R)$ и параметр $\varepsilon > 0$. Обозначим $\sigma = \frac{1}{3}(R - |x_0|)$, $f_m(x) = \sum_{n=m}^{\infty} a_n x^n$. Ранее было доказано (см. Теорему 2.2.21), что

$$|f(x) - f_m(x)| < |x|^{m+1} \frac{1}{\sigma(R - \sigma)^m} \quad (2.54)$$

при $|x| < R - 2\sigma$. Если $|x - x_0| < \sigma$, то $|x| \leq |x - x_0| + |x_0| < \sigma + |x_0| = R - 2\sigma$, поэтому

$$|f(x) - f_m(x)| < \frac{R - 2\sigma}{\sigma} \left(\frac{R - 2\sigma}{R - \sigma} \right)^m \quad \text{при } |x - x_0| < \sigma. \quad (2.55)$$

Возьмем m столь велико, что $\frac{R - 2\sigma}{\sigma} \left(\frac{R - 2\sigma}{R - \sigma} \right)^m < \varepsilon/3$. Поскольку f_m есть непрерывная функция (полином), то найдется $\delta_1 > 0$ такое, что $|f_m(x) - f_m(x_0)| < \varepsilon/3$ при $|x - x_0| < \delta_1$. Положим $\delta = \min(\sigma, \delta_1)$. Окончательно имеем

$$|f(x) - f(x_0)| \leq |f(x) - f_m(x)| + |f_m(x) - f_m(x_0)| + |f_m(x_0) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

для всех $x \in \mathbb{R}$ таких, что $|x - x_0| \leq \delta$. В силу произвольности $\varepsilon > 0$ непрерывность функции f в каждой точке $x_0 \in (-R, R)$ установлена.

Проверим теперь второе утверждение теоремы. В силу равенства $\exp(x) = \frac{1}{\exp(-x)}$ достаточно сделать это только для положительных значений x . Итак, пусть $x = \frac{m}{n}$, $m, n \in \mathbb{N}$. Тогда по основному свойству экспоненты имеем равенства

$$\left(\exp\left(\frac{m}{n}\right) \right)^n = \exp\left(n \frac{m}{n}\right) = \exp(m) = \exp(1)^m = e^m,$$

другими словами, $\exp\left(\frac{m}{n}\right) = e^{\frac{m}{n}}$. □

Для множества $E \subset \mathbb{R}$ всюду в дальнейшем символом $\text{Cl } E$ (от лат. *closure* — замыкание) будет обозначаться множество $\bar{E} \cap \mathbb{R}$, т.е. это означает замыкание множества E относительно $\mathbb{R}$.

ТЕОРЕМА 2.4.20 (О единственности продолжения функций). Пусть $E \subset \mathbb{R}$ — непустое множество, и пусть $f : \text{Cl } E \rightarrow \mathbb{R}$, $g : \text{Cl } E \rightarrow \mathbb{R}$ — непрерывные функции. Тогда, если справедливо тождество $f(x) = g(x)$ для всех $x \in E$, то $f(x) = g(x)$ для всех $x \in \text{Cl } E$.

ДОКАЗАТЕЛЬСТВО. Пусть $x \in \text{Cl } E$. Тогда найдется последовательность точек $x_n \in E$ такая, что $x_n \rightarrow x$ при $n \rightarrow \infty$. Отсюда и из непрерывности функций f, g вытекает, что $f(x) = \lim f(x_n) = \lim g(x_n) = g(x)$. $\square$

В процессе доказательства двух предыдущих утверждений нами была фактически установлена следующая теорема.

ТЕОРЕМА 2.4.21 (Теорема об экспоненте). Существует единственная (совпадающая с $\exp$) функция $f : \mathbb{R} \rightarrow \mathbb{R}$, обладающая свойствами:

- (i) f непрерывна;
- (ii) $f(0) = 1$, $f(1) = e$.
- (iii) $f(x + y) = f(x)f(y)$ для всех $x, y \in \mathbb{R}$.

Более того, эта функция будет строго возрастающей, положительной и $f(x) = e^x$ для всех $x \in \mathbb{Q}$.

ДОКАЗАТЕЛЬСТВО. Равенство $f(x) = e^x$ для всех $x \in \mathbb{Q}$ доказывается также, как и выше для $\exp$. Отсюда вытекает, что $f(x) = \exp(x)$ для всех $x \in \mathbb{Q}$. Следовательно, по Теореме 2.4.20 имеем, что $f(x) = \exp(x)$ для всех $x \in \mathbb{R}$. Тогда остальные утверждения про f вытекают из соответствующих (уже доказанных ранее) свойств $\exp$. $\square$

Определение $\exp(z)$ для $z \in \mathbb{C}$.

Будем записывать упорядоченные пары чисел $(a, b) \in \mathbb{R}^2$ в виде $(a, b) = a + ib$, где i — пока еще формальный символ, роль которого станет видна позднее. Определим операции сложения и умножения на этих парах чисел следующим образом:

$$\begin{aligned} (a + bi) + (c + di) &= (a + b) + (b + d)i, \\ (a + bi) \cdot (c + di) &= (ac - bd) + (ad + bc)i. \end{aligned}$$

Непосредственно проверяется, что эти операции удовлетворяют всем законам сложения и умножения (см. раздел §2.1). Не очевидным является только лишь свойство существования обратного элемента $\frac{1}{z}$. Для $z = a + bi$ полагаем по определению $|z| = \sqrt{a^2 + b^2}$ и $\bar{z} = a - bi$ (величину $|z|$ называют *модулем*, или *абсолютной величиной* числа z , а число $\bar{z}$ называется *сопряженным* к числу z). Тогда прямым вычислением проверяется, что $z\bar{z} = |z|^2$, и $z \cdot \frac{\bar{z}}{|z|^2} = 1$. Таким образом, в качестве обратного элемента можно взять $z^{-1} = \frac{\bar{z}}{|z|^2}$. Нам понадобится еще полезное тождество

$$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2. \quad (2.56)$$

Множество всех таких пар $\{a + ib : a \in \mathbb{R}, b \in \mathbb{R}\}$ со введенными операциями сложения и умножения называется *полем комплексных чисел* и обозначается через $\mathbb{C}$. Теперь можно определить понятие предела для последовательностей из комплексных чисел:

$$x_n + y_n i \rightarrow a + bi \Leftrightarrow x_n \rightarrow a \text{ и } y_n \rightarrow b,$$

а также понятие суммы ряда $f(x) = \sum_{n=0}^{\infty} a_n z^n$, с коэффициентами $a_n \in \mathbb{C}$, причем z также может принимать комплексные значения. Признаки Коши, Даламбера, теорема Мертенса и теорема о радиусе сходимости остаются справедливыми и для комплексных рядов. В частности, ряд для $\exp z$ сходится абсолютно при всех $z \in \mathbb{C}$. Это позволяет определить $\exp z$ для комплексных чисел. При этом по-прежнему будет выполняться тождество

$$\exp(z_1 + z_2) = \exp(z_1) \cdot \exp(z_2) \quad \forall z_1, z_2 \in \mathbb{C}. \quad (2.57)$$

Наша цель в этом параграфе — доказать знаменитую формулу Эйлера.

$$\exp(i\varphi) = \cos \varphi + i \sin \varphi \quad \forall \varphi \in \mathbb{R}. \quad (2.58)$$

Будем исходить из следующего, знакомого со школы, определения синуса и косинуса. Рассмотрим на плоскости (которую мы отождествляем со множеством комплексных чисел $\mathbb{C} = \{a + bi : a \in \mathbb{R}, b \in \mathbb{R}\}$) окружность единичного радиуса $|z| = 1$ с центром в начале координат $z = 0$. Отложим на этой окружности от точки $1 + 0i$ дугу длины x , двигаясь против часовой стрелки, и пусть точка окружности $A = a + bi$ будет являться другим концом этой дуги. Тогда полагают $\cos x = a$ и $\sin x = b$.

Доказательство искомой формулы (2.58) делится на ряд шагов, которые мы оформим в виде лемм.

ЛЕММА 2.4.22. Для каждого $\varphi \in \mathbb{R}$ число $\exp(i\varphi)$ лежит на единичной окружности, т.е. $|\exp(i\varphi)| = 1$.

ДОКАЗАТЕЛЬСТВО. Зафиксируем $\varphi \in \mathbb{R}$ и обозначим $z = i\varphi$. Тогда, очевидно, $\bar{z} = -z$. По определению экспоненты имеем

$$\begin{aligned} \exp z &= 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \cdots + \frac{z^n}{n!} + \cdots, \\ \overline{\exp z} &= 1 + \bar{z} + \frac{\bar{z}^2}{2} + \frac{\bar{z}^3}{6} + \cdots + \frac{\bar{z}^n}{n!} + \cdots = \exp(\bar{z}) = \exp(-z). \end{aligned}$$

Следовательно,

$$|\exp z|^2 = (\exp z) \cdot (\overline{\exp z}) = (\exp z) \cdot (\exp(-z)) \stackrel{(2.57)}{=} \exp(z - z) = 1,$$

что и требовалось доказать. $\square$

ЛЕММА 2.4.23. Справедливо асимптотическое соотношение $\exp z = 1 + z + o(z)$ при $z \rightarrow 0$, т.е. найдется функции $\alpha : (0, +\infty) \rightarrow (0, +\infty)$ такая, что $\alpha(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$ и для каждого $z \in \mathbb{C}$ если $|z| < \varepsilon$, то $|\exp(z) - 1 - z| < \alpha(\varepsilon)$.

ДОКАЗАТЕЛЬСТВО. Данное утверждение мгновенно вытекает из Теоремы 2.2.21 (см. также Замечание 2.2.22). $\square$

Параметризованной плоской кривой будем называть непрерывное отображение $\gamma : [0, T] \subset \mathbb{R} \rightarrow \mathbb{C}$. Длина этой кривой вычисляется по формуле

$$\ell(\gamma) = \sup\{|\gamma(0) - \gamma(t_1)| + |\gamma(t_1) - \gamma(t_2)| + \cdots + |\gamma(t_n) - \gamma(T)| : \quad (2.59)$$

$$0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T\}, \quad (2.60)$$

где супремум берется по всем разбиениям интервала $[0, T]$ указанного вида (2.60).

Из неравенства треугольника легко выводится, что при измельчении разбиения сумма в (2.59) возрастает, поэтому для каждого $\varepsilon > 0$ справедлива формула

$$\ell(\gamma) = \sup\{|\gamma(t_0) - \gamma(t_1)| + |\gamma(t_1) - \gamma(t_2)| + \cdots + |\gamma(t_{n-1}) - \gamma(t_n)| : \quad (2.61)$$

$$0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T; \max_{k=1, \dots, n} |t_k - t_{k-1}| < \varepsilon\}. \quad (2.62)$$

ЛЕММА 2.4.24. Зафиксируем $\varphi > 0$, $\varphi \in \mathbb{R}$ и рассмотрим параметризованную кривую $\gamma : [0, \varphi] \rightarrow \mathbb{C}$, задаваемую формулой $\gamma(t) = \exp(it)$, $t \in [0, \varphi]$. Тогда $\gamma(t)$ обходит единичную окружность против часовой стрелки, причем $\ell(\gamma) = \varphi$.

ДОКАЗАТЕЛЬСТВО. Тот факт, что $\gamma(t)$ лежит на единичной окружности для всех $t \in [0, \varphi]$, мгновенно вытекает из Леммы 2.4.22. Направление обхода против часовой стрелки вытекает из Леммы 2.4.23 при малых t (т.к. по указанной лемме $\gamma(t) - \gamma(0) = \exp(it) - 1 = it + o(t)$), а случай больших t может быть сведен к случаю с малыми t при помощи формулы $\gamma(t + \Delta t) - \gamma(t) = \exp(it)[\exp(i\Delta t) - 1]$.

Посчитаем теперь длину $\ell(\gamma)$. Для $\varepsilon > 0$ возьмем произвольно разбиение $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T$; $\max_{k=1, \dots, n} |t_k - t_{k-1}| < \varepsilon$. Тогда

$$\begin{aligned} |\gamma(t_k) - \gamma(t_{k-1})| &= |\exp(it_k) - \exp(it_{k-1})| = |\exp(it_k)| \cdot |1 - \exp(i(t_{k-1} - t_k))| = \\ &\stackrel{2.4.22}{=} |1 - \exp(i(t_{k-1} - t_k))| \stackrel{2.4.23}{=} (1 + \beta_k)(t_k - t_{k-1}), \end{aligned}$$

где $|\beta_k| < \alpha(\varepsilon)$, и α — функция из Леммы 2.4.23. Таким образом,

$$(1 - \alpha(\varepsilon))(t_k - t_{k-1}) < |\gamma(t_k) - \gamma(t_{k-1})| < (1 + \alpha(\varepsilon))(t_k - t_{k-1}).$$

Производя суммирование по всем $k = 1, \dots, n$, получаем

$$(1 - \alpha(\varepsilon))\varphi = (1 - \alpha(\varepsilon)) \sum_{k=1}^n (t_k - t_{k-1}) < \sum_{k=1}^n |\gamma(t_k) - \gamma(t_{k-1})| < (1 + \alpha(\varepsilon)) \sum_{k=1}^n (t_k - t_{k-1}) = (1 + \alpha(\varepsilon))\varphi.$$

Следовательно,

$$(1 - \alpha(\varepsilon))\varphi \leq \ell(\gamma) \leq (1 + \alpha(\varepsilon))\varphi.$$

Устремляя ε к нулю, получаем, что $\ell(\gamma) = \varphi$. □

Доказанная Лемма 2.4.24, вместе с приведенным выше "школьным" определением тригонометрических функций, влекут справедливость формулы Эйлера (2.58).

Глава 3

Дифференциальное исчисление

§1 Классические теоремы дифференциального исчисления

ТЕОРЕМА 3.1.1 (Теорема Ферма). Если функция $f : I \rightarrow \mathbb{R}$ дифференцируема на интервале $I = (a, b) \subset \mathbb{R}$ и достигает своего максимума (минимума)¹ в точке $x_* \in I$, то справедливо равенство

$$f'(x_*) = 0. \quad (3.1)$$

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы (будем доказывать ее для случая, когда x_* — точка максимума). Тогда

$$f(x_*) \geq f(x) \quad \forall x \in I. \quad (3.2)$$

По определению дифференцируемости существуют конечные пределы слева и справа

$$\lim_{x \rightarrow x_* - 0} \frac{f(x_*) - f(x)}{x_* - x} = \lim_{x \rightarrow x_* + 0} \frac{f(x_*) - f(x)}{x_* - x} = f'(x_*).$$

В силу (3.1) и того факта, что $x_* - x > 0$ при $x \rightarrow x_* - 0$ (слева), первый из указанных пределов $\lim_{x \rightarrow x_* - 0} \frac{f(x_*) - f(x)}{x_* - x}$ является неотрицательным. Аналогичные соображения позволяют заключить, что второй предел (справа) является неположительным. Поскольку эти пределы совпадают с $f'(x_*)$, имеем искомое равенство (3.1). $\square$

ТЕОРЕМА 3.1.2 (Теорема Ролля). Если функция $f : I \rightarrow \mathbb{R}$ дифференцируема на замкнутом промежутке $I = [a, b] \subset \mathbb{R}$ и дифференцируема на открытом интервале (a, b) , причем

$$f(a) = f(b) = 0, \quad (3.3)$$

то найдется точка $\xi \in (a, b)$ такая, что $f'(\xi) = 0$

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. По теореме Вейерштрасса о максимальном значении (см. Теорему 2.4.12) найдутся точки $x_{\max}, x_{\min} \in S$ такие, что $f(x_{\max}) \geq f(x)$ и $f(x_{\min}) \leq f(x)$ для всех $x \in [a, b]$. Если каждая из точек $x_{\max}$ и $x_{\min}$ совпадает с одним из концов a, b промежутка $[a, b]$, то ввиду (3.3) функция f постоянна

¹То есть, $f(x_*) = \max f(I)$ (соответственно, $f(x_*) = \min f(I)$).

на $[a, b]$, и доказывать, собственно говоря, нечего. Если же хотя бы одна из точек $x_{\max}$ или $x_{\min}$ лежит внутри интервала (a, b) , то возьмем ее в качестве ξ , и искомое равенство $f'(\xi) = 0$ мгновенно вытекает из теоремы Ферма 3.1.1. $\square$

ТЕОРЕМА 3.1.3 (Теорема Дарбу о промежуточных значениях производной). Пусть функция $f : I \rightarrow \mathbb{R}$ дифференцируема на открытом интервале $I \subset \mathbb{R}$. Тогда для любой пары точек $a, b \in I$, $a < b$, и для каждого числа y , лежащего между значениями $f'(a)$ и $f'(b)$ (т.е. y принадлежит отрезку с концами $f'(a)$ и $f'(b)$), найдется точка $\xi \in [a, b]$ такая, что $f'(\xi) = y$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Для определенности предположим, что $f'(a) < y < f'(b)$. Прибавляя к f линейную функцию (при этом производная f' изменится на константу), можно свести дело к случаю, когда $y = 0$, т.е. $f'(a) < 0 < f'(b)$. Вследствие определения производной имеем $\lim_{x \rightarrow a+0} \frac{f(x) - f(a)}{x - a} = f'(a) < 0$. Отсюда, ввиду очевидного неравенства $x - a > 0$ при $x \rightarrow a + 0$ (справа), заключаем, что

$$f(x) < f(a) \quad \forall x \in (a, a + \delta_1), \quad (3.4)$$

где $\delta_1 > 0$ — достаточно малый параметр. Аналогично, найдется число $\delta_2 > 0$ такое, что

$$f(x) > f(b) \quad \forall x \in (b - \delta_2, b). \quad (3.5)$$

По теореме Вейерштрасса об экстремальных значениях (см. Теорему 2.4.12) найдется точка $x_{\min} \in [a, b]$ такая, что $f(x_{\min}) \leq f(x)$ для всех $x \in [a, b]$. Вследствие (3.4)–(3.5) точка $x_{\min}$ не может совпадать ни с одним из значений a, b . Поэтому $x_{\min} \in (a, b)$. Тогда по теореме Ферма $f'(x_{\min}) = 0$, так что точку $x_{\min}$ можно взять в качестве искомой точки ξ . $\square$

ТЕОРЕМА 3.1.4 (Теорема Коши о промежуточном значении). Пусть функция $f : I \rightarrow \mathbb{R}$, $g : I \rightarrow \mathbb{R}$ непрерывны на замкнутом отрезке $I = [a, b] \subset \mathbb{R}$ и дифференцируемы на интервале (a, b) , причем $g'(x) \neq 0$ для всех $x \in (a, b)$. Тогда $g(a) \neq g(b)$ и найдется точка $\xi \in (a, b)$ такая, что справедливо равенство

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}. \quad (3.6)$$

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Из условия $g'(x) \neq 0$ для всех $x \in (a, b)$ вытекает, что $g(a) \neq g(b)$ (иначе мы вступаем в противоречие с теоремой Ролля 3.1.2).

Рассмотрим функцию $F : I \rightarrow \mathbb{R}$, заданную формулой $F(x) = [f(x) - f(a)](g(b) - g(a)) - [g(x) - g(a)](f(b) - f(a))$. Непосредственным вычислением проверяется, что $F(a) = F(b) = 0$. Следовательно, по теореме Ролля 3.1.2, найдется точка $\xi \in (a, b)$ такая, что $F'(\xi) = f'(\xi)(g(b) - g(a)) - g'(\xi)(f(b) - f(a)) = 0$. Нетрудно видеть, что последнее равенство, с учетом $g'(\xi) \neq 0 \neq g(b) - g(a)$, эквивалентно искомой формуле (3.6). $\square$

Очевидно, что доказанная формула (3.6) эквивалентна равенству нулю определителя

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

Это равносильно тому, что векторы $(f(b) - f(a), g(b) - g(a))$ и $(f'(\xi), g'(\xi))$ линейно зависимы. Если ввести в рассмотрение новую векторнозначную функцию $h : I \rightarrow \mathbb{R}^2$, определяемую по формуле $h(x) = (f(x), g(x))$, то последней факт можно переформулировать следующим образом.

ТЕОРЕМА 3.1.5. Пусть функция $h : I \rightarrow \mathbb{R}^2$ непрерывна на замкнутом отрезке $I = [a, b] \subset \mathbb{R}$ и дифференцируема на интервале (a, b) . Тогда найдется точка $\xi \in (a, b)$ такая, что вектора $[h(b) - h(a)]$ и $h'(\xi)$ параллельны².

СЛЕДСТВИЕ 3.1.6 (Теорема Лагранжа о промежуточном значении). Пусть функция $f : I \rightarrow \mathbb{R}$ непрерывна на замкнутом отрезке $I = [a, b] \subset \mathbb{R}$ и дифференцируема на интервале (a, b) . Тогда найдется точка $\xi \in (a, b)$ такая, что справедливо равенство

$$\frac{f(b) - f(a)}{b - a} = f'(\xi). \quad (3.7)$$

ДОКАЗАТЕЛЬСТВО. Нужно применить формулу Коши (3.6) для случая, когда $g(x) \equiv x$ и, соответственно, $g'(x) \equiv 1$. $\square$

СЛЕДСТВИЕ 3.1.7 (критерий монотонности). Пусть функция $f : I \rightarrow \mathbb{R}$ дифференцируема на открытом интервале $I \subset \mathbb{R}$, причем $f'(x) \geq 0$ для всех $x \in I$. Тогда функция f является возрастающей. Аналогично, если $f'(x) \leq 0$ на I , то функция f — убывающая.

Ниже (см. Теорему 3.3.7) мы установим аналогичный критерий монотонности при более слабых предположениях (когда функция f дифференцируема в основном).

§2 Формула Тэйлора с остаточным членом в форме Лагранжа

ТЕОРЕМА 3.2.1. Пусть функция $f : I \rightarrow \mathbb{R}$ является $(n + 1)$ -раз дифференцируемой на интервале $I \subset \mathbb{R}$. Тогда для любой пары точек $t, x \in I$ справедлива формула Тейлора с остаточным членом в форме Лагранжа:

$$f(x) = f(t) + f'(t)(x - t) + \frac{f''(t)}{2}(x - t)^2 + \cdots + \frac{f^{(n)}(t)}{n!}(x - t)^n + \frac{f^{(n+1)}(\xi)}{(n + 1)!}(x - t)^{n+1}, \quad (3.8)$$

где ξ лежит между точками t, x .

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Зафиксируем точку $x \in I$. Введем в рассмотрение новую функцию $F : I \rightarrow \mathbb{R}$, определенную по формуле

$$F(t) = f(x) - f(t) - f'(t)(x - t) - \frac{f''(t)}{2}(x - t)^2 - \cdots - \frac{f^{(n)}(t)}{n!}(x - t)^n.$$

Непосредственными вычислениями проверяется, что функция F дифференцируема по t , причем

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

По теореме Коши о среднем значении, для каждой дифференцируемой функции $g : I \rightarrow \mathbb{R}$ найдется точка ξ , лежащая между t, x , такая, что

$$\frac{F(t) - F(x)}{g(t) - g(x)} = \frac{F'(\xi)}{g'(\xi)} = -\frac{f^{(n+1)}(\xi)}{g'(\xi)n!}(x - \xi)^n.$$

²Здесь не исключается случай, когда один из этих векторов нулевой, тогда мы считаем нулевой вектор параллельным любому другому вектору

Возьмем в качестве g функцию $g(t) = (x-t)^{n+1}$. По построению $F(x) = F(t)|_{t=x} = 0 = g(x)$, и $g'(\xi) = -(n+1)(x-\xi)^n$, следовательно, формула Коши принимает вид

$$\frac{F(t)}{(x-t)^{n+1}} = \frac{f^{(n+1)}(\xi)}{(n+1)(x-\xi)^n n!} (x-\xi)^n = \frac{f^{(n+1)}(\xi)}{(n+1)!},$$

что равносильно искомой формуле (3.8). $\square$

§3 Выпуклые функции

Множество $E \subset \mathbb{R}^n$ называется *выпуклым*, если вместе с любой парой своих точек оно содержит и весь отрезок, соединяющий эти точки. Формулами это записывается следующим образом:

$$A, B \in E \Rightarrow tA + (1-t)B \in E \quad \forall t \in [0, 1]. \quad (3.9)$$

Упражнение 3.3.1. С помощью аксиомы непрерывности покажите, что множество $E \subset \mathbb{R}$ является выпуклым тогда и только тогда, когда оно представляет собой интервал, т.е. множество одного из четырех видов (a, b) , $(a, b]$, $[a, b)$, $[a, b]$, где $a, b \in \mathbb{R}$ (причем бесконечные концы не включаются в интервал).

Надграфиком функции $f : I \rightarrow \mathbb{R}$ называется множество $\{(x, y) : x \in I, \mathbb{R} \ni y \geq f(x)\}$. Функция $f : I \rightarrow \mathbb{R}$, заданная на интервале $I \subset \mathbb{R}$, называется *выпуклой*, если ее надграфик представляет собой выпуклое множество.

Говорят, что заданная на интервале функция $f : I \rightarrow \mathbb{R}$ удовлетворяет *неравенству Йенсена*, если для любого набора точек $x_1, \dots, x_n \in I$ и чисел $\tau_1, \dots, \tau_n$ таких, что $\tau_i \geq 0$ и $\sum_{i=1}^n \tau_i = 1$, имеем

$$f\left(\sum_{i=1}^n \tau_i x_i\right) \leq \sum_{i=1}^n \tau_i f(x_i).$$

ТЕОРЕМА 3.3.2. Заданная на интервале функция $f : I \rightarrow \mathbb{R}$ выпукла тогда и только тогда, когда она удовлетворяет неравенству Йенсена.

ДОКАЗАТЕЛЬСТВО. См. [3]. $\square$

ТЕОРЕМА 3.3.3 (Монотонность секущих у выпуклой функции). Если заданная на интервале функция $f : I \rightarrow \mathbb{R}$ выпукла, а точки $x_1, x_2, x_3 \in I$ таковы, что $x_1 \leq x_2 \leq x_3$, то

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}. \quad (3.10)$$

ДОКАЗАТЕЛЬСТВО. Очевидно, что достаточно доказать только первое неравенство: второе вытекает из первого и соображений симметрии. Прибавляя к f линейное отображение³, мы можем считать, не умаляя общности, что

$$f(x_1) = f(x_3) = 0. \quad (3.11)$$

³При добавлении линейного отображения функция f останется выпуклой, а все величины неравенств (3.10) изменятся на одну и ту же константу, поэтому получится эквивалентное неравенство.

Тогда в силу неравенства Йенсена получаем, что $f(x_2) \leq 0$, что влечет справедливость неравенства (3.10) (с учетом (3.11)). $\square$

Зафиксируем точку y , лежащую внутри интервала I , и для $x \in I_y := I \cap (-\infty, y)$ рассмотрим функцию $g(x) = \frac{f(y)-f(x)}{y-x}$. Если $x_1 < x_2 < y$, то, применяя Теорему 3.3.3 к тройке точек $x_1, x_2, x_3 = y$, получаем, что

$$\frac{f(y) - f(x_1)}{y - x_1} \leq \frac{f(y) - f(x_2)}{y - x_2}, \quad (3.12)$$

т.е., функция $g(x)$ является возрастающей на подинтервале I_y . С другой стороны, в силу все той же Теоремы 3.3.3, имеем $g(x) \leq \frac{f(z)-f(y)}{z-y}$ для каждого $z \in I \cap (y, +\infty)$. Итак, функция $g(x)$ является ограниченной и возрастающей на подинтервале I_y , следовательно, она имеет предел при $x \rightarrow y - 0$, который, по определению, совпадает с левой производной функции f . Значит, левая производная $f'_l(x)$ существует в каждой внутренней точке $x \in I$. Повторяя соответствующие рассуждения для правой производной, приходим к выводу, что она тоже существует в каждой внутренней точке $x \in I$. Устремляя в (3.10) $x_1 \rightarrow x_2 - 0$, $x_3 \rightarrow x_2 + 0$, получаем, что $f'_l(x_2) \leq f'_r(x_2)$ для всех $x_2 \in I$. Если теперь заменить обозначение x_2 на x , то

$$f'_l(x) \leq f'_r(x) \quad \forall x \in I \quad (3.13)$$

Далее, устремляя в (3.10) $x_2 \rightarrow x_1 + 0$, получаем, что

$$f'_r(x_1) \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1} \quad (3.14)$$

для всех $x_1 < x_3$ из I . Из неравенств (3.10) и (3.13) вытекает, что

$$f'_r(x_1) \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2} \quad (3.15)$$

для любой тройки чисел $x_1 < x_2 < x_3$ из I . Устремляя теперь в последнем неравенстве $x_2 \rightarrow x_3 - 0$, получаем, что

$$f'_r(x_1) \leq f'_l(x_3) \quad (3.16)$$

для любой пары чисел $x_1 < x_3$ из I . Суммируя сказанное, приходим к следующему утверждению.

ТЕОРЕМА 3.3.4 (Об односторонних производных выпуклой функции). Если заданная на интервале функция $f : I \rightarrow \mathbb{R}$ выпукла, то в каждой внутренней точке ее области определения существует левая $f'_l(x)$ и правая $f'_r(x)$ производные функции f , причем $f'_l(x) \leq f'_r(x)$. Более того, для каждой пары внутренних точек $x < y$ из I имеем цепочку неравенств $f'_l(x) \leq f'_r(x) \leq f'_l(y) \leq f'_r(y)$.

СЛЕДСТВИЕ 3.3.5 (Монотонность односторонних производных). Если заданная на интервале функция $f : I \rightarrow \mathbb{R}$ выпукла, то каждая из ее односторонних производных $f'_l(x)$, $f'_r(x)$ представляет собой возрастающую функцию на I .

Из Теоремы 3.3.4 вытекает, в частности, что точки x_* , где $f'_l(x_*) \neq f'_r(x_*)$, являются точками разрыва каждой из возрастающих функций $f'_l : I \rightarrow \mathbb{R}$, $f'_r : I \rightarrow \mathbb{R}$. А поскольку множество точек разрыва у любой монотонной функции не более чем счетно, получаем

ТЕОРЕМА 3.3.6 (О дифференцируемости выпуклой функции). Если заданная на интервале функция $f : I \rightarrow \mathbb{R}$ выпукла, то она дифференцируема всюду, за исключением не более чем счетного множества, причем производная f' является возрастающей функцией на своей области определения.

Ниже мы покажем, что верно и обратное к Теореме 3.3.6 утверждение.

Нам понадобится следующее определение. Будем говорить, что какое-то свойство $\mathcal{P}(x)$ точек множества X выполнено в основном, если найдется не более чем счетное множество $E \subset X$ такое, что $\mathcal{P}(x)$ истинно для всех $x \in X$.

ТЕОРЕМА 3.3.7 (Монотонность функции, дифференцируемой в основном). Пусть непрерывная на интервале функция $f : I \rightarrow \mathbb{R}$ дифференцируема в основном на I , причем $f'(x) \geq 0$ в основном. Тогда функция f является возрастающей.

ДОКАЗАТЕЛЬСТВО. По условию теоремы, функция f непрерывна на всем интервале I , и, кроме того, и существует не более чем счетное множество $E \subset I$ такое, в каждой точке $x \in I \setminus E$ определена конечная производная $f'(x) \geq 0$.

ШАГ I. Сначала мы рассмотрим упрощенный случай, когда выполнено строгое неравенство

$$f'(x) > 0 \quad \text{для всех } x \in I \setminus E. \quad (3.17)$$

Предположим, что заключение теоремы неверно, т.е. найдутся точки $a, b \in I$ такие, что $a < b$ и $f(a) > f(b)$. Поскольку множество $f(E)$ не более чем счетно, существует число $z \in (f(b), f(a)) \setminus f(E)$. По теореме Больцано–Коши о промежуточных значениях, множество $X = \{x \in [a, b] : f(x) = z\}$ непусто. Положим $x_* = \inf X$ (т.е. x_* — первая точка на оси абсцисс, при движении слева направо, где функция f принимает значение z). Докажем⁴, что $f(x_*) = z$. В самом деле, для каждого $n \in \mathbb{N}$ по свойству инфимума (см. Лемму 1.2.8) найдется точка $x_n \in [x_*, x_* + \frac{1}{n})$ такая, что $x_n \in X$, т.е. $f(x_n) = z$. Поскольку $x_n \rightarrow x_*$ при $n \rightarrow \infty$, то по критерию непрерывности Гейне (см. Теорему 2.4.9) имеем $z = f(x_n) \rightarrow f(x_*)$ при $n \rightarrow \infty$, следовательно (т.к. предел константы есть та же самая константа) верно искомое равенство $f(x_*) = z$. Отсюда и из выбора значения z , в частности, вытекает, что

$$x_* \notin E. \quad (3.18)$$

Докажем, далее, что

$$f(x) > z \quad \forall x \in [a, x_*). \quad (3.19)$$

В самом деле, по построению $f(a) > z$. Если бы имелась точка $b_1 \in [a, x_*)$ такая, что $f(b_1) < z$, то по теореме Коши о промежуточных значениях 2.4.15 нашлась бы точка $x_1 \in [a, b_1]$ такая, что $f(x_1) = z$. Однако это противоречит равенству $x_* = \inf X$, согласно которому $f(x) \neq z$ для всех $x \in [a, x_*)$. Тем самым формула (3.19) доказана.

Вследствие предположения (3.17) и свойства (3.18) существует положительный предел $\lim_{x \rightarrow x_* - 0} \frac{f(x_*) - f(x)}{x_* - x} = f'(x_*) > 0$. Однако это противоречит (3.19) и равенству $f(x_*) = z$. Полученное противоречие завершает доказательство теоремы в первом упрощенном случае (когда выполнены строгие неравенства (3.17)).

ШАГ II. Рассмотрим теперь общий случай, когда выполнено нестрогое неравенство $f'(x) \geq 0$ для всех $x \in I \setminus E$. Мы сведем его к предыдущему. Для параметра $\sigma > 0$

⁴На самом деле доказываемое равенство $f(x_*) = z$ по построению равносильно включению $x_* \in X$ и мгновенно следует из того факта, что X — замкнутое множество (как прообраз замкнутого, см. Теорему 2.4.14) и потому $x_* = \inf X \in \bar{X} = X$ (см. Теорему 2.3.1-(ii)).

рассмотрим функцию $f_\sigma(x) := f(x) + \sigma x$. По построению функция $f_\sigma(x)$ непрерывна и для всех $x \in I \setminus E$ справедлива оценка $f'_\sigma(x) = f'(x) + \sigma > 0$, т.е. выполнены предположения Шага I для функции f_σ . Значит, эта функция является возрастающей, т.е.

$$f_\sigma(x_1) \leq f_\sigma(x_2) \quad (3.20)$$

для любой пары точек $x_1 < x_2$ из I . Устремляя в неравенстве (3.20) параметр σ к нулю и учитывая, что $\lim_{\sigma \rightarrow 0} f_\sigma(x) = \lim_{\sigma \rightarrow 0} [f(x) + \sigma x] = f(x)$ для всех $x \in I$, получаем, что

$$f(x_1) \leq f(x_2) \quad (3.21)$$

для любой пары точек $x_1 < x_2$ из I , т.е. функция f в самом деле является возрастающей. $\square$

Из этой теоремы вытекает часто применяемое в дальнейшем

СЛЕДСТВИЕ 3.3.8. Пусть непрерывная на интервале функция $f : I \rightarrow \mathbb{R}$ дифференцируема в основном на I , причем $f'(x) = 0$ в основном. Тогда функция f постоянна на I .

Доказанная теорема позволяет нам установить следующий красивый критерий выпуклости.

ТЕОРЕМА 3.3.9 (О дифференцируемости выпуклой функции). Пусть функция $f : I \rightarrow \mathbb{R}$ непрерывна на всем интервале $I \subset \mathbb{R}$ и дифференцируема в основном на I , причем производная f' является возрастающей функцией на своей области определения. Тогда функция f выпукла.

ДОКАЗАТЕЛЬСТВО. По условию теоремы, функция f непрерывна на всем интервале I , и, кроме того, существует множество $T \subset I$ такое, разность $I \setminus T$ представляет собой не более чем счетное множество, и в каждой точке $x \in T$ определена конечная производная $f'(x)$, причем

$$f'(x) \leq f'(y) \quad \text{для любой пары точек } x < y \text{ из } T. \quad (3.22)$$

Предположим, что утверждение теоремы не выполнено, т.е. функция f не выпукла, тогда найдутся точки $a, b \in I$ такие, что

(*) отрезок, соединяющий точки плоскости $(a, f(a))$ и $(b, f(b))$, не содержится в надграфике функции f .

Добавляя к f линейное отображение (которое не влияет ни на выпуклость, ни на дифференцируемость в основном, ни на монотонность производной), мы можем считать, не умаляя общности, что $f(a) = f(b) = 0$. Тогда свойство (*) перепишется следующим образом:

(**) существует точка $c \in (a, b)$ такая, что $f(c) > 0$.

Тогда по теореме Вейерштрасса о максимальном значении, найдется точка $c_0 \in (a, b)$ такая, что $f(c_0) = \max f([a, b]) > 0$. Далее, по теореме Вейерштрасса о пределе монотонной функции, существует предел слева

$$A := \lim_{t \rightarrow c_0 - 0} f'(t). \quad (3.23)$$

Докажем, что

$$A \geq 0. \quad (3.24)$$

Предположим противное, т.е. что $A < 0$. Тогда по теореме об устойчивости строгих неравенств при предельном переходе (см. Теорему 2.3.8) найдется точка $t_1 < c_0$ такая, что $f'(x) < 0$ для всех $x \in (t_1, c_0)$. По теореме о монотонности функции, дифференцируемой в основном (см. Теорему 3.3.7, примененную к $-f$), функция f на интервале $[t_1, c_0]$ является убывающей. Отсюда, ввиду того, что c_0 является точкой максимума f , заключаем, что $f(t) \equiv f(c_0)$ для всех $t \in [t_1, c_0]$. Тогда $f'|_{[t_1, c_0]} \equiv 0$, что противоречит предположению $A < 0$. Полученное противоречие завершает доказательство неравенства (3.24).

Из доказанного неравенства (3.24) и возрастания функции $f' : T \rightarrow \mathbb{R}$ вытекает, что $f'(x) \geq 0$ для всех $x \in T \cap [c_0, b]$. Снова применяя Теорему 3.3.7, получаем, что функция f возрастает на промежутке $[c_0, b]$. Однако это противоречит имеющимся соотношениям $f(c_0) > 0 = f(b)$. Полученное противоречие завершает доказательство теоремы. $\square$

§4 Бесконечные произведения

Пусть дана последовательность чисел b_n . Обозначим через P_n произведение $P_n = b_1 b_2 \dots b_n = \prod_{k=1}^n b_k$. Будем говорить, что *бесконечное произведение* $\prod_{k=1}^{\infty} b_k$ *сходится*, если существует *ненулевой конечный предел* $\lim_{n \rightarrow \infty} P_n \in \mathbb{R} \setminus \{0\}$.

ЛЕММА 3.4.1 (Необходимый признак сходимости бесконечного произведения). Если произведение $\prod_{k=1}^{\infty} b_k$ сходится, то $\lim_{n \rightarrow \infty} b_n = 1$.

ДОКАЗАТЕЛЬСТВО. Предоставляется учащимся в качестве упражнения. $\square$

Ввиду последней леммы, у сходящегося произведения все множители b_n , начиная с некоторого, имеют положительный знак. Поэтому, не умаляя общности, обычно мы будем иметь дело с произведениями, у которых все множители положительны.

ТЕОРЕМА 3.4.2 (Логарифмический критерий сходимости бесконечного произведения). Пусть имеется последовательность положительных чисел b_k , $k \in \mathbb{N}$. Тогда бесконечное произведение $\prod_{k=1}^{\infty} b_k$ сходится в том и только том случае, если ряд $\sum_{k=1}^{\infty} \ln b_k$ сходится.

ДОКАЗАТЕЛЬСТВО. ($\Rightarrow$) Пусть бесконечное произведение $\prod_{k=1}^{\infty} b_k$ сходится, т.е. последовательность чисел $P_n = b_1 b_2 \dots b_n$ имеет ненулевой конечный предел, обозначим его символом P . Рассмотрим последовательность $S_n = \ln P_n = \ln b_1 + \dots + \ln b_n$. Поскольку функция $\ln$ непрерывна в точке $P > 0$, и $P_n \rightarrow P$ при $n \rightarrow \infty$, то в силу критерия Гейне (см. Теорему 2.4.9) имеем $S_n = \ln P_n \rightarrow \ln P$ при $n \rightarrow \infty$. Последнее означает, что ряд $\sum_{k=1}^{\infty} \ln b_k$ сходится, что и требовалось доказать.

($\Leftarrow$) Пусть ряд $\sum_{k=1}^{\infty} \ln b_k$ сходится. Это означает, что последовательность $S_n = \ln P_n = \ln b_1 + \dots + \ln b_n$ имеет конечный предел $S_n \rightarrow S \in \mathbb{R}$ при $n \rightarrow \infty$. Поскольку функция $\exp$ непрерывна в точке S , то в силу критерия Гейне (см. Теорему 2.4.9) имеем $P_n = \exp S_n \rightarrow \exp S$ при $n \rightarrow \infty$. Так как $\exp S$ представляет собой положительное число, по определению получаем, что бесконечное произведение P_n также сходится. $\square$

Ввиду доказанной теоремы, нам понадобятся следующие оценки логарифмической функции.

ЛЕММА 3.4.3. Если $|x| < \frac{1}{3}$, то справедлива оценка

$$\ln(1+x) = x(1+\alpha(x)), \quad (3.25)$$

где

$$|\alpha(x)| \leq \frac{3}{8}|x|. \quad (3.26)$$

В частности, $\ln(1+x)$ имеет тот же знак, что и x , причем

$$\frac{5}{8}|x| \leq |\ln(1+x)| \leq \frac{11}{8}|x|. \quad (3.27)$$

ДОКАЗАТЕЛЬСТВО. Напомним, что если функция $f : I \rightarrow \mathbb{R}$ дважды дифференцируема на интервале $I \subset \mathbb{R}$, то для любой пары точек $t, x \in I$ справедлива формула Тейлора с остаточным членом в форме Лагранжа.

$$f(x) = f(t) + f'(t)(x-t) + \frac{f''(\xi)}{2}(x-t)^2, \quad (3.28)$$

где ξ лежит между точками t, x . Применим эту формулу к функции $f(x) = \ln(1+x)$ при $t = 0$. Имеем $f'(x) = \frac{1}{1+x}|_{x=0} = 1$, $f''(x) = -\frac{1}{(1+x)^2}$, поэтому

$$\ln(1+x) = x + \frac{1}{2(1+\xi)^2}x^2, \quad (3.29)$$

где $|\xi| < |x|$. Непосредственным вычислением проверяется, что при $|x| < \frac{1}{3}$ справедлива оценка $\frac{1}{2(1+\xi)^2}x < \frac{3}{8}$, что равносильно искомым соотношениям (3.25)–(3.26). $\square$

ТЕОРЕМА 3.4.4 (Арифметический критерий сходимости бесконечного произведения). Пусть имеется последовательность чисел $\beta_k \neq -1$, $k \in \mathbb{N}$. Тогда, если ряд $\sum_{k=1}^{\infty} |\beta_k|$ сходится, то бесконечное произведение $\prod_{k=1}^{\infty} (1+\beta_k)$ также сходится. Далее, если все числа β_k , начиная с некоторого, одного знака, то верно и обратное: ряд $\sum_{k=1}^{\infty} \beta_k$ сходится в том и только том случае, если бесконечное произведение $\prod_{k=1}^{\infty} (1+\beta_k)$ сходится.

ДОКАЗАТЕЛЬСТВО. ($\Rightarrow$) Докажем первое утверждение теоремы. Пусть ряд $\sum_{k=1}^{\infty} |\beta_k|$ сходится. Тогда $|\beta_k| \leq \frac{1}{3}$ при всех $k \geq k_0$. В силу (3.27) имеем

$$\frac{5}{8}|\beta_k| \leq |\ln(1+\beta_k)| \leq \frac{11}{8}|\beta_k|. \quad (3.30)$$

Значит, по мажорируемому признаку сходимости ряд $\sum_{k=k_0}^{\infty} \ln(1+\beta_k)$ также сходится, что вследствие Теоремы 3.4.2 влечет сходимость бесконечного произведения $\prod_{k=1}^{\infty} (1+\beta_k)$.

($\Leftarrow$) Предположим теперь, что все β_k при $k \geq k_0$ имеют одинаковый знак, и что бесконечное произведение $\prod_{k=1}^{\infty} (1+\beta_k)$ сходится. Требуется доказать, что ряд $\sum_{k=1}^{\infty} \beta_k$ также сходится. В силу необходимого признака сходимости бесконечного произведения (см. Лемму 3.4.1) имеем $\beta_k \rightarrow 0$ при $k \rightarrow \infty$, поэтому можно считать, не умаляя общности, что

$|\beta_k| \leq \frac{1}{3}$ при всех $k \geq k_0$. Тогда вследствие логарифмического признака (Теорема 3.4.1) ряд $\sum_{k=k_0}^{\infty} \ln(1 + \beta_k)$ сходится. Поскольку все β_k имеют одинаковый знак при $k \geq n_0$, то и $\ln(1 + \beta_k)$ имеют тот же знак при $k \geq k_0$, следовательно, ряд $\sum_{k=k_0}^{\infty} |\ln(1 + \beta_k)|$ тоже сходится. Теперь сходимость ряда $\sum_{k=k_0}^{\infty} |\beta_k|$ непосредственно вытекает из первого неравенства в (3.30) и мажорируемого признака сходимости (для рядов). $\square$

Пример 3.4.5. Пусть $p_1, p_2, p_3, \dots$ — совокупность всех простых чисел, занумерованных в порядке возрастания (т.е. $p_1 = 2, p_2 = 3$ и т.д.). Рассмотрим функцию

$$\zeta(s) = \prod_{k=1}^{\infty} \frac{1}{1 - p_k^{-s}}, \quad (3.31)$$

которая называется *дзета-функцией Римана* и играет фундаментальную роль в аналитической теории чисел и в других областях математики. Обозначим $b_k(s) = \frac{1}{1 - p_k^{-s}}$. По формуле геометрической прогрессии при $s > 0$ имеем

$$b_k(s) = 1 + \frac{1}{p_k^s} + \frac{1}{p_k^{2s}} + \frac{1}{p_k^{3s}} + \frac{1}{p_k^{4s}} + \dots$$

Нашей ближайшей целью будет доказательство для $s \geq 0$ следующей классической красивой формулы Эйлера:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (3.32)$$

Обозначим

$$P_m(s) = \prod_{k=1}^m b_k(s) = \prod_{k=1}^m \left[1 + \frac{1}{p_k^s} + \frac{1}{p_k^{2s}} + \frac{1}{p_k^{3s}} + \dots \right]. \quad (3.33)$$

По основной теореме арифметики, всякое натуральное число n однозначно представляется в виде произведения простых чисел $n = p_{i_1}^{\alpha_1} \dots p_{i_k}^{\alpha_k}$, поэтому

(*) при раскрытии скобок в произведении (3.33) получим сумму $-s$ -х степеней всех натуральных чисел, в разложении которых присутствуют числа $p_1, \dots, p_m$ (во всевозможных степенях).

Следовательно,

$$P_m(s) \leq \sum_{n=1}^{\infty} \frac{1}{n^s},$$

и, переходя здесь к пределу при $m \rightarrow \infty$, получаем

$$\zeta(s) \leq \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (3.34)$$

С другой стороны, легко проверяется оценка $p_k \geq k$, поэтому все простые множители чисел $\{1, 2, \dots, m\}$ содержатся во множестве $\{p_1, \dots, p_m\}$, следовательно, привлекая снова (*), получаем

$$P_m(s) \geq \sum_{n=1}^m \frac{1}{n^s},$$

и, переходя опять к пределу при $m \rightarrow \infty$, имеем

$$\zeta(s) \geq \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (3.35)$$

Неравенства (3.34)–(3.35) дают справедливость искомого тождества (3.32).

Из формулы (3.32) вытекает множество следствий. Отметим только одно из них. Мы уже знаем про ряд $\sum_{n=1}^{\infty} \frac{1}{n^s}$, что он сходится при $s > 1$ и расходится при $s \leq 1$ (см. Следствие 2.2.10, вытекающее из телескопического признака). Отсюда сразу же получаем

СЛЕДСТВИЕ 3.4.6. Бесконечное произведение в определении (3.31) дзета функции $\zeta(s)$ сходится при $s > 1$ и расходится при $s \leq 1$.

Попробуем переформулировать последнее следствие, используя арифметический критерий сходимости бесконечных произведений 3.4.4. Обозначим $\beta_k(s) = b_k(s) - 1 = \frac{1}{p_k^s} + \frac{1}{p_k^{2s}} + \frac{1}{p_k^{3s}} + \frac{1}{p_k^{4s}} + \dots$, в частности, $\zeta(s) = \prod_{k=1}^{\infty} b_k(s) = \prod_{k=1}^{\infty} (1 + \beta_k(s))$. Из Теоремы 3.4.4 и Следствия 3.4.6 мгновенно получаем

СЛЕДСТВИЕ 3.4.7. Ряд $\sum_{k=1}^{\infty} \beta_k(s)$ сходится при $s > 1$ и расходится при $s \leq 1$.

Нетрудно доказать, используя формулу для геометрической прогрессии, что $\frac{1}{p_k^s} \leq \beta_k(s) \leq \frac{2}{p_k^s}$. Поэтому формулировку Следствия 3.4.7 можно упростить следующим образом

СЛЕДСТВИЕ 3.4.8. Ряд $\sum_{k=1}^{\infty} \frac{1}{p_k^s}$ сходится при $s > 1$ и расходится при $s \leq 1$. В частности,

$$\sum_{k=1}^{\infty} \frac{1}{p_k} = \infty. \quad (3.36)$$

Данное следствие интересно в том отношении, что, согласно теореме П.Л. Чебышева (основателя Петербургской математической школы)

$$\pi(x) \sim \frac{x}{\ln x}, \quad (3.37)$$

где $\pi(x)$ — количество простых чисел, не превосходящих x . (Отметим, что в завершающей стадии доказательства данной теоремы П.Л. Чебышева приняли участие зарубежные математики Адамар и Валле-Пуссен.) Таким образом, $\frac{\pi(x) \ln x}{x} \rightarrow 1$ при $x \rightarrow \infty$, т.е. плотность множества простых чисел в отрезке $[1, x]$ асимптотически совпадает с величиной $\frac{1}{\ln x}$, в частности, она стремится к нулю при $x \rightarrow \infty$, что делает результат о расходимости ряда (3.36) еще более ценным.

Упражнение 3.4.9. Используя теорему П.Л. Чебышева, докажите следующее утверждение (так называемый *постулат Бертра*на):

для любого $\alpha > 1$ найдется $n_\alpha \in \mathbb{N}$ такое, что для всех натуральных чисел $n \geq n_\alpha$ между n и αn лежит по крайней мере одно простое число.

Упражнение 3.4.10. Выведите расходимость ряда (3.36) из теоремы П.Л. Чебышева, т.е. покажите, что для всякой числовой последовательности $x_k > 0$ такой, что

$$\pi_*(x) \sim \frac{x}{\ln x}, \quad (3.38)$$

где $\pi_*(x)$ — количество чисел x_k , не превосходящих x , справедлива формула

$$\sum_{k=1}^{\infty} \frac{1}{x_k} = \infty. \quad (3.39)$$

При возникновении трудностей с этим упражнением, можно использовать

Указание. Вычислите асимптотику сумм

$$t_n = \sum_{x_k \in [2^n, 2^{n+1})} \frac{1}{x_k}$$

и докажите, что найдутся положительные параметры α, β такие, что

$$\frac{\alpha}{n} \leq t_n \leq \frac{\beta}{n}.$$

§4.1 Приложения к сходимости рядов: модификации мажорируемого признака сходимости

Как известно, согласно признаку Даламбера (2.2.8) ряд с положительными слагаемыми $\sum_{n=1}^{\infty} c_n$ сходится, если $\lim_{n \rightarrow \infty} \frac{c_n}{c_{n+1}} > 1$, и расходится, если $\lim_{n \rightarrow \infty} \frac{c_n}{c_{n+1}} < 1$. В данном разделе мы постараемся "выковать" инструмент, позволяющий исследовать на сходимость ряды, у которых

$$\lim_{n \rightarrow \infty} \frac{c_n}{c_{n+1}} = 1. \quad (3.40)$$

Пусть имеются последовательности чисел a_n, c_n . Ниже мы часто будем налагать на них следующее условие:

(*) $a_n \neq 0, c_n > 0, \lim_{n \rightarrow \infty} \frac{c_n}{c_{n+1}} = 1$ и справедливо тождество

$$\frac{a_n}{a_{n+1}} = \frac{c_n}{c_{n+1}} + \beta_n, \quad (3.41)$$

где величины β_n таковы, что ряд $\sum_{n=1}^{\infty} |\beta_n|$ сходится.

ЛЕММА 3.4.11. При выполнении условия (*) найдется число $\alpha \neq 0$ такое, что $\lim_{n \rightarrow \infty} \frac{a_n}{c_n} = \alpha$.

ДОКАЗАТЕЛЬСТВО. Пусть для последовательностей чисел a_n, c_n выполнено условие (*). Обозначим $\alpha_n = \frac{a_n}{c_n}$. Поделив обе части равенства (3.41) на $\frac{c_n}{c_{n+1}}$, получим

$$\frac{\alpha_n}{\alpha_{n+1}} = 1 + \tilde{\beta}_n, \quad (3.42)$$

где $\tilde{\beta}_n = \beta_n \frac{c_{n+1}}{c_n}$. Поскольку $\frac{c_{n+1}}{c_n} \rightarrow 1$ при $n \rightarrow \infty$, то ряд $\sum_{n=1}^{\infty} |\tilde{\beta}_n|$ также является сходящимся. Так как по построению все числа α_n ненулевые, то $\tilde{\beta}_n \neq -1$. По арифметическому критерию (см. Теорему 3.4.4) получаем, что бесконечное произведение $\prod_{n=1}^{\infty} \frac{\alpha_n}{\alpha_{n+1}}$ сходится, т.е. найдется число $\beta \neq 0$ такое, что

$$\prod_{n=1}^m \frac{\alpha_n}{\alpha_{n+1}} \rightarrow \beta \quad (3.43)$$

при $m \rightarrow \infty$. Непосредственным вычислением проверяется, что

$$\prod_{n=1}^m \frac{\alpha_n}{\alpha_{n+1}} = \frac{\alpha_1}{\alpha_2} \cdot \frac{\alpha_2}{\alpha_3} \cdot \dots \cdot \frac{\alpha_m}{\alpha_{m+1}} = \frac{\alpha_1}{\alpha_{m+1}},$$

так что формулу (3.43) можно переписать в виде

$$\lim_{m \rightarrow \infty} \frac{\alpha_1}{\alpha_{m+1}} = \beta, \quad (3.44)$$

следовательно,

$$\lim_{m \rightarrow \infty} \alpha_m = \frac{\beta}{\alpha_1}, \quad (3.45)$$

что, с учетом равенства $\alpha_m = \frac{a_m}{c_m}$, означает справедливость утверждения доказываемой леммы. $\square$

Из доказанной леммы и мажорируемого признака сходимости рядов легко выводится

ТЕОРЕМА 3.4.12. При выполнении условия (*) ряды $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} c_n$ сходятся или расходятся одновременно.

Рассмотрим в качестве модельных примеров два ряда: $\sum_{n=1}^{\infty} \frac{1}{n^p}$ и $\sum_{n=1}^{\infty} \frac{1}{n \ln^p n}$. Известно, что каждый из этих рядов сходится при $p > 1$ и расходится при $p \leq 1$. Важно оценить асимптотику $\frac{c_n}{c_{n+1}}$.

Упражнение 3.4.13. Докажите, что при $c_n = \frac{1}{n^p}$ справедлива формула

$$\frac{c_n}{c_{n+1}} = 1 + \frac{p}{n} + O\left(\frac{1}{n^2}\right),$$

а при $d_n = \frac{1}{n \ln^p n}$ справедлива формула

$$\frac{d_n}{d_{n+1}} = 1 + \frac{1}{n} + \frac{p}{n \ln n} + O\left(\frac{1}{n^2}\right).$$

Из данных оценок и из Теоремы 3.4.12 легко выводятся

СЛЕДСТВИЕ 3.4.14 (Признак Гаусса). Пусть имеются последовательности чисел $a_n > 0$ и $\beta_n \in \mathbb{R}$, причем ряд $\sum_{n=1}^{\infty} |\beta_n|$ сходится. Тогда, если $\frac{a_n}{a_{n+1}} \leq 1 + \frac{1}{n} + \beta_n$ для всех $n \in \mathbb{N}$, то ряд $\sum_{n=1}^{\infty} a_n$ расходится. Если же найдется $p > 1$ такое, что $\frac{a_n}{a_{n+1}} \geq 1 + \frac{p}{n} + \beta_n$ для всех $n \in \mathbb{N}$, то ряд $\sum_{n=1}^{\infty} a_n$ сходится.

СЛЕДСТВИЕ 3.4.15 (Признак Бертрانا). Пусть имеются последовательности чисел $a_n > 0$ и $\beta_n \in \mathbb{R}$, причем ряд $\sum_{n=1}^{\infty} |\beta_n|$ сходится. Тогда, если $\frac{a_n}{a_{n+1}} \leq 1 + \frac{1}{n} + \frac{1}{n \ln n} + \beta_n$ для всех $n \in \mathbb{N}$, то ряд $\sum_{n=1}^{\infty} a_n$ расходится. Если же найдется $p > 1$ такое, что $\frac{a_n}{a_{n+1}} \geq 1 + \frac{1}{n} + \frac{p}{n \ln n} + \beta_n$ для всех $n \in \mathbb{N}$, то ряд $\sum_{n=1}^{\infty} a_n$ сходится.

Упражнение 3.4.16 (Признак Куммера). Пусть имеются последовательности чисел $a_n > 0$ и $d_n > 0$, причем ряд $\sum_{n=1}^{\infty} \frac{1}{d_n}$ расходится. Предположим, что существует предел $\lim_{n \rightarrow \infty} [d_n \frac{a_n}{a_{n+1}} - d_{n+1}] = K$. Докажите, что ряд $\sum_{n=1}^{\infty} a_n$ сходится при $K > 0$ и расходится при $K < 0$. При возникновении трудностей с этим упражнением, можно использовать

Указание. Докажите, что при $K \neq 0$ последовательность $a_n d_n$ монотонна, и, следовательно, имеет (конечный или бесконечный) предел λ . Покажите, что $\lambda > 0$ при $K < 0$ и выведите отсюда расходимость $\sum_{n=1}^{\infty} a_n$. Покажите, что $\lambda \in \mathbb{R}$ (т.е. указанный предел конечен) при $K < 0$, и оцените a_{n+1} через разность $a_n d_n - a_{n+1} d_{n+1}$.

Упражнение 3.4.17. Выведите признаки сходимости Гаусса и Бертрана из признака Куммера.

§5 Приложения дифференциального исчисления

§5.1 Формула Циолковского

Рассмотрим простейший случай, когда ракета летит по прямой в космическом пространстве (без учета сил тяготения). Начальную скорость и массу ракеты обозначим символами V_0 и M_0 , скорость и масса ракеты в момент времени t будут обозначаться соответственно $V(t)$ и $M(t)$. Изменение скорости достигается за счет реактивного двигателя, извергающего раскаленные газы (полученные путем сжигания топлива) со скоростью ω в сторону, противоположную движению ракеты. Задача состоит в том, чтобы определить скорость ракеты после сжигания заданного количества топлива.

Зафиксируем произвольный момент времени t и рассмотрим, как изменятся интересующие нас величины за малый промежуток времени Δt . Основополагающую роль в приводимых ниже вычислениях играет закон сохранения импульса. Согласно этому закону, изменение импульса ракеты $M(t + \Delta t)V(t + \Delta t) - M(t)V(t)$ совпадает с импульсом испущенных газов. Но этот последний импульс равен $[M(t) - M(t + \Delta t)](\omega - V(t) + o(1))$, где $o(1) \rightarrow 0$ при $\Delta t \rightarrow 0$. Таким образом, получаем соотношение

$$\begin{aligned} [M(t) - M(t + \Delta t)](\omega - V(t) + o(1)) &= M(t + \Delta t)V(t + \Delta t) - M(t)V(t) = \\ &= [M(t + \Delta t) - M(t)]V(t) + [V(t + \Delta t) - V(t)]M(t + \Delta t), \end{aligned}$$

и, после сокращений, получаем

$$[M(t + \Delta t) - M(t)](-\omega + o(1)) = [V(t + \Delta t) - V(t)]M(t + \Delta t).$$

Теперь возьмем в качестве независимой переменной не время, а массу, т.е. рассмотрим зависимость $V = V(M)$. Тогда, поделив выписанное равенство на $\Delta M = [M(t + \Delta t) - M(t)]$, и устремив Δt к нулю, получим

$$V'_M = -\frac{\omega}{M}.$$

Следовательно, $V = \omega \ln \frac{1}{M} + C$, где константа C определяется из начальных условий $V(M_0) = V_0$. Окончательно получаем:

$$V - V_0 = \omega \ln \frac{M_0}{M},$$

это и есть знаменитая формула Циолковского. Ее можно переписать еще в виде:

$$M = M_0 \exp\left(\frac{V_0 - V}{\omega}\right).$$

Эта формула демонстрирует печальную закономерность: при линейном увеличении скорости ракеты, полезная масса (т.е. масса ракеты за вычетом топлива) уменьшается экспоненциально быстро. Впрочем, увеличить полезную массы можно за счет роста скорости сгораемого топлива ω .

Упражнение 3.5.1. Выведите аналог формулы Циолковского в рамках специальной теории относительности, используя релятивистский закон сохранения импульса и соответствующую формулу сложения скоростей.

§5.2 Падение с парашютом

Если тело падает по прямой в вакууме, то его координата $x = x(t)$ подчиняется закону Ньютона $\ddot{x} = \dot{v} = g$, откуда $v = gt + v_0$, $x(t) = \frac{gt^2}{2} + v_0t + x_0$, здесь v_0, x_0 — скорость тела и его положение в момент времени $t = 0$ соответственно, а символ $\dot{v}$ обозначает производную функции $v = v(t)$ по времени t . Если же падение происходит в воздухе, то появляется сила сопротивления, прямо пропорциональная скорости, и уравнение принимает вид:

$$\dot{v} = g - \alpha v, \quad (3.46)$$

где коэффициент $\alpha > 0$ зависит от плотности воздуха и характеристик тела (например, от купола парашюта). Уравнение (3.46) решается в два этапа. Сначала ищется решение так называемого однородного уравнения $\dot{v} = -\alpha v$, легко проверяется, что оно задается формулой $v(t) = Ce^{-\alpha t}$, где $C = \text{const}$. Далее решение исходного уравнения (3.46) ищется в виде

$$v = C(t)e^{-\alpha t}, \quad (3.47)$$

здесь C перестает быть константой, и является уже функцией, зависящей от t , почему этот метод и получил название "метод вариации произвольных постоянных". Дифференцируя формулу (3.47), имеем

$$\dot{v} = \dot{C}e^{-\alpha t} - C\alpha e^{-\alpha t} = \dot{C}e^{-\alpha t} - \alpha C e^{-\alpha t} = \dot{C}e^{-\alpha t} - \alpha v. \quad (3.48)$$

Сравнивая последнюю формулу с уравнением (3.46), получаем уравнение на $C(t)$:

$$\dot{C}e^{-\alpha t} = g \Leftrightarrow \dot{C} = ge^{\alpha t},$$

откуда

$$C(t) = \frac{g}{\alpha} e^{\alpha t} + C_0,$$

где C_0 — произвольная константа. Подставляя последнее равенство в формулу (3.47), приходим к окончательной формуле $v(t) = \frac{g}{\alpha} + C_0 e^{-\alpha t}$, где $C_0 = \text{const}$. Отсюда видно, что с течением времени скорость падения стремится к постоянной величине $\frac{g}{\alpha}$. Уменьшить эту предельную скорость можно за счет увеличения коэффициента α , который зависит от формы тела и т.п.

§5.3 Формула радиоактивного распада. Критический радиус.

Целью данного параграфа является вывод уравнения, моделирующего цепную ядерную реакцию. В ходе этой реакции из атомов радиоактивного вещества высвобождаются нейтроны, обозначим их количество (зависящее от времени) через $N(t)$. Известно, что скорость распада в простейшем случае прямо пропорционально количеству имеющихся свободных нейтронов (которые инициируют распад "соседних" атомов и т.д.), поэтому получаем уравнение $\dot{N} = \alpha N$. Однако необходимо учесть, что часть нейтронов будет "вылетать" за границу тела, и поэтому они перестают участвовать в реакции. Естественным предположить, что количество "улетевших" за единицу времени нейтронов пропорционально величине ρS , где ρ — объемная плотность нейтронов, а S — площадь границы тела. Считая, что тело является шаром радиуса r , получаем, что, что $\rho = \frac{3N}{4\pi r^3}$, $S = 4\pi r^2$, так что количество "улетевших" за единицу времени нейтронов будет пропорционально величине $\frac{N}{r}$. Также будем считать, что для усиления реакции тело подвергается воздействию

нейтронного пучка от внешнего излучателя, скорость "бомбардировки" этим пучком обозначим через n (это количество поступивших за единицу времени частиц извне). В итоге получаем уравнение

$$\dot{N} = \alpha N - \beta \frac{N}{r} + n = \left(\alpha - \frac{\beta}{r}\right)N + n. \quad (3.49)$$

Это уравнение может быть решено тем же способом, что и уравнение падения с парашютом (см. предыдущий параграф). В итоге получаем, что

$$N(t) = \frac{n}{\frac{\beta}{r} - \alpha} + C_0 e^{(\alpha - \frac{\beta}{r})t},$$

где $C_0 = \text{const}$. В зависимости от значения параметров возможны три сценария поведения решения.

СЛУЧАЙ I (радиус меньше критического). Если $r < \frac{\beta}{\alpha}$, то показатель у экспоненты отрицательный, поэтому экспонента стремится к нулю при $t \rightarrow \infty$, и, соответственно, количество нейтронов стремится к постоянной величине: $N(t) \rightarrow \frac{n}{\frac{\beta}{r} - \alpha}$ при $t \rightarrow \infty$.

СЛУЧАЙ II (радиус равен критическому). Если $r = \frac{\beta}{\alpha}$, уравнение (3.49) принимает упрощенный вид $\dot{N} = n$, и его решение является линейной функцией $N(t) = nt + N_0$. Количество нейтронов при этом неограниченно возрастает, но не слишком быстро (с линейной скоростью).

СЛУЧАЙ III (радиус больше критического). Если $r > \frac{\beta}{\alpha}$, то показатель у экспоненты положительный, и, соответственно, количество нейтронов экспоненциально возрастает с увеличением времени. Это может привести к взрыву.

Для урана критический радиус равен приблизительно 8.5 см, а масса такого "критического" шарика составляет примерно 50 кг.

Глава 4

Теория интеграла

Будем говорить, что некоторое свойство $P(x)$ выполняется "в основном" на множестве $E \subset \overline{\mathbb{R}}$, если это свойство выполняется для всех точек $x \in E$ за исключением некоторого не более чем счетного множества Σ . Этот факт будет обозначаться символами $P(x) \forall_0 x \in E$. На языке формул,

$P(x)$ истинно $\forall_0 x \in E \stackrel{\text{def}}{\Leftrightarrow} \exists \Sigma \subset E : \Sigma \text{ не более чем счетно и } P(x) \text{ истинно } \forall x \in E \setminus \Sigma.$

Пример 4.0.2. (i) Если $P(x) \forall x \in E$, то $P(x) \forall_0 x \in E$;

(ii) $(|x|)' = \text{sign } x \quad \forall_0 x \in \mathbb{R}$;

(iii) $(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} \quad \forall_0 x \in [-1, 1]$;

(iv) ввиду счетности множества рациональных чисел, для функции Дирихле

$$\chi(x) = \begin{cases} 1, & \text{если } x \in \mathbb{Q}; \\ 0 & \text{иначе} \end{cases}$$

справедливо утверждение $\chi(x) = 0 \quad \forall_0 x \in \mathbb{R}$.

Функция $F : U \rightarrow \mathbb{R}$ называется *обобщенной первообразной* функции $f : U \rightarrow \mathbb{R}$ на множестве $U \subset \overline{\mathbb{R}}$, если выполнены следующие два условия:

(A) F непрерывна на U ;

(B) F дифференцируема в основном на U причем $F'(x) = f(x) \quad \forall_0 x \in E$.

Соответственно предыдущим примерам, получаем

Пример 4.0.3. (i) Если $F(x)$ является первообразной функции f на U , то она является и обобщенной первообразной;

(ii) $|x|$ является обобщенной первообразной функции $\text{sign } x$ на всем $\mathbb{R}$;

(iii) $\arcsin x$ является обобщенной первообразной функции $\frac{1}{\sqrt{1-x^2}}$ на замкнутом промежутке $[-1, 1]$;

(iv) нулевая функция $F(x) \equiv 0$ является обобщенной первообразной функции Дирихле $\chi(x)$ на всем $\mathbb{R}$.

Тот факт, что F является обобщенной первообразной f на U , будем записывать сокращенно, как $F(x) = \int f(x) dx | U$, или же просто $F(x) = \int f(x) dx$, если из контекста понятно, о каком множестве U идет речь.

ТЕОРЕМА 4.0.4. (i) Если $F(x) = \int f(x) dx \mid U$, то для каждого числа $C \in \mathbb{R}$ имеем $F(x) + C = \int f(x) dx \mid U$.

(ii) Если $F(x) = \int f(x) dx \mid U$ и $G(x) = \int g(x) dx \mid U$, то $\lambda F(x) + \mu G(x) = \int [\lambda f(x) + \mu g(x)] dx \mid U$.

(iii) Если $F(x) = \int f(x) dx \mid U$ и $G(x) = \int g(x) dx \mid U$, то $F(x) - G(x) = \text{const}$ на каждом промежутке $T \subset U$.

ДОКАЗАТЕЛЬСТВО. Единственным нетривиальным моментом в этой теореме является пункт (iii), но и он непосредственно вытекает из Следствия 3.3.8 (т.к. разность $F(x) - G(x)$ представляет собой непрерывную функцию, у которой производная равна нулю в основном). $\square$

Пример 4.0.5. (i) Функция $\frac{1}{s+1}x^{s+1}$ представляет собой о.первообразную для x^s на промежутке $[0, +\infty[$ для $s > -1$ и на промежутке $]0, +\infty]$ для $s < -1$ (при этом договариваются считать, что $\infty^\lambda = 0$ при $\lambda < 0$);

(ii) $\text{arctg } x$ является о. первообразной функции $\frac{1}{1+x^2}$ на всем $\overline{\mathbb{R}}$ (при этом считается, что $\text{arctg}(\pm\infty) = \pm\frac{\pi}{2}$);

(iii) $\arcsin x$ является обобщенной первообразной функции $\frac{1}{\sqrt{1-x^2}}$ на замкнутом промежутке $[-1, 1]$.

(iv) Как видно из предыдущего (см., в частности, Пример 4.0.3(iii)), функция $f(x)$ может иметь о.первообразную на промежутке U , не будучи определенной в некоторых точках этого промежутка. Но сама ее первообразная $F(x)$, конечно же, определена и непрерывна всюду на U .

§1 Определенный интеграл

Будем говорить, что функция $f(x)$ интегрируема (по Ньютону) на промежутке $[a, b] \subset \overline{\mathbb{R}}$, и писать $f \in N[a, b]$, если f имеет обобщенную первообразную на $[a, b]$.

Определенным интегралом функции f по интервалу $[a, b]$ называется приращение ее первообразной $F|_a^b := F(b) - F(a)$ на этом интервале, записывается как

$$F(b) - F(a) = \int_a^b f(x) dx. \quad (4.1)$$

Это знаменитое соотношение носит название "формула Ньютона–Лейбница". Мы записали ее в предположении, что нижний предел интегрирования не превосходит верхнего ($a \leq b$, т.к. $[a, b]$ является промежутком). Однако эту формулу можно распространить и на другие случаи, полагая при $c < d$

$$\int_d^c f(x) dx = F(c) - F(d). \quad (4.2)$$

ТЕОРЕМА 4.1.1 (Свойства определенного интеграла). Предположим, что I — некоторый промежуток в $\overline{\mathbb{R}}$, $f, g \in N(I)$ и $a, b, c \in I$. Тогда

(i) **(линейность)** $\int_a^b [\lambda f(x) + \mu g(x)] dx = \lambda \int_a^b f(x) dx + \mu \int_a^b g(x) dx$ для любой пары чисел $\lambda, \mu \in \mathbb{R}$;

(ii) **(аддитивность)** Если $f \in N[a, b] \cap N[b, c]$, то $f \in N[a, c]$ и $\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$;

- (iii) (**ориентированность**) $\int_a^b f(x) dx = -\int_b^a f(x) dx$;
- (iv) (**монотонность**) Если $a < b$ и $f(x) \leq g(x)$ в основном на $[a, b]$, то $\int_a^b f(x) dx \leq \int_a^b g(x) dx$;
- (v) (**ограниченность**) Если $a < b$, $h \in N[a, b]$ и $|f(x)| \leq h(x)$ в основном на $[a, b]$, то $\left| \int_a^b f(x) dx \right| \leq \int_a^b h(x) dx$;
- (vi) (**устойчивость относительно малых возмущений**) Если $f(x) = g(x) \quad \forall_0 x \in [a, b]$, то $\int_a^b f(x) dx = \int_a^b g(x) dx$;
- (vii) (**точность**) Если $\int_a^b |f(x)| dx = 0$, то $f(x) = 0 \quad \forall_0 x \in [a, b]$.

ДОКАЗАТЕЛЬСТВО. Перечисленные утверждения легко следуют из соответствующих определений, Теоремы 3.3.7 и Следствия 3.3.8, соответствующая проверка оставляется учащимся в качестве упражнения. $\square$

ТЕОРЕМА 4.1.2 (Интеграл как функция переменного верхнего предела). Пусть $[a, b] \subset \mathbb{R}$ и $f \in N[a, b]$. Зададим на промежутке $[a, b]$ функцию $F(x)$ по правилу $F(x) = \int_a^x f(t) dt$. Тогда функция F непрерывна на $[a, b]$, и в каждой точке $x \in (a, b)$, где f непрерывна, функция F дифференцируема, причем $F'(x) = f(x)$.

ДОКАЗАТЕЛЬСТВО. Возьмем какую-нибудь первообразную $G(x)$ функции f на промежутке $[a, b]$. Положим тогда по определению $F(x) = \int_a^x f(t) dt = G(x) - G(a)$. Отсюда сразу же вытекает непрерывность функции F .

Пусть теперь функция f непрерывна в точке $p \in (a, b)$. Возьмем и зафиксируем произвольно $\varepsilon > 0$. По определению непрерывности найдется число $\delta > 0$ такое, что $[p - \delta, p + \delta] \subset [a, b]$ и $|f(x) - f(p)| < \varepsilon$ для всех $x \in [p - \delta, p + \delta]$. Тогда

$$\begin{aligned} \left| \frac{F(x) - F(p)}{x - p} - f(p) \right| &= \frac{1}{|x - p|} \left| \int_p^x f(t) dt - f(p)(x - p) \right| = \frac{1}{|x - p|} \left| \int_p^x [f(t) - f(p)] dt \right| < \\ &< \frac{1}{|x - p|} \left| \int_p^x \varepsilon dt \right| = \varepsilon \end{aligned}$$

для всех $x \in [p - \delta, p + \delta] \setminus \{p\}$. В силу произвольности $\varepsilon > 0$ имеем $\lim_{x \rightarrow p} \frac{F(x) - F(p)}{x - p} = f(p)$, что равносильно искомому равенству $F'(x) = f(p)$. $\square$

§1.1 Формула Тейлора с остаточным членом в интегральной форме

ТЕОРЕМА 4.1.3. Если функция f дифференцируема $(k + 1)$ раз на промежутке $I \subset \mathbb{R}$ и $p, x \in I$, то

$$\begin{aligned} f(x) - f(p) - f'(p)(x - p) - \dots - \frac{f^{(k)}(p)}{k!}(x - p)^k &= \int_p^x \frac{f^{(k+1)}(t)}{k!}(x - t)^k dt = \\ &= -\frac{(x - p)^{k+1}}{(k+1)!} \int_0^1 f^{(k+1)}(y(s)) d(1 - s)^{k+1}, \end{aligned} \quad (4.3)$$

где $y(s) = p + s(x - p)$.

ДОКАЗАТЕЛЬСТВО. Зададим функцию $G(t) = f(x) - f(t) - f'(t)(x-p) - \dots - \frac{f^{(k)}(t)}{k!}(x-t)^k$. Из условий задачи (на гладкость функции f) вытекает, что функция G непрерывна на I . Непосредственной проверкой нетрудно удостовериться, что $G'(t) = -\frac{f^{(k+1)}(t)}{k!}(x-t)^k$, откуда, ввиду непрерывности $G(t)$ и равенства $G(x) = 0$, сразу же вытекает первое равенство в искомой формуле (4.3). Чтобы доказать второе равенство, рассмотрим функцию $R(s) = G(y(s)) = G(p + s(x-p))$. По правилу дифференцирования сложной функции,

$$\begin{aligned} R'(s) &= G'(y(s)) \cdot y'(s) = -\frac{f^{(k+1)}(y(s))}{k!}(x-y(s))^k \cdot (x-p) = -\frac{f^{(k+1)}(y(s))}{k!}(x-p)^{k+1}(1-s)^k = \\ &= -(x-p)^{k+1} \frac{f^{(k+1)}(y(s))}{k!}(1-s)^k. \end{aligned}$$

Тогда, поскольку $R(1) = G(x) = 0$, имеем

$$\begin{aligned} R(0) &= -\int_0^1 R'(s) ds = \frac{(x-p)^{k+1}}{k!} \int_0^1 f^{(k+1)}(y(s))(1-s)^k ds = \\ &= -\frac{(x-p)^{k+1}}{(k+1)!} \int_0^1 f^{(k+1)}(y(s)) d(1-s)^{k+1}, \end{aligned}$$

что равносильно искомому второму равенству в (4.3). $\square$

ТЕОРЕМА 4.1.4 (Ньютоново разложение бинома). При $|x| < 1$ и $\alpha \in \mathbb{R}$ справедлива формула

$$(1+x)^\alpha = \sum_{n=0}^{\infty} C_\alpha^n x^n, \quad (4.4)$$

где $C_\alpha^n = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}$ есть биномальный коэффициент.

ДОКАЗАТЕЛЬСТВО. Легко проверяется, что правая часть в равенстве (4.4) представляет собой ряд Тейлора для функции $f(x) = (1+x)^\alpha$. Таким образом, чтобы доказать сходимость этого ряда к функции f , достаточно проверить, что разность между функцией f и ее полиномом Тейлора степени $k-1$ стремится к нулю при $k \rightarrow \infty$. Для этого воспользуемся предыдущей теоремой 4.1.3. В соответствии с ней, имеем

$$\begin{aligned} \left| f(x) - \sum_{n=0}^{k-1} C_\alpha^n x^n \right| &= \left| \frac{x^k}{k!} \int_0^1 \alpha(\alpha-1)\dots(\alpha-k+1)(1+sx)^{\alpha-k} d(1-s)^k \right| = \\ &= kC_\alpha^k |x|^k \int_0^1 (1+sx)^{\alpha-k} (1-s)^{k-1} ds \leq kC_\alpha^k |x|^k \int_0^1 (1+sx)^{\alpha-1} ds = kC_\alpha^k |x|^k I(\alpha) \end{aligned} \quad (4.5)$$

(в последнем неравенстве мы использовали очевидное соотношение $1+sx > 1-s$, вытекающего из условия $x > -1$). Оценим теперь величину kC_α^k . Возьмем число m такое, что $|\alpha| < m+1$. Тогда

$$\begin{aligned} kC_\alpha^k &= \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{(k-1)!} \leq \frac{(m+1)\dots(m+k)}{(k-1)!} = \frac{(m+k)!}{(k-1)!m!} = \\ &= \frac{k(k+1)\dots(m+k)}{m!} \leq \frac{(m+k)^{m+1}}{m!}. \end{aligned}$$

Поскольку $|x| < 1$, число m не зависит от k (т.е. m — это константа), то легко убедиться, что $|x|^k \frac{(m+k)^{m+1}}{m!} \rightarrow 0$ при $k \rightarrow \infty$, следовательно, $kC_\alpha^k |x|^k I(\alpha) \rightarrow 0$ при $k \rightarrow \infty$, и по формуле (4.5) мы получаем искомую сходимость $\left| f(x) - \sum_{n=0}^{k-1} C_\alpha^n x^n \right| \rightarrow 0$ при $k \rightarrow \infty$. $\square$

§1.2 Иррациональность чисел e^r и π

ЛЕММА 4.1.5. Пусть $r = \frac{p}{q}$ — рациональное число, $p, q \in \mathbb{N}$. Для $n \in \mathbb{N}$ обозначим

$$\Psi_n(x) = \frac{1}{n!}(qx)^n(p - qx)^n.$$

Тогда $D^i\Psi_n(0)$ и $D^i\Psi_n(r)$ представляют собой целые числа для всех $i, n \in \mathbb{N}$ (здесь и в дальнейшем символом $D^i f$ обозначается i -я производная функции f).

ДОКАЗАТЕЛЬСТВО. Если $i < n$, то $D^i\Psi_n(0) = D^i\Psi_n(r) = 0$. Пусть теперь $i \geq n$. Представим полином $\Psi_n(x)$ в виде

$$\Psi_n(x) = \frac{1}{n!} \sum_{k=0}^{2n} a_k (qx)^k,$$

где все коэффициенты a_k являются целыми числами. Тогда

$$D^i\Psi(0) = D^i[a_i(qx)^i]|_{x=0} = \frac{1}{n!} i(i-1)(i-2) \dots 2 \cdot 1 \cdot a_i = \frac{i!}{n!} a_i,$$

(остальные слагаемые, очевидно, занулятся), и потому $D^i\Psi_n(0) \in \mathbb{Z}$. Аналогично,

$$D^i\Psi(r) = \frac{1}{n!} \sum_{k=i}^{2n} k(k-1) \dots (k-i+1) a_k p^{k-i} = \frac{1}{n!} \sum_{k=i}^{2n} \frac{k!}{(k-i)!} a_k p^{k-i} = \sum_{k=i}^{2n} \frac{i!}{n!} \frac{k!}{i!(k-i)!} a_k p^{k-i},$$

и, поскольку $i \geq n$, то $\frac{i!}{n!}$ является целым числом, равно как и $\frac{k!}{i!(k-i)!} = C_k^i$ (биномиальный коэффициент), следовательно, $D^i\Psi_n(r) \in \mathbb{Z}$. $\square$

ТЕОРЕМА 4.1.6. Числа π и e^r при $r \in \mathbb{Q}$ — иррациональны.

ДОКАЗАТЕЛЬСТВО. Докажем второе утверждение. Пусть $r = \frac{p}{q}$, $p, q \in \mathbb{N}$. Предположим, что e^r — рациональное число, т.е. $e^r = \frac{m}{l}$, где $m, l \in \mathbb{N}$. Рассмотрим величину $I_n = l \int_0^r \Psi_n(x) e^x dx$. Многократным применением формулы интегрирования по частям (используя тождество $e^x dx = de^x$) нетрудно убедиться, что величина I_n представима в виде

$$I_n = l \sum_{i=0}^{2n} (-1)^i D^i\Psi_n(x) e^x \Big|_0^r = \sum_{i=0}^{2n} (-1)^i [D^i\Psi_n(r)m - D^i\Psi_n(0)l] \Big|_0^r,$$

следовательно, I_n по предыдущей лемме является целым числом. А поскольку подинтегральная функция в определении I_n положительна при $x \in (0, r)$, то заключаем, что $I_n > 0$, значит, $I_n \geq 1$. С другой стороны, по построению $\sup_{x \in [0, r]} |\Psi_n(x)| \leq \frac{1}{n!} p^{2n} \rightarrow 0$ при $n \rightarrow \infty$. Поэтому $I_n \rightarrow 0$ при $n \rightarrow \infty$. Полученное противоречие завершает доказательство второго утверждения теоремы.

Первое утверждение доказывается аналогично, нужно только, предположив рациональность числа π , т.е. что $\pi = r = \frac{p}{q} \in \mathbb{Q}$, рассмотреть интегралы $J_n = \int_0^r \Psi_n(x) \sin x dx$, и вычислять их по аналогии с предыдущим многократным интегрированием по частям (используя формулы $\sin x dx = -d \cos x$, $\cos x dx = d \sin x$). $\square$

§1.3 Признаки интегрируемости

ТЕОРЕМА 4.1.7 (Критерий интегрируемости на замкнутом промежутке). Пусть $f \in N(a, b)$ и F — некоторая первообразная функции f на открытом интервале (a, b) . Тогда для того, чтобы f обладала первообразной на замкнутом промежутке $[a, b]$, необходимо и достаточно, чтобы существовали пределы $\lim_{x \rightarrow a+0} F(x)$, $\lim_{x \rightarrow b-0} F(x)$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены условия теоремы. Тогда по определению первообразной функция $F(x)$ непрерывна на открытом интервале (a, b) и $F'(x) = f(x)$ в основном на (a, b) . Определим функцию $\tilde{F} : [a, b] \rightarrow \mathbb{R}$ по правилу:

$$\tilde{F}(x) = \begin{cases} F(x), & x \in (a, b); \\ \lim_{t \rightarrow a+0} F(t), & x = a \\ \lim_{t \rightarrow b-0} F(t), & x = b. \end{cases} \quad (4.6)$$

По построению, функция $\tilde{F}(x)$ непрерывна на замкнутом промежутке $[a, b]$ (доказательство оставляется читателю в качестве нетрудного упражнения), и равенство $F'(x) = f(x)$ выполнено в основном на $[a, b]$. Поэтому $\tilde{F}$ — первообразная функции f на всем замкнутом промежутке $[a, b]$, следовательно, $f \in N[a, b]$. $\square$

ТЕОРЕМА 4.1.8 (Критерий существования первообразной). Пусть функция $f : [a, b] \rightarrow \mathbb{R}$ непрерывна в основном на промежутке $[a, b] \subset \mathbb{R}$, и пусть найдется интегрируемая функция $h \in N[a, b]$ такая, что $|f(x)| \leq h(x)$ в основном на (a, b) (в этом случае функцию $h(x)$ называют интегрируемой мажорантой). Тогда $f \in N[a, b]$.

ДОКАЗАТЕЛЬСТВО. Для пары точек $x < y$ из промежутка $[a, b]$ положим $M(x, y) = \{g \in N[x, y] : f(t) \leq g(t) \ \forall t \in [x, y]\}$. Поскольку по условию теоремы

$$-h(t) \leq f(t) \leq h(t) \quad \forall t \in [a, b], \quad (4.7)$$

то $h \in M(x, y) \neq \emptyset$ для любой пары точек $x, y \in [a, b]$, $x < y$. Положим

$$S(x, y) = \inf_{g \in M(x, y)} \int_x^y g(t) dt$$

. Нам понадобятся следующие свойства построенной функции $S(x, y)$.

- (i) $|S(x, y)| \leq \int_x^y h(t) dt$ (ограниченность);
- (ii) если $x, y, z \in [a, b]$, то $S(x, z) = S(x, y) + S(y, z)$ (аддитивность).

В самом деле, из включения $h \in M(x, y)$ немедленно вытекает оценка $S(x, y) \leq \int_x^y h(t) dt$. С другой стороны, в силу левого неравенства в (4.7) имеем $-h(t) \leq f(t) \leq g(t)$ для каждой функции $g \in M(x, y)$. Поэтому $-\int_x^y h(t) dt \leq S(x, y)$. Тем самым свойство (i) установлено.

Далее, если $a \leq x < y < z \leq b$, то очевидна импликация $g \in M(x, z) \Rightarrow g \in M(x, y) \cap M(y, z)$, и $\int_x^z g(t) dt = \int_x^y g(t) dt + \int_y^z g(t) dt \geq S(x, y) + S(y, z)$. Поскольку это неравенство выполняется для всех $g \in M(x, z)$, то получаем оценку снизу

$$S(x, z) = \inf_{g \in M(x, z)} \int_x^z g(t) dt \geq S(x, y) + S(y, z). \quad (4.8)$$

Чтобы получить аналогичную оценку сверху, возьмем произвольно функции $g_1 \in M(x, y)$, $g_2 \in M(y, z)$. Определим новую функцию $g : [x, z] \rightarrow \mathbb{R}$ по правилу

$$g(t) = \begin{cases} g_1(t), & t \in [x, y]; \\ g_2(t), & t \in (y, z]. \end{cases} \quad (4.9)$$

По свойству (ii) Теоремы 4.1.1 и по построению, функция g интегрируема на отрезке $[x, z]$, причем справедливо включение $g \in M(x, z)$. Тогда

$$S(x, z) \leq \int_x^z g(t) dt = \int_x^y g_1(t) dt + \int_y^z g_2(t) dt. \quad (4.10)$$

Итак, мы доказали, что $S(x, z) \leq \int_x^y g_1(t) dt + \int_y^z g_2(t) dt$ для любой пары функций $g_1 \in$

$M(x, y)$, $g_2 \in M(y, z)$. Следовательно, $S(x, z) \leq \inf_{g_1 \in M(x, y)} \int_x^y g_1(t) dt + \inf_{g_2 \in M(y, z)} \int_y^z g_2(t) dt = S(x, y) + S(y, z)$. Таким образом, оценка сверху также получена, и вместе с установленной ранее оценкой снизу (4.8) это приводит к искомому равенству $S(x, z) = S(x, y) + S(y, z)$, т.е. свойство (ii) доказано.

Зададим теперь функцию $F : [a, b] \rightarrow \mathbb{R}$ по правилу $F(x) = S(a, x)$, $x \in [a, b]$. У нее имеются следующие свойства.

(iii) Функция $F(x)$ непрерывна.

Действительно, обозначим через $H(x)$ (обобщенную) первообразную функции h . Зафиксируем точку $x \in [a, b]$, и пусть $y \in [x, b]$. Тогда

$$|F(y) - F(x)| = |S(a, y) - S(a, x)| \stackrel{(ii)}{=} |S(x, y)| \stackrel{(i)}{\leq} |H(y) - H(x)|.$$

Поскольку функция H непрерывна (по определению обобщенной первообразной), имеем $\lim_{y \rightarrow x+} |F(y) - F(x)| = \lim_{y \rightarrow x+} |H(y) - H(x)| = 0$. Аналогично рассматривается симметричный случай $y < x$ и проверяется равенство $\lim_{y \rightarrow x-} |F(y) - F(x)| = 0$. Тем самым непрерывность функции F установлена.

(iv) Если функция f непрерывна в точке $p \in (a, b)$, то $\exists F'(p) = f(p)$.

Докажем это. Возьмем произвольно число $\varepsilon > 0$ и зафиксируем его. Тогда по определению непрерывности найдется число $\delta > 0$ такое, что $[p - \delta, p + \delta] \subset [a, b]$ и $|f(x) - f(p)| < \varepsilon$ для всех $x \in [p - \delta, p + \delta]$. Определим постоянную функцию g_ε по формуле $g_\varepsilon(x) = f(p) + \varepsilon$. По построению справедливо включение $g_\varepsilon \in M(p, p + \delta)$. Поэтому при $y \in (p, p + \delta)$ имеем

$$\frac{F(y) - F(p)}{y - p} \stackrel{(ii)}{=} \frac{S(p, y)}{y - p} \leq \frac{\int_p^y g_\varepsilon(t) dt}{y - p} = f(p) + \varepsilon.$$

С другой стороны, для каждой функции $g \in S(p, y)$ имеем неравенство $g(t) \geq f(t) > f(p) - \varepsilon \quad \forall t \in (p, y)$, следовательно,

$$\frac{F(y) - F(p)}{y - p} \stackrel{(ii)}{=} \frac{S(p, y)}{y - p} = \inf_{g \in M(p, y)} \frac{\int_p^y g(t) dt}{y - p} \geq \frac{\int_p^y [f(p) - \varepsilon] dt}{y - p} = f(p) - \varepsilon.$$

Итак, мы доказали, что $\left| \frac{F(y)-F(p)}{y-p} - f(p) \right| \leq \varepsilon$ для всех $y \in (p, p + \delta)$. Аналогично доказывается то же неравенство для всех $y \in (p - \delta, p)$. В силу произвольности $\varepsilon > 0$ имеем $\lim_{x \rightarrow p} \frac{F(x)-F(p)}{x-p} = f(p)$, что равносильно искомому равенству $F'(x) = f(p)$. $\square$

§1.4 Теоремы о среднем

ТЕОРЕМА 4.1.9 (Первая теорема о среднем). Пусть функции $g(x)$ и $f(x)g(x)$ интегрируемы по промежутку $[a, b] \subset \mathbb{R}$, причем функция g неотрицательна. Обозначим $m = \inf_{x \in (a, b)} f(x)$, $M = \sup_{x \in (a, b)} f(x)$. Тогда

$$m \int_a^b g(x) dx \leq \int_a^b f(x)g(x) dx \leq M \int_a^b g(x) dx. \quad (4.11)$$

Если же вдобавок функция $f(x)$ непрерывна на открытом интервале (a, b) , то найдется точка $\xi \in (a, b)$ такая, что

$$\int_a^b f(x)g(x) dx = f(\xi) \int_a^b g(x) dx. \quad (4.12)$$

ДОКАЗАТЕЛЬСТВО. Пусть выполнены исходные предположения доказываемой теоремы. В частности, справедливы неравенства

$$m g(x) \leq f(x)g(x) \leq M g(x) \quad \forall x \in (a, b). \quad (4.13)$$

Тогда первое утверждение (формула (4.11)) немедленно следует из свойства монотонности определенного интеграла (см. Теорему 4.1.1 (iv)). Мы можем переписать доказанную формулу (4.11) в следующем эквивалентном виде:

$$\exists A \in [m, M] \quad \int_a^b f(x)g(x) dx = A \int_a^b g(x) dx. \quad (4.14)$$

Теперь предположим дополнительно, что функция $f(x)$ непрерывна. Если для параметра A из (4.14) справедливо включение $A \in (m, M)$, то в силу свойств точных граней $\inf$ и $\sup$ (см. Лемму 1.2.8 и Упражнение 1.2.9) найдутся точки $a_1, b_1 \in (a, b)$ такие, что $f(a_1) < A < f(b_1)$. Тогда по теореме Коши о промежуточных значениях (см. Теорему 2.4.15) найдется точка ξ , лежащая между точками a_1, b_1 такая, что $A = f(\xi)$, и справедливость искомой формулы (4.12) установлена.

Рассмотрим теперь пограничный случай, когда параметр A из (4.14) совпадает с одним из крайних значений m или M . Для определенности, пусть $A = m$. Таким образом, имеем

$$m g(x) \leq f(x)g(x) \quad \forall x \in (a, b), \quad (4.15)$$

$$m \int_a^b g(x) dx = \int_a^b f(x)g(x) dx. \quad (4.16)$$

Тогда из свойства точности определенного интеграла (см. Теорема 4.1.1 (vii)) вытекает, что

$$m g(x) \equiv f(x)g(x) \quad \forall x \in (a, b). \quad (4.17)$$

Если найдется точка $\xi \in (a, b)$ такая, что $f(\xi) = m$, то доказывать нечего. Поэтому предположим, далее, что $f(x) \neq m$ для всех $x \in (a, b)$. Тогда вследствие определения $\inf$ получаем, что

$$f(x) > m \quad \forall x \in (a, b). \quad (4.18)$$

Из формул (4.17)–(4.18) немедленно вытекает, что $g(x) = 0$ в основном на (a, b) . Но тогда $\int_a^b f(x)g(x) dx = \int_a^b f(x)g(x) dx = 0$, и в качестве точки ξ из доказываемой формулы (4.12) можно взять любую точку $\xi \in (a, b)$. Теорема 4.1.9 полностью доказана. $\square$

Нам понадобится следующая простая

ЛЕММА 4.1.10. Пусть функция $F(x)$ непрерывна на промежутке $[a, b] \subset \overline{\mathbb{R}}$. Тогда для любого набора точек $\xi_i \in [a, b]$ и для каждой выпуклой комбинации $A = \sum_{i=1}^n \tau_i F(\xi_i)$, $\tau_i \geq 0$, $\sum_{i=1}^n \tau_i = 1$, найдется точка $\xi \in [a, b]$ такая, что $F(\xi) = A$.

ДОКАЗАТЕЛЬСТВО. По теореме Коши о промежуточных значениях, множество $F([a, b])$ выпукло, откуда немедленно вытекает утверждение Леммы. $\square$

ТЕОРЕМА 4.1.11 (Вторая теорема о среднем). Пусть функции $g(x)$ и $f(x)g(x)$ интегрируемы по промежутку $[a, b] \subset \overline{\mathbb{R}}$, причем функция $f : [a, b] \rightarrow \mathbb{R}$ является возрастающей и неотрицательной. Тогда найдется точка $\xi \in [a, b]$ такая, что

$$\int_a^b f(x)g(x) dx = f(b-0) \int_{\xi}^b g(x) dx. \quad (4.19)$$

ДОКАЗАТЕЛЬСТВО. Мы докажем эту теорему при дополнительном предположении, что функция f непрерывна и дифференцируема в основном на $[a, b]$.

Рассмотрим вспомогательную функцию $G(x) = \int_x^b g(t) dt$. По формуле Ньютона–Лейбница и по определению интеграла, функция $G(x)$ непрерывна на $[a, b]$, причем $G'(x) = g(x)$ в основном на $[a, b]$. По формуле интегрирования по частям имеем

$$\int_a^b f(x)g(x) dx = - \int_a^b f(x) dG(x) = f(a)G(a) - f(b)G(b) + \int_a^b G(x) df(x) = f(a)G(a) + \int_a^b G(x) f'(x) dx. \quad (4.20)$$

Учитывая, что $G(b) = 0$, $f'(x) \geq 0$ (в силу предположения о возрастании функции f) и применяя к последнему интегралу первую теорему о среднем, заключаем, что найдется такая точка $\xi_1 \in [a, b]$, для которой справедливо равенство

$$\begin{aligned} \int_a^b f(x)g(x) dx &= f(a)G(a) + G(\xi_1) \int_a^b f'(x) dx \stackrel{\text{формула Н.-Л.}}{=} f(a)G(a) + [f(b) - f(a)]G(\xi_1) = \\ &= f(b)[\tau_1 G(a) + \tau_2 G(\xi_1)], \end{aligned} \quad (4.21)$$

где обозначено $\tau_1 = \frac{f(a)}{f(b)}$ и $\tau_2 = \frac{f(b)-f(a)}{f(b)}$. По построению, $\tau_1 + \tau_2 = 1$ и $\tau_i \geq 0$ (вследствие предположения о возрастании и неотрицательности функции f), поэтому по Лемме 4.1.10 найдется точка $\xi \in [a, b]$ такая, что $G(\xi) = \tau_1 G(a) + \tau_2 G(\xi_1)$. Последнее равенство, вместе с (4.21), влечет справедливость искомой формулы (4.19). $\square$

§2 Вектор-функции одной переменной

ТЕОРЕМА 4.2.1 (О приращениях вектор-функции). Пусть $f : [a, b] \rightarrow \mathbb{R}^m$ и $h : [a, b] \rightarrow \mathbb{R}$ — функции, непрерывные на замкнутом отрезке $[a, b] \subset \mathbb{R}$ и дифференцируемые в основном на (a, b) . Предположим, что $|f'(x)| \leq h(x)$ в основном на (a, b) . Тогда

$$|f(b) - f(a)| \leq h(b) - h(a). \quad (4.22)$$

ДОКАЗАТЕЛЬСТВО. Если $f(b) = f(a)$, то доказывать нечего. Предположим теперь, что $|f(b) - f(a)| \neq 0$. Возьмем вектор $\mathbf{e} = \frac{f(b) - f(a)}{|f(b) - f(a)|}$. Очевидно, что этот вектор имеет единичную длину, $|\mathbf{e}| = 1$, и справедливо тождество

$$|f(b) - f(a)| = \mathbf{e} \cdot [f(b) - f(a)]. \quad (4.23)$$

Рассмотрим функцию $\tilde{f}(x) = f(x) \cdot \mathbf{e} = \sum_{j=1}^m e_j f_j(x)$, где e_j — координаты вектора $\mathbf{e} = (e_1, \dots, e_m)$. Легко проверяется, что эта функция непрерывна на $[a, b]$ и $\tilde{f}'(x) = \sum_{j=1}^m e_j f'_j(x) = f'(x) \cdot \mathbf{e}$ в основном на (a, b) (на самом деле последнее равенство имеет место в каждой точке дифференцируемости функции f). В силу неравенства Коши–Буняковского имеем

$$|\tilde{f}'(x)| = |f'(x) \cdot \mathbf{e}| \leq |f'(x)| \cdot |\mathbf{e}| = |f'(x)|$$

в точках дифференцируемости функции f . Отсюда по предположению доказываемой теоремы получаем, что $|\tilde{f}'(x)| \leq h'(x)$ в основном на (a, b) . Следовательно,

$$|f(b) - f(a)| \stackrel{(4.23)}{=} \tilde{f}(b) - \tilde{f}(a) \stackrel{\text{Теорема 3.3.7}}{\leq} h(b) - h(a).$$

Теорема доказана. $\square$

СЛЕДСТВИЕ 4.2.2. Пусть $f : [a, b] \rightarrow \mathbb{R}^m$ — функция, непрерывная на замкнутом отрезке $[a, b] \subset \mathbb{R}$ и дифференцируемая в основном на (a, b) . Предположим, что $|f'(t)| \in N[a, b]$. Тогда справедливо неравенство

$$|f(b) - f(a)| = \left| \int_a^b f'(t) dt \right| \leq \int_a^b |f'(t)| dt. \quad (4.24)$$

Параметризованной кривой в пространстве $\mathbb{R}^m$ называется непрерывное отображение $f : [a, b] \rightarrow \mathbb{R}^m$, заданное на отрезке $[a, b] \subset \mathbb{R}$ и непостоянное ни на каком подинтервале $(\alpha, \beta) \subset [a, b]$. Говорят, что параметризации $f : [a, b] \rightarrow \mathbb{R}^m$ и $g : [c, d] \rightarrow \mathbb{R}^m$ являются эквивалентными, или, другими словами, что они порождают одну и ту же кривую, если

найдется строго монотонная функция $\psi : [a, b] \rightarrow [c, d]$ такая, что $f(t) \equiv g(\psi(t))$ для всех $t \in [a, b]$. Говорят также, что f и g являются разными параметризациями одной и той же *ориентированной кривой*, если указанная функция ψ является строго возрастающей.

Таким образом, кривая (ориентированная кривая) является классом эквивалентности всевозможных параметризаций по введенному только что отношению.

Пусть Υ представляет собой некоторое разбиение $a = t_0 < t_1 < \dots < t_{n-1} < t_n = b$ отрезка $[a, b]$. Обозначим $\ell(f, \Upsilon) = \sum_{j=1}^n |f(t_j) - f(t_{j-1})|$ — длина ломаной с вершинами в точках $f(t_j)$, $j = 0, \dots, n$. Ясно, что ломаная эта вписана в исходную кривую f . Будем говорить, что разбиение $\Upsilon_2 = \{a = \tau_0, \tau_1, \dots, \tau_N = b\}$ является *измельчением разбиения* $\Upsilon_1 = \{t_0, \dots, t_n\}$ (обозначается символикаой $\Upsilon_1 \leq \Upsilon_2$), если справедливо включение $\Upsilon_1 \subset \Upsilon_2$. Из неравенства треугольника легко выводится, что $\ell(f, \Upsilon_1) \leq \ell(f, \Upsilon_2)$ при $\Upsilon_1 \leq \Upsilon_2$. Множество всех разбиений промежутка $[a, b]$ будем обозначать символом $\mathcal{F}$. Введем следующее важное определение: величина $S(f) := \sup_{\Upsilon \in \mathcal{F}} \ell(f, \Upsilon)$ называется *длиной* (параметризованной) кривой f .

ТЕОРЕМА 4.2.3. Пусть отображение $f : [a, b] \rightarrow \mathbb{R}^m$ непрерывно на замкнутом отрезке $[a, b] \subset \mathbb{R}$ и дифференцируемо в основном на (a, b) , причем $|f'(t)| \in N[a, b]$. Тогда

$$\ell(f) = \int_a^b |f'(t)| dt. \quad (4.25)$$

ДОКАЗАТЕЛЬСТВО. Для точек $x, y \in [a, b]$, $x < y$ через $S_x^y(f)$ будем обозначать длину кривой $f|_{[x, y]}$ между точками x и y . Доказательство теоремы разбивается на несколько шагов.

ШАГ 1 (аддитивность длины кривой). Если $a \leq x < y < z \leq b$, то $S_x^z(f) = S_x^y(f) + S_y^z(f)$ — этот факт легко выводится из соответствующих определений.

ШАГ 2. Если $a \leq x < y \leq b$, то

$$|f(y) - f(x)| = \left| \int_x^y f'(t) dt \right| \leq S_x^y(f) \leq \int_x^y |f'(t)| dt. \quad (4.26)$$

В самом деле для произвольного разбиения $\Upsilon = \{x = t_0, \dots, t_n = y\}$ отрезка $[x, y]$ по формуле (4.24) имеем $|f(t_j) - f(t_{j-1})| \leq \int_{t_{j-1}}^{t_j} |f'(t)| dt$ для каждого $j = 1, \dots, n$. Суммируя по всем таким индексам j и учитывая неравенство треугольника и аддитивность интеграла, получаем

$$|f(y) - f(x)| \leq \ell(f, \Upsilon) = \sum_{j=1}^n |f(t_j) - f(t_{j-1})| \leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} |f'(t)| dt = \int_a^b |f'(t)| dt.$$

Отсюда, в силу произвольности разбиения $\Upsilon \in \mathcal{F}$, получаем искомые неравенства (4.26).

Обозначим $F(x) = S_a^x(f)$.

ШАГ 3. Функция $F(x)$ непрерывна. В самом деле, пусть $G(x)$ — произвольная о. первообразная функции $|f'(x)|$. Она, конечно, непрерывна (по определению первообразной). Тогда для произвольной пары точек $x, y \in [a, b]$, $x < y$ имеем $0 \leq F(y) - F(x) = S_x^y(f) \leq G(y) - G(x)$. Отсюда легко выводится, что $\lim_{y \rightarrow x+} |F(y) - F(x)| = 0$. Аналогично доказывается равенство $\lim_{y \rightarrow x-} |F(y) - F(x)| = 0$.

ШАГ 4. Функция $F(x)$ дифференцируема в основном на (a, b) , причем $F'(x) = |f(x)| \forall x \in (a, b)$. В самом деле, пусть функции f, G дифференцируемы в точке $x \in (a, b)$, и выполнено равенство $G'(x) = |f'(x)|$. Тогда при $y \in (x, b)$ имеем

$$\begin{aligned} \frac{|f(y) - f(x)|}{y - x} &= \frac{|f'(x)(y - x) + o(y - x)|}{y - x} = |f'(x)| + o(1) \stackrel{\text{Шаг 2}}{\leq} \frac{S_x^y(f)}{y - x} \stackrel{\text{Шаг 2}}{\leq} \frac{G(y) - G(x)}{y - x} = \\ &= |f'(x)| + o(1), \end{aligned}$$

что ввиду равенства $S_x^y(f) = F(y) - F(x)$ (см. Шаг 1) влечет справедливость предельного соотношения $\lim_{y \rightarrow x+} \frac{F(y) - F(x)}{y - x} = |f'(x)|$. Аналогично доказывается равенство $\lim_{y \rightarrow x-} \frac{F(y) - F(x)}{y - x} = |f'(x)|$.

Из результатов шагов 3-4 вытекает, что функция $F(x)$ является обобщенной первообразной функции $|f'(x)|$, что по построению означает справедливость искомого утверждения доказываемой теоремы. $\square$

§2.1 Об аддитивных функция отрезка

Пусть дан некоторый замкнутый промежуток $[a, b] \subset \mathbb{R}$ и предположим, что имеется отображение $S(\cdot)$, сопоставляющее каждому отрезку $[x, y] \subset [a, b]$ некоторое число $S([x, y])$. Тогда S называют *функцией отрезка*. Для удобства значение $S([x, y])$ будем обозначать символом S_x^y . Говорят, что функция отрезка S аддитивна, если для любой тройки точек $x < y < z$, лежащих в $[a, b]$, справедливо тождество $S_x^z = S_x^y + S_y^z$.

ТЕОРЕМА 4.2.4. Пусть S — аддитивная функция отрезков, содержащихся в замкнутом промежутке $[a, b] \subset \mathbb{R}$. Предположим, далее, что имеется функция $f : [a, b] \rightarrow \mathbb{R}$ со свойствами: f ограничена на всем $[a, b]$ и непрерывна в основном на $[a, b]$, и для каждого отрезка $[x, y] \subset [a, b]$ и произвольной пары чисел h_1, h_2 таких, что $h_1 \leq f(t) \leq h_2 \forall t \in [x, y]$, справедливы неравенства

$$h_1(y - x) \leq S_x^y \leq h_2(y - x). \quad (4.27)$$

Тогда $S_x^y = \int_x^y f(t) dt$ для каждого отрезка $[x, y] \subset [a, b]$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Обозначим $F(x) = S_a^x$. Фактически, нам нужно доказать, что введенная таким образом функция $F : [a, b] \rightarrow \mathbb{R}$ будет обобщенной первообразной для f , т.е. F непрерывна и существует $F'(x) = f(x)$ в основном на $[a, b]$. По построению и в силу аддитивности S имеем $F(y) - F(x) = S_x^y$ для любой пары точек $x, y \in [a, b]$, $x < y$. В силу предположения об ограниченности f , найдется константа $M \in \mathbb{R}$ такая, что $|f(x)| \leq M$ для всех $x \in [a, b]$. Тогда из (4.27) получаем $|F(y) - F(x)| \leq M|y - x|$ для любой пары точек $x, y \in [a, b]$, откуда немедленно вытекает непрерывность функции F .

Осталось проверить, что в существует производная $F'(p) = f(p)$ в каждой точке p непрерывности функции f . При доказательстве этого утверждения используется тот же прием, что применялся ранее при доказательстве Теорем 4.1.2, 4.1.8 шаг (iv) и в других местах, поэтому мы оставляем его учащимся в качестве упражнения. $\square$

§2.2 Площадь криволинейного сектора

Точку на плоскости удобно задавать в виде комплексного числа $z = re^{i\theta} = r(\cos \theta + i \sin \theta)$, где (r, θ) — полярные координаты точки z .

Пусть имеется плоская параметризованная кривая, у которой в качестве параметра выступает угловая координата θ : $[a, b] \ni \theta \mapsto r(\theta)e^{i\theta}$, где $r(\theta)$ — некоторая непрерывная на $[a, b] \subset [0, 2\pi]$ неотрицательная функция. Через Λ_φ^ψ будем обозначать криволинейный сектор $\Lambda_\varphi^\psi = \bigcup_{\theta \in [\varphi, \psi]} \{\rho e^{i\theta} : \rho \in [0, r(\theta)]\}$, а через S_φ^ψ — его площадь. Мы предполагаем

известным из школы следующие свойства площади:

- (i) Площадь криволинейного сектора является аддитивной функцией угла, т.е. при $a \leq \varphi < \psi < \phi \leq b$ справедливо равенство $S_\varphi^\psi + S_\psi^\phi = S_\varphi^\phi$;
- (ii) Площадь кругового сектора (когда $r(\theta) \equiv R = \text{const}$) вычисляется по формуле $S_\varphi^\psi = \frac{1}{2}R^2[\psi - \varphi]$.
- (iii) Монотонность площади: если криволинейный сектор Λ_φ^ψ содержит в себе круговой сектор Λ_0 и сам, в свою очередь, содержится в другом круговом секторе Λ_1 , то справедливы неравенства $S_0 \leq S_\varphi^\psi \leq S_1$, где через S_0, S_1 обозначены площади круговых секторов Λ_0, Λ_1 соответственно.

ТЕОРЕМА 4.2.5. Пусть $r(\theta)$ — непрерывная неотрицательная функция, заданная на интервале $[a, b] \subset [0, 2\pi]$. Тогда для любой пары углов $\varphi < \psi$ из $[a, b]$ площадь соответствующего криволинейного сектора Λ_φ^ψ (между этими углами) вычисляется по формуле

$$S_\varphi^\psi = \frac{1}{2} \int_{\varphi}^{\psi} r^2(\theta) d\theta. \quad (4.28)$$

ДОКАЗАТЕЛЬСТВО. Нам понадобятся три известные из элементарной геометрии факта:

- (i) Площадь есть аддитивная функция, т.е. $S_\varphi^\psi + S_\psi^\gamma = S_\varphi^\gamma$ при $0 \leq \varphi < \psi < \gamma$.
- (ii) У круга радиуса R площадь сектора с углом α равна $\frac{1}{2}\alpha R^2$.
- (iii) Площадь монотонна, т.е. если задающая криволинейный сектор функция $r(\theta)$ удовлетворяет¹ неравенствам $R_1 \leq r(\theta) \leq R_2$ при $\theta \in [\varphi, \psi]$, то

$$\frac{1}{2}[\psi - \varphi]R_1^2 \leq S_\varphi^\psi \leq \frac{1}{2}[\psi - \varphi]R_2^2.$$

Теперь формула (4.28) легко выводится из этих фактов и Теоремы 4.2.4. $\square$

¹Это означает, что соответствующий криволинейный сектор содержит в себе круговой сектор радиуса R_1 и одновременно содержится в круговом секторе с радиусом R_2 .

§2.3 Объем тела вращения

Пусть $f : [a, b] \rightarrow \mathbb{R}$ — неотрицательная функция, заданная на конечном отрезке $[a, b] \subset \mathbb{R}$. Рассмотрим поверхность, образованную вращением графика $\{(x, y) \in \mathbb{R}^2 : y = f(x), x \in [a, b]\}$ функции f относительно оси O_x . Пусть Λ_t^u означает тело, ограниченное этой поверхностью и плоскостями $x = t$ и $x = u$, т.е. $\Lambda_t^u = \{(x, y, z) : x \in [t, u], y^2 + z^2 \leq f^2(x)\}$. Задача состоит в нахождении объема этого тела. Обозначим этот объем через V_t^u . Из школьной геометрии предполагаются известными следующие свойства объема:

- (i) Объем есть аддитивная функция, т.е. $V_t^u + V_u^w = V_t^w$ при $a \leq t < u < w \leq b$.
- (ii) У цилиндра радиуса R и высоты h объем равен $\pi R^2 h$.
- (iii) Объем представляет собой монотонную функцию относительно теоретико-множественного включения, т.е. если функция $f(x)$ удовлетворяет² неравенствам $R_1 \leq f(x) \leq R_2$ при $x \in [t, u]$, то

$$\pi[u - t]R_1^2 \leq V_t^u \leq \pi[u - t]R_2^2.$$

Из этих элементарных фактов и Теоремы 4.2.4 легко выводится

ТЕОРЕМА 4.2.6. Пусть $f(x)$ — непрерывная неотрицательная функция, заданная на интервале $[a, b] \subset \mathbb{R}$. Тогда для любой пары значений $t < u$ из $[a, b]$ объем соответствующего тела вращения Λ_t^u (между плоскостями $x = t$ и $x = u$) вычисляется по формуле

$$V_t^u = \pi \int_t^u f^2(x) dx. \quad (4.29)$$

§2.4 Площадь поверхности вращения

Из школьной геометрии известно, что объем цилиндра высотой h , основанием которого служит плоская фигура площади S , вычисляется по формуле Sh . Поэтому естественно предположить, что для данного тела вращения Λ , если мы добавим к нему совокупность точек, отстоящих от него на расстояние $\leq h$, то приращение объема $V_h - V$ составит величину порядка

$$V_h - V = Sh + o(h), \quad (4.30)$$

где S — площадь поверхности вращения. Далее, чтобы получить тело, содержащее точки, отстоящие от Λ на расстояние $\leq h$ (обозначим это новое тело символом Λ_h), нужно от точек поверхности, ограничивающей Λ , обозначим эту поверхность Γ , отложить отрезки длиной h , перпендикулярные Γ .

У тела вращения (которое было построено в предыдущем параграфе) граничная поверхность Γ образуется вращением графика функции f вокруг оси O_x . Соответственно, направление перпендикуляра, лежащего в плоскости xy , будет $(-f'(x), 1)$ (т.к. касательная ко графику параллельна вектору $(1, f'(x))$). Длина вертикальной гипотенузы прямоугольного треугольника в плоскости xy , у которого катет длины h направлен параллельно вектору $(-f'(x), 1)$, равна $h\sqrt{1 + (f'(x))^2}$ (проверьте с помощью теоремы Пифагора!). Следовательно, поверхность, ограничивающее тело Λ_h , порождается вращением графика функции $f_h(x) = f(x) + h\sqrt{1 + (f'(x))^2} + o(h)$.

²Это означает, что указанное тело вращения между плоскостями $x = t$ и $x = u$ содержит в себе цилиндр радиуса R_1 (между теми же плоскостями) и одновременно содержится в цилиндре с радиусом R_2 .

Теперь помощью Теоремы 4.2.6 мы можем посчитать объем V_h тела Λ_h :

$$V_h = \pi \int_a^b f_h^2(x) dx = \pi \int_a^b [f^2(x) + 2f(x)\sqrt{1 + (f'(x))^2}h + o(h)] dx = V + 2\pi h \int_a^b f(x)\sqrt{1 + (f'(x))^2} dx + o(h),$$

где через V обозначен объем исходного тела (полученного вращением подграфика функции f). Сравнивая эту величину с формулой (4.30), получаем, что

$$S = 2\pi \int_a^b f(x)\sqrt{1 + (f'(x))^2} dx.$$

Это и есть формула площади поверхности, образованной вращением графика функции $y = f(x)$ вокруг оси O_x (и содержащейся между плоскостями $x = a$ и $x = b$).

Глава 5

Метрические пространства

Пусть X — некоторое множество. Функция $\rho : X \times X$ называется *метрическим пространством*, если для всех точек $x, y, z \in X$ выполняются следующие свойства:

- (i) $\rho(x, y) \geq 0$, причем $\rho(x, y) = 0 \Leftrightarrow x = y$;
- (ii) $\rho(x, y) = \rho(y, x)$ (симметричность);
- (iii) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ (неравенство треугольника).

Упражнение 5.0.7. Пусть (X, ρ) — произвольное метрическое пространство. Докажите *неравенство четырехугольника*: для произвольных точек $x, y, z, t \in X$ справедлива оценка $|\rho(x, y) - \rho(t, z)| \leq \rho(x, t) + \rho(y, z)$.

Пример 5.0.8. 1) Пусть $X = \mathbb{R}^m$, $m \in \mathbb{N}$. Для $p \in [1, +\infty)$ и $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ положим $|x|_p = \left(\sum_{i=1}^m x_i^p \right)^{1/p}$. Тогда для точек $x, y \in \mathbb{R}^m$ можно определить расстояние (метрику) по правилу $\rho_p(x, y) := |x - y|_p$. С помощью неравенства Минковского легко проверяется, что введенная таким образом функция ρ_p удовлетворяет всем аксиомам (i)–(iii) метрического пространства. Обычно, когда речь идет о пространстве $\mathbb{R}^m$, по умолчанию предполагается (если не оговорено противное), что оно снабжено метрикой ρ_2 (т.е. $p = 2$), которую еще называют *евклидовой метрикой*.

2) На пространстве $\mathbb{R}^m$ метрику можно ввести и другим способом, например, для $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ полагая $|x|_\infty = \max\{x_1, \dots, x_m\}$ и $\rho_\infty(x, y) := |x - y|_\infty$.

3) На расширенной числовой прямой $\overline{\mathbb{R}}$ можно ввести такую метрику: $\bar{\rho}(x, y) = \arctg |y - x|$, $\forall x, y \in \overline{\mathbb{R}}$.

Упражнение 5.0.9. Проверьте, что заданные в приведенных выше примерах 1)–3) функции действительно удовлетворяют всем аксиомам (i)–(iii) метрического пространства. Докажите неравенства:

$$\forall p, q \in [1, +\infty] \quad q \leq p \Rightarrow |x|_p \leq |x|_q \quad \forall x, y \in \mathbb{R}^m, \quad (5.1)$$

$$\bar{\rho}(x, y) \leq |x - y| \quad \forall x, y \in \mathbb{R}. \quad (5.2)$$

§1 Предел последовательности

Пусть (X, ρ) — метрическое пространство, и пусть имеется последовательность $x_n \in X$. Говорят, что $x_n \rightarrow a \in X$, если для каждого $\varepsilon > 0$ найдется номер $N = N(\varepsilon) \in \mathbb{N}$ такой, что $\rho(x_n, a) < \varepsilon$ для всех $n \geq N$.

Другими словами,

$$\lim_{n \rightarrow \infty} x_n = a \Leftrightarrow \lim_{n \rightarrow \infty} \rho(x_n, a) = 0 \Leftrightarrow \forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N}: \rho(x_n, a) < \varepsilon \quad \forall n \geq N.$$

Пусть $\varepsilon > 0$ и $x \in X$. Множество $U_\varepsilon(x, X) = \{y \in X : \rho(x, y) < \varepsilon\}$ называется ε -шаром или ε -окрестностью точки x . Обычно пишут $U_\varepsilon(x)$ вместо $U_\varepsilon(x, X)$, когда понятно, в каком именно объемлющем пространстве X задается шар.

Тогда определение предела последовательности можно переформулировать таким образом:

$$\lim_{n \rightarrow \infty} x_n = a \Leftrightarrow \forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N}: x_n \in U_\varepsilon(a) \quad \forall n \geq N.$$

Множество $E \subset X$ называется *ограниченным*, если найдется точка $a \in X$ и число $\varepsilon > 0$ такие, что $E \subset U_\varepsilon(a)$.

ТЕОРЕМА 5.1.1 (Координатный признак сходимости). Пусть $X = \mathbb{R}^m$, $m \in \mathbb{N}$, и пусть $\rho(x, y)$ представляет собой стандартную евклидову метрику $\rho(x, y) = |x - y| = \sqrt{\sum_{i=1}^m |x^i - y^i|^2}$, где x^i, y^i — координаты точек $x = (x^1, \dots, x^m)$ и $y = (y^1, \dots, y^m)$. Тогда для сходимости последовательности $x_n \rightarrow a = (a^1, \dots, a^m)$ при $n \rightarrow \infty$ необходимо и достаточно, чтобы имела место сходимость по каждой координате:

$$\forall i = 1, \dots, m \quad x_n^i \rightarrow a^i \quad \text{при } n \rightarrow \infty. \quad (5.3)$$

ДОКАЗАТЕЛЬСТВО. $\Rightarrow$ Пусть $x_n \rightarrow a$, т.е. $|x_n - a| \rightarrow 0$ при $n \rightarrow \infty$. Для каждого координатного индекса $i = 1, \dots, m$ имеем очевидные неравенства $0 \leq |x_n^i - a^i| \leq |x_n - a|$. Тогда по теореме о двух милиционерах $|x_n^i - a^i| \rightarrow 0$ при $n \rightarrow \infty$.

$\Leftarrow$ Пусть

$$\lim_{n \rightarrow \infty} |x_n^i - a^i| = 0 \quad \text{для всех } i = 1, \dots, m. \quad (5.4)$$

Тогда нетрудно проверяется, что

$$0 \leq |x_n - a| = \sqrt{\sum_{i=1}^m |x_n^i - a^i|^2} \leq \sqrt{m \max_{i=1, \dots, m} |x_n^i - a^i|^2} \leq \sqrt{m} \max_{i=1, \dots, m} |x_n^i - a^i|. \quad (5.5)$$

В силу сходимости (5.4) имеем $\max_{i=1, \dots, m} |x_n^i - a^i| \rightarrow 0$ при $n \rightarrow \infty$, следовательно, $\sqrt{m} \max_{i=1, \dots, m} |x_n^i - a^i| \rightarrow 0$ при $n \rightarrow \infty$, и по теореме о двух милиционерах из (5.5) получаем искомую сходимость $|x_n - a| \rightarrow 0$ при $n \rightarrow \infty$. $\square$

Доказанная теорема позволяет переносить множество фактов из теории предела последовательностей вещественных чисел на случай многомерных пространств $\mathbb{R}^n$. В частности, справедлива следующая

ТЕОРЕМА 5.1.2 (Вейерштрасса о сходящихся подпоследовательностях). Если последовательность $x_n \in \mathbb{R}^m$ ограничена, то из нее можно извлечь сходящуюся подпоследовательность.

ДОКАЗАТЕЛЬСТВО. Пусть последовательность $x_n = (x_n^1, \dots, x_n^m) \in \mathbb{R}^m$ ограничена. Из определения ограниченных множеств легко выводится, что найдется число $M > 0$ со свойством $|x_n| \leq M$ для всех $n \in \mathbb{N}$. Тогда аналогичное неравенство справедливо и для каждой

координаты: $|x_n^i| \leq M$ для всех $n \in \mathbb{N}$ и $i = 1, \dots, m$. Рассмотрим поведение по первой координате $i = 1$. Так как последовательность вещественных чисел x_n^1 ограничена, то по "одномерной" теореме Вейерштрасса, доказанной в первом семестре, найдется сходящаяся подпоследовательность $x_{n_k}^1$, имеющая конечный предел: $x_{n_k}^1 \rightarrow a^1 \in \mathbb{R}$ при $k \rightarrow \infty$.

Теперь рассмотрим соответствующую последовательность вторых координат $x_{n_k}^2$. Поскольку она ограничена, то существует последовательность индексов n_k^2 , являющаяся подпоследовательностью n_k и такая, что последовательность $x_{n_k^2}^2$ сходится. Продолжая в том же духе, мы построим m последовательностей индексов $n_k = n_k^1, n_k^2, n_k^3, \dots, n_k^m$, каждая из которых является подпоследовательностью предыдущей последовательности, причем соответствующие последовательности координат $x_{n_k^i}^i$ сходятся. Тогда по построению последовательность $x_{n_k^m}^m$ сходится при всех $i = 1, \dots, m$. Следовательно, по доказанной теореме последовательность $x_{n_k^m} = (x_{n_k^m}^1, x_{n_k^m}^2, \dots, x_{n_k^m}^m)$ сходится при $k \rightarrow \infty$. $\square$

Последовательность $x_n \in X$ называется *фундаментальной*, если $\forall \varepsilon > 0$ найдется число $\exists N = N(\varepsilon) \in \mathbb{N}$ такое, что $\rho(x_n, x_k) < \varepsilon$ для любой пары номеров $n, k \in \mathbb{N}$. Метрическое пространство называется *полным*, если каждая фундаментальная последовательность в нем имеет предел.

ТЕОРЕМА 5.1.3 (О полноте евклидовых пространств). *Пространство $\mathbb{R}^m$ — полное для каждого $m \in \mathbb{N}$.*

ДОКАЗАТЕЛЬСТВО. Пусть $x_n = (x_n^1, \dots, x_n^m) \in \mathbb{R}^m$ — фундаментальная последовательность точек пространства $\mathbb{R}^m$. В силу очевидного неравенства $|x_n^i - x_k^i| \leq |x_n - x_k|$, справедливого для всех $i = 1, \dots, m$ и $n, k \in \mathbb{N}$, координатная последовательность $(x_n^i)_{n \in \mathbb{N}}$ является фундаментальной для каждого фиксированного $i = 1, \dots, m$. Поэтому согласно критерию Коши для $\mathbb{R}$, каждая координатная последовательность $(x_n^i)_{n \in \mathbb{N}}$ является сходящейся, т.е. найдутся числа $a^1, \dots, a^m \in \mathbb{R}$ такие, что $x_n^i \rightarrow a^i$ при $n \rightarrow \infty$. Но тогда по координатному критерию сходимости 5.1.1 имеем $x_n \rightarrow a = (a^1, \dots, a^m) \in \mathbb{R}^m$ при $n \rightarrow \infty$. Теорема доказана. $\square$

§2 Открытые и замкнутые множества

Пусть (X, ρ) — метрическое пространство. Напомним, что ε -шаром или ε -окрестностью точки $x \in X$ называется множество $U_\varepsilon(x, X) = \{y \in X : \rho(x, y) < \varepsilon\}$. Обычно пишут $U_\varepsilon(x)$ вместо $U_\varepsilon(x, X)$, когда понятно, в каком именно объемлющем пространстве X задается шар.

Следующее понятие является одним из базовых для последующего изложения. Множество $V \subset X$ называется *открытым в X* , если для каждой точки $x \in V$ найдется $\varepsilon > 0$ такое, что $U_\varepsilon(x) \subset V$. Другими словами, если точка x содержится в множестве V , то в нем содержится и некоторый шар положительного радиуса с центром в той же самой точке x .

Очевидно, что данное определение зависит от объемлющего метрического пространства (X, ρ) . Для краткости ниже мы предполагаем, что это пространство (относительно которого изучается открытость и замкнутость множеств) фиксировано.

ТЕОРЕМА 5.2.1 (Свойства открытых множеств). (i) Множества $\emptyset$ и X являются открытыми;

- (ii) Если имеется произвольное семейство $\{V_\alpha\}_{\alpha \in A}$ открытых множеств, то его объединение $V = \bigcup_{\alpha \in A} V_\alpha$ также представляет собой открытое множество;
- (iii) Если имеется конечное семейство $\{V_i\}_{i=1, \dots, n}$ открытых множеств, то его пересечение $V = \bigcap_{i=1}^n V_i$ также представляет собой открытое множество;
- (iv) Для всех $x \in X$ и $\varepsilon > 0$ шар $U_\varepsilon(x)$ представляет собой открытое множество.

ДОКАЗАТЕЛЬСТВО. (i)–(ii) почти мгновенно вытекают из определений. Докажем (iii). Пусть имеется конечное семейство $\{V_i\}_{i=1, \dots, n}$ открытых множеств. Возьмем произвольно точку $x \in \bigcap_{i=1}^n V_i$. Тогда для всякого $i = 1, \dots, n$, поскольку $x \in V_i$, то найдется положительное число ε_i такое, что $U_{\varepsilon_i} \subset V_i$. Положим $\varepsilon = \min_{i=1, \dots, n} \varepsilon_i$. Тогда по построению $\varepsilon > 0$ и $U_\varepsilon(x) \subset U_{\varepsilon_i}(x) \subset V_i$ для всех $i = 1, \dots, n$. Следовательно, $U_\varepsilon \subset V$.

Докажем (iv). Возьмем произвольный шар $U_\varepsilon(x)$ с центром $x \in X$ и радиусом $\varepsilon > 0$. Пусть $y \in U_\varepsilon(x)$. Обозначим $r = \varepsilon - \rho(y, x)$, тогда по построению $r > 0$. Возьмем произвольно точку $z \in U_r(y)$. Тогда в силу неравенства треугольника и определения шаров имеем

$$\rho(x, z) \leq \rho(x, y) + \rho(y, z) < \rho(x, y) + r = (\varepsilon - r) + r = \varepsilon,$$

следовательно, $z \in U_\varepsilon(x)$. Тем самым в силу произвольности $z \in U_r(y)$ установлено включение $U_r(y) \subset U_\varepsilon(x)$. Итак, мы доказали, что для всякой точки $y \in U_\varepsilon(x)$ найдется шар $U_r(y)$ с $r > 0$ такой, что $U_r(y) \subset U_\varepsilon(x)$. Значит, множество $U_\varepsilon(x)$ открытое. $\square$

Множество U называется *окрестностью* точки $x \in X$ (обозначается $U \in \mathcal{N}(x)$), если найдется число $\varepsilon > 0$ такое, что $U_\varepsilon(x) \subset U$. Таким образом, данное выше определение открытых множеств можно переформулировать следующим образом: *множество открыто тогда и только тогда, когда оно является окрестностью каждой своей точки*, на языке формул: $U \in \mathcal{N}(x) \quad \forall x \in U$.

Пусть имеется непустое множество $S \subset X$. Говорят, что точка $x \in X$ является *точкой прикосновения* множества S , если $\forall \varepsilon > 0 \quad U_\varepsilon(x) \cap S \neq \emptyset$.

На языке окрестностей это равносильно тому, что каждая окрестность точки x имеет непустое пересечение с S , т.е. $\forall U \in \mathcal{N}(x) \quad U \cap S \neq \emptyset$.

Совокупность точек прикосновения множества S называется *замыканием множества S* (относительно X) и обозначается символом $\bar{S}$. Очевидно, что всегда выполняется включение $\bar{S} \supset S$. Если $\bar{S} = S$, то множество S называется *замкнутым в X* .

ТЕОРЕМА 5.2.2 (Свойства замыкания). Для всякого непустого множества $S \subset X$ справедливы следующие утверждения.

- (i) $(\bar{S}) = \bar{S}$ (другими словами, какое бы ни было множество $S \subset X$, множество $\bar{S}$ всегда замкнуто);
- (ii) $x \in \bar{S}$ тогда и только тогда, когда найдется последовательность элементов $x_n \in S$ такая, что $x_n \rightarrow x$ при $n \rightarrow \infty$.

ДОКАЗАТЕЛЬСТВО.

(i) Включение $(\bar{S}) \supset \bar{S}$ очевидно. Докажем обратное включение. Пусть $x \in (\bar{S})$. Возьмем произвольно $\varepsilon > 0$. Тогда по определению множества $(\bar{S})$ найдется точка $y \in U_\varepsilon(x) \cap \bar{S}$. Согласно Теореме 5.2.1 (iv), найдется шар $U_r(y)$ с радиусом $r > 0$ такой, что

$$U_r(y) \subset U_\varepsilon(x). \quad (5.6)$$

Поскольку $y \in \bar{S}$, то по определению точек прикосновения найдется точка $z \in U_r(y)$ такая, что $z \in U_r(y) \cap S$. Ввиду формулы (5.6) имеем $z \in U_\varepsilon(x) \cap S$. Итак, доказано, что для всякого $\varepsilon > 0$ найдется точка $z \in U_\varepsilon(x) \cap S$. Следовательно, $x \in \bar{S}$, что и требовалось доказать.

(iii) Пусть $S \ni x_n \rightarrow x$. Тогда по определению предела для каждого шара $U_\varepsilon(x)$ найдется номер N такой, что $x_n \in U_\varepsilon(x)$ при $n \geq N$, в частности, $x_N \in U_\varepsilon(x) \cap S \neq \emptyset$, следовательно $x \in \bar{S}$. Докажем в обратную сторону. Пусть $x \in \bar{S}$. Возьмем последовательность шаров $U_{\frac{1}{n}}(x)$ и точек $x_n \in U_{\frac{1}{n}}(x) \cap S$ (существование таких точек вытекает из непустоты множества $U_{\frac{1}{n}}(x) \cap S$). Утверждается, что имеет место сходимости $x_n \rightarrow x$. В самом деле, для произвольного $\varepsilon > 0$ возьмем номер $N \in \mathbb{N}$ такой, что, $\frac{1}{N} < \varepsilon$. Тогда по построению справедливы включения

$$x_n \in U_{\frac{1}{n}}(x) \subset U_{\frac{1}{N}}(x) \subset U_\varepsilon(x)$$

для всех $n \geq N$, откуда, согласно определению предела, заключаем, что $x_n \rightarrow x$ при $n \rightarrow \infty$. $\square$

СЛЕДСТВИЕ 5.2.3. Множество $S \subset X$ замкнуто тогда и только тогда, когда для всякой последовательности элементов $x_n \in S$ из сходимости $x_n \rightarrow a$ при $n \rightarrow \infty$ вытекает включение $a \in S$.

Если (X, ρ) — метрическое пространство, и $S \subset X$, то сужение функции ρ на $S \times S$ по прежнему является метрикой, и (S, ρ) также представляет собой метрическое пространство (иногда при этом говорят, что метрика на S индуцируется пространством X).

ТЕОРЕМА 5.2.4. Пусть (X, ρ) — метрическое пространство, и $S \subset X$. Тогда, если метрическое пространство (S, ρ) полное, то множество S замкнуто в X .

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения Теоремы. Воспользуемся критерием Следствия 5.2.3 и возьмем произвольную сходящуюся последовательность $S \ni x_n \rightarrow a \in X$. Тогда эта последовательность является, конечно же, фундаментальной. В силу полноты (S, ρ) существует предел последовательности x_n , лежащий в множестве S . Вследствие единственности предела $a \in S$, что и требовалось доказать. $\square$

Оказывается, утверждение последней теоремы допускает обращение (при дополнительном условии полноты самого пространства X).

ТЕОРЕМА 5.2.5. Пусть (X, ρ) — полное метрическое пространство, и $S \subset X$. Тогда множество S замкнуто в X в том и только том случае, когда метрическое пространство (S, ρ) полное.

ДОКАЗАТЕЛЬСТВО. $\Leftarrow$ Непосредственно вытекает из уже доказанной Теоремы 5.2.4.

$\Rightarrow$ Пусть множество S замкнуто в X . Возьмем любую фундаментальную последовательность $x_n \in S$. По определению полного пространства, эта последовательность является сходящейся в X , т.е. $x_n \rightarrow a \in X$. В силу замкнутости множества S и Следствия 5.2.3 справедливо включение $a \in S$, т.е. последовательность x_n является сходящейся в метрическом пространстве (S, ρ) . Отсюда ввиду произвольности фундаментальной последовательности x_n вытекает полнота S . $\square$

ТЕОРЕМА 5.2.6. Множество $E \subset X$ замкнуто в X тогда и только тогда, когда его дополнение $X \setminus E$ открыто.

ДОКАЗАТЕЛЬСТВО. $\Rightarrow$ Пусть множество E замкнуто. Рассмотрим его дополнение $V = X \setminus E$. Нужно доказать открытость V . Предположим, что это не так, тогда найдется точка $x \in V$ такая, что $U_\varepsilon(x) \not\subset V$ для любого $\varepsilon > 0$. Это означает, что для каждого $\varepsilon > 0$ найдется точка $y_\varepsilon \in U_\varepsilon(x) \setminus V$. Но по закону Де Моргана $U_\varepsilon(x) \setminus V = U_\varepsilon(x) \cap E$, следовательно, $y_\varepsilon \in U_\varepsilon(x) \cap E \neq \emptyset$. По определению, точка x является точкой прикосновения множества E . Значит, в силу замкнутости множества E имеем $x \in E$. Но это противоречит предположению $x \in V$.

$\Leftarrow$ Предположим теперь, что множество $V = X \setminus E$ открыто. Нужно показать, что исходное множество E замкнуто. Допустим, что это не так, тогда найдется последовательность $x_n \in E$ такая, что $x_n \rightarrow x \in X \setminus E = V$. Рассмотрим шар $U_\varepsilon(x) \subset V$ (существование такого шара с некоторым $\varepsilon > 0$ вытекает из открытости V). Но тогда по определению предела найдется номер $N \in \mathbb{N}$ такой, что $x_n \in U_\varepsilon(x) \subset V$ для всех $n \geq N$. Получили противоречие со "стартовым" включением $x_n \in E$. $\square$

ТЕОРЕМА 5.2.7 (Свойства замкнутых множеств). (i) Множества $\emptyset$ и X являются замкнутыми;

(ii) Если имеется произвольное семейство $\{E_\alpha\}_{\alpha \in A}$ замкнутых множеств, то его пересечение $V = \bigcap_{\alpha \in A} E_\alpha$ также представляет собой замкнутое множество;

(iii) Если имеется конечное семейство $\{E_i\}_{i=1, \dots, n}$ замкнутых множеств, то его объединение $V = \bigcup_{i=1}^n V_i$ также представляет собой замкнутое множество;

(iv) Для всех $x \in X$ и $\varepsilon > 0$ замкнутый шар $B_\varepsilon(x) = \{y \in X : \rho_X(y, x) \leq \varepsilon\}$ представляет собой замкнутое множество.

ДОКАЗАТЕЛЬСТВО. Очевидно, в доказательстве нуждается только последнее, четвертое свойство, а все остальные легко следуют из Теоремы 5.2.1 и законов Де Моргана о связи операций объединения, пересечения и дополнения множеств.

Итак, возьмем произвольно шар $B_\varepsilon(x)$. Для доказательства его замкнутости воспользуемся критерием из Следствия 5.2.3. Пусть $B_\varepsilon(x) \ni x_n \rightarrow a$. Тогда

$$\rho(x, x_n) \leq \varepsilon \quad \forall n \in \mathbb{N} \quad (5.7)$$

и

$$\lim_{n \rightarrow \infty} \rho(a, x_n) = 0. \quad (5.8)$$

Согласно неравенству треугольника

$$\rho(a, x) \leq \rho(a, x_n) + \rho(x_n, x) \stackrel{(5.7)}{\leq} \rho(a, x_n) + \varepsilon. \quad (5.9)$$

Переходя в последнем неравенстве к пределу и используя сходимость (5.8), получаем $\rho(a, x) \leq \varepsilon$, что равносильно искомому включению $a \in B_\varepsilon(x)$. $\square$

§3 Внутренность, замыкание и граница множеств

Пусть (X, ρ) — метрическое пространство. Напомним, что *замыканием* $\bar{S}$ множества $S \subset X$ называется множество всех точек прикосновения S . Иногда замыкание обозначается символом $\text{Cl}_X S$, или, короче $\text{Cl } S$ (от английского слова *closure*).

ТЕОРЕМА 5.3.1. $\bar{S}$ представляет собой наименьшее (относительно включения) замкнутое множество в X , содержащее S . Другими словами, $\bar{S}$ представляет собой пересечение всех замкнутых подмножеств X , содержащих S .

ДОКАЗАТЕЛЬСТВО. Докажем первое утверждение теоремы. Пусть множество E замкнуто и $E \supset S$. Возьмем произвольно точку $x \in \bar{S}$. Тогда по Теореме 5.2.2 (ii) найдется последовательность точек $S \ni x_n \rightarrow x$. По построению, $x_n \in E$, значит, в силу замкнутости E , имеем $x \in E$. Тем самым установлена импликация $x \in \bar{S} \Rightarrow x \in E$, которая равносильна искомому включению $\bar{S} \subset E$.

Докажем второе утверждение Теоремы. Обозначим через F пересечение всех замкнутых множеств, лежащих в X и содержащих S . Тогда в силу свойства (ii) из Теоремы 5.2.7 множество F является замкнутым и $S \subset F$. Ввиду только что доказанного первого утверждения Теоремы 5.3.1 справедливо включение $\bar{S} \subset F$. Обратное включение $\bar{S} \supset F$ вытекает из построения множества F и замкнутости множества $\bar{S}$. Тем самым равенство $\bar{S} = F$ установлено и Теорема 5.3.1 доказана. $\square$

Точка x называется *внутренней точкой* множества $S \subset X$, если найдется такое $\varepsilon > 0$, что $U_\varepsilon(x) \subset S$. Совокупность внутренних точек называется *внутренностью* множества S , обозначается $\text{Int } S$ (от английского слова *interior*), на языке формул:

$$\text{Int } S = \{x \in S : \exists \varepsilon > 0 \ U_\varepsilon(x) \subset S\}.$$

Когда хотят уточнить, относительно какого именно метрического пространства $X \supset S$ вычисляется внутренность (и берутся соответствующие шары), то употребляют обозначение $\text{Int}_X S$, т.е.

$$\text{Int}_X S = \{x \in S : \exists \varepsilon > 0 \ U_\varepsilon(x, X) \subset S\}.$$

ТЕОРЕМА 5.3.2. $\text{Int } S$ представляет собой наибольшее (относительно включения) открытое множество, содержащееся в S . Другими словами, $\text{Int } S$ представляет собой объединение всех открытых подмножеств S .

ДОКАЗАТЕЛЬСТВО. Пусть $S \subset X$. Докажем открытость множества $\text{Int } S$. Возьмем произвольно $x \in \text{Int } S$. По определению внутренности, найдется $\varepsilon > 0$ такое, что

$$U_\varepsilon(x) \subset S. \quad (5.10)$$

Докажем, что на самом деле справедливо включение

$$U_\varepsilon(x) \subset \text{Int } S. \quad (5.11)$$

Возьмем произвольную точку $y \in U_\varepsilon(x)$. Поскольку по Теореме 5.2.1 (iv) сам шар $U_\varepsilon(x)$ представляет собой открытое множество, то найдется $\delta > 0$ такое¹ что $U_\delta(y) \subset U_\varepsilon(x)$.

¹Как явствует из доказательства Теоремы 5.2.1 (iv), в качестве параметра δ можно взять число $\delta = \varepsilon - \rho(x, y)$

Тогда вследствие (5.10) имеем включение $U_\delta(y) \subset S$, и, по определению внутренности, $y \in \text{Int } S$. Ввиду произвольности $y \in U_\varepsilon(x)$, тем самым установлено искомое включение (5.11), откуда, в свою очередь, ввиду произвольности $x \in \text{Int } S$, вытекает открытость $\text{Int } S$.

Остальные утверждения теоремы легко вытекают из только что доказанной открытости $\text{Int } S$ и определений и свойств открытых множеств (см. Теорему 5.2.1). $\square$

Точка $x \in X$ называется *граничной* точкой множества $S \subset X$, если каждый шар с центром в x имеет непустые пересечения как с S , так и с его дополнением. Множество граничных точек называется *границей* множества S и обозначается символом ∂S , на языке формул:

$$\partial S = \{x \in X : \forall \varepsilon > 0 \quad U_\varepsilon(x) \cap S \neq \emptyset \neq U_\varepsilon(x) \setminus S\}.$$

Когда хотят уточнить, относительно какого именно метрического пространства $X \supset S$ вычисляется граница, то употребляют обозначение $\partial_X S$, т.е.

$$\partial_X S = \{x \in X : \forall \varepsilon > 0 \quad U_\varepsilon(x, X) \cap S \neq \emptyset \neq U_\varepsilon(x, X) \setminus S\}.$$

ТЕОРЕМА 5.3.3. *Справедливо равенство*

$$\partial S = \overline{S} \setminus \text{Int } S. \quad (5.12)$$

ДОКАЗАТЕЛЬСТВО. Докажем сначала включение

$$\partial S \subset \overline{S} \setminus \text{Int } S. \quad (5.13)$$

Возьмем произвольно точку $x \in \partial S$. Тогда из условия $\forall \varepsilon > 0 \quad U_\varepsilon(x, X) \cap S \neq \emptyset$ непосредственно вытекает, что x является точкой прикосновения множества S , следовательно, $x \in \overline{S}$. С другой стороны, ни для какого шара $U_\varepsilon(x)$ при $\varepsilon > 0$ не выполняется включение $U_\varepsilon(x) \subset S$, значит, $x \notin \text{Int } S$, и формула (5.13) доказана. Обратное включение $\overline{S} \setminus \text{Int } S \subset \partial S$ доказывается аналогично. $\square$

ТЕОРЕМА 5.3.4. *Для любого множества $S \subset X$ его граница ∂S представляет собой замкнутое множество.*

ДОКАЗАТЕЛЬСТВО. По Теореме 5.3.2, множество $\text{Int } S$ открыто, тогда по Теореме 5.2.6, множество $X \setminus \text{Int } S$ замкнуто. Далее, по Теореме 5.2.2, множество $\overline{S}$ тоже замкнуто. Тогда их пересечение $\overline{S} \cap (X \setminus \text{Int } S)$ тоже замкнуто (см. свойства (ii) из Теоремы 5.2.7), и теперь искомое утверждение о замкнутости ∂S вытекает из очевидного равенства $\partial S \stackrel{5.12}{=} \overline{S} \setminus \text{Int } S = \overline{S} \cap (X \setminus \text{Int } S)$. $\square$

Пример 5.3.5. В приведенных ниже примерах роль объемлющего метрического пространства X играет вещественная прямая $X = \mathbb{R}$.

- 1) Справедливы формулы $\text{Int } \mathbb{Q} = \emptyset$, $\text{Cl } \mathbb{Q} = \partial \mathbb{Q} = \mathbb{R}$.
- 2) Если $S = (a, b)$, $a < b$, то $\text{Int } S = S = (a, b)$ (т.е. S — открытое множество), $\text{Cl } S = [a, b]$, $\partial S = \{a, b\}$.
- 3) Если $S = [a, b]$, $a \leq b$, то $\text{Int } S = (a, b)$, $\text{Cl } S = S = [a, b]$ (т.е. S — замкнутое множество), $\partial S = \{a, b\}$.

Множество $E \subset F$ называется *плотным* в F , если $F \subset \text{Cl } E$. Из приведенных примеров видно, что множество рациональных чисел $\mathbb{Q}$ плотно в $\mathbb{R}$, а множество точек открытого интервала (a, b) плотно в $[a, b]$.

§4 Относительная внутренность и замыкание множеств

Пусть (X, ρ) — метрическое пространство, и $S \subset E \subset X$. Множество S называется *открытым* (соотв., *замкнутым*) *относительно* E , если оно открыто (соотв., замкнуто) в метрическом пространстве (E, ρ) . Аналогично определяются понятия относительной внутренности, относительного замыкания и т.п.

ТЕОРЕМА 5.4.1. Пусть (X, ρ) — метрическое пространство, и $S \subset X$. Тогда множество $E \subset S$ является открытым (соотв., замкнутым) относительно S , если найдется открытое (соотв., замкнутое) в X множество E_1 такое, что

$$E = E_1 \cap S. \quad (5.14)$$

ДОКАЗАТЕЛЬСТВО. $\Rightarrow$ Пусть множество $E \subset S$ открыто относительно S . Это означает, по определению, что для всякой точки $x \in E$ найдется число $\varepsilon_x > 0$ такое, что

$$U_{\varepsilon_x}(x, S) := \{y \in S : \rho(y, x) < \varepsilon_x\} \subset E. \quad (5.15)$$

Положим

$$E_1 = \bigcup_{x \in E} U_{\varepsilon_x}(x, X).$$

Множество E_1 представляет собой объединение семейства открытых шаров в X , следовательно, E_1 — открыто в X (это вытекает из свойств (ii), (iv) Теоремы 5.2.1). Далее, по построению имеем

$$E \subset S \cap E_1 = S \cap \left(\bigcup_{x \in E} U_{\varepsilon_x}(x, X) \right) = \bigcup_{x \in E} [U_{\varepsilon_x}(x, X) \cap S] = \bigcup_{x \in E} U_{\varepsilon_x}(x, S) \stackrel{(5.15)}{\subset} E.$$

Итак, нами доказаны включения $E \subset E_1 \subset E$, откуда вытекает искомое равенство (5.14).

$\Leftarrow$ Докажем в другую сторону. Пусть найдется открытое в X множество E_1 такое, что $E_1 \cap S = E$. Возьмем произвольно точку $x \in E$. Поскольку $E \subset E_1$, то $x \in E_1$, и в силу определения открытых множеств найдется $\varepsilon > 0$ такое, что $U_\varepsilon(x, X) \subset E_1$. Тогда $U_\varepsilon(x, X) \cap S \subset E_1 \cap S = E$. Отсюда, в силу очевидного равенства $U_\varepsilon(x, X) \cap S = U_\varepsilon(x, S)$ получаем искомое включение $U_\varepsilon(x, S) \subset E$, которое, ввиду произвольности точки $x \in E$, означает открытость E относительно S .

Утверждения про относительную замкнутость и т.п. доказываются аналогично. $\square$

§5 Пределы отображений метрических пространств

Пусть (X, ρ) , (Y, ϱ) — метрические пространства, и пусть имеется отображение $f : S \subset X \rightarrow Y$. Говорят, что $f(x)$ *имеет предел* $a \in Y$ *по множеству* S *при* $x \rightarrow p \in X$ (обозначается $\lim_{x \rightarrow p|S} f(x) = a$), если для каждого $\varepsilon > 0$ найдется $\delta > 0$ такое, что $\varrho(f(x), a) < \varepsilon$ для всякого $x \in S \cap U_\delta(p, X)$, на языке формул:

$$\lim_{x \rightarrow p|S} f(x) = a \Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 : f(x) \in U_\varepsilon(a, Y) \quad \forall x \in S \cap U_\delta(p, X).$$

ТЕОРЕМА 5.5.1 (О пределе композиции). Пусть (X, ρ) , (Y, ϱ) , (Z, σ) — метрические пространства, $S \subset X$, $T \subset Y$, и пусть имеются отображения $f : S \rightarrow T$, $g : T \rightarrow Z$. Предположим, что $\lim_{x \rightarrow p|S} f(x) = a$, $\lim_{y \rightarrow a|T} g(y) = b$ для некоторых точек $p \in X$, $a \in Y$, $b \in Z$.

Тогда

$$\lim_{x \rightarrow p|S} h(x) = b, \quad (5.16)$$

где через h обозначена композиция: $h(x) = g(f(x))$ для $x \in S$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Возьмем произвольно $\varepsilon > 0$. Тогда из равенства $\lim_{y \rightarrow a|T} g(y) = b$ и определения предела вытекает, что найдется $\delta > 0$ такое, что

$$g(y) \in U_\varepsilon(b, Z) \quad \forall y \in T \cap U_\delta(a, Y). \quad (5.17)$$

Далее, из предположения $\lim_{x \rightarrow p|S} f(x) = a$ получаем, что существует $\gamma > 0$ такое, что

$$f(x) \in U_\delta(a, Y) \quad \forall x \in S \cap U_\gamma(p, X).$$

Вспоминая, что f есть отображение из S в T , мы можем уточнить последнюю формулу следующим образом:

$$f(x) \in T \cap U_\delta(a, Y) \quad \forall x \in S \cap U_\gamma(p, X). \quad (5.18)$$

Из (5.17)–(5.18) непосредственно вытекает, что

$$h(x) = g(f(x)) \in U_\varepsilon(b, Z) \quad \forall x \in S \cap U_\gamma(p, X),$$

что, в силу произвольности $\varepsilon > 0$ равносильно искомому равенству (5.16). $\square$

§6 Непрерывные отображения метрических пространств

Пусть (X, ρ) , (Y, ϱ) — метрические пространства, и пусть имеется отображение $f : S \subset X \rightarrow Y$. Говорят, что f непрерывно в точке $p \in S$, если $\lim_{x \rightarrow p|S} f(x) = f(p)$. Отображение f называется непрерывным, если оно непрерывно в каждой точке своей области определения.

ТЕОРЕМА 5.6.1 (О непрерывности композиции). Пусть (X, ρ) , (Y, ϱ) , (Z, σ) — метрические пространства, $S \subset X$, $T \subset Y$, и пусть имеются непрерывные отображения $f : S \rightarrow T$, $g : T \rightarrow Z$. Тогда их композиция $h : S \rightarrow Z$, $h(x) = g(f(x))$ для $x \in S$, представляет собой непрерывное отображение.

ДОКАЗАТЕЛЬСТВО. Возьмем произвольно точку $p \in S$ и обозначим $a = f(p)$. Вследствие предположений о непрерывности имеем $\lim_{x \rightarrow p} f(x) = a$, $\lim_{y \rightarrow a} g(y) = g(a)$. По предыдущей Теореме 5.5.1 имеем $\lim_{x \rightarrow p} h(x) = g(a) = g(f(p)) = h(p)$, что равносильно непрерывности h в точке p . $\square$

ТЕОРЕМА 5.6.2 (Топологический критерий глобальной непрерывности). Пусть $(X, \rho), (Y, \varrho)$ — метрические пространства и $S \subset X$. Тогда для всякого отображения $f : S \rightarrow Y$ следующие условия эквивалентны:

- (i) f непрерывно;
- (ii) для каждого открытого множества $V \subset Y$ прообраз $f^{-1}(V) = \{x \in S : f(x) \in V\}$ представляет собой открытое множество относительно S ;
- (iii) для каждого замкнутого множества $T \subset Y$ прообраз $f^{-1}(T)$ представляет собой замкнутое множество относительно S .

ДОКАЗАТЕЛЬСТВО. (i) $\Rightarrow$ (ii) Пусть функция $f : S \rightarrow Y$ непрерывна и $V \subset Y$ — открытое множество. Обозначим $E = f^{-1}(V)$. Если $E = \emptyset$, то доказывать нечего. Предположим теперь, что множество E непусто и возьмем произвольную точку $p \in E$. Обозначим $a = f(p)$. По построению, $a \in V$. Вследствие открытости V найдется $\varepsilon > 0$ такое, что

$$U_\varepsilon(a, Y) \subset V. \quad (5.19)$$

В силу непрерывности f справедливо равенство $\lim_{x \rightarrow p|S} f(x) = a$, поэтому найдется $\delta > 0$ такое, что

$$f(x) \in U_\varepsilon(a, Y) \quad \forall x \in S \cap U_\delta(p, X).$$

Тогда с учетом (5.19) имеем

$$f(x) \in V \quad \forall x \in S \cap U_\delta(p, X).$$

Следовательно, $S \cap U_\delta(p, X) \subset f^{-1}(V)$. Таким образом, установлено, что каждая точка p множества $f^{-1}(V)$ является внутренней его точкой относительно S , следовательно, множество $f^{-1}(V)$ открыто относительно S .

(ii) $\Rightarrow$ (i) Пусть выполнено предположение (ii). Возьмем произвольную точку $p \in S$ и число $\varepsilon > 0$. Обозначим $a = f(p)$ и рассмотрим шар $U = U_\varepsilon(a, Y)$. Этот шар представляет собой открытое множество, поэтому его прообраз $f^{-1}(U)$ также есть открытое относительно S множество, причем $p \in f^{-1}(U)$. Поэтому найдется $\delta > 0$ такое, что $S \cap U_\delta(p, X) \subset f^{-1}(U)$, другими словами, $f(x) \in U = U_\varepsilon(a, Y) \quad \forall x \in S \cap U_\delta(p, X)$. Отсюда, ввиду произвольности $\varepsilon > 0$, вытекает искомое равенство $\lim_{x \rightarrow p|S} f(x) = a = f(p)$.

(ii) $\Rightarrow$ (iii) Пусть выполнено условие (ii). Возьмем произвольное замкнутое множество $T \subset Y$. По Теореме 5.2.6, множество $V = Y \setminus T$ открыто. По предположению (ii) множество $f^{-1}(V)$ также открыто относительно S . Тогда все по той же Теореме 5.2.6 множество $S \setminus f^{-1}(V)$ замкнуто относительно S . Легко проверяется равенство $S \setminus f^{-1}(V) = S \setminus f^{-1}(Y \setminus T) = f^{-1}(T)$, из которого вытекает искомая замкнутость множества $f^{-1}(T)$ относительно S .

(iii) $\Rightarrow$ (ii) Доказывается применением Теоремы 5.2.6 повторением рассуждений предыдущего пункта (только при этом замкнутые и открытые множества меняются ролями). $\square$

Из координатного признака сходимости (Теорема 5.1.1) вытекает

ТЕОРЕМА 5.6.3 (Координатный признак сходимости). Пусть (X, ρ) — метрическое пространство, и пусть имеется отображение $f : X \rightarrow \mathbb{R}^m$. Тогда для непрерывности $f = (f_1, \dots, f_m)$ необходимо и достаточно, чтобы каждая из координатных функций f_i была непрерывна.

Из последних двух теорем получаем

СЛЕДСТВИЕ 5.6.4 (О задании открытых и замкнутых множеств с помощью неравенств). Пусть (X, ρ) — метрическое пространство, и пусть имеется конечный набор вещественных непрерывных функций $f_i : X \rightarrow \mathbb{R}, i = 1, \dots, m$. Тогда множество $A = \{x \in X : f_i(x) > 0 \ \forall i = 1, \dots, m\}$ открыто, а множество $B = \{x \in X : f_i(x) \geq 0 \ \forall i = 1, \dots, m\}$ замкнуто.

§7 Гомеоморфизмы

Пусть (X, ρ) и (Y, ϱ) — метрические пространства. Отображение $f : X \rightarrow Y$ называется *гомеоморфизмом*, если оно непрерывно, биективно, и обратное к нему отображение $f^{-1} : Y \rightarrow X$ также непрерывно.

ТЕОРЕМА 5.7.1 (Топологический критерий гомеоморфности). Пусть $(X, \rho), (Y, \varrho)$ — метрические пространства, и пусть отображение $f : X \rightarrow Y$ биективно. Тогда следующие утверждения эквивалентны:

- (i) f — гомеоморфизм;
- (ii) множество $U \subset X$ открыто в X в том и только том случае, когда множество $f(U)$ открыто в Y .
- (iii) множество $U \subset X$ замкнуто в X в том и только том случае, когда множество $f(U)$ замкнуто в Y .

Итак, гомеоморфизмы сохраняют структуру открытых и замкнутых множеств. А какие еще свойства пространства сохраняются (инвариантны) при гомеоморфизмах? Наука, занимающаяся поиском таких свойств, именуется *топологией*, а сами эти свойства — *топологическими инвариантами*. Мы опишем здесь только один из простейших топологических инвариантов — *связность*.

Множество $E \subset X$ называется *связным*², если его нельзя представить в виде $E = E_1 \cup E_2, E_i \neq \emptyset, E_i \cap \text{Cl } E_j = \emptyset$ при $i \neq j$. Последнее условие означает, что каждое из множеств E_i является замкнутым относительно E . Поскольку же эти множества дополняют друг друга, каждое из них является также и открытым относительно E . Таким образом, определение связности может быть переформулировано так: множество E связно, если его нельзя представить в виде объединения двух непустых замкнутых (= "открытых") относительно E множеств.

ТЕОРЕМА 5.7.2 (связность отрезка). Множество $E \subset \mathbb{R}$ связно тогда и только тогда, когда оно выпукло (т.е. представляет собой промежуток $E = \langle a, b \rangle, a, b \in \mathbb{R}$).

ДОКАЗАТЕЛЬСТВО. Предположим сначала, что множество E не выпукло, т.е. найдутся вещественные числа $x < y < z$ такие, что $x, z \in E$ и $y \notin E$. Обозначим $E_1 = \{t \in E : t < y\}, E_2 = \{t \in E : t > y\}$. Ясно, что $E_1 \cap \text{Cl } E_2 \subset E_1 \cap [y, +\infty) = \emptyset$, аналогично, $E_2 \cap \text{Cl } E_1 = \emptyset$. Поэтому множество E несвязно.

Докажем в другую сторону. Возьмем $E = [0, 1]$, и предположим, что E несвязно, т.е. $E = E_1 \cup E_2, E_i \neq \emptyset, E_i \cap \text{Cl } E_j = \emptyset$ при $i \neq j$. Тогда каждое из множеств E_i является замкнутым (в $\mathbb{R}$). Зададим функцию $f : [0, 1] \rightarrow \mathbb{R}$ по правилу

$$f(x) = \begin{cases} 1, & x \in E_1; \\ 2, & x \in E_2. \end{cases} \quad (5.20)$$

²Здесь приводится определение связности в смысле понятий общей топологии. Имеются и другие типы связности; один из них, так называемая *линейная связность*, будет описан чуть ниже.

Данная функция непрерывна по теореме 5.6.2 (т.к. прообраз каждого замкнутого множества замкнут). Но это противоречит теореме Коши о промежуточных значениях. $\square$

ТЕОРЕМА 5.7.3. Пусть $(X, \rho), (Y, \varrho)$ — метрические пространства, и пусть отображение $f : X \rightarrow Y$ непрерывно. Тогда для всякого связного множества $E \subset X$ его образ $f(E)$ также представляет собой связное множество.

ДОКАЗАТЕЛЬСТВО. Данная теорема мгновенно вытекает из определений и Теоремы 5.6.2 (iii). $\square$

Множество $E \subset X$ называется *линейно связным*, если для каждой пары точек $x, y \in E$ найдется непрерывное отображение $f : [0, 1] \rightarrow E$ такое, что $f(0) = x$ и $f(1) = y$.

ТЕОРЕМА 5.7.4. Пусть (X, ρ) — метрическое пространство. Тогда всякое линейно связное множество $E \subset X$ связно.

ДОКАЗАТЕЛЬСТВО. Данная теорема почти мгновенно вытекает из Теорем 5.7.2–5.7.3. В самом деле, предположим противное, что линейно связное множество E несвязно. Тогда оно представляется в виде $E = E_1 \cup E_2$, $E_i \neq \emptyset$, $E_i \cap \text{Cl } E_j = \emptyset$ при $i \neq j$. Возьмем произвольно пару точек $x \in E_1$, $y \in E_2$, и непрерывное отображение $g : [0, 1] \rightarrow E$ с $g(0) = x$, $g(1) = y$ (см. определение линейной связности). Тогда множество $g([0, 1])$ связно по теоремам 5.7.2–5.7.3, с другой стороны, его можно представить в виде объединения $g([0, 1]) = [g([0, 1]) \cap E_1] \cup [g([0, 1]) \cap E_2]$, причем каждое из множеств $g([0, 1]) \cap E_i$ непусто и замкнуто относительно $g([0, 1])$. Получили противоречие (со связностью $g([0, 1])$), которое и завершает доказательство теоремы. $\square$

Пусть теперь $E \subset X$ — произвольное непустое множество. Зафиксируем произвольно точку $x \in E$. Объединение всех связных подмножеств множества E , содержащих x , называется *компонентой связности множества E , содержащей x* , и обозначается $\text{comp}_x E$.

ТЕОРЕМА 5.7.5. Пусть (X, ρ) — метрическое пространство, $E \subset X$ и $x \in E$. Тогда множество $\text{comp}_x E$ связно.

ДОКАЗАТЕЛЬСТВО. Оставляется учащимся в качестве упражнения (можно действовать по аналогии с доказательством предыдущей теоремы 5.7.4). $\square$

Таким образом, каждая компонента связности E представляет собой связное множество, т.е. является максимальным связным подмножеством E , содержащим данную точку x . Легко видеть также, что компоненты связности либо не пересекаются, либо совпадают.

Следующая теорема иногда помогает проверить гомеоморфность пространств.

ТЕОРЕМА 5.7.6. Пусть $(X, \rho), (Y, \varrho)$ — метрические пространства, и пусть отображение $f : X \rightarrow Y$ — гомеоморфизм. Тогда для всякого непустого множества $E \subset X$ количество компонент связности E совпадает с количеством компонент связности $f(E)$. В частности, множество $E \subset X$ связно в том и только том случае, когда $f(E)$ связно.

ДОКАЗАТЕЛЬСТВО. Легко выводится из соответствующих определений. $\square$

Упражнение 5.7.7. Доказать, что промежутки $(a, b) \subset \mathbb{R}$ и $(c, d) \subset \mathbb{R}$ гомеоморфны тогда и только тогда, когда они принадлежат к одному типу. В частности никакая пара промежутков $(0, 1)$, $[0, 1)$, $[1, 1]$ не гомеоморфны друг другу³.

³Посчитать количество компонент связности, возникающих послед удаления из указанных промежутков крайних точек.

§8 Равномерно непрерывные, липшицевы, сжимающие отображения

Пусть (X, ρ) и (Y, ϱ) — метрические пространства. Отображение $f : X \rightarrow Y$ называется *равномерно непрерывным*, если найдется функция $\omega : [0, +\infty) \rightarrow [0, +\infty)$ такая, что $\omega(\varepsilon) \rightarrow \omega(0) = 0$ при $\varepsilon \rightarrow 0$, и $\rho(f(x), f(y)) \leq \omega(\varepsilon)$ если $\varrho(x, y) < \varepsilon$. В этом случае функция $\omega = \omega_f$ называется *модулем непрерывности отображения* f .

Упражнение 5.8.1. Доказать, что всякое равномерно непрерывное отображение является непрерывным. Обратное неверно: привести соответствующий контрпример⁴.

Если найдется такая константа $L \geq 0$, что для каждой пары точек $x, y \in X$ выполнено неравенство $\rho(f(x), f(y)) \leq L\varrho(x, y)$, то отображение f называется *липшицевым* (по фамилии известного немецкого математика Lipschitz). Легко видеть, что липшицево отображение является равномерно непрерывным, причем в качестве модуля непрерывности можно взять соответствующую линейную функцию $\omega_f(\varepsilon) = L\varepsilon$.

Упражнение 5.8.2. Доказать, что отображение $f(x) = x^3$, заданное на интервале $[0, 1]$, является липшицевым с константой $L = 3$. С другой стороны, то же отображение $f(x) = x^3$ с областью определения $\mathbb{R}$ не является липшицевым (оно даже не равномерно непрерывно).

Отображение $f : X \rightarrow Y$ называется *сжимающим*, если его константа липшицевости меньше единицы, т.е. если найдется константа $c < 1$ такая, что $\rho(f(x), f(y)) \leq c\varrho(x, y)$ для всех $x, y \in X$. Следующая важная теорема, принадлежащая знаменитому польскому математику Стефану Банаху, носит фундаментальный характер и имеет множество приложений.

ТЕОРЕМА 5.8.3 (Принцип сжимающих отображений). Пусть (X, ρ) — полное метрическое пространство, и пусть имеется сжимающее отображение $f : X \rightarrow X$. Тогда у отображения f существует, и притом единственная, неподвижная точка:

$$\exists! x_* \in X \quad f(x_*) = x_*.$$

ДОКАЗАТЕЛЬСТВО. Возьмем произвольно $x_0 \in X$ и зададим последовательность $x_n \in X$, используя рекуррентное соотношение $x_{n+1} = f(x_n)$. Использование определения сжимающего отображения и элементарных оценок частичных сумм геометрической прогрессии легко проверяется, что эта последовательность является фундаментальной в смысле Коши. В самом деле, по индукции $\rho(x_n, x_{n+1}) \leq c^n \rho(x_0, x_1)$, откуда, применяя неравенство треугольника, получаем

$$\rho(x_n, x_{n+k}) \leq \rho(x_n, x_{n+1}) + \rho(x_{n+1}, x_{n+2}) + \dots + \rho(x_{n+k-1}, x_{n+k}) \leq c^n \rho(x_0, x_1) (1 + c + \dots + c^{k-1}) \leq \frac{c^n}{1-c} \rho(x_0, x_1). \quad (5.21)$$

⁴**Ответ.** В качестве контрпримера рассмотрим простейшее отображение $f : \mathbb{R} \rightarrow \mathbb{R}$, задаваемое формулой $f(x) = x^2$. Для каждого $\varepsilon > 0$ найдутся точки x, y такие, что $y - x = \frac{\varepsilon}{2}$, но $f(y) - f(x) = y^2 - x^2 > 1$ (например, можно взять $x = \frac{1}{\varepsilon}$, $y = \frac{1}{\varepsilon} + \frac{\varepsilon}{2}$). Это означает, что отображение $f(x) = x^2$, заданное на всем $\mathbb{R}$, не является равномерно непрерывным. С другой стороны, если мы рассмотрим его сужение $f_1 = f|_{[0,1]}$ на конечный интервал $[0, 1]$, то это сужение уже будет равномерно непрерывным. В самом деле, для $x, y \in [0, 1]$, $0 < y - x < \varepsilon$ имеем $|f(y) - f(x)| = y^2 - x^2 = (y - x)(y + x) < \varepsilon(y + x) < 2\varepsilon$, так что здесь в качестве модуля непрерывности ω можно взять $\omega(\varepsilon) = 2\varepsilon$.

Поэтому, ввиду полноты пространства X , $\exists \lim_{n \rightarrow \infty} x_n$, обозначим этот предел символом x_* . Вследствие непрерывности отображения f имеем

$$\lim_{n \rightarrow \infty} f(x_n) = f(x_*). \quad (5.22)$$

С другой стороны, из определения последовательности x_n получаем

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} x_n = x_*. \quad (5.23)$$

Из последних двух равенств вытекает искомое соотношение $f(x_*) = x_*$.

Докажем единственность точки x_* . Предположим, что существуют две различные точки $x_{*1}, x_{*2} \in X$ со свойствами $f(x_{*1}) = x_{*1}$, $f(x_{*2}) = x_{*2}$. Тогда $\rho(f(x_{*1}), f(x_{*2})) = \rho(x_{*1}, x_{*2}) > 0$. С другой стороны, в силу определения сжимающего отображения $\rho(f(x_{*1}), f(x_{*2})) \leq c \rho(x_{*1}, x_{*2})$ для некоторого $c \in (0, 1)$. Получили очевидное противоречие, которое завершает доказательство Теоремы (5.8.3). $\square$

Пример 5.8.4. Отметим, что условие сжимаемости отображения f в предыдущей теореме нельзя заменить более слабым условием $\rho(f(x), f(y)) < \rho(x, y)$. Это вытекает из следующего характерного контрпримера. Рассмотрим $f(x) = x - \arctg x - 2$. Тогда $|f(x) - f(y)| < |x - y|$ для любой пары различных чисел $x, y \in \mathbb{R}$ (докажите!)⁵, в то же время уравнение $\arctg x + 2 = 0$ не имеет решений в $\mathbb{R}$, следовательно, отображение $f(x)$ не имеет неподвижных точек.

§9 Компактность

Пусть (X, ρ) — метрическое пространство, и $E \subset X$. Семейство множеств $(U_\alpha)_{\alpha \in A}$ называется *покрытием* множества E , если $E \subset \bigcup_{\alpha \in A} U_\alpha$, т.е. для каждого $x \in E$ найдется $\alpha \in A$ такое, что $x \in U_\alpha$. Покрытие называется *открытым*, если для каждое из множеств U_α является открытым.

Множество $K \subset X$ называется *компактным*, если из каждого его открытого покрытия можно извлечь конечное подпокрытие, т.е. для любого набора открытых множеств V_α таких, что $K \subset \bigcup_{\alpha \in A} V_\alpha$, найдется конечный набор индексов $\alpha_1, \dots, \alpha_n$ таких, что $K \subset$

$$\bigcup_{i=1}^n V_{\alpha_i}.$$

Нам понадобится еще одно понятие. Множество $E \subset K$ называется ε -*сетью* для множества K , если $K \subset \bigcup_{a \in E} U_\varepsilon(a)$.

ТЕОРЕМА 5.9.1 (Секвенциальный критерий компактности). Пусть (X, ρ) — метрическое пространство, и пусть $K \subset X$. Тогда следующие условия эквивалентны:

- (i) множество K компактно;
- (ii) из любой последовательности $x_n \in K$ можно извлечь сходящуюся подпоследовательность x_{n_m} , предел которой лежит в K : $x_{n_m} \rightarrow a \in K$ при $m \rightarrow \infty$;
- (iii) пространство (K, ρ) полное и для любого $\varepsilon > 0$ существует конечная ε -сеть $E \subset K$.

⁵При возникновении затруднений с этим упражнением, можно воспользоваться следующим указанием. Прямым вычислением $0 \geq f'(x) = 1 - \frac{1}{1+x^2} = \frac{x^2}{1+x^2} < 1$, далее нужно воспользоваться теоремой Лагранжа.

ДОКАЗАТЕЛЬСТВО. (i) $\Rightarrow$ (ii) Пусть множество K компактно, и пусть имеется произвольная последовательность $x_n \in K$. Предположим, что утверждение теоремы неверно, т.е. не существует подпоследовательности x_{n_m} , сходящейся к элементу из K . Тогда для каждой точки $a \in K$ найдется $\varepsilon_a > 0$ такое, что шар $U_{\varepsilon_a}(a)$ содержит лишь конечное число⁶ элементов последовательности x_n . Теперь рассмотрим совокупность всех шаров $\{U_{\varepsilon_a}(a)\}_{a \in K}$. Очевидно, что это семейство шаров образует открытое покрытие множества K (т.к. каждый элемент множества K является центром одного из построенных шаров). Поэтому по определению компактности найдется конечная совокупность таких шаров, покрывающая все множество K . Однако в каждой из шаров содержится лишь конечное число элементов последовательности x_n . В итоге получаем, что во всей последовательности x_n лишь конечное число элементов, — очевидное противоречие.

(ii) $\Rightarrow$ (iii) Пусть из любой последовательности $x_n \in K$ можно извлечь сходящуюся подпоследовательность. Докажем, что тогда для всякого $\varepsilon > 0$ найдется конечная ε -сеть. В самом деле, предположим, что это неверно, т.е. для некоторого $\varepsilon > 0$ не существует конечной ε -сети. Тогда по индукции зададим последовательность x_n следующим образом. В качестве x_1 возьмем произвольный элемент множества K . Пусть уже определены $x_1, \dots, x_n$. Тогда в качестве x_{n+1} возьмем произвольный элемент из множества

$$x_{n+1} \in K \setminus \left(\bigcup_{i=1}^n U_{\varepsilon}(x_i) \right).$$

Если указанное множество вдруг оказалось пустым, то $\{x_1, \dots, x_n\}$ является конечной ε -сетью для K , что противоречит нашему предположению. Таким образом, по указанному правилу можно определить x_n для всех $n \in \mathbb{N}$. По построению, $x_n \notin U_{\varepsilon}(x_k)$ при $n > k$, т.е. $\rho(x_n, x_k) \geq \varepsilon$ для любой пары различных $n, k \in \mathbb{N}$. Но тогда из последовательности x_n , очевидно, нельзя извлечь сходящуюся подпоследовательность.

(iii) $\Rightarrow$ (i). Пусть выполнены предположения (iii). Тогда для каждого $m \in \mathbb{N}$ существует конечная ε_m -сеть $E_m \subset K$ с $\varepsilon = \frac{1}{m}$. Возьмем произвольное открытое покрытие $\bigcup_{\alpha \in A} V_{\alpha} \supset K$.

Предположим, что из него нельзя извлечь конечное подпокрытие.

Множество K представимо в виде конечного объединения шаров $K = \bigcup_{e \in E_1} U_1(e, K)$.

Ясно, что $(V_{\alpha})_{\alpha \in A}$ образуют покрытие каждого из этих шаров. Но тогда хотя бы один из шаров $U_{\varepsilon}(e, K)$ с $e \in E_1$ нельзя покрыть *конечным* набором множеств V_{α} , обозначим этот шар B_1 .

Далее, шар B_1 , в свою очередь, представим в виде конечного объединения шаров $B_1 = \bigcup_{e \in E_2} U_{\varepsilon_2}(e, B_1)$. По аналогичным причинам, хотя бы один из этих шаров $U_{\varepsilon_2}(e, B_1)$ с $e \in E_2$ нельзя покрыть *конечным* набором множеств V_{α} , обозначим этот шар B_2 .

Продолжая этот процесс по индукции, получим последовательность шаров B_n таких, что $B_{n+1} = U_{\varepsilon_{n+1}}(e_n, B_n)$ при некотором $e_n \in E_n$, и, кроме того, ни один из шаров B_n нельзя покрыть конечным набором множеств V_{α} . Отсюда, в частности, получаем, что эти шары вложены друг в друга: $B_{n+1} \subset B_n$ для всех $n \in \mathbb{N}$ и, следовательно, $B_k \subset B_n$ для всех $k \geq n$. В частности, $\rho(e_n, e_k) \leq \varepsilon_n = \frac{1}{n}$ для всех $k \geq n$. Отсюда немедленно получаем, что последовательность центров e_n является фундаментальной, значит, она имеет предел:

⁶В самом деле, если бы для некоторого $a \in K$ и для любых $\varepsilon > 0$ и $N \in \mathbb{N}$ нашелся бы элемент $x_n \in U_{\varepsilon}(a)$ с $n > N$, то, взяв последовательность стягивающихся шаров $U_{\frac{1}{m}}(a)$ и соответствующих элементов $x_{n_m} \in U_{\frac{1}{m}}(a)$, получили бы противоречащую предположению сходимости $x_{n_m} \rightarrow a$.

$e_n \rightarrow b \in K$ при $n \rightarrow \infty$. Возьмем множество V_α такое, что $b \in V_\alpha$. В силу открытости этого множества найдется число $\varepsilon > 0$ такое, что $U_{2\varepsilon}(b) \subset V_\alpha$. Но тогда существует номер $N \in \mathbb{N}$ такой, что $e_n \in U_\varepsilon(b)$ при всех $n \geq N$. Тогда при $e_n < \varepsilon$ имеем $U_{\varepsilon_n}(e_n) \subset U_{\varepsilon+\varepsilon_n}(b) \subset U_{2\varepsilon}(b) \subset V_\alpha$. Что противоречит условию при построении B_n . $\square$

СЛЕДСТВИЕ 5.9.2 (О замкнутости компактного множества). Пусть (X, ρ) — метрическое пространство, и пусть множество $K \subset X$ компактно. Тогда K замкнуто в X и ограничено.

ДОКАЗАТЕЛЬСТВО. Пусть, в самом деле, множество $K \subset X$ компактно. Возьмем произвольную последовательность элементов $x_n \in K$, сходящуюся к некоторой точке $a \in X$. Для проверки замкнутости достаточно установить включение $a \in K$. С этой целью отметим, что по Теореме 5.9.1 (ii) из последовательности x_n можно извлечь сходящуюся подпоследовательность x_{n_k} , предел которой лежит в K . Но очевидно, что предел подпоследовательности x_{n_k} совпадает с a . Следовательно $a \in K$, что и требовалось доказать.

Можно установить искомое включение $a \in K$ и другим способом. В самом деле, если последовательность $x_n \in K$ сходится в пространстве (X, ρ) , то эта последовательность является фундаментальной. А поскольку пространство (K, ρ) полное (см. Теорему 5.9.1 (iii)), то и предел последовательности x_n должен лежать в K .

Второе утверждение доказываемой теоремы об ограниченности множества K мгновенно вытекает из Теоремы 5.9.1 (iii): существование конечных ε -сетей гарантирует, что найдется шар с конечным радиусом, который будет содержать все K . $\square$

Оказывается, что для конечномерных пространств (и только для них!) утверждение последней теоремы допускает обращение.

СЛЕДСТВИЕ 5.9.3 (Компактные множества в $\mathbb{R}^n$). Множество $K \subset \mathbb{R}^n$ компактно в том и только том случае, когда оно замкнуто и ограничено.

ДОКАЗАТЕЛЬСТВО. $\Rightarrow$ Непосредственно вытекает из предыдущего Следствия 5.9.2.

$\Leftarrow$ Пусть, в самом деле, множество $K \subset \mathbb{R}^n$ замкнуто и ограничено. Возьмем произвольную последовательность элементов $x_n \in K$. По теореме Вейерштрасса (см. Теорему 5.1.2) из нее можно извлечь сходящуюся подпоследовательность. В силу замкнутости множества K предел этой последовательности снова лежит в K . Отсюда и из Теоремы 5.9.1 (ii) вытекает компактность множества K . $\square$

§9.1 Непрерывные отображения компактных множеств

ТЕОРЕМА 5.9.4 (Теорема Вейерштрасса о максимальном значении). Пусть метрическое пространство (K, ρ) компактно и пусть имеется непрерывное отображение $f : K \rightarrow \mathbb{R}$. Тогда найдутся точки $a, b \in K$ такие, что

$$f(a) = \max_{x \in K} f(x); \quad f(b) = \min_{x \in K} f(x). \quad (5.24)$$

ДОКАЗАТЕЛЬСТВО. Обозначим $A = \sup\{f(x) : x \in K\}$. Возьмем последовательность чисел $A_n \in \mathbb{R}$ таких, что $A_n < A$ и $A_n \rightarrow A$ при $n \rightarrow \infty$. По свойству супремума (см. Лемму 1.2.8), для каждого $n \in \mathbb{N}$ найдется точка $x_n \in K$ такая, что $f(x_n) \in (A_n, A]$. По принципу двух милиционеров, $f(x_n) \rightarrow A$ при $n \rightarrow \infty$. По Теореме 5.9.1 (ii), найдется

сходящаяся подпоследовательность $x_{n_m} \rightarrow a \in K$. Вследствие непрерывности f , имеем $\lim_{m \rightarrow \infty} f(x_{n_m}) = f(a)$, тогда по построению $f(a) = A$. $\square$

ТЕОРЕМА 5.9.5 (Основная теорема алгебры). Пусть $P(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0$ — полином степени n с комплексными коэффициентами $a_m \in \mathbb{C}$, зависящий от комплексной переменной $z \in \mathbb{C}$. Тогда уравнение

$$P(z) = 0 \quad (5.25)$$

имеет в точности n корней (с учетом кратности).

ДОКАЗАТЕЛЬСТВО. Утверждение Теоремы очевидно при $n = 1$. Напомним, что по теореме Безу, если z_* — корень уравнения (5.25), то полином $P(z)$ делится без остатка на $(z - z_*)$, т.е. $P(z) = (z - z_*)P_1(z)$, где $P_1(z)$ — некоторый полином степени $(n-1)$. Ввиду этого факта, достаточно проверить, что у полинома $P(z)$ степени $n \geq 1$ всегда найдется *хотя бы один* комплексный корень z_* (если мы это докажем, то полное утверждение Теоремы 5.9.5, о наличии n корней легко устанавливается методом математической индукции).

Предположим, что это не так, т.е. что имеется полином $P(z)$ степени $n \geq 1$, у которого нет *ни одного* корня, т.е. что

$$P(z) \neq 0 \quad \forall z \in \mathbb{C}. \quad (5.26)$$

Легко проверяется, что найдется радиус $R > 0$ такой, что $|P(z)| > |P(0)|$ при $|z| \geq R$ (это вытекает из соотношения $|P(z)| \rightarrow \infty$ при $|z| \rightarrow \infty$, которое, в свою очередь, следует из того очевидного факта, что $P(z) \sim z^n$ при $z \rightarrow \infty$, т.е. при больших значениях $|z|$ при вычислении $P(z)$ слагаемые вида $a_m z_m$, $m < n$, пренебрежимо малы по сравнению с первым слагаемым z^n). По теореме Вейерштрасса, найдется точка $z_1 \in B_R(0) = \{z \in \mathbb{C} : |z| \leq R\}$ такая, что $|P(z_1)| \leq |P(z)|$ для всех $z \in B_R(0)$. В силу выбора R имеем $|P(z_1)| \leq |P(z)|$ для всех $z \in \mathbb{C}$, т.е. z_1 доставляет минимум модуля полинома P на всей комплексной плоскости. Заменой координат (рассматривая вместо исходного полинома $P(z + z_1)$) можно свести дело к случаю, когда $z_1 = 0$ и

$$P(0) = 1 = \min_{z \in \mathbb{C}} |P(z)|. \quad (5.27)$$

В рассматриваемой ситуации имеет место представление $P(z) = 1 + a_k z^k + \dots + a_n z^n$, где $a_k \neq 0$. Будем считать, что $a_k = r e^{i\varphi}$. Возьмем $z = \rho e^{i[\pi + \varphi] \frac{1}{k}}$, где параметр ρ будет выбран позднее. Тогда

$$|P(z)| \leq 1 - \rho^k [1 - |a_{k+1}| \rho - \dots - |a_n| \rho^{n-k}].$$

Из приведенной оценки вытекает, что $|P(z)| < 1 = |P(0)|$ при достаточно малых ρ . Это противоречит предположению (5.27). Полученное противоречие завершает доказательство Теоремы 5.9.5. $\square$

ТЕОРЕМА 5.9.6 (эквивалентность норм в $\mathbb{R}^n$). Пусть $|\cdot|_*$ — некоторая норма в пространстве $\mathbb{R}^n$. Тогда найдутся положительные константы C_1, C_2 такие, что неравенства

$$C_1|x| \leq |x|_* \leq C_2|x| \quad (5.28)$$

справедливы для всех $x \in \mathbb{R}^n$, здесь $|x| = \left(\sum_{k=1}^n |x_k|^2 \right)^{\frac{1}{2}}$ — стандартная евклидова норма в $\mathbb{R}^n$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим единичную сферу $S = \{x \in \mathbb{R}^n : |x| = 1\}$. Это множество, очевидно, ограничено. Более того, по Следствию 5.6.4 это множество замкнуто. Значит, в силу Следствия 5.9.3, множество S компактно.

Докажем, что функция $|\cdot|_*$ непрерывна относительно стандартной евклидовой нормы $|\cdot|$. Обозначим $a_k = |\mathbf{e}_k|_*$, где $\mathbf{e}_k$ — стандартные единичные вектора, параллельные координатным осям: $\mathbf{e}_1 = (1, 0, \dots, 0) \in \mathbb{R}^n$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, $\dots$, $\mathbf{e}_n = (0, \dots, 0, 1)$. Тогда для каждого элемента $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ имеем

$$|x|_* = |x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n|_* \leq |x_1 \mathbf{e}_1| + \dots + |x_n \mathbf{e}_n|_* = |x_1| a_1 + \dots + |x_n| a_n \leq |x| \cdot |a|,$$

где на последнем шаге мы воспользовались неравенством Коши–Буняковского $x \cdot y \leq |x| \cdot |y|$, здесь $|a| = \left(\sum_{k=1}^n a_k^2\right)^{\frac{1}{2}}$ — стандартная евклидова норма вектора $a = (a_1, \dots, a_n)$.

Итак, установлена оценка $|x|_* \leq C_2 |x|$ для всех $x \in \mathbb{R}^n$ с константой $C_2 = |a|$. Отсюда легко выводится, что функция $|\cdot|_*$ липшицева:

$$\left| |x|_* - |y|_* \right| \stackrel{\text{неравенство треугольника для } |\cdot|_*}{\leq} |x - y|_* \leq C_2 |x - y|.$$

Следовательно, функция $|\cdot|_*$ непрерывна в пространстве $\mathbb{R}^n$, снабженным стандартной евклидовой метрикой. По теореме Вейерштрасса, она принимает свое минимальное значение на сфере S : $\exists C_1 = \min_{y \in S} |y|_*$. По построению, это значение положительно (т.к. все значения нормы $|x|_*$ положительны при $x \neq 0 \notin S$). Тогда для каждого ненулевого вектора $x \in \mathbb{R}^n$ имеем $y = \frac{1}{|x|} x \in S$ и потому $|y|_* \geq C_1$ и

$$|x|_* = \left| |x| \cdot y \right|_* \stackrel{\text{положительная однородность нормы } |\cdot|_*}{=} |x| \cdot |y|_* \geq |x| C_1.$$

Теорема доказана. □

Глава 6

Дифференциальное исчисление в многомерном случае

§1 Предел функции в точке и предел по направлению

Пусть V — открытое множество в $\mathbb{R}^n$, и пусть $\mathbf{e} \in \mathbb{R}^n$ — ненулевой вектор. Говорят, что функция $f : V \rightarrow \mathbb{R}^m$ имеет предел $A \in \mathbb{R}^m$ в точке $p \in V$ по направлению $\mathbf{e}$, если $\lim_{t \rightarrow 0+} f(p + t\mathbf{e}) = A$.

При взятии предела по направлению рассматриваемое отображение, фактически, превращается в функцию одного переменного (заданную на положительном луче вещественной прямой).

Характерны случаи, когда $\mathbf{e}$ совпадает с одним из координатных векторов.

Важно отчетливо уяснить себе разницу между понятием предела в точке (см. раздел §5) и предела по направлению.

ТЕОРЕМА 6.1.1 (Предел в точке и предел по направлению). Пусть V — открытое множество в $\mathbb{R}^n$, $p \in V$ и пусть имеется отображение $f : V \rightarrow \mathbb{R}^m$. Тогда, если $\lim_{x \rightarrow p} f(x) = A$, то $\lim_{t \rightarrow 0} f(p + t\mathbf{e}) = A$ для каждого направления $\mathbf{e} \in \mathbb{R}^n$, $\mathbf{e} \neq \mathbf{0}$.

ДОКАЗАТЕЛЬСТВО. Возьмем произвольно $\varepsilon > 0$. По определению предела, найдется $\delta > 0$ такое, что $|f(x) - A| < \varepsilon$ для всех $x \in V \cap B(p, \delta)$, здесь и далее через $B(p, \delta)$ мы обозначаем открытый шар $B(p, \delta) = \{x \in \mathbb{R}^n : |x - p| < \delta\}$. Поскольку множество V открыто, можно, беря δ достаточно малым, считать, не умаляя общности, что $B(p, \delta) \subset V$. Теперь, если $|t| < \frac{\delta}{|\mathbf{e}|}$, то $|t\mathbf{e}| < \delta$, и, соответственно, $p + t\mathbf{e} \in B(p, \delta)$, следовательно, $|f(p + t\mathbf{e}) - A| < \varepsilon$ при всех $|t| < \frac{\delta}{|\mathbf{e}|}$. В силу произвольности ε это означает, что $\lim_{t \rightarrow 0} f(p + t\mathbf{e}) = A$, что и требовалось доказать. $\square$

Обратное, вообще говоря, неверно: из существования предела по направлению не следует существования предела в точке.

Пример 6.1.2. (i) Рассмотрим функцию

$$f(x, y) = \frac{xy}{x^2 + y^2}, \quad x^2 + y^2 \neq 0,$$

полагая $f(0, 0) = 0$. Эта функция имеет в начале координат предел по каждому направлению $\mathbf{e} \in \mathbb{R}^2$, но эти пределы различны, поэтому предела в точке $(0, 0)$ не существует.

(ii) Рассмотрим функцию

$$f(x, y) = \begin{cases} 1, & y = x^2; \\ 0, & y \neq x^2. \end{cases} \quad (6.1)$$

Эта функция имеет один и тот же нулевой предел по каждому направлению $\mathbf{e} \in \mathbb{R}^2$, но предела в точке $(0, 0)$ не существует.

§2 Частные производные и дифференцируемость

В этом разделе всюду символом V обозначается открытое множество в $\mathbb{R}^n$.

Пусть $\mathbf{e} \in \mathbb{R}^n$ — ненулевой вектор. Говорят, что функция $f : V \rightarrow \mathbb{R}^m$ имеет производную A в точке $p \in V$ по направлению $\mathbf{e}$, если $\lim_{t \rightarrow 0+} \frac{f(p+t\mathbf{e})-f(p)}{t} = A$. Предел по направлению $\mathbf{e}$ обозначается символом $\frac{\partial f}{\partial \mathbf{e}}$, или, короче, $\partial_{\mathbf{e}} f$, т.е.

$$\partial_{\mathbf{e}} f(p) = \lim_{t \rightarrow 0+} \frac{f(p+t\mathbf{e}) - f(p)}{t}.$$

Особую роль играют так называемые *частные производные* по переменным $x_1, \dots, x_n$, когда в качестве направления берутся единичные вектора, параллельные координатным осям: $\mathbf{e}_1 = (1, 0, \dots, 0) \in \mathbb{R}^n$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, $\dots$, $\mathbf{e}_n = (0, \dots, 0, 1)$. Эти частные производные обозначаются символами $\frac{\partial f}{\partial x_j} = \frac{\partial f}{\partial \mathbf{e}_j}$, или, короче, $\partial_j f$. Таким образом, для точки $p = (p_1, \dots, p_n)$ имеем

$$\partial_j f(p) = \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{e}_j) - f(p)}{t} = \lim_{t \rightarrow 0} \frac{f(p_1, \dots, p_{j-1}, p_j + t, p_{j+1}, \dots, p_n) - f(p)}{t}.$$

Частная производная фактически является производной функции одного переменного. В самом деле, рассмотрим функцию $g(t) = f(p + t\mathbf{e}_j)$. Тогда $g'(0) = \partial_j f(p)$.

На практике при вычислении частных производных употребляется следующий прием: все переменные, кроме j -й, "замораживаются", т.е. считаются константами, и функция становится зависящей как бы от одной переменной — от x_j , по которой и берется производная.

Пример 6.2.1. (i) Имеем $\frac{\partial(x^y)}{\partial x} = yx^{y-1}$ (при "при замороженном" y функция x^y превращается в степенную функцию относительно переменной x); с другой стороны, $\frac{\partial(x^y)}{\partial y} = x^y \ln x$ (при "при замороженном" x функция x^y превращается в показательную функцию относительно переменной y).

(ii) Отметим одно геометрическое отличие: при вычислении производной по направлению берутся только положительные значения t , а при вычислении частной производной они могут быть и отрицательными, т.е. производная по направлению считается *по лучу*, а частная производная — *по всей прямой (с проколотой точкой)*. Это отличие проявляется в случае функции $f(x) = |x| = \sqrt{x_1^2 + \dots + x_n^2}$. Для произвольного ненулевого вектора $\mathbf{e} \in \mathbb{R}^n$ имеем

$$\partial_{\mathbf{e}} f(0, 0) = \lim_{t \rightarrow 0+} \frac{f(t\mathbf{e})}{t} = \lim_{t \rightarrow 0+} \frac{|t\mathbf{e}|}{t} = |\mathbf{e}|.$$

Таким образом, для каждого вектора $\mathbf{e} \in \mathbb{R}^n$ производная по направлению $\mathbf{e}$ функции $f(x) = |x|$ существует и равна $|\mathbf{e}|$. Однако ни одной частной производной в начале координат не существует, в самом деле, "замораживая" переменные $x_2 = \dots = x_n = 0$, имеем

$f(x_1, 0, \dots, 0) = |x_1|$. Однако функция $\mathbb{R} \ni x_1 \mapsto |x_1|$ не имеет производной в нуле, поэтому частной производной по x_1 у функции f в начале координат не существует.

(iii) Рассмотрим функцию

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & x^2 + y^2 > 0; \\ 0, & x = y = 0, \end{cases} \quad (6.2)$$

По построению, что $f(x, 0) = 0 \ \forall x \in \mathbb{R}$ и $f(0, y) = 0 \ \forall y \in \mathbb{R}$. Тогда по определению частных производных легко вычисляется, что

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x} = 0.$$

Аналогично,

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y} = 0.$$

Таким образом, данная функция имеет конечные частные производные в каждой¹ точке $(x, y) \in \mathbb{R}^2$. Однако эта функция не является даже непрерывной в точке $(0, 0)$ (это видно хотя бы из того, что на луче $x = y > 0$ функция f тождественно равна $\frac{1}{2} \neq 0$, так что в точке $(0, 0)$ она имеет различные пределы по различным направлениям).

Последний пример функции, обладающей в каждой точке из $\mathbb{R}^2$ конечными частными производными, и тем не менее разрывной, показывает, что понятие дифференцируемости для многомерного случая нуждается в некотором уточнении.

Пусть V — открытое множество в $\mathbb{R}^n$, и $p \in V$. Говорят, что *функция* $f : V \rightarrow \mathbb{R}^m$ *дифференцируема в точке* $p \in V$, если существует линейное отображение $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$, для которого справедливо представление

$$f(x) = f(p) + L(x - p) + o(|x - p|) \quad \text{при } x \rightarrow p. \quad (6.3)$$

Здесь и далее символом Lx обозначено значение линейного отображения L на векторе $x \in \mathbb{R}^n$.

Линейное отображение L из (6.3) называется *дифференциалом отображения* f , и обозначается символом $df(p)$, соответственно, его действие на вектор $v \in \mathbb{R}^n$ обозначается символом $df(p) \langle v \rangle$.

ТЕОРЕМА 6.2.2 (Регулярность дифференцируемых функций). Пусть V — открытое множество в $\mathbb{R}^n$, $p \in V$, и пусть отображение $f : V \rightarrow \mathbb{R}^m$ дифференцируемо в точке p . Тогда f непрерывно в точке p и обладает в этой точке частными производными по каждому направлению $e \in \mathbb{R}^n$, причем справедливо равенство

$$\partial_e f(p) = df(p) \langle v \rangle. \quad (6.4)$$

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения доказываемой теоремы. Обозначим через L соответствующее линейное отображение $L = df(p)$. Из курса алгебры и аналитической геометрии известно, что $|Lz| \leq \|L\| \cdot |z|$ для каждого вектора $z \in \mathbb{R}^n$, где символом $\|L\|$ обозначена норма оператора (линейного отображения) L . Поэтому в силу (6.3) имеем

$$|f(x) - f(p)| \leq \|L\| \cdot |x - p| + o(|x - p|).$$

¹Вне начала координат эти производные вычисляются по обычным формулам $\frac{\partial f}{\partial x}(x, y) = \frac{y(y^2 - x^2)}{(x^2 + y^2)^2}$; $\frac{\partial f}{\partial y}(x, y) = \frac{x(x^2 - y^2)}{(x^2 + y^2)^2}$.

Очевидно, правая часть стремится к нулю при $x \rightarrow p$. Значит, и $|f(x) - f(p)| \rightarrow 0$ при $x \rightarrow p$, что означает непрерывность отображения f .

Докажем второе утверждение, о наличии частных производных. Возьмем произвольный ненулевой вектор $\mathbf{e} \in \mathbb{R}^n$ и зафиксируем его. Тогда подставляя в формулу (6.3) выражение $x = p + t\mathbf{e}$ при малых $t \in \mathbb{R}$ имеем

$$f(p + t\mathbf{e}) - f(p) = L(t\mathbf{e}) + o(|t\mathbf{e}|) = t[L\mathbf{e} + o(1)].$$

Поэтому

$$\partial_{\mathbf{e}} f(p) = \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{e}) - f(p)}{t} = L\mathbf{e}.$$

□

Как известно, всякое линейное отображение конечномерных пространств задается умножением вектора на матрицу. Именно, будем считать элементы пространств $\mathbb{R}^n$, $\mathbb{R}^m$ вектор-столбцами:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \in \mathbb{R}^m.$$

Тогда для всякого линейного отображения $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ найдется матрица $A = A_L = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ такая, что $Lx = Ax$ для всех $x \in \mathbb{R}^n$. Здесь символом Lx обозначено значение линейного отображения L в точке x , а символом Ax — результат умножения матрицы A на вектор-столбец x справа.

Очевидно, если $x = \mathbf{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ (единица на j -й позиции), то $A\mathbf{e}_j$ совпадает с j -м столбцом матрицы A :

$$L\mathbf{e}_j = A\mathbf{e}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Будем сокращенно обозначать этот столбец a_j , т.е. $A = (a_1 | a_2 | \dots | a_n)$. В интересующем нас случае, когда линейное отображение L является дифференциалом $L = df(p)$, по формуле (6.4) получаем

$$\frac{\partial f}{\partial x_j} = \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(p) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(p) \end{pmatrix} = a_j.$$

Таким образом, элементы матрицы A здесь задаются формулой $a_{ij} = \frac{\partial f_i}{\partial x_j}(p)$. Эта матрица называется *матрицей Якоби* отображения f . В последующем она будет обозначаться символом $f'(p)$.

Нашей ближайшей целью будет дальнейшее выяснение связи между частными производными и дифференцируемостью.

ТЕОРЕМА 6.2.3 (Липшицевость функций с ограниченными частными производными). Пусть V — открытое множество в $\mathbb{R}^n$, и пусть имеется куб $Q = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset V$. Предположим, что для отображения $f : V \rightarrow \mathbb{R}^m$ найдется константа $c \in \mathbb{R}$ такая, что на Q все частные производные функции f существуют и выполнена оценка

$$\left| \frac{\partial f}{\partial x_j}(p) \right| \leq c \quad \forall j = 1, \dots, n \quad (6.5)$$

для всех $p \in Q$. Тогда для любой пары точек $x, y \in Q$ справедливо неравенство

$$|f(x) - f(y)| \leq c\sqrt{n}|x - y|. \quad (6.6)$$

ДОКАЗАТЕЛЬСТВО. Пусть дана произвольная пара точек $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, лежащих в Q . Зададим конечную последовательность точек $A_0, A_1, A_2, \dots, A_n \in Q$ по правилу $A_0 = x$, $A_n = y$, $A_i = (y_1, \dots, y_i, x_{i+1}, \dots, x_n)$. Тогда точки A_i, A_{i-1} будут отличаться только одной (i -й) координатой, соответственно, и отрезок $[A_i, A_{i-1}]$, их соединяющий, будет параллелен i -й координатной оси. Соответственно, сужение функции f на этот отрезок будет функцией одной переменной, у которой производная ограничена числом c . По одномерной формуле конечных приращений Лагранжа (см. Следствие 3.1.6) справедлива оценка

$$|f(A_i) - f(A_{i-1})| \leq c|x_i - y_i|. \quad (6.7)$$

Тогда в силу неравенства треугольника получаем

$$\begin{aligned} |f(x) - f(y)| &\leq \sum_{i=1}^n |f(A_i) - f(A_{i-1})| \leq c \sum_{i=1}^n |x_i - y_i| \stackrel{\text{нер-во Гельдера}}{\leq} \left(\sum_{i=1}^n 1^2 \right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}} = \\ &= c\sqrt{n}|x - y|. \end{aligned}$$

□

СЛЕДСТВИЕ 6.2.4. Пусть V — открытое выпуклое множество в $\mathbb{R}^n$, и пусть для отображения $f : V \rightarrow \mathbb{R}^m$ найдется константа $c \in \mathbb{R}$ такая, что на V все частные производные функции f существуют и выполнена оценка (6.5) для всех $p \in V$. Тогда для любой пары точек $x, y \in V$ справедливо неравенство (6.6).

ДОКАЗАТЕЛЬСТВО.

□

ТЕОРЕМА 6.2.5 (Дифференцируемость функций с непрерывными частными производными). Пусть V — открытое множество в $\mathbb{R}^n$, и пусть у отображения $f : V \rightarrow \mathbb{R}^m$ все частные производные $\frac{\partial f}{\partial x_j}$ существуют и непрерывны в точке $p \in V$. Тогда f дифференцируемо в p .

ДОКАЗАТЕЛЬСТВО. Зафиксируем произвольно точку $p \in V$ и число $\varepsilon > 0$. Предположим сначала, что все частные производные в точке p зануляются. Тогда в силу их непрерывности найдется $\delta > 0$ такое, что

$$\left| \frac{\partial f}{\partial x_j}(x) \right| < \frac{\varepsilon}{\sqrt{n}} \quad (6.8)$$

для всех $x \in Q$, где Q представляет собой n -мерный куб с центром в p и длиной стороны δ . По доказанной только что теореме 6.2.3 справедлива оценка

$$|f(x) - f(p)| \leq \varepsilon |x - p| \quad (6.9)$$

для всех $x \in Q$. Ввиду произвольности $\varepsilon > 0$ отсюда вытекает, что $|f(x) - f(p)| = o(|x - p|)$, т.е. f дифференцируема в точке p с нулевым дифференциалом.

Рассмотрим теперь общий случай, когда $f'(p) \neq 0$. Возьмем линейное отображение $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ у которого все частные производные совпадают с соответствующими частными производными в точке p функции f . Тогда у отображения $\tilde{f}(x) = f(x) - Lx$ все частные производные будут по-прежнему непрерывны, причем в точке p все эти частные производные занулятся. По доказанному выше, отображение $\tilde{f}$ будет дифференцируемо в точке p с нулевым дифференциалом. Тогда отображение $f = \tilde{f}(x) + Lx$ будет дифференцируемо в p с дифференциалом L . $\square$

Отображения, у которых все частные производные непрерывны на всей области определения (представляющей собой открытое множество), называются *непрерывно дифференцируемыми*.

Пример 6.2.6. Отметим, что условие непрерывности частных производных в последней теореме нельзя заменить условием существования производных по всем направлениям. А именно, у функции

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2}, & x^2 + y^2 > 0; \\ 0, & x = y = 0, \end{cases} \quad (6.10)$$

в каждой точке плоскости существуют производные по всем направлениям. Они непрерывны всюду, кроме начала координат, причем в начале координат все производные по всем направлениям равны нулю (можно посчитать их вдоль прямой $x = 0$ и на прямых вида $y = kx$). Но данная функция не только не дифференцируема — она даже разрывна в нуле. Это вытекает из того очевидного факта, что $f(x, y) = \frac{1}{2}$ при $y = x^2$.

ТЕОРЕМА 6.2.7 (Дифференциал композиции). Пусть $V \subset \mathbb{R}^n$, $U \subset \mathbb{R}^m$ — открытые множества, и пусть отображения $f : V \rightarrow U$, $g : U \rightarrow \mathbb{R}^k$ удовлетворяют следующим условиям: f дифференцируемо в $p \in V$, а g дифференцируемо в $q = f(p)$. Тогда композиция $g \circ f = h : V \rightarrow \mathbb{R}^k$, вычисляемая по формуле $h(x) = g(f(x))$, дифференцируема в точке p , причем справедливы равенства

$$dh(p) = dg(q) \circ df(p). \quad (6.11)$$

$$h'(p) = g'(q) \cdot f'(p) \quad (6.12)$$

(в последней формуле символом $g'(q) \cdot f'(p)$ обозначается произведение соответствующих матриц Якоби).

ДОКАЗАТЕЛЬСТВО. Пусть выполнены все предположения доказываемой теоремы. Обозначим $L = df(p)$, $T = dg(q)$, $A = f'(p)$, $B = g'(q)$. Тогда

$$f(x) = f(p) + L(x - p) + o(|x - p|) = q + L(x - p) + o(|x - p|), \quad \text{при } x \rightarrow p, \quad (6.13)$$

$$g(y) = g(q) + T(y - q) + o(|y - q|), \quad \text{при } y \rightarrow q. \quad (6.14)$$

Поскольку $f(x) - f(p) = f(x) - q = O(|x - p|) \rightarrow 0$ при $x \rightarrow p$, подставляя $y = q + L(x - p) + o(|x - p|)$ в формулу (6.14), имеем

$$\begin{aligned} h(x) &= g(f(x)) = g(q + L(x - p) + o(|x - p|)) = g(q) + T[L(x - p) + o(|x - p|)] + o(|L(x - p) + o(|x - p|)|) = \\ &= g(q) + T[L(x - p)] + O(1) \cdot o(|x - p|) + o(O(|x - p|)) = g(q) + T[L(x - p)] + o(|x - p|). \end{aligned}$$

Таким образом, отображение h дифференцируемо в p , и его дифференциал совпадает с композицией дифференциалов f и g , т.е. формула (6.11) доказана. Теперь второе из искомых соотношений (6.12) вытекает из уже доказанной формулы (6.11) и того факта, что матрица композиции линейных отображений равна произведению соответствующих матриц. $\square$

Из доказанной теоремы вытекает элегантное правило вычисления частных производных у композиции функций. В условиях предыдущей теоремы, обозначим $A = f'(p)$, $B = g'(q)$, $C = BA = h'(p)$. Пусть a_{ij} , b_{li} , c_{lj} — элементы матриц A , B , C соответственно (здесь индексы пробегают значения $i = 1, \dots, m$, $j = 1, \dots, n$, $l = 1, \dots, k$ соответственно). Поскольку $C = BA$, по правил вычисления произведения матриц имеем

$$c_{lj} = \sum_{i=1}^n b_{li} a_{ij}.$$

Вспоминая, что по определению матриц Якоби $a_{ij} = \frac{\partial f_i}{\partial x_j}(p)$, $b_{li} = \frac{\partial g_l}{\partial y_i}(q)$, $c_{lj} = \frac{\partial h_l}{\partial x_j}(p)$, получаем следующую важную формулу:

$$\frac{\partial (g \circ f)_l}{\partial x_j}(p) = \sum_{i=1}^n \frac{\partial g_l}{\partial y_i}(f(p)) \frac{\partial f_i}{\partial x_j}(p), \quad (6.15)$$

следовательно, для самой вектор-функции $g = (g_1, \dots, g_k)$ имеем

$$\frac{\partial (g \circ f)}{\partial x_j}(p) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f(p)) \frac{\partial f_i}{\partial x_j}(p), \quad (6.16)$$

Обозначим через dx_j дифференциал линейного отображения $x = (x_1, \dots, x_n) \mapsto x_j$. Поскольку дифференциал линейного отображения совпадает с самим этим отображением, то, как нетрудно видеть,

$$dx_j \langle v \rangle = v_j \quad \text{для всякого вектора } v = (v_1, \dots, v_n) \in \mathbb{R}^n. \quad (6.17)$$

Тогда дифференциал функции $f(x)$ можно записать в виде

$$df(p) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(p) dx_j. \quad (6.18)$$

В самом деле, если мы подействуем правой частью на базисный вектор $\mathbf{e}_i \in \mathbb{R}^n$, то получим $\frac{\partial f}{\partial x_i}(p)$ — а это как раз значение $df(p) \langle \mathbf{e}_i \rangle$ (см. (6.4)). Но если у двух линейных отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$ совпадают действия на каждый базисный вектор из $\mathbb{R}^n$, то эти отображения тождественно равны, откуда и вытекает искомое равенство (6.18).

Из доказанной теоремы о композиции вытекает важное свойство, которое называется *инвариантностью первого дифференциала*. Обозначим $h = g \circ f$. Тогда по формулам (6.18), (6.16) получаем

$$dh(p) = \sum_{j=1}^n \frac{\partial h}{\partial x_j}(p) dx_j = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f(p)) \frac{\partial f_i}{\partial x_j}(p) dx_j = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f(p)) \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(p) dx_j. \quad (6.19)$$

Учитывая еще раз (6.18), получаем красивое соотношение

$$d(g \circ f)(p) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(f(p)) df_i(p). \quad (6.20)$$

Запишем для сравнения формулу дифференциала самой функции g :

$$dg(q) = \sum_{i=1}^n \frac{\partial g}{\partial y_i}(q) dy_i. \quad (6.21)$$

Сходство очевидно: формулу для дифференциала композиции можно получить, формально подставляя в формулу дифференциала "внешней" функции g вместо независимых переменных y_i соответствующие компоненты промежуточной вектор-функции f_i .

Интересно, что дифференциал второго порядка уже **НЕ** инвариантен (см. следующий раздел).

§3 Градиент

Пусть имеется дифференцируемое вещественнозначное отображение $f : V \rightarrow \mathbb{R}$ открытого множества $V \subset \mathbb{R}^n$. Дифференциалом функции f в точке $p \in V$ является некоторое линейное отображение $df(p) : \mathbb{R}^n \rightarrow \mathbb{R}$, $df(p) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p) dx_i$. Согласно сказанному выше, для каждого вектора $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ имеем

$$df(p) \langle v \rangle = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p) v_i. \quad (6.22)$$

Обозначим через w вектор с компонентами $w = (\frac{\partial f}{\partial x_1}(p), \frac{\partial f}{\partial x_2}(p), \dots, \frac{\partial f}{\partial x_n}(p)) \in \mathbb{R}^n$. Тогда формулу (6.22) можно переписать в виде

$$df(p) \langle v \rangle = \langle w, v \rangle \quad \forall v \in \mathbb{R}^n, \quad (6.23)$$

где символом $\langle w, v \rangle$ обозначено, как обычно, скалярное произведение в пространстве $\mathbb{R}^n$:

$$\langle w, v \rangle = \sum_{i=1}^n w_i v_i.$$

Вектор w , для которого выполнено тождество (6.23), называется *градиентом функции f в точке p* и обозначается символом $\nabla f(p)$. Таким образом, справедливы формулы

$$\nabla f(p) = (\frac{\partial f}{\partial x_1}(p), \frac{\partial f}{\partial x_2}(p), \dots, \frac{\partial f}{\partial x_n}(p)), \quad (6.24)$$

$$df(p) \langle v \rangle = \langle \nabla f(p), v \rangle \quad \forall v \in \mathbb{R}^n, \quad (6.25)$$

В этой связи отметим, что имеет место классическая

ТЕОРЕМА 6.3.1 (Теорема Рисса о представлении). Пусть $L : \mathbb{R}^n \rightarrow \mathbb{R}$ — линейное отображение. Тогда существует единственный вектор $w \in \mathbb{R}^n$ такой, что

$$Lv = \langle w, v \rangle \quad \forall v \in \mathbb{R}^n. \quad (6.26)$$

Особый интерес представляет *геометрический смысл градиента*. Обозначим через $\mathbb{S}$ единичную сферу в $\mathbb{R}^n$: $\mathbb{S} = \{x \in \mathbb{R}^n : |x| = 1\}$. Выясним, во-первых, каково направление *наибольшего роста функции* f . А именно, для какого единичного вектора $u \in \mathbb{S}$, справедливо равенство

$$df(p) \langle u \rangle = \max_{v \in \mathbb{S}} df(p) \langle v \rangle \quad ? \quad (6.27)$$

Ввиду (6.25), последнюю формулу можно переписать в следующем эквивалентном виде:

$$df(p) \langle u \rangle = \max_{v \in \mathbb{S}} \langle \nabla f(p), v \rangle. \quad (6.28)$$

Теперь решение очевидно:

$$u = \frac{1}{|\nabla f(p)|} |\nabla f(p)|.$$

Выясним, далее, как связан градиент со множествами уровня функции f . А именно, зафиксируем точку $p \in V$ и рассмотрим множество уровня $M = \{x \in V : f(x) = f(p)\}$. Возьмем произвольную последовательность точек $x_k \in M \setminus \{p\}$, $x_k \rightarrow p$ при $k \rightarrow \infty$. Тогда по построению и по определению дифференцируемости имеем

$$f(x_k) = f(p) + df(p) \langle x_k - p \rangle + o(|x_k - p|) = f(p),$$

следовательно

$$df(p) \langle x_k - p \rangle = o(|x_k - p|).$$

Учитывая (6.25), перепишем последнюю формулу в виде

$$\langle \nabla f(p), x_k - p \rangle = o(|x_k - p|).$$

Поделив равенство на $|x_k - p|$, получаем

$$\left\langle \nabla f(p), \frac{x_k - p}{|x_k - p|} \right\rangle = o(1). \quad (6.29)$$

Поскольку последовательность векторов $\frac{x_k - p}{|x_k - p|}$ принадлежит единичной сфере, которая, напомним, является компактным множеством (см. Следствие 5.9.3), из нее можно извлечь сходящуюся подпоследовательность, обозначим ее снова через $\frac{x_k - p}{|x_k - p|} \rightarrow u \in \mathbb{S}$. Переходя к пределу в равенстве (6.29), получаем

$$\langle \nabla f(p), u \rangle = 0. \quad (6.30)$$

Приведенные выкладки имеют простой геометрический смысл. Обозначим через $T_p M$ множество

$T_p M = \{u \in \mathbb{R}^n : \exists \text{ последовательность } M \ni x_k \rightarrow p \text{ такая, что } |u| \frac{x_k - p}{|x_k - p|} \rightarrow u \text{ при } k \rightarrow \infty\}$. Множество $T_p M$ называется *касательным пространством* к поверхности M в точке p . Выше мы доказали, фактически, что $T_p M \subset \{u \in \mathbb{R}^n : \langle \nabla f(p), u \rangle = 0\}$. В регулярном случае справедливо и обратное включение, а именно, справедлива

ТЕОРЕМА 6.3.2 (Касательное пространство и градиент). Пусть имеется непрерывно дифференцируемое вещественнозначное отображение $f : V \rightarrow \mathbb{R}$ открытого множества $V \subset \mathbb{R}^n$. Предположим, что $p \in V$ и $\nabla f(p) \neq 0$. Тогда

$$T_p M = \{u \in \mathbb{R}^n : \langle \nabla f(p), u \rangle = 0\}.$$

Это утверждение нетрудно вывести из теоремы о неявной функции (см. Теорему 7.4.1).

Помимо касательного пространства, в геометрии часто фигурирует множество $p + T_p M$, которое называют *контингенцией*, или, в просторечии, *касательной* к поверхности M в точке p (если поверхность S одномерна, говорят о *касательной прямой*, если она двумерна, то о *касательной плоскости* и т.д.). По построению

$$p + T_p M = \{v \in \mathbb{R}^n : \langle \nabla f(p), v - p \rangle = 0\}. \quad (6.31)$$

Пример 6.3.3. Рассмотрим эллипсоид в $\mathbb{R}^3$:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

где $a, b, c \neq 0$. Фактически, этот эллипсоид представляет собой множество уровня

$$M = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = 1, \}$$

где $f = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$. Найдем уравнение его касательной плоскости $p + T_p M$ в точке $p = (x_0, y_0, z_0)$. Прямым вычислением $\nabla f(p) = 2\left(\frac{x_0}{a^2}, \frac{y_0}{b^2}, \frac{z_0}{c^2}\right)$. Поэтому уравнение касательного пространства имеет вид

$$T_p M = \{(x, y, z) \in \mathbb{R}^3 : \frac{x \cdot x_0}{a^2} + \frac{y \cdot y_0}{b^2} + \frac{z \cdot z_0}{c^2} = 0\}.$$

Чтобы найти касательную плоскость, необходимо в последнем равенстве вместо нуля поставить какую-то константу, так, чтобы точка $p = (x_0, y_0, z_0)$ принадлежала заданному множеству. Эта константа легко находится (например, с помощью формулы (6.31)), так что в итоге получаем

$$p + T_p M = \{(x, y, z) \in \mathbb{R}^3 : \frac{x \cdot x_0}{a^2} + \frac{y \cdot y_0}{b^2} + \frac{z \cdot z_0}{c^2} = 1\}.$$

Это и есть уравнение *касательной плоскости* к эллипсоиду.

§4 Старшие производные

Всюду в этом разделе V — открытое множество в $\mathbb{R}^n$. Многократная дифференцируемость определяется по индукции следующим образом. Функция $f : V \rightarrow \mathbb{R}^m$ называется k раз дифференцируемой в точке $p \in V$, $k > 1$, если

- (i) f дифференцируема в каждой точке из V ;
- (ii) Частные производные функции f дифференцируемы $k - 1$ раз в точке $p \in V$.

ТЕОРЕМА 6.4.1 (Теорема о равенстве смешанных производных). Пусть V — открытое множество в $\mathbb{R}^n$. Тогда, если функция $f : V \rightarrow \mathbb{R}^m$ дважды дифференцируема в точке $p \in V$, то ее смешанные производные равны:

$$\partial_i(\partial_j f)(p) = \partial_j(\partial_i f)(p) \quad \forall i, j = 1, \dots, n. \quad (6.32)$$

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения доказываемой теоремы. Возьмем число $t > 0$ и рассмотрим точки $A_t = p + t\mathbf{e}_i$, $B_t = p + t\mathbf{e}_i + t\mathbf{e}_j$, $C_t = p + t\mathbf{e}_j$. Основная идея доказательства очень простая: от точки p к точке B_t можно пойти двумя способами: сначала от p до A_t по координатной линии $\mathbf{e}_i$, а потом от A_t к B_t по координатной линии $\mathbf{e}_j$, а можно и по другому — сначала от p до C_t по координатной линии $\mathbf{e}_j$, а потом от C_t к B_t по координатной линии $\mathbf{e}_i$. Соответственно, разность $f(B_t) - f(p)$ может быть вычислена двумя различными способами интегрированием частных производных по указанным координатным линиям, и, поскольку разность эта одна и та же, в итоге будет получено искомое тождество (6.32).

Перейдем к точным формулировкам. По формуле Ньютона — Лейбница имеем

$$f(B_t) - f(p) = [f(B_t) - f(A_t)] + [f(A_t) - f(p)] = \int_0^t \partial_i f(p + s\mathbf{e}_i) ds + \int_0^t \partial_j f(p + t\mathbf{e}_i + s\mathbf{e}_j) ds. \quad (6.33)$$

Аналогично, двигаясь по другому маршруту от p к B_t , получаем

$$f(B_t) - f(p) = [f(B_t) - f(C_t)] + [f(C_t) - f(p)] = \int_0^t \partial_j f(p + s\mathbf{e}_j) ds + \int_0^t \partial_i f(p + t\mathbf{e}_j + s\mathbf{e}_i) ds. \quad (6.34)$$

Поскольку каждая из частных производных $\partial_i f$, $\partial_j f$ дифференцируема в точке p , справедливо представление

$$\partial_i f(p + s\mathbf{e}_i + u\mathbf{e}_j) = \partial_i f(p) + s\partial_{ii} f(p) + u\partial_{ji} f(p) + o(|s| + |u|),$$

$$\partial_j f(p + s\mathbf{e}_i + u\mathbf{e}_j) = \partial_j f(p) + s\partial_{ij} f(p) + u\partial_{jj} f(p) + o(|s| + |u|).$$

Подставляя это в формулы (6.33)–(6.34), получаем

$$\begin{aligned} f(B_t) - f(p) &= t[\partial_i f(p) + \partial_j f(p)] + \frac{1}{2}t^2[\partial_{ii} f(p) + 2\partial_{ij} f(p) + \partial_{jj} f(p)] + o(t^2) = \\ &= t[\partial_j f(p) + \partial_i f(p)] + \frac{1}{2}t^2[\partial_{jj} f(p) + 2\partial_{ji} f(p) + \partial_{ii} f(p)] + o(t^2). \end{aligned}$$

Тем самым приходим к искомому равенству (6.32). $\square$

Пример 6.4.2. Если условие двукратной дифференцируемости заменить более слабым предположением об однократной дифференцируемости и существовании всех частных производных второго порядка, то утверждение последней теоремы (о равенстве смешанных производных) может оказаться неверным. Это вытекает из следующего контрпримера. Функция $f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}$, $f(0, 0) = 0$ дифференцируема в каждой точке и обладает всеми четырьмя вторыми производными, причем на проколотой окрестности $\mathbb{R}^2 \setminus (0, 0)$ эта функция бесконечно дифференцируема. Тем не менее,

$$\left(\frac{\partial}{\partial x} \frac{\partial}{\partial y} f \right) \Big|_{(0,0)} \neq \left(\frac{\partial}{\partial y} \frac{\partial}{\partial x} f \right) \Big|_{(0,0)}$$

(проверка этого факта прямым вычислением оставляется читателям в качестве упражнения; при этом следует рационально организовать процесс вычислений, отбрасывая на заключительном этапе слагаемые, которые заведомо обращаются в ноль и не влияют на значение искомых смешанных производных при $x = y = 0$).

§5 Мультииндексный формализм

Мультииндекс — это упорядоченный набор натуральных чисел $\mu = (\mu_1, \dots, \mu_n) \in \mathbb{N}^n$. Если $\mu = (\mu_1, \dots, \mu_n)$, $\nu = (\nu_1, \dots, \nu_n)$ — мультииндексы, то их сумма $\mu + \nu = (\mu_1 + \nu_1, \dots, \mu_n + \nu_n)$ — тоже мультииндекс.

Используются обозначения $|\mu| = \mu_1 + \dots + \mu_n$ — абсолютная величина мультииндекса, $\mu! = \mu_1! \cdot \mu_2! \cdot \dots \cdot \mu_n!$.

Для $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ используется обозначение $x^\mu = x_1^{\mu_1} \dots x_n^{\mu_n}$. Соответственно,

$$\partial^\mu f = \partial_1^{\mu_1} \dots \partial_n^{\mu_n} f = \overbrace{\frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_1} \left(\dots \frac{\partial}{\partial x_1} \right) \dots \right)}^{\mu_1 \text{ раз}} \dots \overbrace{\frac{\partial}{\partial x_n} \left(\frac{\partial}{\partial x_n} \left(\dots \frac{\partial}{\partial x_n} \right) \dots \right)}^{\mu_n \text{ раз}} f.$$

ЛЕММА 6.5.1. Для всякой функции $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ класса C^k (здесь U — открытое в $\mathbb{R}^n$ множество) и любой пары мультииндексов μ, ν таких, что $|\mu + \nu| \leq k$ имеет место равенство

$$\partial^\mu (\partial^\nu f) = \partial^\nu (\partial^\mu f) = \partial^{\mu+\nu} f. \quad (6.35)$$

ДОКАЗАТЕЛЬСТВО. Данная лемма легко доказывается применением теоремы 6.4.1 (о равенстве смешанных частных производных). $\square$

ЛЕММА 6.5.2. Для любой пары мультииндексов μ, ν справедливо равенство

$$\partial^\mu (x^\nu)|_{x=0} = \mu! \delta_{\mu,\nu} = \begin{cases} \mu!, & \mu = \nu, \\ 0, & \mu \neq \nu. \end{cases} \quad (6.36)$$

Если же $|\mu| \geq |\nu|$, то данное равенство справедливо не только в точке $x = 0$, но и всюду в $\mathbb{R}^n$:

$$|\mu| \geq |\nu| \Rightarrow \partial^\mu (x^\nu) = \mu! \delta_{\mu,\nu}. \quad (6.37)$$

ДОКАЗАТЕЛЬСТВО. Проверяется прямым вычислением. $\square$

Из последней леммы вытекает, что коэффициенты полиномов могут быть однозначно восстановлены по их производным в начале координат.

ТЕОРЕМА 6.5.3 (Многомерный аналог формулы бинома Ньютона). Имеет место тождество

$$(x_1 + \dots + x_n)^k = \sum_{|\mu|=k} \frac{k!}{\mu!} x^\mu. \quad (6.38)$$

ДОКАЗАТЕЛЬСТВО. Очевидно, что $(x_1 + \dots + x_n)^k = \sum_{|\mu|=k} a_\mu x^\mu$, где a_μ — некоторые числа (коэффициенты), которые нам предстоит вычислить. Возьмем мультииндекс ν порядка $|\nu| = k$ и посчитаем производную ∂^ν от левой и правой части последнего равенства. Тогда по формуле (6.37) получим $k! = a_\nu \nu!$, откуда вытекает искомое равенство $a_\nu = \frac{k!}{\nu!}$. $\square$

Пусть $P(x) = \sum_{|\mu| \leq k} a_\mu x^\mu$ — некоторый полином с постоянными коэффициентами $a_\mu \in \mathbb{R}$.

Тогда выражение $P(\partial) = \sum_{|\mu| \leq k} a_\mu \partial^\mu$ называется *линейным дифференциальным оператором с постоянными коэффициентами*. Действие этого оператора на функцию f обозначается

символом $P(\partial)f$, так что $P(\partial)f = \sum_{|\mu| \leq k} a_\mu \partial^\mu f$. Как известно, пару полиномов можно перемножить, а от пары операторов — взять композицию. Оказывается, что эти операции прекрасно согласуются.

ТЕОРЕМА 6.5.4 (Теорема о композиции дифференциальных операторов). Пусть $P(x) = \sum_{|\mu| \leq k} a_\mu x^\mu$, $R(x) = \sum_{|\nu| \leq m} a_\nu x^\nu$. Обозначим через $T(x) = P(x)R(x)$ произведение этих полиномов. Тогда имеет место тождество

$$T(\partial) = P(\partial) \circ R(\partial), \quad (6.39)$$

т.е. композиция дифференциальных операторов, порожденных парой полиномов, совпадает с оператором, порожденным произведением этих полиномов.

ДОКАЗАТЕЛЬСТВО. Любой полином может быть получен как сумма мономов, помноженных на некоторые числа (коэффициенты). А потому, поскольку операции взятия композиции и произведения хорошо согласуются (коммутируют) с операциями суммы и умножения на число, то достаточно доказать утверждение теоремы только для мономов $P(x) = x^\mu$, $R(x) = x^\nu$. Произведение этих мономов вычисляется очевидным образом: $T(x) = x^\mu \cdot x^\nu = x^{\mu+\nu}$. С другой стороны, $P(\partial) = \partial^\mu$, $R(\partial) = \partial^\nu$, и действие композиции этих операторов на функцию f по Лемме 6.5.1 равно $\partial^\mu \circ \partial^\nu[f] = \partial^\mu(\partial^\nu f) = \partial^{\mu+\nu} f$, т.е. $\partial^\mu \circ \partial^\nu = \partial^{\mu+\nu} = T(\partial)$, и формула (6.39) доказана. $\square$

§6 Дифференциал k -го порядка

Пусть V — открытое множество в $\mathbb{R}^n$, функция f дифференцируема k раз на V , и пусть имеется вектор $v = (v_1, \dots, v_n) \in \mathbb{R}^n$. Напомним, что через $\partial_v f$ обозначается производная по направлению

$$\partial_v(p)f = \lim_{t \rightarrow 0} \frac{f(p + te) - f(p)}{t} = df(p) \langle v \rangle = \sum_{j=1}^n \partial_j(p) v_j. \quad (6.40)$$

Обозначим через $\partial_v^k f$ операцию взятия k раз производной по направлению v :

$$\partial_v^k f = \overbrace{\partial_v(\partial_v(\dots(\partial_v f))\dots)}^{k \text{ раз}}.$$

Введем в рассмотрение полином $P(x) = v_1 x_1 + \dots + v_n x_n$, тогда операция взятия производной по направлению v совпадает с действием соответствующего дифференциального оператора $P(\partial) = v_1 \partial_1 + \dots + v_n \partial_n$. Тогда по построению и по Теореме 6.5.4 получаем

$$\partial_v^k f = \overbrace{P(\partial)(\dots(P(\partial)f))\dots}^{k \text{ раз}} = P^k(\partial)f.$$

По формуле бинома Ньютона, $P^k(x) = \sum_{|\mu|=k} \frac{k!}{\mu!} v^\mu x^\mu$, соответственно, $P^k(\partial) = \sum_{|\mu|=k} \frac{k!}{\mu!} v^\mu \partial^\mu$, таким образом,

$$\partial_v^k f(p) = \sum_{|\mu|=k} \frac{k!}{\mu!} v^\mu \partial^\mu f(p).$$

Обозначим через $d^k f(p)$ дифференциальную форму $d^k f(p) = \sum_{|\mu|=k} \frac{k!}{\mu!} \partial^\mu f(p) dx^\mu$. Напомним, что $dx^\mu = (dx_1)^{\mu_1} \dots (dx_n)^{\mu_n}$, причем элементарная форма dx_i действует на вектор $v = (v_1, \dots, v_n)$ по правилу: $dx_i(v) = v_i$, так что $dx^\mu \langle v \rangle = v^\mu$. В итоге получаем

$$\partial_v^k f(p) = d^k f(p) \langle v \rangle. \quad (6.41)$$

Форма $d^k f(p)$ называется *дифференциалом функции f k -го порядка в точке p* , ее аналитическая природа иллюстрируется последним равенством.

Пример 6.6.1. Дифференциал второго порядка функции $f(x)$ задается формулой

$$d^2 f(p) = \sum_{i=1}^n \partial_{ii} f(p) dx_i^2 + 2 \sum_{i=2}^n \sum_{j=1}^{i-1} \partial_{ij} f(p) dx_i dx_j. \quad (6.42)$$

Если задать симметричную матрицу второго порядка $H = (\partial_{ij} f(p))_{i,j=1,\dots,n}$, и считать dx столбцом $dx = \begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix}$, то получим представление

$$d^2 f(p) = (dx)^t H dx, \quad (6.43)$$

где символом $(dx)^t$ обозначена вектор-строка, полученная из столбца dx транспонированием: $(dx)^t = (dx_1, \dots, dx_n)$. Матрица H называется *матрицей Гессе*. Отображение $\mathbb{R}^n \ni v \mapsto d^2 f(p) \langle v \rangle = v^t H v$ называют *гессианом*, иногда его значение обозначают сокращенно Hv .

Интересно, что дифференциал второго порядка уже **НЕ** инвариантен: если отображения $f : V \subset \mathbb{R}^n \rightarrow U \subset \mathbb{R}^m$, $g : U \rightarrow \mathbb{R}^l$ — дважды дифференцируемы, то для $h = g \circ f$ справедливы тождества, то $dh(p) = \sum_{i=1}^m \partial_{ig}(f(p)) \cdot df_i(p)$, и

$$d^2 h(p) = \sum_{i=1}^m \partial_{ii} g(f(p)) df_i^2 + 2 \sum_{i=2}^m \sum_{j=1}^{i-1} \partial_{ij} g(f(p)) df_i \cdot df_j + \sum_{i=1}^m \partial_{ig}(f(p)) \cdot d^2 f_i(p).$$

Видно, что в сравнении с (6.41) появляется дополнительное (последнее) слагаемое. Оно исчезает, когда $d^2 f_i \equiv 0$, т.е. функция f *линейная*, в этом случае имеет место искомая инвариантность второго дифференциала.

§7 Формула Тейлора

Напомним, что 1-норма $|x|_1$ вектора $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ определяется по формуле $|x|_1 = |x_1| + \dots + |x_n|$. Справедлива двухсторонняя оценка

$$|x| \leq |x|_1 \leq \sqrt{n} |x|, \quad (6.44)$$

где, как обычно, $|x| = \sqrt{x_1^2 + \dots + x_n^2}$.

ТЕОРЕМА 6.7.1 (Теорема о степенной оценке приращения). Пусть V — открытое множество в $\mathbb{R}^n$ и пусть имеется шар $B = B_r(p) \subset V$. Предположим, что для отображения $f : V \rightarrow \mathbb{R}^m$ найдется константа $c \in \mathbb{R}$ такая, что на B все частные производные функции f существуют и выполнена оценка

$$\left| \frac{\partial f}{\partial x_j}(x) \right| \leq c|x - p|_1^s \quad \forall j = 1, \dots, n \quad (6.45)$$

для всех $x \in B$. Тогда для любой точки $x \in B$ справедливо неравенство

$$|f(x) - f(p)| \leq \frac{c}{s+1}|x - p|_1^{s+1} \leq \frac{c}{s+1}(\sqrt{n}|x - p|)^{s+1}. \quad (6.46)$$

ДОКАЗАТЕЛЬСТВО. Данная теорема доказывается также, как теорема 6.2.3, подробности мы опускаем (остаются в качестве упражнения). $\square$

ТЕОРЕМА 6.7.2 (Формула Тейлора с остаточным членом в форме Пеано). Пусть V — открытое множество в $\mathbb{R}^n$. Тогда, если функция $f : V \rightarrow \mathbb{R}^m$ дифференцируема k в точке $p \in V$, то

$$f(x) = f(p) + df(p)\langle x - p \rangle + \frac{1}{2}d^2f(p)\langle x - p \rangle^2 + \dots + \frac{1}{k!}d^kf(p)\langle x - p \rangle^k + o(|x - p|^k) \quad (6.47)$$

при $x \rightarrow p$.

ДОКАЗАТЕЛЬСТВО. Будем доказывать эту теорему индукцией по k . При $k = 1$ ее справедливость следует из определения дифференциала. Предположим, что она верна для $k - 1$. Рассмотрим функцию $g(x) = f(x) - f(p) - df(p)\langle x - p \rangle - \frac{1}{2}d^2f(p)\langle x - p \rangle^2 - \dots - \frac{1}{(k-1)!}d^{k-1}f(p)\langle x - p \rangle^{k-1}$. Прямым вычислением проверяется, что все частные производные этой функции до порядка k включительно в точке p равны нулю. Это означает, что все частные производные у функций $\partial_i g$ до порядка $k - 1$ включительно в точке p равны нулю. Отсюда, применяя индукционное предположение для функций $\partial_i g$, получаем, что $|\partial_i g(x)| = o(|x - p|^{k-1})$ при $x \rightarrow p$. Тогда по теореме (6.45) верна оценка $g(x) = o(|x - p|^k)$, которая равносильна искомому неравенству (6.47). $\square$

Нашей следующей целью будет обобщение формулы Тейлора с остаточным членом в интегральной форме и в форме Лагранжа.

ТЕОРЕМА 6.7.3 (Формула Тейлора с точными выражениями для остаточного члена). Пусть V — открытое множество в $\mathbb{R}^n$. Тогда, если функция $f : V \rightarrow \mathbb{R}^m$ дифференцируема $k + 1$ раз на V , и если отрезок $[p, x] = \{p + t(x - p) : t \in [0, 1]\}$ содержится в V , то

$$f(x) = f(p) + df(p)\langle x - p \rangle + \frac{1}{2}d^2f(p)\langle x - p \rangle^2 + \dots + \frac{1}{k!}d^kf(p)\langle x - p \rangle^k + \frac{1}{(k+1)!}d^{k+1}f(\xi)\langle x - p \rangle^{k+1}, \quad (6.48)$$

где ξ лежит на отрезке $[p, x]$. Справедливо также интегральное соотношение

$$\begin{aligned} f(x) - f(p) - df(p)\langle x - p \rangle - \frac{1}{2}d^2f(p)\langle x - p \rangle^2 - \dots - \frac{1}{k!}d^kf(p)\langle x - p \rangle^k = \\ = \frac{1}{k!} \int_0^1 df^{(k+1)}(y(s))\langle x - p \rangle^k \cdot (1 - s)^k ds \end{aligned} \quad (6.49)$$

где $y(s) = p + s(x - p)$, и интегрирование ведется по (одномерной) переменной s .

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения доказываемой теоремы. Зададим вектор $v = x - p$. Рассмотрим функцию одного переменного $g(t) = f(p + tv)$ для $t \in [0, 1]$. Эта функция является $k + 1$ раз дифференцируемой, причем $g^{(l)}(t) = \partial_v^l f(p + tv) = d^l f(p + tv) \langle v \rangle$ для всех порядков $l = 1, \dots, k + 1$. Тогда по формуле (3.8) получаем

$$f(x) = g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + \dots + \frac{g^{(k)}(0)}{k!} + \frac{g^{(k+1)}(\lambda)}{(k+1)!} =$$

$$f(p) + df(p) \langle x - p \rangle + \frac{1}{2}d^2 f(p) \langle x - p \rangle + \dots + \frac{1}{k!}d^k f(p) \langle x - p \rangle + \frac{1}{(k+1)!}d^{k+1} f(\xi) \langle x - p \rangle,$$
(6.50)

где λ есть некоторое число из интервала $(0, 1)$, и $\xi = p + \lambda(x - p)$. Последнее равенство эквивалентно первой из искомых формул (6.48). Аналогично получается вторая искомая формула (6.49) (из формулы (4.3)). $\square$

§8 Локальный экстремум

Пусть $V \subset \mathbb{R}^n$ — открытое множество. Точка $p \in V$ называется *точкой локального максимума* отображения f , если найдется шар $B_r(p) \subset V$ такой, что $f(p) = \max_{x \in B_r(p)} f(x)$.

Аналогично определяется точка *локального минимума*². Точки локального максимума и локального минимума называются точками *локального экстремума*.

ТЕОРЕМА 6.8.1 (Теорема Ферма — необходимое условие экстремума). Пусть V — открытое множество в $\mathbb{R}^n$, и пусть точка $p \in V$ является точкой локального экстремума у отображения $f : V \rightarrow \mathbb{R}$. Тогда, если f дифференцируемо в точке p , то $f'(p) = 0$.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы и $p \in V$ является точкой локального максимума функции f . Поскольку множество V открыто, найдется радиус $r > 0$ такой, что $B(p, r) = \{x \in \mathbb{R}^n : |p - x| < r\} \subset V$. Докажем, что производная $\frac{\partial f}{\partial x_1}(p)$ равна нулю. "Заморозим" переменные $x_2 = p_2, \dots, x_n = p_n$. Тогда функция $g : (p_1 - r, p_1 + r) \rightarrow \mathbb{R}$, определяемая формулой $g(t) = f(t, p_2, \dots, p_n)$ по построению имеет локальный максимум в точке $t = p_1$. Следовательно, по "одномерной" теореме Ферма (см. Теорему 3.1.1) справедливо равенство $g'(t)|_{t=p_1} = \frac{\partial f}{\partial x_1}(p) = 0$. Аналогично доказываются равенства $\frac{\partial f}{\partial x_2}(p) = \dots = \frac{\partial f}{\partial x_n}(p) = 0$. $\square$

ТЕОРЕМА 6.8.2 (Достаточные условия локального экстремума). Пусть V — открытое множество в $\mathbb{R}^n$, и пусть отображение $f : V \rightarrow \mathbb{R}^m$ дважды дифференцируемо в точке $p \in V$, причем $df(p) = 0$. Тогда,

(i) если гессиан $H(v) := d^2 f(p) \langle v \rangle$ является положительно (отрицательно) определенной квадратичной формой, то p есть точка строго локального минимума (максимума) функции f .

(ii) Если же существуют вектора $v, w \in \mathbb{R}^n$ такие, что $H(v) > 0$, $H(w) < 0$, то точка p не является точкой локального экстремума (хотя бы и нестрого) функции f .

ДОКАЗАТЕЛЬСТВО. Докажем сначала второе утверждение. Пусть $H(v) > 0$, $H(w) < 0$. Введем в рассмотрение функции одного переменного $\varphi(t) = f(p + tv)$, $\psi(t) = f(p + tw)$.

²а также точки *строго локального максимума*, *строго локального минимума* и т.п.

Тогда $\varphi'(0) = \psi'(0) = 0$ и $\varphi''(0) = H(v) > 0$, $\psi''(0) = H(w) < 0$. Поэтому 0 является точкой минимума для функции φ , т.е. сужение функции f на прямую $p + tv$ имеет строгий минимум в точке p . Поэтому p заведомо не может быть максимумом f в (многомерной) окрестности p . Аналогично, поскольку 0 представляет собой точку максимума для ψ , p заведомо не может быть локальным минимумом для f .

Докажем утверждение (i). Пусть $H(v) > 0$ для всех $v \in \mathbb{R}^n \setminus \{0\}$. Обозначим $\sigma = \min_{v \in \mathbb{R}^n, |v|=1} H(v)$. Поскольку единичная сфера в $\mathbb{R}^n$ является компактным множеством, по теореме Вейерштрасса указанный минимум достигается, и $\sigma > 0$. Так как гессиан $H(v)$ является квадратичной функцией, $H(tv) = t^2 H(v)$ для любого $t \in \mathbb{R}$, беря в качестве t величину $t = |v|^{-1}$, сразу же получаем неравенство $|v|^{-2} H(v) = H(v/|v|) \geq \sigma$, т.е. $H(v) \geq \sigma |v|^2$ для всех $v \in \mathbb{R}^n$. Поставляя это соотношение в формулу Тэйлора с остаточным членом в форме Пеано (6.47) при $v = x - p$ и учитывая равенство $df(p) = 0$, получаем $f(x) - f(p) = \frac{1}{2} H(v) + o(|v|^2) \geq |v|^2 (\frac{1}{2} \sigma + o(1)) > 0$ при $0 < |v| < \delta$, где δ — достаточно малый параметр. Это означает, что p является точкой минимума. Утверждение про точку максимума доказывается аналогично. $\square$

Пример 6.8.3. Найдем точки локального экстремума функции $f(x, y) = \sin x + \cos y + \cos(x - y)$ при ограничениях $0 \leq x \leq \frac{\pi}{2}$, $0 \leq y \leq \frac{\pi}{2}$. Для внутренних точек квадрата согласно теореме Ферма имеем систему уравнений

$$\partial_x f = \cos x - \sin(x - y) = 0,$$

$$\partial_y f = -\sin y + \sin(x - y) = 0,$$

откуда получаем $\cos x = \sin y = \sin(x - y)$. С учетом ограничений (принадлежности (x, y) квадрату $[0, \pi/2]^2$), заключаем, что $x - y = y = \frac{\pi}{2} - x$. Тогда $x = \frac{\pi}{3}$, $y = \frac{\pi}{6} = x - y$. Найдем гессиан H (матрицу вторых производных). Прямым вычислением

$$H = \begin{pmatrix} \partial_{xx} f & \partial_{xy} f \\ \partial_{xy} f & \partial_{yy} f \end{pmatrix} \Big|_{x=\frac{\pi}{3}, y=\frac{\pi}{6}} = \begin{pmatrix} -\sqrt{3} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\sqrt{3} \end{pmatrix}$$

Тогда $-H = \begin{pmatrix} \sqrt{3} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \sqrt{3} \end{pmatrix}$ является положительно определенной матрицей (по критерию Сильвестра, т.к. все ее главные миноры положительны), следовательно, сама H является отрицательно определенной матрицей, значит, по теореме 6.8.2 получаем, что точка $A = (\frac{\pi}{3}, \frac{\pi}{6})$ является точкой строгого локального максимума.

Интересно проверить, будет ли найденная точка A точкой *глобального* максимума функции f на квадрате $Q = \{(x, y) : 0 \leq x \leq \frac{\pi}{2}; 0 \leq y \leq \frac{\pi}{2}\}$. Внутри квадрата Q производная f' не зануляется в точках, отличных от A , поэтому других экстремумов (кроме точки A) там быть не может. Осталось проверить максимальное значение функции f на границе ∂Q , т.е. когда $x = 0$, или $x = \frac{\pi}{2}$, или $y = 0$, или $y = \frac{\pi}{2}$. Прямым вычислением устанавливается, что

$$\max_{(x,y) \in \partial Q} f(x, y) = f\left(\frac{\pi}{2}, \frac{\pi}{4}\right) = 1 + \sqrt{2} < \frac{3\sqrt{3}}{2} = f(A)$$

(проверьте!!). Следовательно, A является точкой глобального максимума f на Q . Для полноты картины отметим, что

$$\min_{(x,y) \in Q} f(x, y) = \min_{(x,y) \in \partial Q} f(x, y) = f(0, 0) = 0.$$

Упражнение 6.8.4. Найдите максимальное и минимальное значение функции f из приведенного выше примера на всей плоскости $\mathbb{R}^2$.

Глава 7

Теорема об обратном отображении

§1 Предварительные сведения

ТЕОРЕМА 7.1.1. Если множество $E \subset \mathbb{R}^n$ — компактно, а отображение $f : E \rightarrow K \subset \mathbb{R}^m$ — непрерывно и инъективно, $K = f(E)$, то обратное отображение $f^{-1} : K \rightarrow E$ также непрерывно, т.е. $f : E \rightarrow K$ представляет собой гомеоморфизм.

ТЕОРЕМА 7.1.2 (Теорема Брауэра об инвариантности области). Если множество $U \subset \mathbb{R}^n$ открыто в $\mathbb{R}^n$, и отображение $f : U \rightarrow V \subset \mathbb{R}^n$ непрерывно и инъективно, $f(U) = V$, то множество V также открыто в $\mathbb{R}^n$, и обратное отображение $f^{-1} : V \rightarrow U$ также непрерывно, т.е. f представляет собой гомеоморфизм.

СЛЕДСТВИЕ 7.1.3. Если отображение $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ непрерывно и инъективно, и $p \in \text{Int } U$, то $f(p) \in \text{Int } f(U)$.

СЛЕДСТВИЕ 7.1.4. Если $m < n$, $E \subset \mathbb{R}^n$ и $\text{Int } E \neq \emptyset$, то не существует непрерывного инъективного отображения $f : E \rightarrow \mathbb{R}^m$.

§2 Теорема о липшицевом гомеоморфизме

Следующая теорема, несмотря на устрашающее название, является весьма важной и полезной, а кроме того, у нее очень красивое доказательство.

ТЕОРЕМА 7.2.1. Пусть $V \subset \mathbb{R}^n$ — открытое множество, и пусть $f : V \rightarrow \mathbb{R}^n$ — сжимающее отображение. Тогда отображение $\varphi : V \ni x \mapsto x - f(x)$ является липшицевым, его образ $U = \varphi(V)$ представляет собой открытое множество в $\mathbb{R}^n$, и обратное отображение $\varphi^{-1} : U \rightarrow V$ также удовлетворяет условию Липшица.

ДОКАЗАТЕЛЬСТВО. Пусть отображение f является сжимающим с коэффициентом $c < 1$, т.е. $|f(x) - f(y)| \leq c|x - y|$. Тогда отображение φ липшицево с константой $(1 + c)$, т.к. $|\varphi(x) - \varphi(y)| = |x - y - (f(x) - f(y))| \leq |x - y| + |f(x) - f(y)| \leq (1 + c)|x - y|$. По тем же причинам легко видеть, что отображение φ инъективно:

$$|\varphi(x) - \varphi(y)| = |x - y - (f(x) - f(y))| \geq |x - y| - |f(x) - f(y)| \geq (1 - c)|x - y| > 0 \quad (7.1)$$

при $x \neq y$.

Возьмем произвольную точку $p \in V$. В силу условия открытости, найдется число $r > 0$ такое, что $\bar{B}_r(p) = \{x \in \mathbb{R}^n : |p - x| \leq r\} \subset V$. Обозначим $q = \varphi(p)$. Утверждается, что тогда справедливо включение

$$\bar{B}_{(1-c)r}(q) \subset U. \quad (7.2)$$

В самом деле, возьмем произвольно точку $y \in \bar{B}_{(1-c)r}(q)$ и зафиксируем ее. По построению

$$(1-c)r \geq |y - q| = |y - p + f(p)|. \quad (7.3)$$

При фиксированном y рассмотрим отображение $g(x) = f(x) + y$. Это отображение (полученное из f сдвигом на постоянный вектор y) будет, конечно, сжимающим с тем же самым коэффициентом c . Утверждается, что

$$g(x) \in \bar{B}_r(p) \quad \forall x \in \bar{B}_r(p). \quad (7.4)$$

В самом деле, если $|x - p| \leq r$, то в силу условия сжимаемости $|f(x) - f(p)| \leq cr$ и

$$|g(x) - p| = |f(x) + y - p| = |f(x) - f(p) + y - p + f(p)| \leq |f(x) - f(p)| + |y - p + f(p)| \stackrel{(7.3)}{\leq} cr + (1-c)r = r.$$

Таким образом, g есть сжимающее отображение шара $\bar{B}_r(p)$ в себя. Поэтому g имеет в точности одну неподвижную точку: $\exists! x_* \in \bar{B}_r(p) \quad g(x_*) = x_*$. Другими словами, $f(x_*) + y = x_* \Leftrightarrow y = x_* - f(x_*) = \varphi(x_*)$. Значит, $y \in \varphi(V) = U$. Искомое включение (7.2) доказано. Вместе с тем доказана открытость множества U . Лишшицевость обратного отображения $\varphi^{-1} : U \rightarrow V$ с коэффициентом $\frac{1}{1-c}$ вытекает из неравенства (7.1). $\square$

§3 Теорема об обратной функции. Диффеоморфизмы

Отображение $\varphi : V \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ называется C^1 -гладким, если все его частные производные непрерывны. Отображение $\varphi : V \subset \mathbb{R}^n \rightarrow U \subset \mathbb{R}^n$ называется диффеоморфизмом, если и оно само, и обратное к нему отображение $\varphi^{-1} : U \rightarrow V$ являются C^1 -гладкими. Из этого определения мгновенно вытекает, что отображение представляет собой диффеоморфизм тогда и только тогда, когда обратное к нему — диффеоморфизм.

Упражнение 7.3.1. Докажите, что линейное отображение $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ является диффеоморфизмом тогда и только тогда, когда его матрица невырождена (т.е. ее детерминант отличен от нуля).

ТЕОРЕМА 7.3.2. Пусть V, U, W — открытые множества в $\mathbb{R}^n$, и пусть $f : V \rightarrow U$, $g : U \rightarrow W$ — диффеоморфизмы. Тогда их композиция $g \circ f : V \rightarrow W$ также является диффеоморфизмом.

Данная теорема почти мгновенно вытекает из соответствующих определений, поэтому ее доказательство мы опускаем.

ТЕОРЕМА 7.3.3. Пусть V, U — открытые множества в $\mathbb{R}^n$, и $\varphi : V \rightarrow U$ — C^1 -гладкий гомеоморфизм, причем $\det \varphi'(p) \neq 0$ для каждого $p \in V$. Тогда φ представляет собой диффеоморфизм, т.е. обратное к нему отображение также C^1 -гладко.

ДОКАЗАТЕЛЬСТВО. Обозначим через $\psi : U \rightarrow V$ отображение, обратное к φ . Нам требуется проверить гладкость ψ . Зафиксируем произвольно точку $p \in V$, и обозначим $q = \varphi(p)$, $L = d\varphi(p)$. Тогда по определению дифференцируемости имеет место представление

$$\varphi(x) - q = L(x - p) + o(|x - p|)$$

при $x \rightarrow p$. Обозначим через T линейное преобразование пространства $\mathbb{R}^n$, обратное к L . Подействуем оператором T на левую и правую части последнего равенства. Тогда получим $T(\varphi(x) - q) = x - p + o(|x - p|)$. Обозначим $y = \varphi(x)$, тогда $\psi(y) = x$, $\psi(q) = p$, и предыдущее равенство можно переписать в виде

$$T(y - q) = \psi(y) - \psi(q) + o(|\psi(y) - p|) \quad (7.5)$$

при $y \rightarrow q$ (в самом деле, сходимости $x \rightarrow p$ и $y \rightarrow q$ равносильны в силу непрерывности φ и ψ). Беря норму от обеих частей данного равенства, и учитывая неравенство треугольника, имеем

$$|\psi(y) - \psi(q)|(1 + o(1)) = |T(y - q)| \leq \|T\| \cdot |y - q|$$

(последнее неравенство мгновенно вытекает из определения операторной нормы). Отсюда следует, что $|\psi(y) - \psi(q)| \leq C|y - q|$ для некоторой положительной константы C при $y \rightarrow q$. Значит, выполнено соотношение $o(|\psi(y) - p|) = o(|y - q|)$ при $y \rightarrow q$, тогда формулу (7.5) можно переписать в виде

$$T(y - q) = \psi(y) - \psi(q) + o(|y - p|) \Leftrightarrow \psi(y) = \psi(q) + T(y - q) + o(|y - p|) \quad \text{при } y \rightarrow q, \quad (7.6)$$

которая равносильна искомой дифференцируемости функции $\psi(y)$ в точке q , причем ее дифференциал равен $d\psi(q) = T = L^{-1} = (df(p))^{-1}$. Соответственно, имеет место тождество $\psi'(q) = [\varphi'(p)]^{-1}$. Отсюда с помощью стандартных формул Крамера легко выводится непрерывность отображения $U \ni q \mapsto \psi'(q)$, т.е. все частные производные функции ψ непрерывны на U . Теорема доказана. $\square$

ТЕОРЕМА 7.3.4 (Теорема об обратной функции). Пусть $V \subset \mathbb{R}^n$ — открытое множество, и пусть $\varphi : V \rightarrow \mathbb{R}^n$ — C^1 -гладкое отображение. Тогда, если в точке $p \in V$ выполнено условие $\det \varphi'(p) \neq 0$, то существует шар $B = B_\delta(p)$ такой, что множество $W = \varphi(B)$ открыто и сужение $\varphi : B \rightarrow W$ является диффеоморфизмом (т.е. оно биективно и обратное к нему отображение $\varphi^{-1} : W \rightarrow B$ также C^1 -гладкое).

ДОКАЗАТЕЛЬСТВО. Предположим сначала, что $\varphi'(p)$ совпадает с единичной матрицей, т.е. что линейное отображение $d\varphi(p)$ — тождественное, $d\varphi(p) = \text{Id}$. Рассмотрим отображение $f = \text{Id} - \varphi$. По построению, $f'(p) = 0$. Тогда найдется $\delta > 0$ такое, что $|\partial_i f(x)| \leq \frac{1}{3\sqrt{n}}$ для всех $x \in B = B_\delta(p)$. Отсюда по Теореме 6.2.3 вытекает оценка $|f(x) - f(y)| \leq \frac{1}{3}|x - y|$ для всех $x, y \in B$, т.е. отображение $f : B \rightarrow \mathbb{R}^n$ является сжимающим с коэффициентом $\frac{1}{3}$. Обозначим $W = \varphi(B)$, $q = \varphi(p)$. Тогда по Теореме 7.2.1 множество W открыто, сужение $\varphi : B \rightarrow W$ является непрерывным, инъективным, и обратное к нему $\varphi^{-1} : W \rightarrow B$ также непрерывно и инъективно. Следовательно, по предыдущей Теореме 7.3.3 отображение $\varphi : B \rightarrow W$ является диффеоморфизмом. Теорема доказана для случая $df(p) = \text{Id}$.

В общем случае, обозначим через L линейное отображение $L = df(p)$, и рассмотрим функцию $\phi = L^{-1} \circ \varphi$. По теореме 6.2.7, отображение ϕ является C^1 -гладким, и $d\phi(p) = \text{Id}$, т.е. мы попали в уже только что разобранный случай. По доказанному выше, найдется

шар $B = B_\delta(p)$ такой, что отображение $\phi : B \rightarrow W_1$ с $W_1 = \phi(B)$ является диффеоморфизмом. С другой стороны, по построению $\varphi = L \circ \phi$, причем L также является диффеоморфизмом (см. упражнение 7.3.1). Тогда по Теореме 7.3.2 отображение $\varphi : B \rightarrow W$ с $W = \varphi(B) = L(W_1)$ — диффеоморфизм (поскольку φ удалось представить как композицию диффеоморфизмов). $\square$

Замечание 7.3.5 (О классах C^k). Отображение $f : U \rightarrow \mathbb{R}^m$ открытого множества $U \subset \mathbb{R}^n$ принадлежит классу C^k , если все частные производные f непрерывны и принадлежат классу C^{k-1} . Отображение называется диффеоморфизмом класса C^k , если оно само и обратное к нему принадлежит классу C^k . Предыдущие теоремы обобщаются на случай отображений класса C^k , а именно: композиция отображений класса C^k принадлежит классу C^k , если гомеоморфизм принадлежит классу C^k и его дифференциал в каждой точке невырожден, то и обратное отображение принадлежит классу C^k , наконец, если отображение принадлежит классу C^k и его дифференциал невырожден, то такое отображение локально представляет собой диффеоморфизм класса C^k (т.е. у любой точки из области определения найдется шаровая окрестность такая, что сужение отображения на эту окрестность представляет собой диффеоморфизм класса C^k).

Следствие 7.3.6. Пусть $V \subset \mathbb{R}^n$ — открытое множество, и пусть $\varphi : V \rightarrow \mathbb{R}^n$ — отображение класса C^k , $k \geq 1$. Предположим, что $\det \varphi'(p) \neq 0$ в каждой точке $p \in V$. Тогда множество $U = \varphi(V)$ открыто. Далее, если отображение $\varphi : V \rightarrow U$ вдобавок инъективно, то оно представляет собой диффеоморфизм класса C^k .

Пример 7.3.7. Отождествляя точку плоскости $(x, y) \in \mathbb{R}^2$ с комплексным числом $z = x + iy$, рассмотрим отображение $f : \mathbb{C} \rightarrow \mathbb{C}$, задаваемое равенством $f(z) = u + iv = e^z$. Как известно, это отображение определяется с помощью формул Эйлера $e^z = e^x(\cos y + i \sin y)$. На языке вещественных функций, $f(x, y) = (u(x, y), v(x, y)) = e^x(\cos y, \sin y)$. Тогда соответствующая матрица Якоби будет иметь вид

$$f'(x, y) = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = e^x \begin{pmatrix} \cos y & -\sin y \\ \sin y & \cos y \end{pmatrix}.$$

Очевидно, что она обратима в каждой точке $(x, y) \in \mathbb{R}^2$, причем обратная матрица имеет вид

$$[f'(x, y)]^{-1} = e^{-x} \begin{pmatrix} \cos y & \sin y \\ -\sin y & \cos y \end{pmatrix}. \quad (7.7)$$

Таким образом, выполнены все условия теоремы об обратной функции. Отсюда получаем, в частности, что множество $f(\mathbb{R}^2)$ должно быть открыто, и отображение f должно быть локально инъективно. Справедливость этих утверждений более чем очевидно.

Интересно, что глобальной инъективности нет, т.к. $f(x, y + 2\pi) \equiv f(x, y)$.

Интересно также, что матрица Якоби $g'(u, v)$ обратного отображения $g = f^{-1}$ может быть подсчитана двумя способами. Можно воспользоваться равенством $g'(u, v) = [f'(x, y)]^{-1}$, и в формуле (7.7) перейти к переменным (u, v) , используя равенства $e^x \cos y = u$, $e^x \sin y = v$. Можно сделать и по-другому: выразив x, y через u, v непосредственно $x = \frac{1}{2} \ln(u^2 + v^2)$, $y = \arctg \frac{v}{u}$, продифференцировать последние два равенства и посчитать матрицу Якоби $g'(u, v) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$ напрямую. Разумеется, в обоих случаях получим

одинаковый результат

$$g'(u, v) = \frac{1}{u^2 + v^2} \begin{pmatrix} u & v \\ -v & u \end{pmatrix}.$$

§4 Теорема о неявной функции

ТЕОРЕМА 7.4.1 (Теорема о неявной функции). Пусть $W \subset \mathbb{R}^{n+m}$ — открытое множество, и пусть $f : W \rightarrow \mathbb{R}^m$ — отображение класса C^k , $k \geq 1$, $f = f(x, y)$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. Предположим, что множество $M = \{(x, y) \in W : f(x, y) = \mathbf{0} \in \mathbb{R}^m\}$ непусто, и что точка $(p, q) \in M$ такова, что

$$\det \frac{\partial f}{\partial y}(p, q) \neq 0, \quad (7.8)$$

где через $\frac{\partial f}{\partial y}$ обозначена квадратная $m \times m$ матрица с элементами $\frac{\partial f_i}{\partial y_j}$, $i, j = 1, \dots, m$. Тогда найдется окрестности U, V точек p, q соответственно и C^k -гладкая функция $g : U \rightarrow V$ такие, что множество $(U \times V) \cap M$ совпадает с графиком функции g , т.е.

$$\forall (x, y) \in U \times V \quad [f(x, y) = 0 \Leftrightarrow y = g(x)]. \quad (7.9)$$

ДОКАЗАТЕЛЬСТВО. Рассмотрим отображение $\varphi : W \rightarrow \mathbb{R}^{n+m}$, задаваемое формулой $\varphi(x, y) = (x, f(x, y))$. Нетрудно видеть, что квадратная $(n + m) \times (n + m)$ матрица Якоби этого отображения имеет вид

$$\varphi'(x, y) = \begin{pmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} E & 0 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} \quad (7.10)$$

где через E обозначена единичная $n \times n$ матрица, а через 0 — нулевая $n \times m$ матрица. Отсюда $\det \varphi'(p, q) = \det \frac{\partial f}{\partial y}(p, q) \neq 0$. Значит, по теореме об обратной функции, найдется шар $B = B_\delta(p, q)$ такой, что множество $\varphi(B)$ открыто и сужение $\varphi : B \rightarrow \varphi(B)$ с является диффеоморфизмом, т.е. оно биективно и обратное к нему отображение $\varphi^{-1} : \varphi(B) \rightarrow B$ также C^k -гладкое. Обозначем $\psi = \varphi^{-1}$. По построению, $\varphi(p, q) = (p, 0)$. Тогда найдутся шары U, V с центрами в точках p, q соответственно такие, что

$$U \times V \subset B, \quad U \times \{0\} \subset \varphi(B). \quad (7.11)$$

Рассмотрим отображение $g : U \rightarrow \mathbb{R}^m$, определяемое формулой $g(x) = \text{Pr}_y(\psi(x, 0))$, где Pr_y — проекция на пространство переменных $y \in \mathbb{R}^m$, определяемая формулой $\text{Pr}_y(x, y) = y$ при $x \in \mathbb{R}^n, y \in \mathbb{R}^m$.

По построению, если $y = g(x)$, то $(x, y) = \psi(x, 0)$, следовательно, $(x, f(x, y)) = \varphi(x, y) = \varphi(\psi(x, 0)) = (x, 0)$, значит, $f(x, y) = 0$. Обратно, пусть $f(x, y) = 0$, $x \in U$, $y \in V$. По построению (см. (7.11)), $(x, y) \in B$, значит, $(x, y) = \psi(\varphi(x, y)) = \psi(x, f(x, y)) = \psi(x, 0)$. Тогда $y = \text{Pr}_y(x, y) = \text{Pr}_y \psi(x, 0) = g(x)$, что и требовалось доказать.

В дальнейшем нам понадобятся формулы для вычисления старших производных функции g . Поскольку для всех $x \in U$ справедливо тождество $f(x, g(x)) \equiv 0$, то $\frac{\partial f(x, g(x))}{\partial x} = 0$, поэтому, дифференцируя $f(x, g(x))$ как сложную функцию, получаем

$$\frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x)) \frac{\partial g}{\partial x}(x) = 0,$$

где через $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial g}{\partial x}$ обозначены соответствующие матрицы Якоби (матрицы частных производных) размерности $m \times n$, $m \times m$ и $m \times n$ соответственно. Тогда

$$\frac{\partial g}{\partial x}(x) = - \left[\frac{\partial f}{\partial y}(x, g(x)) \right]^{-1} \frac{\partial f}{\partial x}(x, g(x)) \quad \forall x \in U, \quad (7.12)$$

где через $\left[\frac{\partial f}{\partial y} \right]^{-1}$ обозначена $m \times m$ матрица, обратная к $\frac{\partial f}{\partial y}$. Расписывая j -й столбец из последнего матричного тождества, получаем

$$\frac{\partial g}{\partial x_j}(x) = - \left[\frac{\partial f}{\partial y}(x, g(x)) \right]^{-1} \frac{\partial f}{\partial x_j}(x, g(x)) \quad \forall x \in U. \quad (7.13)$$

□

Пример 7.4.2. Состояние газа характеризуется формулой Менделеева–Клайперона

$$pV = \nu RT, \quad (7.14)$$

где p, V, T — давление, объем и температура изучаемого газа соответственно. К этому соотношению присоединяется еще и уравнение состояния

$$pV^\gamma = \text{const}, \quad (7.15)$$

где γ — некоторая постоянная (отличная от единицы). Таким образом, три величины p, V, T связаны двумя соотношениями, это дает надежду, взяв одну из них в качестве независимой переменной, выразить через нее две остальные. Чтобы удостовериться в возможности такого выражения, требуется проверить выполнение условий теоремы о неявной функции. Положим $f_1(p, V, T) = pV - \nu RT$, $f_2(p, V, T) = pV^\gamma - \text{const}$, тогда соответствующая матрица Якоби будет иметь вид

$$\begin{pmatrix} \frac{\partial f_1}{\partial p} & \frac{\partial f_1}{\partial V} & \frac{\partial f_1}{\partial T} \\ \frac{\partial f_2}{\partial p} & \frac{\partial f_2}{\partial V} & \frac{\partial f_2}{\partial T} \end{pmatrix} = \begin{pmatrix} V & p & -\nu R \\ V^\gamma & \gamma p V^{\gamma-1} & 0 \end{pmatrix}.$$

Нетрудно проверить, что у этой матрицы любой минор 2×2 является невырожденным. Поэтому выполнены предположения о неявной функции и, действительно, любые две из этих величин можно гладким отображением выразить через третью, например, давление и объем выразить через температуру. Отметим, что теорема о неявной функции утверждает *локальную* возможность такого выражения; нетрудно видеть, однако, что в данном случае $p(T)$, $V(T)$ можно выразить из уравнений (7.14)–(7.15) и глобально. Возможность аналогично с помощью этих уравнений выразить p, T через V или V, T через p еще более очевидна.

Пример 7.4.3. Состояние двойного математического маятника задается координатами двух грузов (x, y) и (u, v) соответственно. Первый груз соединен с подвесом (находящимся в начале координат) стержнем длины R , и также оба груза соединены друг с другом стержнем длины r . Получаем два уравнения состояния

$$x^2 + y^2 = R^2, \quad (7.16)$$

$$(x - u)^2 + (y - v)^2 = r^2. \quad (7.17)$$

Таким образом, четыре величины x, y, u, v связаны двумя соотношениями, и можно попробовать, взяв две из них в качестве независимых переменных, выразить через них остальные две. Проверим, для каких пар и когда выполняются условия теоремы о неявной функции. Положим $f_1(x, y, u, v) = x^2 + y^2 - R^2$, $f_2(x, y, u, v) = (x - u)^2 + (y - v)^2 - r^2$, тогда соответствующая матрица Якоби будет иметь вид

$$\begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{pmatrix} = 2 \begin{pmatrix} x & y & 0 & 0 \\ (x - u) & (y - v) & (u - x) & (v - y) \end{pmatrix}.$$

1) Нетрудно видеть, что минор $\frac{\partial(f_1, f_2)}{\partial(u, v)}$ (образованный последними двумя столбцами) тождественно равен нулю. Это означает, что теорема о неявной функции не позволяет выразить u, v через x, y . Данное утверждение более чем очевидно геометрически: если мы фиксируем положение первого груза, то ничто не мешает второму грузу двигаться в произвольном направлении по соответствующей окружности радиуса r .

2) Посчитаем теперь минор $\frac{\partial(f_1, f_2)}{\partial(x, y)}$, образованный первыми двумя столбцами. Он равен $x(y - v) - y(x - u) = uy - vx$. Данный минор зануляется, когда вектора (x, y) и (u, v) линейно зависимы, т.е. когда оба груза лежат на одной прямой. Действительно, в этом случае нетрудно проверить, что даже локально нельзя выразить (x, y) через (u, v) . А именно, если мы сдвинем второй груз вдоль той же самой прямой на какое-то расстояние по направлению к первому грузу, то у того появится два варианта для смещения (две точки, лежащие симметрично по левую и по правую стороны от указанной прямой, соединяющей точки подвеса и второго груза). В остальных случаях, в соответствии с теоремой о неявной функции, локально положение первого груза (x, y) однозначно и гладко восстанавливается по положению второго груза (u, v) . Впрочем, глобально такой однозначности нет (в силу все той же симметрии).

3) Менее очевидны варианты с остальными парами. Например, посчитаем минор $\frac{\partial(f_1, f_2)}{\partial(y, u)}$, образованный вторым и третьим столбцами. Он равен $y(u - x)$. Следовательно, когда $y \neq 0 \neq u - x$, то локально можно выразить (y, u) через (x, v) с помощью гладкой функции, а вот если $y = 0$ или $u = x$, то возможность такого однозначного гладкого выражения сомнительна. И в самом деле, при $y = 0 \neq x - u$ можно, оставляя неизменной координату v , сдвинуть первый груз в одну из двух точек, симметрично расположенных относительно прямой $y = 0$. У обеих точек сдвига координата x одинаковая (вследствие симметрии), а координата v второго груза, повторимся, может остаться неизменной (это означает, что второй груз движется по прямой $v = \text{const}$). Таким образом, система может принимать два разных положения при одинаковых значениях (x, v) (причем локально сколь угодно близких к начальному положению), что и следовало ожидать. Рекомендуются разобрать оставшиеся ситуации с $x = u$, а также оставшиеся пары координат.

§5 Правило множителей Лагранжа для поиска условного экстремума

ТЕОРЕМА 7.5.1 (Необходимый признак условного экстремума). Пусть $W \subset \mathbb{R}^{n+m}$ — открытое множество, и пусть $f : W \rightarrow \mathbb{R}^m$ — отображение класса C^k , $k \geq 1$, $f =$

$(f_1, \dots, f_m)$. Предположим, что множество $M = \{z \in W : f(z) = 0\}$ непусто, и что некоторая C^1 -гладкая функция $h : W \rightarrow \mathbb{R}$ достигает локального минимума в точке $z_* \in M$ на множестве M , т.е.¹

$$\exists r > 0 \forall z \in M \quad |z - z_*| < r \Rightarrow f(z) \geq f(z_*). \quad (7.18)$$

Тогда, если выполнено условие регулярности

$$\text{rank } f'(p, q) = m, \quad (7.19)$$

то найдутся числа $\lambda_1, \dots, \lambda_m$ такие, что функция $\tilde{h}(z) = h(z) - \lambda_1 f_1(z) - \dots - \lambda_m f_m(z)$ удовлетворяет условию

$$\tilde{h}'(z_*) = 0, \quad (7.20)$$

т.е. все частные производные функции $\tilde{h}$ в точке z_* равны нулю.

ДОКАЗАТЕЛЬСТВО. Будем представлять переменную $z \in \mathbb{R}^{n+m}$ в виде $z = (x, y)$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. Аналогично, $z_* = (p, q)$, $p \in \mathbb{R}^n$, $q \in \mathbb{R}^m$. Не умаляя общности, можно считать, что

$$\det \frac{\partial f}{\partial y}(p, q) \neq 0, \quad (7.21)$$

где через $\frac{\partial f}{\partial y}$ обозначена квадратная $m \times m$ матрица с элементами $\frac{\partial f_i}{\partial y_j}$, $i, j = 1, \dots, m$. Тогда по теореме о неявной функции найдутся окрестности U, V точек p, q соответственно и C^k -гладкая функция $g : U \rightarrow V$ такие, что множество $(U \times V) \cap M$ совпадает с графиком функции g , т.е.

$$\forall (x, y) \in U \times V \quad [f(x, y) = 0 \Leftrightarrow y = g(x)]. \quad (7.22)$$

Рассмотрим функцию $\omega : U \rightarrow \mathbb{R}$, вычисляемую по формуле $\omega(x) = h(x, g(x))$. Тогда по построению функция ω будет C^1 -гладкой, причем p является точкой локального минимума для ω :

$$\exists r > 0 \forall x \in U \quad |x - p| < r \Rightarrow \omega(x) \geq \omega(p). \quad (7.23)$$

Тогда по теореме Ферма 6.8.1 имеем $\frac{\partial \omega}{\partial x}(p) = 0$, где через $\frac{\partial \omega}{\partial x}(p)$ обозначен вектор пространства $\mathbb{R}^n$, состоящий из частных производных $\frac{\partial \omega}{\partial x}(p) = \left(\frac{\partial \omega}{\partial x_1}, \frac{\partial \omega}{\partial x_2}, \dots, \frac{\partial \omega}{\partial x_n} \right)$. Напомним, что по теореме о неявной функции 7.4.1 производные функции g вычисляются по формуле

$$\frac{\partial g}{\partial x}(x) = - \left[\frac{\partial f}{\partial y}(x, g(x)) \right]^{-1} \frac{\partial f}{\partial x}(x, g(x)) \quad \forall x \in U, \quad (7.24)$$

где через $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial g}{\partial x}$ обозначены соответствующие матрицы Якоби (матрицы частных производных) размерности $m \times n$, $m \times m$ и $m \times n$ соответственно. Тогда по формулам для производных сложной функции $\omega(x) = h(x, g(x))$ получаем

$$0 = \frac{\partial \omega}{\partial x}(p) = \frac{\partial h}{\partial x}(p, q) + \frac{\partial h}{\partial y}(p, q) \frac{\partial g}{\partial x}(p) = \frac{\partial h}{\partial x}(p, q) - \frac{\partial h}{\partial y}(p, q) \left[\frac{\partial f}{\partial y}(p, q) \right]^{-1} \frac{\partial f}{\partial x}(p, q). \quad (7.25)$$

Обозначим

$$\lambda = \frac{\partial h}{\partial y}(p, q) \left[\frac{\partial f}{\partial y}(p, q) \right]^{-1}.$$

¹Это так называемый *условный экстремум*, т.е. при условии $f(z) = 0$.

Поскольку матрица $\frac{\partial h}{\partial y}$ имеет размерность $1 \times m$, а матрица $\left[\frac{\partial f}{\partial y}(p, q)\right]^{-1}$ — соответственно размерность $m \times m$, то их произведение λ имеет размерность $1 \times m$, т.е. является вектор-строкой пространства $\mathbb{R}^m$: $\lambda = (\lambda_1, \dots, \lambda_m)$. Тогда формула (7.25) предстанет в виде

$$0 = \frac{\partial \omega}{\partial x}(p) = \frac{\partial h}{\partial x}(p, q) - \lambda \frac{\partial f}{\partial x}(p, q) = \frac{\partial h}{\partial x}(p, q) - \frac{\partial(\lambda \cdot f)}{\partial x}(p, q). \quad (7.26)$$

Введем в рассмотрение новую функцию $\tilde{h}(z) = h(z) - \lambda \cdot f(z) = h(z) - \lambda_1 f_1(z) - \dots - \lambda_m f_m(z)$. Тогда формула (7.26) означает, что

$$\frac{\partial \tilde{h}}{\partial x}(p, q) = 0. \quad (7.27)$$

По построению,

$$\tilde{h}(x, y) = h(x, y) - \frac{\partial h}{\partial y}(p, q) \left[\frac{\partial f}{\partial y}(p, q)\right]^{-1} f(x, y).$$

Дифференцируя это равенство по y , получаем прямым вычислением

$$\frac{\partial \tilde{h}}{\partial y}(p, q) = \frac{\partial h}{\partial y}(p, q) - \frac{\partial h}{\partial y}(p, q) \left[\frac{\partial f}{\partial y}(p, q)\right]^{-1} \frac{\partial f}{\partial y}(p, q) = \frac{\partial h}{\partial y}(p, q) - \frac{\partial h}{\partial y}(p, q) = 0. \quad (7.28)$$

Равенства (7.27)–(7.28) приводят нас к искомой формуле (7.20), т.е. все частные производные функции $\tilde{h}$ в точке (p, q) равны нулю, что и требовалось доказать. $\square$

Отметим, что построенная функция $\tilde{h}(z) = h(z) - \lambda_1 f_1(z) - \dots - \lambda_m f_m(z)$ часто называется *функцией Лагранжа*. Доказанная теорема дает рецепт поиска "подозрительных" точек, которые могут оказаться точками экстремума (это критические точки функции Лагранжа). Следующая теорема в некоторых случаях позволяет удостовериться, что эти "подозрительные" точки в самом деле являются экстремальными.

ТЕОРЕМА 7.5.2 (Достаточный признак условного экстремума). Пусть $W \subset \mathbb{R}^{n+m}$ — открытое множество, и пусть $f : W \rightarrow \mathbb{R}^m$ — отображение класса C^k , $k \geq 1$, $f = (f_1, \dots, f_m)$. Предположим, что множество $M = \{z \in W : f(z) = 0\}$ непусто, и что некоторая C^2 -гладкая функция $h : W \rightarrow \mathbb{R}$ имеет нулевой дифференциал² в точке $z_* \in M$, т.е.

$$h'(z_*) = 0. \quad (7.29)$$

Обозначим через $T_p M$ линейное подпространство

$$T_p M = \{v \in \mathbb{R}^n : df_1(z_*)\langle v \rangle = df_2(z_*)\langle v \rangle = \dots = df_m(z_*)\langle v \rangle = 0\}.$$

Предположим, далее, что

$$H(v) > 0 \quad \forall v \in T_p M \setminus \{0\}, \quad (7.30)$$

где $H(v)$ — гессиан (квадратичная форма, порожденная симметричной матрицей вторых производных) функции h в точке z_* . Тогда точка z_* является точкой локального минимума для функции h на M .

²Обычно в качестве h берется функция Лагранжа $\tilde{h}$, см. предыдущую теорему.

ДОКАЗАТЕЛЬСТВО. Пусть выполнены предположения теоремы. Допустим, что ее заключение неверно. Тогда найдется последовательность точек $z_j \in M$ таких, что $z_j \neq z_*$,

$$z_j \rightarrow z_* \quad \text{при } j \rightarrow \infty, \quad (7.31)$$

и

$$h(z_j) \leq h(z_*) \quad \forall j \in \mathbb{N}. \quad (7.32)$$

Не умаляя общности, можно считать, что $h(z_*) = 0$. По формуле Тейлора с остаточным членом в форме Пеано имеем

$$h(z_j) = H(z_j - z_*) + o(|z_j - z_*|^2) \leq 0 \quad (7.33)$$

(здесь мы учли равенство нулю первого дифференциала h в точке z_*). Обозначим $t_j = |z_j - z_*|$, $v_j = \frac{z_j - z_*}{|z_j - z_*|}$. Тогда

$$|v_j| \equiv 1 \quad (7.34)$$

и $t_j^2 [H(v_j) + o(1)] \leq 0$, деля последнее равенство на t_j^2 , получаем

$$H(v_j) + o(1) \leq 0 \quad \text{при } j \rightarrow \infty. \quad (7.35)$$

Поскольку последовательность векторов v_j равномерно ограничена, то из нее можно по теореме Вейерштрасса выделить сходящуюся подпоследовательность $v_j \rightarrow v_*$. Переходя к пределу в соотношениях (7.34)–(7.35), получаем

$$|v_*| = 1; \quad H(v_*) \leq 0. \quad (7.36)$$

По построению имеем $f(z_*) = f(z_j) = 0$. Применяя еще раз формулу Тейлора для f , получаем

$$0 = f(z_j) = df(z_*) < z_j - z_* > + o(|z_j - z_*|) = t_j [df(z_*) < v_j > + o(1)],$$

следовательно,

$$0 = df(z_*) < v_j > + o(1) \quad \text{при } j \rightarrow \infty.$$

Переходя в последнем равенстве к пределу, получаем

$$0 = df(z_*) < v_* >,$$

что равносильно (по определению) включению $v_* \in T_p M$. Последнее включение, вместе с формулой (7.36), вступает в противоречие с предположением (7.30). Полученное противоречие завершает доказательство теоремы. $\square$

Пример 7.5.3. I. Найдем точки условного экстремума функции $h(x, y) = xy$ при условии $x + y = 1$. Здесь условие связи³ задается функцией $f(x, y) = x + y$, соответственно, функция Лагранжа имеет вид $L = h - \lambda f = xy - \lambda(x + y)$. Запишем необходимое условие локального экстремума, в соответствии с Теоремой 7.5.1.

$$f(x, y) = \begin{cases} \frac{\partial L}{\partial x} = y - \lambda = 0; \\ \frac{\partial L}{\partial y} = x - \lambda = 0; \\ f(x, y) = x + y = 1. \end{cases} \quad (7.37)$$

³В Теореме 7.5.1 фигурировало условие связи $f(x, y) = 0$, а в данном случае оно задается соотношением $f(x, y) = 1$. Это различие несущественно, т.к. сдвиг функции на константу не влияет на нахождение производных и на поиск экстремума.

Мы видим, что решением этой системы является точка $x_* = y_* = \frac{1}{2}$. Соответственно, $\lambda = \frac{1}{2}$, $L = xy - \frac{1}{2}(x+y)$. Найдем пространство $T_p M$ из предыдущей теоремы. Для вектора $v = (v_1, v_2)$ имеем $df\langle v \rangle = v_1 + v_2$, поэтому $T_p M = \{(v_1, v_2) \in \mathbb{R}^2 : v_1 = -v_2\}$. Матрица Гессе функции L имеет вид

$$H = \begin{pmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial y} \\ \frac{\partial^2 L}{\partial x \partial y} & \frac{\partial^2 L}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (7.38)$$

соответственно, ее действие на вектор $v = (v_1, v_2) \in T_p M$ вычисляется по формуле

$$H(v) = (v_1, v_2)H \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 2v_1v_2 = -2v_1^2 < 0$$

при $v \neq 0$, значит, найденная точка $x_* = y_* = \frac{1}{2}$ является точкой локального максимума. Нетрудно проверить, что этот максимум будет и глобальным.

II. Найдем точки условного экстремума функции $h(x, y) = \frac{x}{a} + \frac{y}{b}$ при условии $x^2 + y^2 = 1$. Функция Лагранжа имеет вид $L = h - \lambda f = \frac{x}{a} + \frac{y}{b} - \lambda(x^2 + y^2)$. Запишем необходимое условие локального экстремума, в соответствии с Теоремой 7.5.1.

$$f(x, y) = \begin{cases} \frac{\partial L}{\partial x} = \frac{1}{a} - 2\lambda x = 0; \\ \frac{\partial L}{\partial y} = \frac{1}{b} - 2\lambda y = 0; \\ f(x, y) = x^2 + y^2 = 1. \end{cases} \quad (7.39)$$

Из первых двух равенств получаем дополнительное соотношение $\frac{1}{xa} = \frac{1}{yb}$, откуда $x = \frac{b}{a}$, и по третьему уравнению получаем $x^2 + y^2 = y^2(1 + \frac{b^2}{a^2}) = y^2 \frac{a^2 + b^2}{a^2} = 1$. Обозначим $c = \sqrt{a^2 + b^2}$. У приведенной системы уравнений два решения: $y = \pm \frac{a}{c}$, $x = \pm \frac{b}{c}$, соответственно $\lambda = \pm \frac{c}{2ab}$. Легко видеть также, что знак гессиана у функции L совпадает со знаком $-\lambda$, поэтому при $\lambda > 0$ будет точка максимума, а при $\lambda < 0$ — соответственно точка минимума.

III. Найдем точки условного экстремума функции $h(x, y, z) = x^2 + y^2 + z^2$ на эллипсоиде $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, $a > b > c > 0$. Функция Лагранжа имеет вид $L = h - \lambda f = x^2 + y^2 + z^2 - \lambda(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2})$. Запишем необходимое условие локального экстремума, в соответствии с Теоремой 7.5.1.

$$f(x, y) = \begin{cases} \frac{\partial L}{\partial x} = 2x - 2\lambda \frac{x}{a^2} = 0; \\ \frac{\partial L}{\partial y} = 2y - 2\lambda \frac{y}{b^2} = 0; \\ 2z - 2\lambda \frac{z}{c^2} = 0. \end{cases} \quad (7.40)$$

У этой системы три основных решения

III(a). $\lambda = a^2$, $y = z = 0$, $x = a$. Тогда функция Лагранжа имеет вид $L = (1 - \frac{a^2}{b^2})y^2 + (1 - \frac{a^2}{c^2})z^2$. Поскольку в критической точке для вектора $v = (v_1, v_2, v_3)$ имеем $df\langle v \rangle = \frac{2v_1}{a}$, то легко видеть, что соответствующее пространство $T_p M$ определяется соотношением $T_p M = \{(v_1, v_2, v_3) \in \mathbb{R}^3 : v_1 = 0\}$. Очевидно, на этом пространстве гессиан функции L (который совпадает с точностью до множителя с самой этой функцией) отрицателен. Поэтому указанная точка есть точка максимума для L .

III(b). $\lambda = b^2$, $x = z = 0$, $y = b$. Тогда функция Лагранжа имеет вид $L = (1 - \frac{b^2}{a^2})x^2 + (1 - \frac{b^2}{c^2})z^2$. Поскольку в критической точке для вектора $v = (v_1, v_2, v_3)$ имеем $df\langle v \rangle = \frac{2v_2}{b}$, то легко видеть, что соответствующее пространство $T_p M$ определяется соотношением $T_p M =$

$\{(v_1, v_2, v_3) \in \mathbb{R}^3 : v_2 = 0\}$. Поскольку L представляет собой квадратичную форму, для ее гессиана $H(v)$ справедливо равенство $H(v) = 2L(v)$. На пространстве $T_p M$ квадратичная форма L принимает как положительные, так и отрицательные значения (т.к. один из ее коэффициентов $(1 - \frac{b^2}{a^2})$ положителен, а другой $(1 - \frac{b^2}{c^2})$ отрицателен). Поэтому данная точка **НЕ** является точкой локального экстремума.

III(c). $\lambda = c^2$, $x = y = 0$, $z = c$. Тогда функция Лагранжа имеет вид $L = (1 - \frac{c^2}{a^2})x^2 + (1 - \frac{c^2}{b^2})y^2$. Поскольку в критической точке для вектора $v = (v_1, v_2, v_3)$ имеем $df_{<v>} = \frac{2v_3}{c}$, то легко видеть, что соответствующее пространство $T_p M$ определяется соотношением $T_p M = \{(v_1, v_2, v_3) \in \mathbb{R}^3 : v_3 = 0\}$. Очевидно, на этом пространстве гессиан функции L (который совпадает с точностью до множителя с самой этой функцией) положителен. Поэтому указанная точка есть точка минимума для L .

Остальные критические точки отличаются от выписанных только знаками, а поскольку исследуемая функция четна по всем переменным, то тип критической точки не меняется при перемене знака одной из координат.

Литература

- [1] ФИХТЕНГОЛЬЦ Г. М., Дифференциальное и интегральное исчисление, тт. I и II.
- [2] РЕШЕТНЯК Ю.Г., Курс математического анализа. Новосибирск. Изд-во Ин-та математики, 2000. Ч. 1, кн. 1-2.
- [3] ШВЕДОВ И. А., Компактный курс математического анализа. Ч. 1 и 2. Новосибирский гос. ун-т. Новосибирск, 2003.
- [4] ДЕМИДОВИЧ Б. П., Сборник задач по математическому анализу. М.: Наука, 1990.