МАТЕМАТИКА В ЭКОНОМИКЕ

Часть 2
МАТЕМАТИЧЕСКИЙ АНАЛИЗ

Издание третье, переработанное и дополненное

Рекомендовано
Министерством образования Российской Федерации
в качестве учебника
для студентов экономических специальностей
высших учебных заведений

МОСКВА
"ФИНАНСЫ И СТАТИСТИКА"
2013

УДК 330.4(075.8)
ББК 65в6я73
М34

АВТОРЫ:

**А.С. Солодовников, В.А. Бабайцев,
А.В. Браилов, И.Г. Шандра**

РЕЦЕНЗЕНТЫ:

Е.Г. Гольштейн,
доктор физико-математических наук, профессор,
заведующий лабораторией ЦЭМИ РАН;

Э.М. Карташов,
доктор физико-математических наук, профессор,
заведующий кафедрой высшей и прикладной математики
Московской государственной академии тонкой
химической технологии (МИТХТ) им. М.В. Ломоносова

М34 **Математика** в экономике: учебник. Ч. 2. Математический анализ / А.С. Солодовников, В.А. Бабайцев, А.В. Браилов, И.Г. Шандра. – 3-е изд., перераб. и доп. – М.: Финансы и статистика, 2013. – 560 с.: ил.

ISBN 978-5-279-03489-5

В части 1 изложены вопросы линейной алгебры и линейного программирования, часть 2 посвящена математическому анализу функций одной и нескольких переменных, выпуклому анализу, рядам и дифференциальным уравнениям. В отличие от второго издания (2003 г.) в новом выпуске учебника уточнены некоторые формулировки.

Для преподавателей и студентов экономических вузов, бизнес-школ, а также для всех, кто интересуется математическими приложениями в экономике.

УДК 330.4(075.8)
ББК 65в6я73

ISBN 978-5-279-03489-5

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Вторая часть учебника «Математика в экономике» посвящена различным вопросам, так или иначе связанным с математическим анализом. Математический анализ есть в первую очередь учение о скоростях протекания процессов, и это определяет огромную сферу его приложений, в том числе и в экономике. Динамика финансовых потоков и цен, анализ производственных функций и функций полезности, спроса и предложения – вот лишь некоторые из этих приложений.

Авторы стремились использовать максимально доступные, а там, где это оказалось возможным, – и оригинальные схемы изложения. Это относится, например, к теории кратного интеграла, к элементам выпуклого анализа и выпуклого программирования, к анализу уравнений Слуцкого, различным моделям экономической динамики.

Как и в первой части, особое внимание уделено примерам с экономическим содержанием. Последние присутствуют практически в каждой главе.

Второе издание отличается от первого значительно большей широтой изложения. Включена новая глава о кратных интегралах, а также существенно расширены главы, посвященные дифференциальному исчислению функций нескольких переменных и дифференциальным уравнениям.

Из других значимых добавлений отметим: теорему Куна – Таккера в форме, допускающей как нелинейные ограничения в виде неравенств, так и линейные ограничения в виде равенств; применение квадратичных форм в формулировке необходимых и достаточных условий локального экстремума; формулу Тейлора для функций нескольких переменных; вывод уравнений Слуцкого в общем случае; уравнения в полных дифференциалах и интегрирующий множитель; ряд новых экономических примеров.

В третьем издании уточнены некоторые формулировки и исправлены замеченные во втором издании опечатки.

ГЛАВА 1

ВВЕДЕНИЕ В АНАЛИЗ

§ 1.1. Понятие функции. Числовые функции и графики. Обратная и сложная функции

Понятие функции – одно из наиболее важных в математике и ее приложениях. В самом общем понимании функция – это *зависимость* между двумя переменными. Уточнением этой идеи является следующее.

Определение. *Пусть имеются два множества X и Y. Пусть, далее, указано правило, по которому каждому элементу $x \in X$ сопоставляется некоторый (единственный) элемент $y \in Y$. Тогда говорят, что задано* ***отображение*** *или, по-другому,* ***функция*** *из X в Y* (рис. 1.1). *При этом множество X называется* ***областью определения*** *функции.*

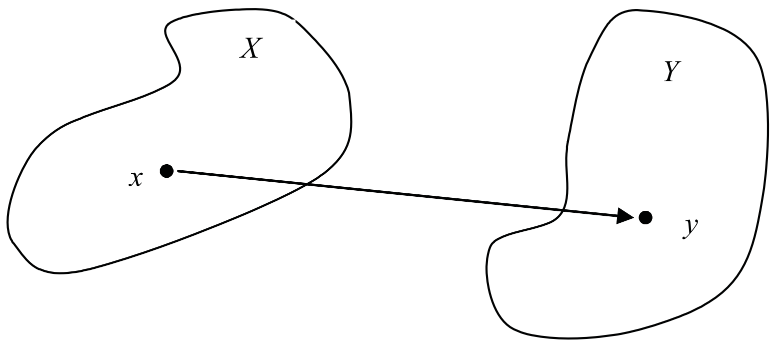

Рис. 1.1

Функции (как, впрочем, и другие объекты) в математике обозначают буквами. Чаще всего для этого используются буквы f, g, h (латинские) или φ, ψ (греческие). Обозначим данную нам функцию, например, буквой f. Обычно пишут так:

$$f: X \to Y \quad (\text{или } X \xrightarrow{f} Y),$$

что означает: f есть отображение множества X в множество Y. Для соответствующих элементов x и y используют запись

$$y = f(x)$$

("элемент y есть функция f от элемента x").

В курсе математического анализа изучают главным образом *числовые функции*. Числовая функция характеризуется тем, что оба множества X и Y состоят из чисел, т.е. $X \subset \mathbb{R}$, $Y \subset \mathbb{R}$. При этом переменный элемент x из X называется *аргументом* (или значением аргумента), а соответствующий элемент y – *функцией* (значением функции).

Наглядное представление о числовой функции дает ее *график*. Это некоторое множество точек на координатной плоскости, обычно некоторая линия.

Определение. ***Графиком функции*** *f с областью определения X называется множество*

$$\Gamma = \{(x, f(x)) \mid x \in X\}.$$

Отличительной чертой каждого графика является то, что любая прямая, параллельная оси ординат, т.е. всякая прямая $x = \text{const}$ пересекает его либо в единственной точке (если $x \in X$), либо вовсе не пересекает (если $x \notin X$), как это показано на рис. 1.2.

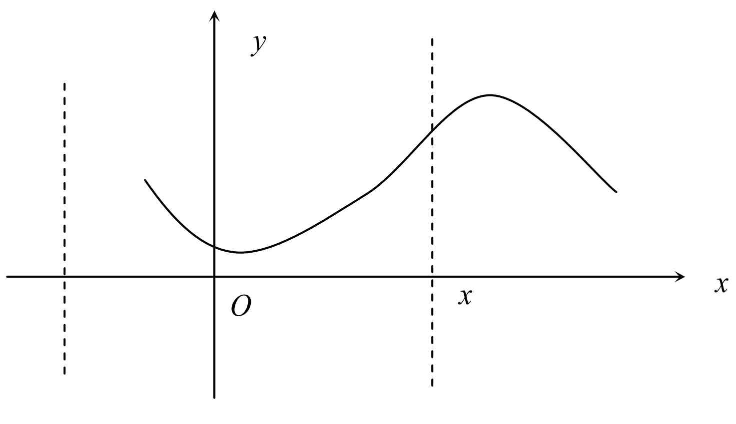

Рис. 1.2

Остановимся кратко на *способах задания* функции.

В принципе задать функцию означает: указать множество X (область определения функции) и описать правило, позволяющее по данному значению $x \in X$ аргумента находить соответствующее значение функции. Наиболее употребительными являются три способа задания функции.

1. *Табличный*. Используется, когда область определения состоит из конечного множества чисел: $X = \{x_1, x_2, \ldots, x_n\}$. Тогда для задания функции проще всего указать таблицу:

x	x_1	x_2	…	x_n
y	y_1	y_2	…	y_n

В качестве примера приведем таблицу, содержащую данные о числе жителей, населяющих земной шар, в отдельные годы:

Год	1800	1930	1960	1975	1987	2000
Млрд	1	2	3	4	5	6

Другие примеры: расписание движения поездов, таблица выигрышей в лотерее и т.п.

2. *Аналитический*, т.е. задание с помощью формулы, указывающей, какие действия нужно выполнить над x, чтобы получить y. Например,

$$y = -x^2 + 1, \quad y = [x] \text{ (целая часть } x\text{)}.$$

Замечание. Если функция задана формулой, а область ее определения не оговорена, то обычно подразумевается, что областью определения является множество всех значений x, для которых написанная формула имеет смысл (в школьной практике это множество называется ОДЗ – область допустимых значений для x). Например, в случае функции

$$y = \sqrt{1 - x^2}$$

областью определения является множество всех значений x, удовлетворяющих условию $1 - x^2 \geq 0$, т.е. промежуток $[-1; 1]$.

При аналитическом способе не исключается и такое положение, когда функция задается не одной формулой, а с помощью нескольких. Примером может служить функция

$$y = \begin{cases} +1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0, \end{cases}$$

обозначаемая $\operatorname{sgn}(x)$ (от латинского signum – знак).

3. *Графический*. На практике этот способ используется, например, в физических измерениях, когда зависимость y от x вычерчивается прибором (самописцем). Примером может служить снятие ЭКГ с больного.

Помимо "арифметических" действий над функциями, например, сложения, умножения и т.д., существует еще несколько операций, позволяющих по данным функциям строить новые. Наиболее важными из них являются две операции.

1. *Построение обратной функции*. Пусть $y = f(x)$ – функция с областью определения X. Обозначим множество всех значений функции (т.е. множество всех чисел $f(x)$, где $x \in X$) через Y. Пишут обычно: $D(f) = X$, $E(f) = Y$. Предположим дополнительно, что *разным* значениям x отвечают *разные* значения y:

$$x_1 \neq x_2 \Rightarrow y_1 \neq y_2$$

(ясно, что не всякая функция удовлетворяет такому условию: примером может служить функция $y = x^2,\ x \in \mathbb{R}$). Тогда для каждого значения $y \in Y$ существует только одно $x \in X$, такое, что $f(x) = y$. Если сопоставим каждому $y \in Y$ именно такое x, то получим отображение множества Y в множество X. Это отображение называется *обратным* к данному отображению f и обозначается f^{-1}. Таким образом,

$$f\colon X \to Y,$$

$$f^{-1}\colon Y \to X,$$

причем связь между f и f^{-1} устанавливается соглашением:

если $f(x) = y$, *то* $f^{-1}(y) = x$.

Итак, обратная функция для $y = f(x)$ есть $x = f^{-1}(y)$. Обычно аргумент в обратной функции обозначают снова буквой x, а значение функции – буквой y. Тогда обратная функция запишется: $y = f^{-1}(x)$.

Пример 1.1. Пусть $y = x^2$. Если за область определения $D(y)$ принять множество всех действительных чисел, то $E(y) = [0; \infty)$, и обратной функции не существует, т.к. $x = \pm\sqrt{y}$ (два значения, а не одно). Если же принять $D(y) = [0; \infty)$, то будем иметь обратную функцию $x = \sqrt{y}$, определенную на множестве $[0; \infty)$ и принимающую значения в том же множестве. Меняя обозначения, получим запись обратной функции в виде $y = \sqrt{x}$. На рис. 1.3 приведены графики исходной функции $y = x^2$, $X = [0; \infty)$, и обратной функции (с измененными обозначениями) $y = \sqrt{x}$, $x \in [0; \infty)$. Эти графики взаимно симметричны относительно "биссектрисы" I и III координатных четвертей.

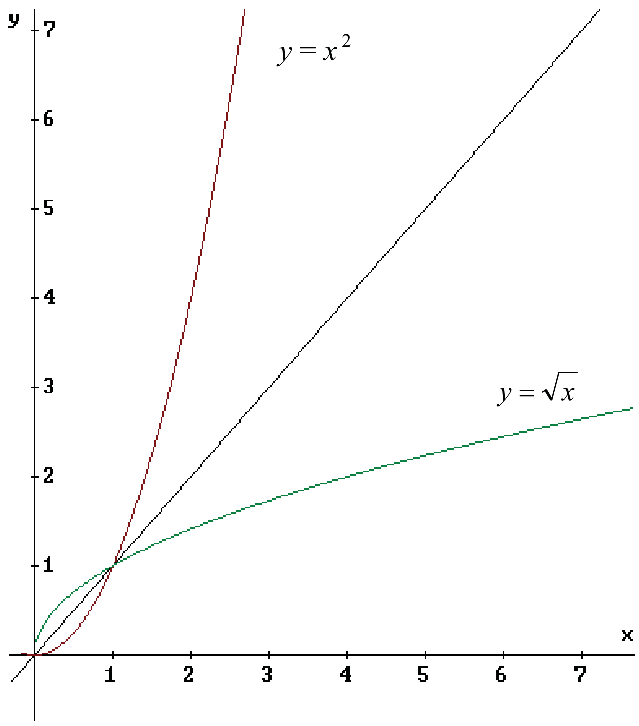

Рис. 1.3

Вообще графики исходной функции $y = f(x)$ и обратной $y = f^{-1}(x)$ всегда взаимно симметричны относительно указанной биссектрисы (почему?).

Для иллюстрации приведем еще один пример: график функции $y = 2^x$ ($D(x) = (-\infty; \infty)$, $E(y) = (0; \infty)$) и обратной функции $y = \log_2 x$ ($D(x) = (0; \infty)$, $E(y) = (-\infty; \infty)$) (рис.1.4).

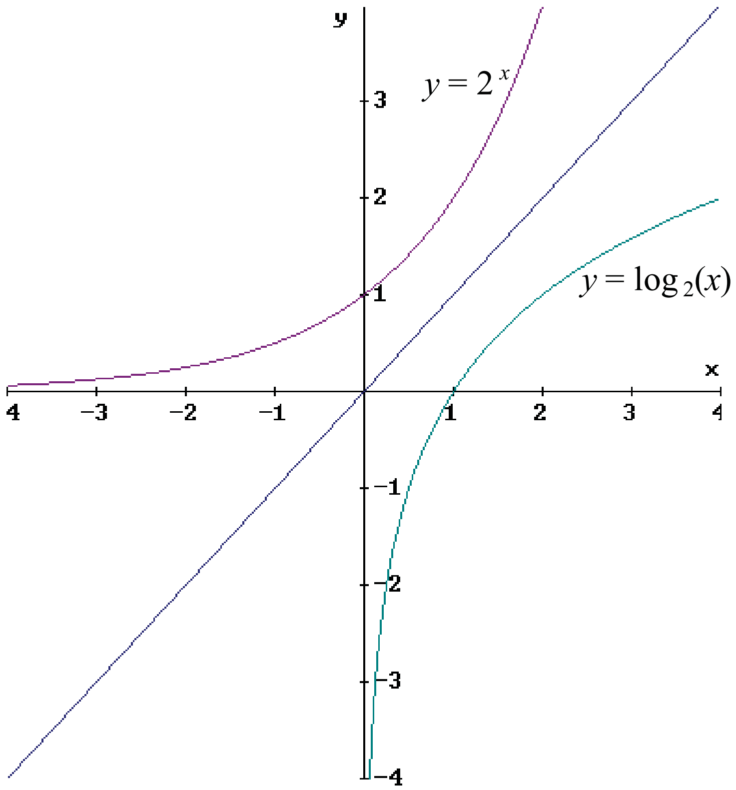

Рис.1.4

2. *Построение сложной функции*. Пусть даны две функции: $x = h(t)$ с областью определения T и множеством значений X и $y = g(x)$ с областью определения X и множеством значений Y.

Тогда "цепное" правило

$$t \xrightarrow{h} x \xrightarrow{g} y$$

определяет новую функцию с областью определения T. Эта новая функция обозначается

$$y = g(h(t))$$

и называется *сложной функцией*, или *функцией от функции*.

Например, две функции $y = \lg x$ и $x = t + 2$ определяют сложную функцию $y = \lg(t + 2)$ с областью определения $t + 2 > 0$.

Элементарные функции

Наиболее простые приложения математического анализа ограничиваются кругом так называемых *элементарных* функций.

Перечислим сначала *основные* элементарные функции.

♦ *Степенные функции* $y = x^a$,

где a – любое постоянное число. Областью определения считается промежуток $x > 0$, хотя в некоторых случаях, например, если a – натуральное число, функция определена для всех x.

♦ *Показательные функции* $y = a^x$,

где $a > 0$, $a \neq 1$. Область определения – множество всех действительных чисел.

♦ *Логарифмические функции* $y = \log_a x$,

где $a > 0$, $a \neq 1$. Область определения: $x > 0$.

♦ *Тригонометрические функции*: $y = \sin x$, $y = \cos x$, $y = \operatorname{tg} x$.

Область определения для $\sin x$ и $\cos x$ – множество действительных чисел, $x \neq \pi/2 + \pi k$ – для $\operatorname{tg} x$.

♦ *Обратные тригонометрические функции*: $y = \arcsin x$, $y = \arccos x$, $y = \operatorname{arctg} x$.

Область определения: $x \in [-1; 1]$ для $\arcsin x$ и $\arccos x$, множество действительных чисел для $\operatorname{arctg} x$.

Укажем теперь действия над функциями, которые мы будем называть *допустимыми*:

• все арифметические действия ($f + g$, $f - g$, $f \cdot g$, f/g);

• построение сложной функции.

Определение. ***Элементарными*** *называются такие функции, которые получаются из основных с помощью допустимых действий.*

Например, $y = \sqrt{1 - x^2} = \left(1 - x^2\right)^{1/2}$, $y = 1 + \sqrt{\lg \sin 2\pi x}$ и т.п. Во всех указанных примерах областью определения функции можно считать ОДЗ (или часть ОДЗ).

§ 1.2. Предел числовой последовательности

1. Определения. Примеры

Определение. ***Числовой последовательностью*** *называется числовая функция, определенная на множестве натуральных чисел.*

Задать числовую последовательность – значит задать отображение $n \to x(n)$, где $n \in \mathbb{N}$, т.е. задать бесконечный ряд чисел

$$x_1 = x(1), x_2 = x(2), x_3 = x(3), \ldots .$$

Числовую последовательность с членами x_1, x_2, … записывают коротко, $\{x_n\}$, $n \in \mathbb{N}$, или просто $\{x_n\}$.

В основе всего здания математического анализа лежит понятие *предела* числовой последовательности.

Определение. *Число a называется **пределом последовательности** $\{x_n\}$, если для любого положительного числа ε существует такой номер n_0 (зависящий, вообще говоря, от выбранного ε), начиная с которого все члены последовательности отличаются от a по модулю меньше, чем на ε:*

$$(\forall \varepsilon > 0)(\exists n_0) : n \geq n_0 \Rightarrow |x_n - a| < \varepsilon.$$

Если последовательность $\{x_n\}$ имеет предел a, то она называется *сходящейся* (к числу a) и мы пишем

$$\lim_{n \to \infty} x_n = a. \tag{1.1}$$

Заметим, что иногда вместо (1.1) пишут просто

$$x_n \to a$$

(словами: x_n стремится к a при $n \to \infty$).

Таким образом, равенство (1.1) означает, что *как бы ни было мало положительное число ε*, все члены последовательности, начиная с некоторого номера n_0, удовлетворяют неравенству $|x_n - a| < \varepsilon$. Проще говоря, числа x_n с ростом *n неограниченно приближаются* к a.

Как известно, множество всех чисел x, удовлетворяющих условию $|x - a| < \varepsilon$ (где $\varepsilon > 0$), называется *ε-окрестностью* точки a. Это множество – интервал $(a - \varepsilon, a + \varepsilon)$ (рис. 1.5). Согласно определению предела, неравенство (1.1) означает, что какова бы ни была окрестность точки a, найдется такой номер n_0, начиная с которого все члены последовательности будут содержаться в этой окрестности. Следовательно, вне ε-окрестности может находиться лишь конечное число членов последовательности.

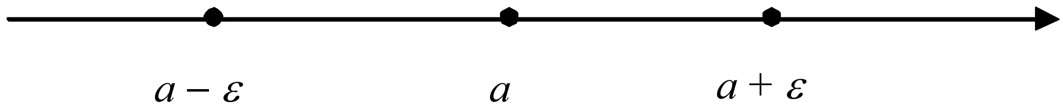

Рис. 1.5

Пример 1.2. Рассмотрим последовательность $1, \frac{1}{2}, \frac{1}{3}, \ldots,$ для которой $x_n = \frac{1}{n}$ $(n = 1, 2, \ldots)$. Числа x_n неограниченно приближаются к 0, поэтому естественно ожидать, что $\lim\limits_{n\to\infty} \frac{1}{n} = 0.$ Проверим, что такое предположение согласуется с формальным определением предела. Возьмем любое $\varepsilon > 0$ (например, $\varepsilon = 0{,}001$). Неравенство $|x_n - 0| < \varepsilon$, или, что то же, $\frac{1}{n} < \varepsilon$, выполняется для всех $n > \frac{1}{\varepsilon}$. Поэтому, приняв за n_0 ближайшее к $\frac{1}{\varepsilon}$ справа целое число (в нашем примере $n_0 = 1001$), для всех $n \geq n_0$ имеем неравенство $|x_n - 0| < \varepsilon$, следовательно, $\lim\limits_{n\to\infty} x_n = 0.$

Пример 1.3. Пусть $x_n = \frac{n-1}{n+1}$ $(n \in \mathbb{N})$, покажем, что $\lim\limits_{n\to\infty} x_n = 1.$ Взяв произвольное $\varepsilon > 0$, составим неравенство $|x_n - 1| < \varepsilon$ и выясним, для каких n оно справедливо. Имеем:

$$\left|\frac{n-1}{n+1} - 1\right| < \varepsilon \Leftrightarrow \frac{2}{n+1} < \varepsilon \Leftrightarrow n + 1 > \frac{2}{\varepsilon} \Leftrightarrow n > \frac{2}{\varepsilon} - 1.$$

Таким образом, приняв за n_0 ближайшее к $\dfrac{2}{\varepsilon}-1$ справа положительное целое число, будем иметь $|x_n - 1| < \varepsilon$ для всех $n \geq n_0$, что означает справедливость равенства $\lim\limits_{n\to\infty} x_n = 1$.

2. Основные свойства пределов последовательностей

1. *Сходящаяся последовательность имеет только один предел.*

Действительно, предположим, рассуждая от противного, что некоторая последовательность $\{x_n\}$ имеет два *различных* предела a и b, $a \neq b$.

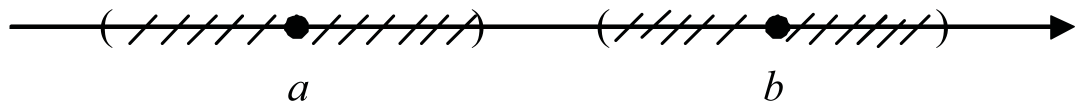

Рис. 1.6

Выберем столь малые окрестности точек a и b (рис. 1.6), чтобы они не имели общих точек. Поскольку $\lim x_n = a$, все x_n, начиная с некоторого номера n_1, содержатся в выбранной окрестности точки a; точно так же из $\lim x_n = b$ следует, что все x_n, начиная с некоторого номера n_2, содержатся в выбранной окрестности точки b. Положим $n_0 = \max\{n_1, n_2\}$. Тогда числа x_n с номерами $n \geq n_0$ должны принадлежать как к первой, так и ко второй окрестности, что невозможно, так как окрестности не имеют общих точек.

2. *Сходящаяся последовательность ограничена.*

Напомним, что множество чисел X называется *ограниченным*, если существует такой отрезок $[a, b]$ числовой оси, который содержит все числа из X. Или (что эквивалентно): множество X ограничено, если существует такое $A > 0$, что для всех чисел x из X имеем $|x| < A$.

Доказательство свойства 2 пределов не составляет труда. Пусть $\lim x_n = a$. Положим $\varepsilon = 1$ и найдем номер n_0, начиная с которого $|x_n - a| < 1$, т.е. $-1 < x_n - a < 1$ для $n \geq n_0$.

Отсюда следует $a - 1 < x_n < a + 1$ для всех $n \geq n_0$. Заменим отрезок $[a - 1, a + 1]$ таким (быть может, бóльшим) отрезком $[A, B]$, чтобы в него попали не только числа x_n, $n \geq n_0$, но и все числа x_1, x_2, …, x_{n_0}. Тогда будем иметь $x_n \in [A, B]$ для *всех* $n \in \mathbb{N}$, что означает ограниченность множества $\{x_n\}$.

3. *Если члены сходящейся последовательности* $\{x_n\}$ *удовлетворяют неравенству* $x_n \geq b$, *то и* $\lim x_n \geq b$.

Доказательство. Пусть $\lim x_n = a$. Предположим, рассуждая от противного, что $a < b$. Выберем окрестность точки a так (рис. 1.7), чтобы все числа из этой окрестности были строго меньше числа b. В этой окрестности, по определению предела, должны содержаться все числа x_n, начиная с некоторого номера n_0. Но это невозможно, ибо все эти числа по условию не меньше b. Получили противоречие.

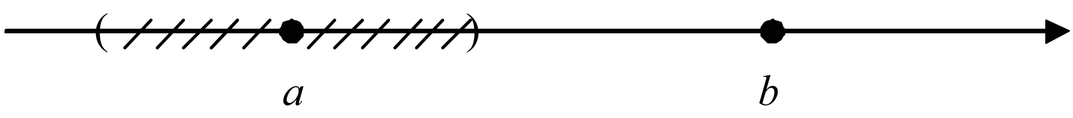

Рис. 1.7

Замечание. Пусть все x_n удовлетворяют *строгому* неравенству $x_n > b$. Напрашивается вывод, что и $\lim x_n > b$. Это не всегда так; можно лишь утверждать, что $\lim x_n \geq b$.

Пример: для последовательности $x_n = \dfrac{1}{n}$ $(n \in \mathbb{N})$, все члены которой строго больше 0, предел равен 0.

4. *Если члены двух сходящихся последовательностей* $\{x_n\}$ *и* $\{y_n\}$ *связаны неравенствами* $x_n \geq y_n$ $(n = 1, 2,\ldots)$, *то и* $\lim x_n \geq \lim y_n$.

Доказательство проведите самостоятельно.

5. *Если члены трех последовательностей* $\{x_n\}$, $\{y_n\}$, $\{z_n\}$ *связаны неравенствами* $x_n \geq y_n \geq z_n$ $(n = 1, 2, \ldots)$ *и при этом последовательности* $\{x_n\}$ *и* $\{z_n\}$ *имеют один и тот же предел* a, *то и последовательность* $\{y_n\}$ *имеет предел* a.

Это свойство пределов иногда называют "леммой о двух милиционерах" (сообразите, с каким сюжетом связано такое название).

Доказательство. Выберем любую окрестность точки a. Существуют номер n_1, начиная с которого все числа x_n принадлежат этой окрестности, и номер n_2, начиная с которого все z_n принадлежат той же окрестности. Ввиду того что $x_n \geq y_n \geq z_n$, все числа y_n, начиная с номера $n_0 = \max\{n_1, n_2\}$, также принадлежат выбранной окрестности. Это и доказывает, что $\lim y_n = a$.

3. Общие правила нахождения пределов

Кроме рассмотренных выше свойств 1 … 5 пределов, имеется несколько правил, постоянно применяемых при нахождении пределов. Перечислим эти правила.

Пусть $\lim x_n = a$ и $\lim y_n = b$. Тогда:

1. $\lim(x_n + y_n) = a + b$

(словами: "предел суммы равен сумме пределов");

2. $\lim(x_n \cdot y_n) = a \cdot b$

(словами: "предел произведения равен произведению пределов");

3. $\lim \dfrac{1}{y_n} = \dfrac{1}{b}$,

если все числа y_n, а также и сам предел b отличны от нуля;

4. $\lim \dfrac{x_n}{y_n} = \dfrac{a}{b}$,

если все y_n, а также b не равны нулю.

Ограничимся доказательством только правила 3, предоставив случаи 1, 2, 4 читателю. Пусть для определенности $b > 0$. Начиная с некоторого номера n_1, имеем $y_n > \dfrac{b}{2}$ (достаточно рассмотреть окрестность $\left(\dfrac{b}{2}, \dfrac{3b}{2}\right)$ точки b), следовательно, $by_n > \dfrac{b^2}{2}$. Оценим теперь разность между $\dfrac{1}{y_n}$ и предполагаемым пределом $\dfrac{1}{b}$: при $n \geq n_1$ имеем

$$\left|\frac{1}{y_n} - \frac{1}{b}\right| = \frac{|y_n - b|}{by_n} < \frac{2}{b^2}|y_n - b|.$$

Пусть теперь $\varepsilon > 0$ – произвольное положительное число. Положим $\varepsilon' = \dfrac{b^2}{2}\varepsilon$ и, пользуясь определением предела, найдем такой номер n_2, начиная с которого $|y_n - b| < \varepsilon'$. Тогда для любого номера n, большего, чем n_1 и n_2, будет

$$\left|\frac{1}{y_n} - \frac{1}{b}\right| < \frac{\varepsilon'}{b^2/2} = \varepsilon.$$

Отсюда следует, что $\lim \dfrac{1}{y_n} = \dfrac{1}{b}$.

4. Монотонные последовательности и их пределы

Определение. *Последовательность* $\{x_n\}$ *называется*:

возрастающей, *если* $x_n < x_{n+1}$ *для всех* n,

неубывающей, *если* $x_n \le x_{n+1}$ *для всех* n,

убывающей, *если* $x_n > x_{n+1}$ *для всех* n,

невозрастающей, *если* $x_n \ge x_{n+1}$ *для всех* n.

Все такие последовательности называются *монотонными*. Справедлива следующая фундаментальная

Теорема 1.1. *Любая монотонная и ограниченная последовательность имеет предел.*

Разумеется, в случае возрастающей (или неубывающей) последовательности ограниченность означает фактически ограниченность сверху, т.е. существование такого числа A, что все $x_n < A$; действительно, ограниченность такой последовательности снизу очевидна (все $x_n \ge x_1$). Точно так же в случае убывающей (невозрастающей) последовательности ограниченность означает ограниченность снизу.

Строгое доказательство теоремы требует уточнения самого понятия действительного числа, заниматься этим мы здесь не будем. Геометрический же смысл теоремы весьма прост: если, например, последовательность $\{x_n\}$ возрастает (рис. 1.8) и при этом ограничена сверху, то это означает, что с ростом n точки x_n на числовой оси смещаются вправо, но при этом *не переходят* через некоторый рубеж A. Геометрически ясно, что в этом случае числа x_n должны накапливаться к некоторому числу a, которое и будет, таким образом, пределом последовательности $\{x_n\}$.

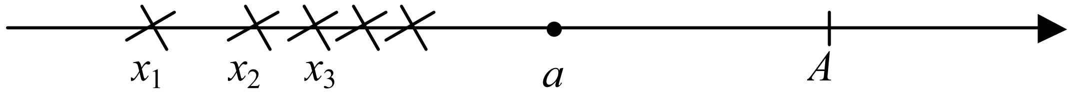

Рис. 1.8

5. Бесконечные пределы

Определение. *Говорят, что последовательность* $\{x_n\}$ *имеет своим пределом* $+\infty$, *если для любого* (*сколь угодно большого*) *числа* $A > 0$ *существует такой номер* n_0, *начиная с которого все числа* x_n *больше* A:

$$(\forall A > 0)(\exists n_0)\, n \geq n_0 \Rightarrow x_n > A.$$

В этом случае пишут $\lim\limits_{n\to\infty} x_n = +\infty.$

Например, каждая из последовательностей

$$\left\{\sqrt{n}\right\},\ \{n\},\ \left\{n^2\right\},\ \left\{\frac{n^2}{n+1}\right\}$$

имеет пределом $+\infty$ (докажите).

Аналогично определяется предел $-\infty$: запись $\lim\limits_{n\to\infty} x_n = -\infty$ означает, что

$$(\forall A > 0)(\exists n_0)\, n \geq n_0 \Rightarrow x_n < -A.$$

Очевидно, если $\lim x_n = +\infty$ или $\lim x_n = -\infty$, то $\lim x_n^{-1} = 0$ (предполагается, что все $x_n \neq 0$).

Пределы $+\infty$ и $-\infty$ называются *бесконечными* (или *несобственными*) пределами.

§ 1.3. Число *e*

Рассмотрим последовательность, состоящую из чисел

$$x_n = \left(1 + \frac{1}{n}\right)^n, \; n = 1, 2, \ldots . \tag{1.2}$$

Следующая таблица дает представление о некоторых из чисел x_n:

n	1	2	3	4	5	10	100	1000	10000
x_n	2	2,25	2,37	2,44	2,49	2,59	2,70	2,717	2,718

Видно, что числа x_n растут, но этот рост постепенно замедляется. Напрашивается предположение, что последовательность $\{x_n\}$ *сходится*.

Для доказательства сходимости рассмотрим другую последовательность

$$y_n = \left(1 + \frac{1}{n}\right)^{n+1}$$

и установим, что:

а) последовательность $\{y_n\}$ убывает;

б) последовательность $\{y_n\}$ ограничена снизу числом 2, т.е. $y_n \geq 2$ для всех n.

Предварительно докажем так называемое неравенство Бернулли.

Неравенство Бернулли. *Для любого $\alpha > -1$ и натурального p справедливо неравенство*

$$(1 + \alpha)^p \geq 1 + p\alpha. \tag{1.3}$$

Доказательство. Убедимся, что если неравенство (1.3) верно для некоторого показателя степени p, то оно верно и при показателе $p + 1$. Имеем

$$(1 + \alpha)^{p+1} = (1 + \alpha)^p (1 + \alpha) \geq (1 + p\alpha)(1 + \alpha) =$$
$$= 1 + (p + 1)\,\alpha + p\alpha^2 \geq 1 + (p + 1)\alpha.$$

Далее заметим, что неравенство (1.3) безусловно верно для $p = 1$; в силу доказанного выше оно будет тогда верно и для $p = 2$, затем для $p = 3$ и т.д. Значит, неравенство (1.3) верно при любом натуральном p.

Теперь легко получаем неравенство $y_n \geq 2$:

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

Осталось доказать, что числа y_n убывают. Оценим для этого отношение $\dfrac{y_n}{y_{n-1}}$:

$$\frac{y_n}{y_{n-1}} = \frac{\left(1+\dfrac{1}{n}\right)^{n+1}}{\left(1+\dfrac{1}{n-1}\right)^{n}} = \frac{\left(1+\dfrac{1}{n}\right)^{n+1}}{\left(\dfrac{n}{n-1}\right)^{n}} = \left(1+\frac{1}{n}\right)^{n+1}\left(\frac{n-1}{n}\right)^{n} = \left(1+\frac{1}{n}\right)\left(1+\frac{1}{n}\right)^{n}\left(1-\frac{1}{n}\right)^{n} =$$

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

и, поскольку из неравенства Бернулли следует, что $1 + n \cdot \dfrac{1}{n^2} \leq \left(1 + \dfrac{1}{n^2}\right)^{n}$, имеем

$$\frac{y_n}{y_{n-1}} \leq \left(1 + \frac{1}{n^2}\right)^{n}\left(1 - \frac{1}{n^2}\right)^{n} = \left(1 - \frac{1}{n^4}\right)^{n} < 1.$$

Отсюда вытекает $y_n < y_{n-1}$ при любом n, т.е. числа y_n убывают.

Итак, последовательность $\{y_n\}$ убывает и ограничена снизу (числом 2). По теореме 1.1 отсюда следует, что эта последовательность имеет предел $\lim y_n$.

Учитывая, что

$$x_n = \frac{y_n}{1 + \dfrac{1}{n}},$$

можем записать

$$\lim x_n = \lim \frac{1}{1+\frac{1}{n}} \cdot \lim y_n = \lim y_n,$$

т.е. предел x_n существует и совпадает с пределом y_n.

Определение. *Числом **e** называется предел*

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n.$$

Подсчитано, что начальный отрезок десятичного разложения для e имеет вид 2,718281828.

§ 1.4. Предел функции

1. Определение и примеры

Пусть функция $f(x)$ определена в некоторой окрестности точки x_0, исключая, быть может, саму точку x_0. Поскольку такая ситуация будет встречаться у нас многократно, введем для нее подходящее название. Окрестность $(x_0 - \varepsilon\,;\, x_0 + \varepsilon)$ точки x_0, из которой исключена сама точка x_0 (т.е. множество $(x_0 - \varepsilon\,;\, x_0) \cup (x_0;\, x_0 + \varepsilon)$), назовем *проколотой окрестностью* точки x_0.

Итак, пусть $f(x)$ определена в проколотой окрестности x_0. Выберем в этой окрестности какую-нибудь последовательность x_1, x_2, …, сходящуюся к точке x_0: $\lim\limits_{n\to\infty} x_n = x_0$. Значения функции в выбранных точках образуют последовательность $f(x_1)$, $f(x_2)$,…, и можно ставить вопрос о существовании предела этой последовательности.

Определение. *Число a называется **пределом функции f(x) в точке x_0*** (*или* ***пределом при $x \to x_0$***), *если для любой сходящейся к x_0 последовательности значений аргумента, отличных от x_0, со-*

ответствующая последовательность значений функции сходится к числу a, *т.е.*

$$\lim x_n = x_0\ (x_n \neq x_0) \Rightarrow \lim f(x_n) = a.$$

В этом случае пишут

$$\lim_{x \to x_0} f(x) = a. \tag{1.4}$$

Пример 1.4. Постоянная функция $f(x) = c$ имеет предел в любой точке x_0, причем этот предел равен c. Действительно, в этом случае все числа $f(x_1), f(x_2), \dots$ равны c, поэтому $\lim f(x_n) = c$.

Пример 1.5. Функция $f(x) = x$ в произвольной точке x_0 имеет предел x_0. Действительно, пусть последовательность $\{x_n\}$ сходится к x_0. Последовательность $\{f(x_n)\}$ совпадает с $\{x_n\}$, поэтому ее предел также равен x_0.

Пример 1.6. Пусть $f(x) = x^2$. Тогда $\lim\limits_{x \to x_0} f(x) = x_0^2$. Действительно, пусть $\{x_n\}$ сходится к x_0. Соответствующая последовательность значений функции будет $x_1^2, x_2^2, \dots$; ее предел равен x_0^2, так как

$$\lim x_n^2 = \lim (x_n \cdot x_n) = \lim x_n \cdot \lim x_n = x_0 \cdot x_0 = x_0^2.$$

2. Основные свойства пределов функции

Эти свойства аналогичны соответствующим свойствам пределов последовательностей.

1. *Функция* $f(x)$ *в точке* x_0 *может иметь только один предел.*

Доказательство. Пусть $\lim\limits_{x \to x_0} f(x_n) = a$ и одновременно $\lim\limits_{x \to x_0} f(x_n) = b$, где $a \neq b$. Тогда для любой последовательности $\{x_n\}$, сходящейся к x_0 (где все $x_n \neq x_0$), мы должны иметь

$$\lim_{n \to \infty} f(x_n) = a \text{ и } \lim_{n \to \infty} f(x_n) = b,$$

что невозможно, так как последовательность $\{f(x_n)\}$ может иметь только один предел.

2. *Если функция* $f(x)$ *имеет предел в точке* x_0, *то в некоторой окрестности этой точки функция ограничена, т.е. существует такая* (*проколотая*) *окрестность точки* x_0 *и такое число* $A > 0$, *что* $|f(x)| \leq A$ *для всех* x *из этой окрестности.*

Доказательство проведем от противного. Пусть в *любой* окрестности точки x_0 функция $f(x)$ не ограничена. Возьмем какую-либо окрестность $(x_0 - \varepsilon,\ x_0 + \varepsilon)$; ввиду неограниченности $f(x)$ в этой окрестности должна найтись точка x_1 $(x_1 \neq x_0)$, такая, что $|f(x_1)| > 1$. Уменьшим вдвое эту окрестность, т.е. рассмотрим окрестность $\left(x_0 - \frac{\varepsilon}{2}, x_0 + \frac{\varepsilon}{2}\right)$; в ней снова должна найтись точка x_2 $(x_2 \neq x_0)$, такая, что $|f(x_2)| > 2$. Продолжив это рассуждение, получим последовательность $x_1, x_2, \ldots$, сходящуюся, очевидно, к точке x_0. Соответствующая последовательность $f(x_1)$, $f(x_2)$, … должна сходиться к числу $a = \lim\limits_{x \to x_0} f(x)$. Но эта последовательность не является ограниченной (ибо $|f(x_n)| > n$ при любом n), поэтому сходиться не может (§ 1.2, п.2). Получили противоречие.

3. *Если для всех точек* x *некоторой* (*проколотой*) *окрестности точки* x_0 *выполняется неравенство* $f(x) \geq b$, *то и* $\lim\limits_{x \to x_0} f(x) \geq b$, *если только указанный предел существует.*

Доказательство проведите самостоятельно. Воспользуйтесь соответствующим свойством предела числовой последовательности.

4. *Если в некоторой* (*проколотой*) *окрестности точки* x_0 *имеем* $f(x) \geq g(x)$, *то и* $\lim\limits_{x \to x_0} f(x) \geq \lim\limits_{x \to x_0} g(x)$, *если только указанные пределы существуют.*

Доказательство также проведите самостоятельно.

5. *Пусть в некоторой* (*проколотой*) *окрестности точки* x_0 *выполняются неравенства*

$$f(x) \geq g(x) \geq h(x),$$

причем пределы $f(x)$ и $h(x)$ при $x \to x_0$ существуют и равны между собой:

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = a.$$

Тогда предел $g(x)$ при $x \to x_0$ также существует и $\lim\limits_{x\to x_0} g(x) = a.$

И в этом случае рекомендуем провести доказательство самостоятельно, воспользовавшись соответствующим свойством предела числовой последовательности.

3. Общие правила нахождения пределов функций

Пусть существуют пределы

$$\lim_{x\to x_0} f(x) = a \ \text{ и } \ \lim_{x\to x_0} g(x) = b.$$

Тогда:

1. $\lim\limits_{x\to x_0} \left(f(x) + g(x)\right) = a + b$;
2. $\lim\limits_{x\to x_0} \left(f(x) \cdot g(x)\right) = a \cdot b$;
3. $\lim\limits_{x\to x_0} \dfrac{1}{g(x)} = \dfrac{1}{b}$, если $g(x) \neq 0$ в окрестности точки x_0, а также $b \neq 0$;
4. $\lim\limits_{x\to x_0} \dfrac{f(x)}{g(x)} = \dfrac{a}{b}$, если $g(x) \neq 0$ в окрестности точки x_0, а также $b \neq 0$.

Каждое из этих правил вытекает из соответствующего правила для пределов числовых последовательностей (§ 1.2, п.3); предоставим читателю провести необходимые рассуждения.

Пример 1.7. Найти предел функции

$$f(x) = \frac{x^3 - 2}{x^2 + x - 5}$$

в какой-нибудь точке x_0, где знаменатель отличен от нуля.

Р е ш е н и е. На основании правил 1 и 2, а также учитывая, что предел постоянной функции равен ей самой, имеем:

$$\lim_{x\to x_0}\left(x^3-2\right)=\lim_{x\to x_0}x_0^3-\lim_{x\to x_0}2=x_0^3-2\,,$$

$$\lim_{x\to x_0}\left(x^2+x-5\right)=\lim_{x\to x_0}x^2+\lim_{x\to x_0}x-\lim_{x\to x_0}5=x_0^2+x_0-5\,,$$

после чего правило 4 дает

$$\lim_{x\to x_0}f(x)=\frac{x_0^3-2}{x_0^2+x_0-5}=f\left(x_0\right).$$

Мы видим, что в данном примере оказалось

$$\lim_{x\to x_0}f\left(x\right)=f\left(x_0\right).$$

Это важное свойство функции $f(x)$, называемое *непрерывностью* (в точке x_0), будет обсуждаться в § 1.8; пока лишь заметим, забегая вперед, что им обладают *все элементарные функции*.

4. Более общий подход к понятию предела функции

Пусть X – некоторое множество точек на числовой прямой.

Определение. *Точка x_0 называется* ***предельной*** *для множества X, если существует последовательность точек x_1, x_2,… из множества X, отличных от x_0 и сходящихся к x_0.*

При этом сама точка x_0 может как принадлежать, так и не принадлежать множеству X.

Пример 1.8. Для множества $X = (0;\ 1)$ (интервала) точки 0 и 1 являются предельными, хотя сами не принадлежат X.

Определение. *Пусть функция $y = f(x)$ определена на множестве X и пусть x_0 – предельная точка для X. Число a называется* ***пределом функции $f(x)$ в точке x_0*** *(или при $x \to x_0$), если для любой последовательности x_1, x_2,… точек из X, отличных от x_0 и сходя-*

щихся к x_0, *соответствующая последовательность значений функции* $f(x_1), f(x_2), \ldots$ *сходится к числу* a.

Если функция $f(x)$ определена справа от x_0, точнее, в интервале вида $(x_0, x_0 + h)$, где $h > 0$, то предел $f(x)$ при $x \to x_0$ (если он существует) называют *правосторонним пределом* $f(x)$ в точке x_0 и обозначают $\lim\limits_{x \to x_0,\, x > x_0} f(x)$, а также $\lim\limits_{x \to x_0 + 0} f(x)$ или даже еще короче $f(x_0 + 0)$.

Аналогично определяется левосторонний предел. Если в точке x_0 существуют как предел $f(x)$ справа, так и предел слева, причем оба пределы равны, то существует и просто предел $f(x)$ при $x \to x_0$, совпадающий с обоими односторонними пределами. И обратно, если существует $\lim\limits_{x \to x_0} f(x) = a$, то существуют и оба односторонних предела, которые совпадают с a (докажите самостоятельно).

Пример 1.9. Для функции

$$f(x) = \operatorname{sgn} x = \begin{cases} +1, & \text{если } x > 0, \\ 0, & \text{если } x = 0, \\ -1, & \text{если } x < 0, \end{cases}$$

имеем: $\lim\limits_{x \to 0+0} f(x) = 1, \quad \lim\limits_{x \to 0-0} f(x) = -1.$

5. Предел при $x \to +\infty$ или $x \to -\infty$. Бесконечный предел

Данное нами ранее определение предела функции придает смысл равенству $\lim\limits_{x \to x_0} f(x) = a.$ При этом предполагается, что x_0 и a – обычные (т.е. конечные) числа. Между тем часто приходится рассматривать случаи, когда x_0 есть $+\infty$ или $-\infty$, а также когда a есть $+\infty$ или $-\infty$. Для таких случаев мы сохраняем сформулированное выше определение предела функции. А именно, равенство $\lim\limits_{x \to x_0} f(x) = a$ во всех случаях означает следующее: для любой последовательности $\{x_n\}$, сходящейся к x_0, соответствующая последовательность $\{f(x_n)\}$ сходится к a.

В частности, например, равенство

$$\lim_{x\to+\infty} f(x) = -\infty$$

означает, что для любой последовательности $\{x_n\}$, такой, что $\lim_{n\to\infty} x_n = +\infty$, должно выполняться равенство $\lim_{n\to\infty} f(x_n) = -\infty$.

Пример 1.10. а) $\lim_{x\to 0} \frac{1}{x^2} = +\infty$, б) $\lim_{x\to-\infty} \frac{x-1}{x+1} = 1$.

§ 1.5. Два замечательных предела

Докажем равенство

$$\lim_{x\to 0} \frac{\sin x}{x} = 1. \tag{1.5}$$

Этот предел обычно называют *первым замечательным пределом*.

Доказательство. Так как функция $(\sin x)/x$ – четная, то достаточно показать, что предел при $x \to 0$ справа равен 1. На рис. 1.9 показан угол AOM с радианной мерой $x > 0$ ($AO = MO = 1$).

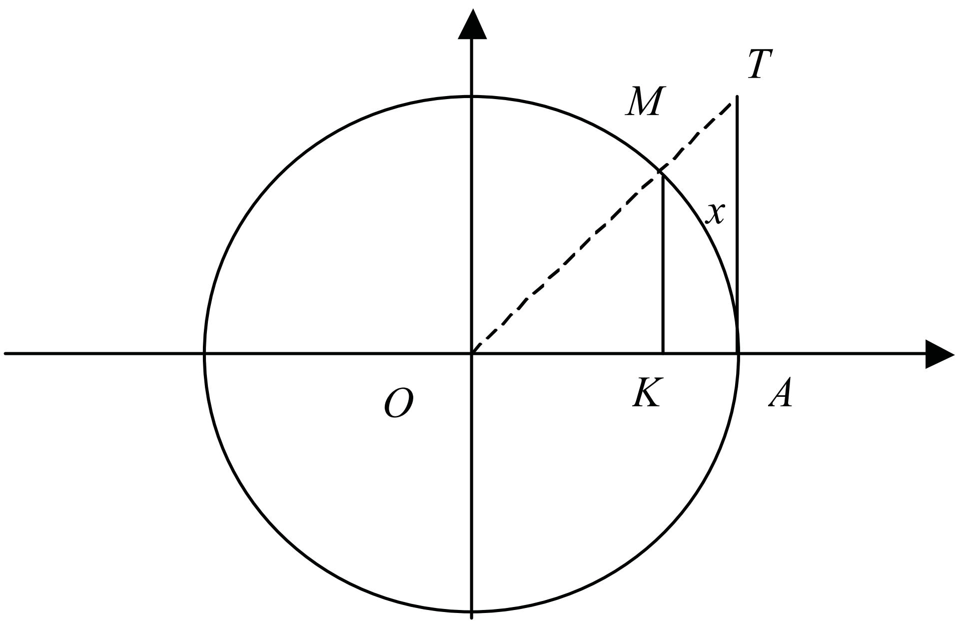

Рис. 1.9

Из сравнения площадей треугольников OAM, OAT и сектора круга следуют неравенства:

$$\frac{1}{2}OA\cdot MK < \frac{1}{2}OA\cdot A\breve{M} < \frac{1}{2}OA\cdot AT \text{ или } \sin x < x < \operatorname{tg} x.$$

Деля на $\sin x$, получим

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}, \text{ или } 1 > \frac{\sin x}{x} > \cos x.$$

При $x \to 0$ (справа), очевидно, имеем $\lim \cos x = 1$. Значит, пределы обеих функций 1 и $\cos x$ при $x \to 0$ (справа) равны 1. Отсюда по одному из свойств предела функции (§ 1.4, п. 2, свойство 5) получим требуемое равенство (1.5).

В дополнение к установленному ранее пределу

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e \tag{1.6}$$

докажем еще два равенства:

$$\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = e, \quad \lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = e. \tag{1.7}$$

Очевидно, что равенство (1.6) есть как бы частный случай первого из равенств (1.7), когда x стремится к $+\infty$, принимая только целые значения.

Д о к а з а т е л ь с т в о. Сначала докажем первое из равенств (1.7). Пусть x – любое число. Найдем такое целое число n, чтобы выполнялись неравенства

$$n \le x < n + 1; \tag{1.8}$$

число n называют обычно целой частью числа x и обозначают $[x]$. Будем считать $x > 1$, тогда, очевидно, $n > 0$.

Из (1.8) следует

$$1+\frac{1}{n}\geq 1+\frac{1}{x}>1+\frac{1}{n+1}. \qquad (1.9)$$

Поскольку $n + 1 > x \geq n$, то из (1.9) следует

$$\left(1+\frac{1}{n}\right)^{n+1}\geq\left(1+\frac{1}{x}\right)^{x}>\left(1+\frac{1}{n+1}\right)^{n},$$

или

$$f(x)\geq\left(1+\frac{1}{x}\right)^{x}>g(x),$$

где

$$f(x)=\left(1+\frac{1}{n}\right)\left(1+\frac{1}{n}\right)^{n},\;\; g(x)=\frac{1}{1+\dfrac{1}{n+1}}\left(1+\frac{1}{n+1}\right)^{n+1}.$$

При $x \to +\infty$, очевидно, n также стремится к $+\infty$, поэтому $f(x)$ и $g(x)$ будут иметь общий предел e. Согласно одному из свойств предела функции (§ 1.4, п. 2, свойство 5), предел функции $\left(1+\frac{1}{x}\right)^{x}$. при $x \to +\infty$ также будет равен e, что и требовалось получить.

Докажем второе равенство (1.7). Пусть $x = -t$, где $t > 0$. Имеем

$$\left(1+\frac{1}{x}\right)^{x}=\left(1-\frac{1}{t}\right)^{-t}=\left(\frac{t-1}{t}\right)^{-t}=\left(\frac{t}{t-1}\right)^{t}=$$

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

Теперь ясно, что при $x \to -\infty$, т.е. когда $t \to +\infty$, выражение в правой части будет иметь предел, равный

$$\lim_{t\to+\infty}\left(1+\frac{1}{t-1}\right)^{t-1}\cdot\lim_{t\to+\infty}\left(1+\frac{1}{t-1}\right)=e\cdot 1=e,$$

что завершает доказательство второго равенства (1.7).

Любую из формул (1.7) называют *вторым замечательным пределом*.

§ 1.6. Формула непрерывных процентов

Пусть к началу года в нашем распоряжении имеется сумма A_0 рублей. Как добиться к концу года максимального роста этой суммы? Один из возможных способов – воспользоваться услугами банка. Предположим, что банк дает 100% годовых; это означает, что за год хранения вклад возрастет на 100%, за любой меньший срок вклад возрастет пропорционально этому сроку (например, за один месяц хранения прирост составит $\frac{100}{12}$ %).

Итак, после года хранения вклад станет $A_0 + A_0 = 2A_0$, т.е. удвоится. Можно, однако, добиться большего эффекта, если по истечении 0,5 года закрыть счет и тут же открыть его снова на очередные полгода. В этом случае к концу первого полугодия вклад станет равным $A_0+\frac{1}{2}A_0=A_0\left(1+\frac{1}{2}\right)$, а к концу года будет равным $A_0\left(1+\frac{1}{2}\right)^2=2{,}25A_0$. Если операцию по закрытию и открытию счета проводить чаще, то получим еще больший эффект: например, если эту операцию проводить в конце каждого месяца, то к концу года будем иметь $A_0\left(1+\frac{1}{12}\right)^{12}$ = $2{,}613A_0$, а если открывать и закрывать счет каждый день, то конечная сумма составит $A_0\left(1+\frac{1}{365}\right)^{365}=$ $= 2{,}715A_0$ руб. Если представить себе (что, конечно, является абстракцией), что операция открытия-закрытия проводится непрерывно, то в итоге к концу года вклад составит

$$A_0 \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = A_0 \cdot e = 2{,}718\ldots A_0$$

рублей. Таким образом, при номинальной ставке 100% эффективная ставка может составить 171,…%, что существенно лучше.

Аналогичное рассуждение можно провести для случая, когда номинальная ставка банка будет p% (вместо 100%). Тогда (теоретически) возможная конечная сумма вклада будет

$$A_0 \lim_{n\to\infty}\left(1+\frac{p}{100n}\right)^n = A_0 \lim_{n\to\infty}\left[\left(1+\frac{p}{100n}\right)^{\frac{100n}{p}}\right]^{\frac{p}{100}}.$$

Согласно (1.7) выражение, заключенное в квадратные скобки, при $n \to \infty$ будет иметь пределом число e, так что конечная сумма будет $A_0 e^{\frac{p}{100}}$.

Рассмотрим еще более общую ситуацию, когда сумма A_0, вложенная в банк под p% годовых, хранится не один год, а любое количество t лет. Деля промежуток $[0,\ t]$ на n равных периодов и затем устремляя n к бесконечности, получим теоретически возможную конечную сумму

$$A_0 \lim_{n\to\infty}\left(1+\frac{p}{100n}t\right)^n = A_0 \lim_{n\to\infty}\left[\left(1+\frac{pt}{100n}\right)^{\frac{100n}{pt}}\right]^{\frac{pt}{100}} = A_0 e^{\frac{pt}{100}}.$$

Формулу

$$A = A_0 e^{\frac{p}{100}t} \tag{1.10}$$

для конечной суммы вклада называют *формулой непрерывных процентов*.

Например, при годовой ставке 100% можно к концу второго года получить $A_0e^2 = 7{,}414\ldots A_0$, т.е. увеличить начальный капитал более чем в семь раз.

§ 1.7. Бесконечно малые и бесконечно большие функции

Определение. *Функция $f(x)$ называется **бесконечно малой при** $x \to x_0$, если $\lim\limits_{x\to x_0} f(x) = 0$, и **бесконечно большой при** $x \to x_0$, если $\lim\limits_{x\to x_0} f(x) = \pm\infty$.*

Справедливы следующие две теоремы.

Теорема 1.2. *Сумма и произведение двух бесконечно малых функций (при $x \to x_0$) снова являются бесконечно малыми функциями (при $x \to x_0$).*

Для доказательства следует воспользоваться общими правилами нахождения пределов (§ 1.4, п. 2, свойства 1 и 2).

Теорема 1.3. *Произведение бесконечно малой функции на ограниченную функцию есть снова бесконечно малая функция.*

Доказательство. Пусть $f(x)$ – бесконечно малая функция при $x \to x_0$: $\lim\limits_{x\to x_0} f(x) = 0$, а $g(x)$ – ограниченная функция: $|g(x)| \le A$ ($A = \text{const}$) при любом x. Запишем

$$0 \le |f(x)\cdot g(x)| \le A|f(x)|.$$

Первый и третий члены неравенства (т.е. 0 и $A|f(x)|$) имеют при $x \to x_0$ предел 0. По известному положению о пределе функции второй член, т.е. $|f(x)\cdot g(x)|$, также имеет предел 0. Отсюда очевидным образом следует $\lim\limits_{x\to x_0} f(x)g(x) = 0$.

Особое значение имеет вопрос о поведении *частного* $f(x)/g(x)$ двух бесконечно малых (при $x \to x_0$) функций. В общем случае о пределе такого частного ничего определенного сказать нельзя. Вот несколько примеров:

$$\lim_{x\to 0}\frac{x^2}{x} = 0,\ \lim_{x\to 0}\frac{x}{x^2} = \pm\infty,\ \lim_{x\to 0}\frac{x\sin\frac{1}{x}}{x}\ \text{не существует.}$$

Заметим, что исследование отношения двух бесконечно малых функций представляет чрезвычайную важность для всех приложений математического анализа, поскольку на рассмотрении предела такого отношения основывается важнейшее понятие математического анализа – понятие *производной* функции.

Между понятиями бесконечно малой и бесконечно большой функций существует очевидная связь: если $f(x)$ – бесконечно большая (при $x \to x_0$), то $\dfrac{1}{f(x)}$ – бесконечно малая (при $x \to x_0$).

При сравнении бесконечно малых функций часто используют термин "бесконечно малая более высокого порядка".

Определение. *Пусть $f(x)$ и $g(x)$ – две бесконечно малые функции при $x \to x_0$. Мы говорим, что $f(x)$ есть **бесконечно малая более высокого порядка, чем** $g(x)$, если*

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0.$$

Например, при $x \to 0$ функция x^2 есть бесконечно малая более высокого порядка, чем x.

§ 1.8. Непрерывность функции

Понятие непрерывности, как и понятие предела, – одно из центральных в математическом анализе.

1. Непрерывность функции в точке

Определение. *Функция $f(x)$, определенная в некоторой окрестности точки x_0, называется **непрерывной в этой точке**, если предел функции в точке x_0 существует и равен значению в этой точке*:

$$\lim_{x\to x_0} f(x) = f(x_0). \tag{1.11}$$

Геометрический смысл (1.11): при стремлении x к x_0 ордината графика функции приближается к числу $f(x_0)$, т.е. график в точке x_0 "не рвется" (рис.1.10).

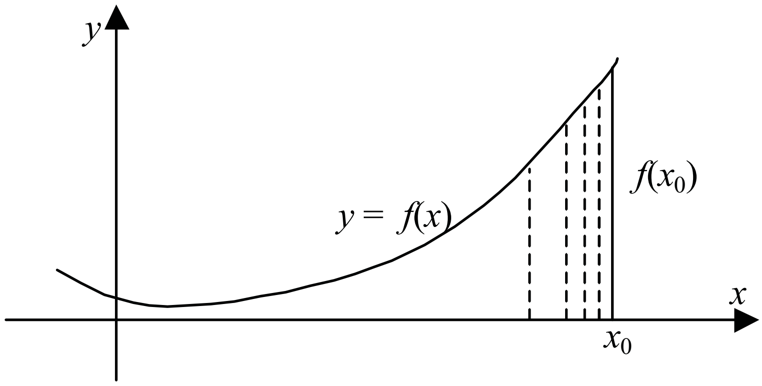

Рис. 1.10

Учитывая определение предела функции, можно дать и более развернутое определение непрерывности функции в точке: $f(x)$ непрерывна в точке x_0, если для любой последовательности x_1, x_2, …, сходящейся к x_0, соответствующая последовательность $f(x_1)$, $f(x_2)$, … сходится к числу $f(x_0)$.

Приведем еще одно определение непрерывности, которое, по существу, перефразирует данное выше определение. Пусть x – любая точка из окрестности точки x_0. Разность $f(x) - f(x_0)$, как известно, называется *приращением* функции. Обычно пишут $x - x_0 = \Delta x$, $f(x) - f(x_0) = \Delta y$. Таким образом,

$$\Delta y = f(x_0 + \Delta x) - f(x_0),$$

т.е. Δy является некоторой функцией от Δx. Тогда определение непрерывности $f(x)$ в точке x_0 можно записать в виде условия

$$\lim_{\Delta x \to 0} \Delta y = 0. \qquad (1.12)$$

Действительно, равенство (1.12) означает

$$\lim_{\Delta x \to 0} f(x_0 + \Delta x) - \lim_{\Delta x \to 0} f(x_0) = 0, \text{ или } \lim_{\Delta x \to 0} f(x_0 + \Delta x) = f(x_0),$$

что лишь способом записи отличается от (1.11).

Итак, *функция* $f(x)$ *непрерывна в точке* x_0, *если приращение функции стремится к нулю, когда приращение аргумента стремится к нулю*.

Укажем еще одно истолкование непрерывности. Равенство (1.11) можно записать в виде

$$\lim_{x \to x_0} f(x) = f\left(\lim_{x \to x_0} x\right),$$

т.е. функция $f(x)$ непрерывна, если *можно менять местами символы* f *и* lim (операцию нахождения функции и операцию предельного перехода).

Пусть функция $f(x)$ задана на отрезке $[a, b]$. Можно ли говорить о непрерывности $f(x)$ в точке a (или в точке b)? Исходя из данного выше определения непрерывности, конечно, нельзя, так как функция не определена слева от a. Поэтому введем понятие односторонней непрерывности.

Определение. *Функция* $f(x)$, *определенная на отрезке* $[a, b]$, *называется непрерывной в точке а* ***справа****, если*

$$\lim_{x \to a+0} f(x) = f(a).$$

Аналогично определяется непрерывность $f(x)$ в точке b слева:

$$\lim_{x \to b-0} f(x) = f(b).$$

Например, функция

$$f(x) = \begin{cases} +1, & \text{если } x \geq 0, \\ -1, & \text{если } x < 0, \end{cases}$$

в точке $x_0 = 0$ непрерывна справа, так как $\lim\limits_{x \to 0+0} f(x) = \lim 1 = 1;$ однако та же функция в точке 0 не является непрерывной слева, так как $\lim\limits_{x \to 0-0} f(x) = \lim(-1) = -1,$ в то время как $f(0) = +1$.

2. Арифметические операции над непрерывными функциями

Для решения вопроса о непрерывности тех или иных конкретных функций полезно знать те операции над функциями, которые сохраняют непрерывность, т.е., будучи примененными к непрерывным функциям, дают снова непрерывные функции. К числу таких операций относятся прежде всего арифметические.

Теорема 1.4. *Если функции $f(x)$ и $g(x)$ непрерывны в точке x_0, то каждая из функций*

$$f(x)+g(x),\ f(x)-g(x),\ f(x)\cdot g(x),\ \frac{f(x)}{g(x)} \tag{1.13}$$

(*последняя при* $g(x_0)\neq 0$) *также непрерывна в точке* x_0.

Докажем, например, непрерывность $\dfrac{f(x)}{g(x)}$ в точке x_0 (при условии $g(x_0)\neq 0$). Мы имеем

$$\lim_{x\to x_0}\frac{f(x)}{g(x)}=\frac{\lim\limits_{x\to x_0} f(x)}{\lim\limits_{x\to x_0} g(x)}=\frac{f(x_0)}{g(x_0)},$$

что и требовалось получить. Аналогично доказывается непрерывность остальных функций из ряда (1.13).

Доказанная теорема позволяет установить непрерывность значительного класса функций. Мы исходим из того, что каждая из функций

$$f(x)=c=\text{const},\quad f(x)=x$$

непрерывна в любой точке. Применяя теорему 1.4, получим тогда, что любая функция вида

$$P(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n$$

(т.е. любой многочлен от x) непрерывна в любой точке x_0, а любая функция вида $\frac{P(x)}{Q(x)}$, где $P(x)$ и $Q(x)$ – два многочлена, непрерывна в любой точке x_0, где $Q(x) \neq 0$.

Например, функция

$$\frac{x^3-2}{x^2+4x-5}$$

непрерывна в любой точке x_0, за исключением $x_0 = 1$ и $x_0 = -5$ (корни знаменателя).

Другие виды операций, сохраняющих непрерывность, будут указаны в § 1.15.

3. Постоянство знака непрерывной функции

Пусть $f(x)$ непрерывна в точке x_0 и $f(x_0) > 0$. Покажем, что *тогда существует такая окрестность точки* x_0, *во всех точках которой* $f(x) > 0$.

Будем рассуждать от противного. Допустим, что в любой окрестности точки x_0 есть точки, где $f \leq 0$. Очевидно, что в этом случае можно найти последовательность $x_1, x_2, \ldots$, сходящуюся к x_0, причем $f(x_i) \leq 0$ для всех i. Но в таком случае $\lim\limits_{i\to\infty} f(x_i)$ также не больше 0, а это противоречит равенству $\lim\limits_{x\to x_0} f(x) = f(x_0) > 0.$ Полученное противоречие доказывает, что существует окрестность точки x_0, в которой $f(x) > 0$.

Аналогично, если $f(x)$ непрерывна в точке x_0 и $f(x_0) < 0$, то существует окрестность точки x_0, в которой $f(x) < 0$.

Итак, *если функция* $f(x)$ *непрерывна в точке* x_0, *причем* $f(x_0) \neq 0$, *то существует окрестность точки* x_0, *во всех точках которой* $f(x)$ *имеет тот же знак*, *что и в точке* x_0.

4. Расширение понятия непрерывности функции в точке

Равенство

$$\lim_{x \to x_0} f(x) = f(x_0), \tag{1.14}$$

являющееся определением непрерывности функции в точке, можно понимать и в более широком смысле. При этом необязательно требовать, чтобы функция $f(x)$ была определена во всех точках из некоторой окрестности точки x_0. Можно заменить это требование более слабым, а именно – точка x_0 принадлежит области определения X функции $f(x)$ и является *предельной точкой* для X. Например, если множество X есть отрезок $[a, b]$, то концы отрезка принадлежат X и являются предельными точками для X.

Определение. *Пусть функция $f(x)$ определена на множестве X и пусть x_0 есть точка, принадлежащая множеству X и являющаяся предельной для X. Мы скажем, что $f(x)$* ***непрерывна*** *в точке x_0, если предел $\lim\limits_{x \to x_0} f(x)$ существует и равен $f(x_0)$.*

Часто приходится рассматривать случай, когда функция $f(x)$ задана лишь с одной стороны от точки x_0, например справа: это означает, что $f(x)$ определена на отрезке вида $[x_0, x_1]$, где $x_1 > x_0$. Если в этом случае $f(x)$ непрерывна в точке x_0, то говорят о непрерывности $f(x)$ в точке x_0 *справа*. Если $f(x)$ определена слева от x_0, то в случае непрерывности $f(x)$ в точке x_0 говорят о непрерывности *слева*.

Впрочем, имеется еще один случай, на который мы хотели бы распространить понятие непрерывности. Это случай, когда функция определена в точке x_0, но не определена в близких точках. Примем на этот счет такое определение.

Определение. *Точка $x_0 \in X$ называется* ***изолированной*** *точкой множества X, если существует такой интервал $(x_0 - \varepsilon, x_0 + \varepsilon)$, где $\varepsilon > 0$, в котором нет ни одной точки из X, кроме самой точки x_0.*

Нетрудно видеть, что любая точка множества X либо является предельной для X, либо является изолированной точкой множества X. Действительно, если точка x_0 не изолированная, то в любой близости от нее найдутся точки из X, отличные от x_0, а это и означает, что x_0 – предельная точка для X.

В дополнение к данному выше определению примем следующее соглашение.

Если функция $f(x)$ *определена на множестве* X *и* x_0 *– изолированная точка* X, *то будем считать, что* $f(x)$ *непрерывна в точке* x_0.

Ясно, что в этом случае апеллировать к равенству (1.14) мы не можем, поскольку теряет всякий смысл выражение, стоящее в левой части.

В заключение данного параграфа примем еще одно определение.

Определение. *Функция* $f(x)$, *определенная на множестве* X, *называется* ***непрерывной*** *на этом множестве, если она непрерывна в каждой точке из* X.

§ 1.9. Теорема о стягивающихся отрезках. Точные границы числового множества

Для дальнейшего, более углубленного изучения непрерывных функций, необходимо сделать отступление в сторону и доказать одну теорему, означающую, в известном смысле, непрерывность самого множества действительных чисел.

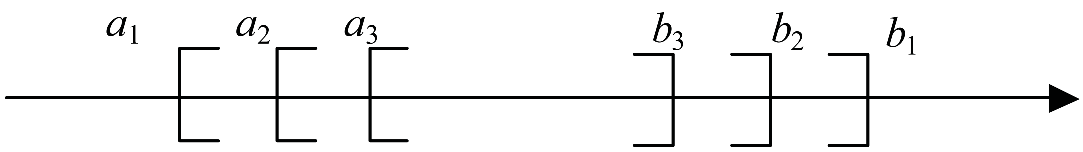

Рис. 1.11

Определение. *Последовательность отрезков*

$$[a_1, b_1],\ [a_2, b_2], \ldots$$

называется ***стягивающейся системой отрезков*** (рис. 1.11), *если выполнены два условия*:

1) *каждый следующий отрезок содержится в предыдущем, т.е.* $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$ $(n = 1, 2, \ldots)$;

2) *длина отрезка* $[a_n, b_n]$ *с ростом его номера n стремится к нулю, т.е.* $\lim\limits_{n\to\infty}(b_n - a_n) = 0$.

Поскольку числовая прямая $\mathbb{R}$ представляется нам непрерывной, не имеющей "пробелов", то интуитивно ясно, что система отрезков, удовлетворяющая условиям 1) и 2), стягивается обязательно к *некоторой точке* числовой прямой. Дадим более четкую формулировку этого положения.

Теорема 1.5 (*о стягивающейся системе отрезков*). *Для любой стягивающейся системы отрезков существует точка, притом единственная, принадлежащая всем этим отрезкам.*

Доказательство. Последовательность $a_1, a_2, \ldots$ левых концов отрезков является, очевидно, монотонной (неубывающей), поскольку $a_1 \le a_2 \le \ldots$. Так как при этом все числа $a_1, a_2, \ldots$ не превосходят, например, b_1, то эта последовательность является также ограниченной. По теореме 1.1 существует число $a = \lim\limits_{n\to\infty} a_n$. Так как при любом n имеем $a_n \ge a_1$, то и $a \ge a_1$. Аналогично из $a_n \ge a_2$, при $n = 2, 3, \ldots$ следует $a \ge a_2$. Точно так же получим $a \ge a_3$, $a \ge a_4$ и т.д. С другой стороны, поскольку все члены последовательности $\{a_n\}$ удовлетворяют условию $a_n \le b_1$, то и $a \le b_1$; точно так же получим $a \le b_2$, $a \le b_3$ и т.д. Таким образом, имеем $a_n \le a \le b_n$ при любом n, т.е. точка a принадлежит всем отрезкам $[a_n, b_n]$. Остается показать, что других точек, принадлежащих всем $[a_n, b_n]$, не существует.

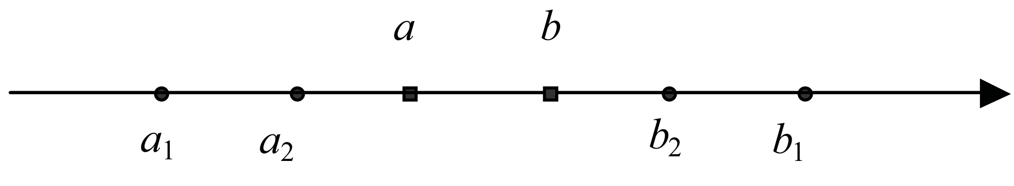

Рис. 1.12

Предположим, рассуждая от противного, что имеется точка b, отличная от a и принадлежащая всем $[a_n, b_n]$. Тогда имеем, очевидно, $b_n - a_n \ge d$, где d – расстояние между точками a и b (рис. 1.12). Переходя в этом неравенстве к пределу при $n \to \infty$, получим $0 \ge d$, т.е. $d = 0$, что невозможно, так как точки a и b различны. Теорема доказана.

Наглядный смысл теоремы о стягивающейся системе отрезков состоит, как уже отмечалось, в том, что числовая прямая $\mathbb{R}$ "не имеет пробелов": действительно, если бы точка, принадлежащая всем отрезкам системы, не существовала, то это означало бы наличие "пробела" в $\mathbb{R}$.

Пусть M – ограниченное сверху множество чисел. Любое число b, такое, что $x \le b$ для всех $x \in M$, называется *верхней границей* множества M.

Теорема 1.6. *Если множество M ограничено сверху, то среди всех его верхних границ имеется наименьшая.*

Доказательство проведем от противного. Допустим, что среди верхних границ множества M нет наименьшей. Пусть b – какая-нибудь верхняя граница для M. Рассмотрим любое число a из M. Имеем $a \le b$. Равенство $a = b$ невозможно, так как это означало бы, что b – наименьшая из верхних границ множества M. Следовательно, $a < b$.

Разделим отрезок $[a, b]$ пополам точкой c. Если c – верхняя граница для множества M, то выберем левую половину отрезка. Если же c не является верхней границей, то выберем правую половину. В обоих случаях выбранную половину отрезка $[a, b]$ обозначим $[a_1, b_1]$. Итак, правый конец отрезка $[a_1, b_1]$ является верхней границей для M, в самом же отрезке $[a_1, b_1]$ есть точки из M. Разделим отрезок $[a_1, b_1]$ пополам и из двух половин снова выберем такую половину $[a_2, b_2]$, что правее нее нет точек из M, а в самом отрезке $[a_2, b_2]$ точки из M есть. Этот процесс можно продолжать неограниченно, в итоге получим стягивающуюся систему отрезков

$$[a, b], [a_1, b_1], [a_2, b_2], \ldots,$$

таких, что в любом из этих отрезков имеются точки из M, а правее отрезка точек из M нет. По теореме 1.5 существует точка p, принадлежащая всем отрезкам. Правее точки p нет точек M, так как их нет правее любого отрезка $[a_n, b_n]$; таким образом, p – одна из верхних границ.

Осталось показать, что p – *наименьшая* верхняя граница.

Предположим, что это не так, и имеется другая верхняя граница t, $t < p$. Ввиду того, что отрезки $[a_n, b_n]$ стягиваются к p, имеем $p = \lim a_n$ $(n \to \infty)$. Так как при этом $t \geq a_n$ (в противном случае на отрезке $[a_n, b_n]$ не могло бы быть точек из M), то $t \geq \lim a_n$, т.е. $t \geq p$. Неравенства $t < p$ и $t \geq p$ противоречат друг другу. Теорема доказана.

Определение. *Если множество M ограничено сверху, то наименьшая из его верхних границ называется* ***точной верхней границей*** *множества M.*

Аналогично определяются *нижняя граница* и *точная нижняя граница* множества M. Нижняя граница есть такое число c, что $x \geq c$ для всех $x \in M$; точная нижняя граница есть наибольшая из всех нижних границ. Если множество M ограничено снизу, то его точная нижняя граница всегда существует – это можно доказать по аналогии с теоремой о точной верхней границе, но можно поступить по-другому, заменив множество M на $-M$ (меняем знаки у всех чисел из M), тогда нижние границы для M превратятся в верхние границы для $-M$.

Если множество M ограничено и сверху, и снизу (т.е. просто ограничено), то для него существуют обе точные границы – верхняя p и нижняя q. Отрезок $[p, q]$ – наименьший из всех отрезков, содержащих целиком все множество M.

Пример 1.11. Множество $M = (0; 1)$ ограничено; его точной верхней границей является число 1, точной нижней границей – число 0.

§ 1.10. Свойства функций, непрерывных на отрезке

Напомним, что согласно данному ранее определению функция $f(x)$, определенная на отрезке $[a, b]$, называется непрерывной на этом отрезке, если она непрерывна в каждой точке отрезка $[a, b]$.

Функции, непрерывные на отрезке, обладают рядом важных свойств. Ниже даются формулировки и доказательства этих свойств. Еще раз обратим внимание на необходимость безусловного знания свойств непрерывных функций на отрезке.

Свойство 1. Теорема о существовании корня

Если функция $f(x)$ *непрерывна на отрезке* $[a, b]$ *и на концах отрезка принимает значения разных знаков* (т.е. $f(a)\cdot f(b) < 0$), *то существует точка* $c \in (a, b)$, *такая что* $f(c) = 0$ (рис. 1.13).

Доказательство. Предположим, для определенности, что $f(a) < 0$, $f(b) > 0$. Разделим отрезок $[a, b]$ пополам точкой c_1. Если $f(c_1) = 0$, то доказывать нечего. Если же $f(c_1) \neq 0$, то обозначим через $[a_1, b_1]$ ту из половин отрезка $[a, b]$, для которой $f(a_1) < 0$, $f(b_1) > 0$ (в случае $f(c_1) > 0$ – это левая половина, при $f(c_1) < 0$ – правая). Отрезок $[a_1, b_1]$ также разделим пополам точкой c_2. Тогда либо $f(c_2) = 0$, и доказывать нечего, либо $f(c_2) \neq 0$, и тогда обозначим $[a_2, b_2]$ ту из половин отрезка $[a_1, b_1]$, для которой $f(a_2) < 0$, $f(b_2) > 0$.

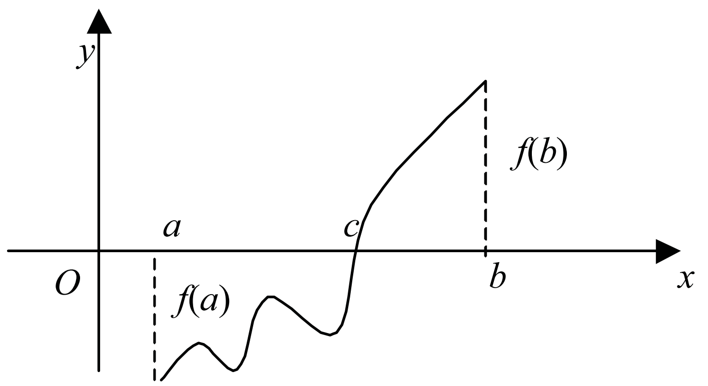

Рис. 1.13

Повторяя это рассуждение, получим либо точку c_n, в которой $f(c_n) = 0$, либо бесконечную последовательность вложенных друг в друга отрезков:

$$[a_1, b_1] \supset [a_2, b_2] \supset \ldots,$$

причем длины отрезков стремятся к 0, поскольку каждый следующий отрезок является одной из половин предыдущего отрезка. По теореме о стягивающейся системе существует точка c, принадлежащая всем отрезкам:

$$a_n \leq c \leq b_n, \qquad n = 1, 2, \ldots$$

Имеем $c = \lim\limits_{n\to\infty} a_n$, а также $c = \lim\limits_{n\to\infty} b_n$; поскольку функция $f(x)$ непрерывна в точке c, то должно быть

$$f(c) = \lim_{n\to\infty} f(a_n),\ f(c) = \lim_{n\to\infty} f(b_n).$$

Но $f(a_n) < 0$ для всех n, откуда следует $f(c) \le 0$; аналогично из $f(b_n) > 0$ следует $f(c) \ge 0$. Выполнение обоих неравенств $f(c) \le 0$ и $f(c) \ge 0$ означает, что $f(c) = 0$. Свойство 1 доказано.

Свойство 2. Теорема о промежуточном значении

Пусть $f(x)$ *непрерывна на* $[a, b]$ *и* $f(a) = A$, $f(b) = B$, $A \ne B$. *Тогда для любого числа* C, *расположенного между* A *и* B, *найдется такая точка* $c \in (a, b)$, *что* $f(c) = C$.

Иначе говоря, непрерывная функция на отрезке принимает все значения, промежуточные между ее значениями на концах отрезка.

Доказательство. Пусть для определенности $A < B$. Рассмотрим функцию $\varphi(x) = f(x) - C$. Функция $\varphi(x)$ непрерывна вслед за $f(x)$, причем $\varphi(a) = A - C < 0$ и $\varphi(b) = B - C > 0$. Согласно свойству 1 найдется точка $c \in (a, b)$, для которой $\varphi(c) = 0$, т.е. $f(c) = C$.

Свойство 3. Ограниченность непрерывной функции

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

Иначе говоря, существует такое число $A > 0$, что $|f(x)| \le A$ для всех x из $[a, b]$.

Доказательство проведем от противного. Допустим, что $f(x)$ не ограничена на $[a, b]$. Положим $a_1 = a$, $b_1 = b$ и разделим отрезок $[a_1, b_1]$ пополам. Хотя бы на одной из половин функция $f(x)$ будет неограничена; зафиксируем одну из таких половин и обозначим ее $[a_2, b_2]$. Снова разделим отрезок $[a_2, b_2]$ пополам и обозначим $[a_3, b_3]$ ту из половин, где функция неограничена. И так далее. В итоге получим бесконечную последовательность вложенных друг в друга отрезков:

$$[a_1, b_1] \supset [a_2, b_2] \supset \dots,$$

длины которых стремятся к нулю; при этом на каждом из отрезков функция $f(x)$ неограничена. По теореме о стягивающейся системе отрезков существует точка x_0, принадлежащая всем отрезкам.

Так как $f(x)$ неограничена на $[a_1, b_1]$, то найдется точка $x_1 \in [a_1, b_1]$, для которой $|f(x_1)| > 1$. Так как $f(x)$ неограничена на $[a_2, b_2]$, то найдется точка $x_2 \in [a_2, b_2]$, для которой $|f(x_2)| > 2$. Продолжая это рассуждение, получим бесконечную последовательность точек $x_1, x_2, \ldots$, таких что $x_n \in [a_n, b_n]$ и $|f(x_n)| > n$ ($n = 1, 2, \ldots$). Очевидно, что $\lim\limits_{n\to\infty} x_n = x_0$. Ввиду непрерывности функции $f(x)$ в точке x_0 отсюда должно следовать $\lim\limits_{n\to\infty} f(x_n) = f(x_0)$, т.е. последовательность чисел $f(x_1), f(x_2), \ldots$ должна быть сходящейся. Но в таком случае эта последовательность обязательно ограничена (свойство 2 п.2 §1.2). Между тем неравенства $|f(x_1)| > 1$, $|f(x_2)| > 2$, … показывают, что эта последовательность неограничена. Получили противоречие.

Свойство 4. Достижение крайних значений

Для всякой функции $f(x)$, *непрерывной на отрезке* $[a, b]$, *существует точка* $c \in [a, b]$, *такая, что* $f(c) \geq f(x)$ *для всех* $x \in [a, b]$ *и точка* $d \in [a, b]$, *такая, что* $f(d) \leq f(x)$ *для всех* $x \in [a, b]$.

Иначе говоря, функция, непрерывная на отрезке, достигает на нем как своего наибольшего, так и своего наименьшего значений.

Доказательство. Обозначим $E(f)$ множество значений функции $f(x)$ на отрезке $[a, b]$, т.е.

$$E(f) = \{f(x) \mid x \in [a, b]\}.$$

Согласно свойству 3 множество $E(f)$ ограниченное, поэтому оно имеет (см. конец § 1.9) точную нижнюю границу m и точную верхнюю границу M. Докажем, что существуют точка $c \in [a, b]$, для которой $f(c) = M$, и точка $d \in [a, b]$, для которой $f(d) = m$. Этим, очевидно, и будет доказано свойство 4.

Предположим, что точки c с таким свойством не существует, т.е. что $f(x) \neq M$ для всех $x \in [a, b]$. Рассмотрим тогда вспомогательную функцию

$$F(x) = \frac{1}{M - f(x)}.$$

Знаменатель этой дроби – непрерывная на $[a, b]$ функция (вслед за $f(x)$), не обращающаяся ни в одной точке в нуль, а потому, ввиду $f(x) \le M$, строго положительная. Согласно свойству 3 непрерывных функций $F(x)$ ограничена на $[a, b]$. Следовательно, найдется такое число $K > 0$, что $F(x) \le K$ для всех $x \in [a, b]$.

Итак, имеем

$$\frac{1}{M - f(x)} \le K,$$

или, учитывая, что $M - f(x) > 0$,

$$M - f(x) \ge \frac{1}{K}.$$

Следовательно, $f(x) \le M - \dfrac{1}{K}$, что противоречит выбору M как *точной* верхней границы множества $E(f)$.

Полученное противоречие доказывает, что точка $c \in [a, b]$, для которой $f(c) = M$, существует. Аналогично доказывается существование точки $d \in [a, b]$, для которой $f(d) = m$. Свойство 4 доказано.

Свойство 5. Множество значений непрерывной функции

Если функция $f(x)$ *непрерывна на отрезке* $[a, b]$, *то множество* $E(f)$ *значений функции также является отрезком*.

Доказательство. По свойству 4 существуют $c, d \in [a, b]$, такие, что $f(c) = M = \max f$ на $[a, b]$, $f(d) = m = \min f$ на $[a, b]$.

Если $m = M$, то $f(x)$ – постоянная функция, и $E(f)$ есть отрезок, состоящий из единственной точки m.

В противном случае (т.е. при $m < M$) любое число C, заключенное между m и M, также является (по свойству 2) значением функции, т.е. принадлежит $E(f)$. Следовательно, $E(f)$ есть отрезок $[m, M]$. Свойство 5 доказано.

Таким образом, *функция, непрерывная на отрезке* $[a, b]$, *отображает его снова на некоторый отрезок*.

На рис. 1.14 множество $E(f)$ значений функции на $[a, b]$ выделено двойной чертой (на оси Oy). Это некоторый отрезок $[m, M]$.

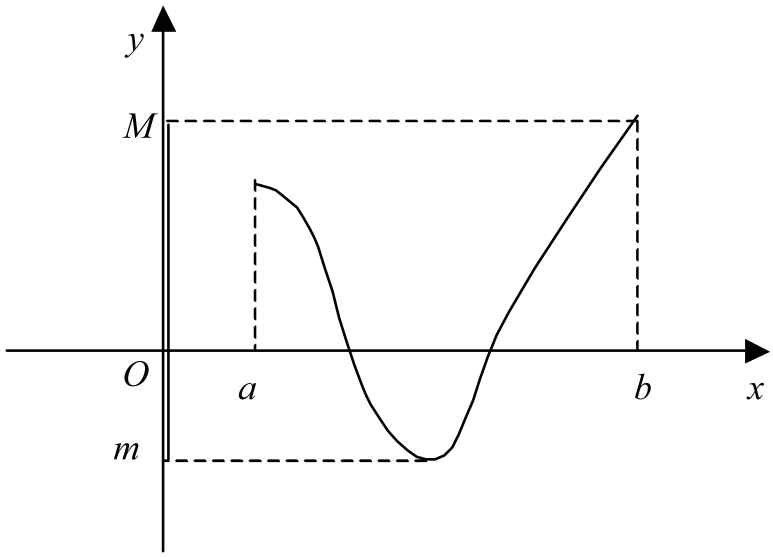

Рис. 1.14

Следующее свойство функции, непрерывной на отрезке, носит название *равномерной непрерывности*. Оно будет использовано в главе 4.

Свойство 6. Равномерная непрерывность

Если функция $f(x)$ *непрерывна на отрезке* $[a, b]$, *то для любого положительного числа* ε *найдется такое положительное* δ, *что разность значений функции в любых двух точках отрезка, отстоящих друг от друга меньше, чем на* δ, *по модулю меньше, чем* ε:

$$|x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon.$$

Д о к а з а т е л ь с т в о дается в приложении 3 к данной главе.

§ 1.11. Непрерывность сложной и обратной функций. Непрерывность элементарных функций

Ранее (в § 1.9) уже отмечалось, что арифметические операции над функциями сохраняют непрерывность. Теперь добавим к числу операций, сохраняющих непрерывность, еще две: *составление сложной функции* и *переход от данной функции к обратной*.

Теорема 1.7 (***о непрерывности сложной функции***). *Пусть даны две непрерывные функции*:

$x = h(t)$, *отображающая множество* T *в множество* X;

$y = g(x)$, *отображающая множество* X *в множество* Y.

Тогда сложная функция $y = g(h(t))$ *непрерывна на множестве* T (*и отображает его на* Y).

Доказательство. Пусть $t_0 \in T$. Если t_0 – изолированная точка множества T, то функция $y = g(h(t))$ непрерывна в этой точке по определению (§ 1.8 п.4). Пусть теперь t_0 – предельная точка для T. Рассмотрим произвольную последовательность $t_1, t_2, \ldots$ точек из T, сходящуюся к t_0. Вследствие непрерывности $h(t)$ последовательность $h(t_1), h(t_2), \ldots$ сходится к числу $h(t_0)$. Положим

$$x_1 = h(t_1), x_2 = h(t_2), \ldots; x_0 = h(t_0).$$

Ввиду непрерывности $g(x)$ последовательность $g(x_1), g(x_2), \ldots$ должна сходиться к числу $g(x_0)$. Итак, последовательность значений сложной функции

$$g(h(t_1)), g(h(t_2)), \ldots$$

сходится к числу $g(h(t_0))$, что и доказывает непрерывность сложной функции в точке t_0.

Теорема 1.8 (*о непрерывности обратной функции*). *Пусть функция* $y = f(x)$ *непрерывна и возрастает на отрезке* $[a, b]$. *Тогда обратная функция* $x = f^{-1}(y)$ *также непрерывна и возрастает на* $[f(a), f(b)]$.

(Аналогичное утверждение справедливо для непрерывной *убывающей* функции.)

Доказательство. Возрастание f^{-1} непосредственно следует из возрастания f. Докажем непрерывность f^{-1} в любой точке $y_0 \in [f(a), f(b)]$.

Положим $x_0 = f^{-1}(y_0)$, т.е. $f(x_0) = y_0$. Возьмем произвольную последовательность $y_1, y_2, \ldots$ из $[f(a), f(b)]$, сходящуюся к y_0; для простоты будем считать, что эта последовательность возрастающая, т.е. $y_1 < y_2 < \ldots$ (рис. 1.15). Пусть $x_1, x_2, \ldots$ – соответствующая последовательность значений f^{-1}, тогда $x_1 < x_2 < \ldots$. Ввиду того, что последовательность $x_1, x_2, \ldots$ – возрастающая и ограниченная (все члены последовательности содержатся в $[a, b]$), существует предел $c = \lim_{n\to\infty} x_n$. Учитывая непрерывность функции f в точке c, будем иметь $f(c) = \lim f(x_n) = \lim y_n = y_0$; сопоставляя $f(x_0) = y_0$ и $f(c) = y_0$, получаем $x_0 = c$.

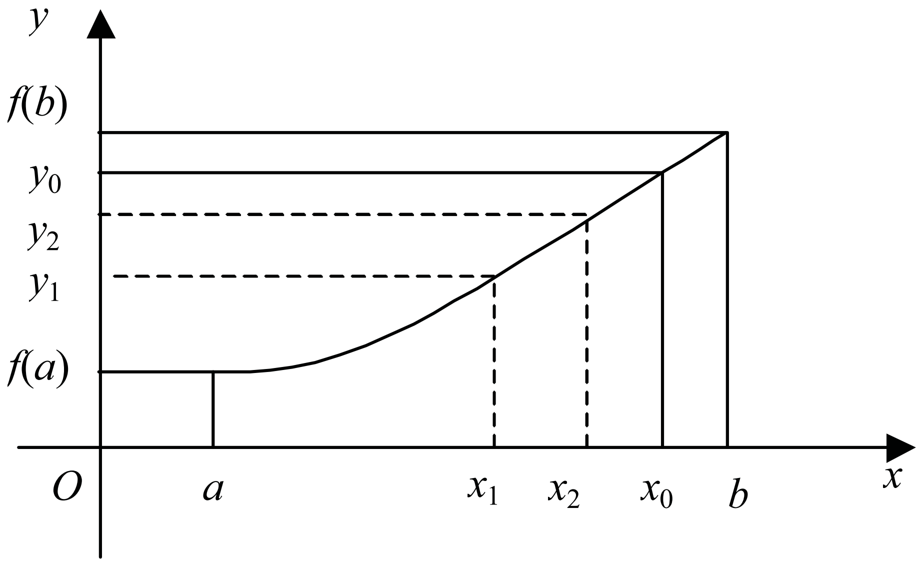

Рис. 1.15

Итак, $\lim x_n = x_0$, или

$$\lim f^{-1}(y_n) = f^{-1}(y_0), \tag{1.15}$$

что и доказывает непрерывность функции f^{-1} в точке y_0.

Если последовательность $y_1, y_2, \ldots$ не является монотонной, то равенство (1.15) так просто не получается. Однако, если из этой последовательности выбрать любую *монотонную* подпоследовательность $y_{n_1}, y_{n_2}, \ldots,$ то по доказанному получаем, что последовательность чисел

$$f^{-1}(y_{n_1}), f^{-1}(y_{n_2}), \ldots$$

сходится к числу $f^{-1}(y_0)$. Примем без доказательства следующий факт: если последовательность чисел такова, что любая ее монотонная подпоследовательность сходится к одному и тому же числу A, то и вся последовательность сходится к A. Отсюда следует (1.15), что и завершает доказательство теоремы.

В своих рассуждениях, носивших общий характер, мы несколько отдалились от тех конкретных функций, с которыми приходится иметь дело на практике, точнее – от *элементарных функций*. Как же обстоит дело с непрерывностью элементарных функций? Одна из весьма важных теорем математического анализа утверждает следующее.

Теорема 1.9. *Любая элементарная функция непрерывна в любой точке, принадлежащей области определения функции.*

Для доказательства достаточно проверить непрерывность основных элементарных функций:

$$y = x^a,\ y = a^x,\ y = \log_a x\ (a > 0, a \neq 1),$$
$$y = \sin x,\ y = \cos x,\ y = \operatorname{tg} x,$$
$$y = \arcsin x,\ y = \arccos x,\ y = \operatorname{arctg} x, \tag{1.16}$$

поскольку любая другая элементарная функция получается из этих с помощью операций, сохраняющих непрерывность: арифметических, а также составления сложной функции.

Непрерывность каждой из функций (1.16) доказывается с помощью своего специального рассуждения. Заниматься этим здесь не будем, апеллируя к наглядному представлению о графиках функций (1.16).

Пример 1.12. Функция $y = \arcsin\dfrac{1}{x}$ непрерывна во всей области ее существования, т.е. при $\dfrac{1}{x} \in [-1; 1]$, или, что то же, при $x \in$ $\in (-\infty, -1] \cup [1, \infty)$.

§ 1.12. Паутинные модели рынка

Свойство 1 непрерывной функции (теорема о существовании корня) находит неожиданное применение в математических *моделях рынка*. Как известно, две основные категории рыночных отношений – *спрос* и *предложение*. И то, и другое зависит от многих факторов, среди которых главный – это *цена* товара.

Обозначим цену товара p, объем спроса D, величину предложения S (от первых букв английских слов *Price* – цена, *Demand* – спрос, *Supply* – предложение). При малых p имеем $D(p) - S(p) > 0$ (спрос превышает предложение), при больших p, наоборот, $D(p) - S(p) < 0$. Считая $D(p)$ и $S(p)$ непрерывными функциями, приходим к заключению, что существует такая цена p^*, для которой $D(p^*) = S(p^*)$, т.е. спрос равен предложению (рис. 1.16). Цена p^* называется *равновесной*, спрос и предложение в таком случае также называются *равновесными*.

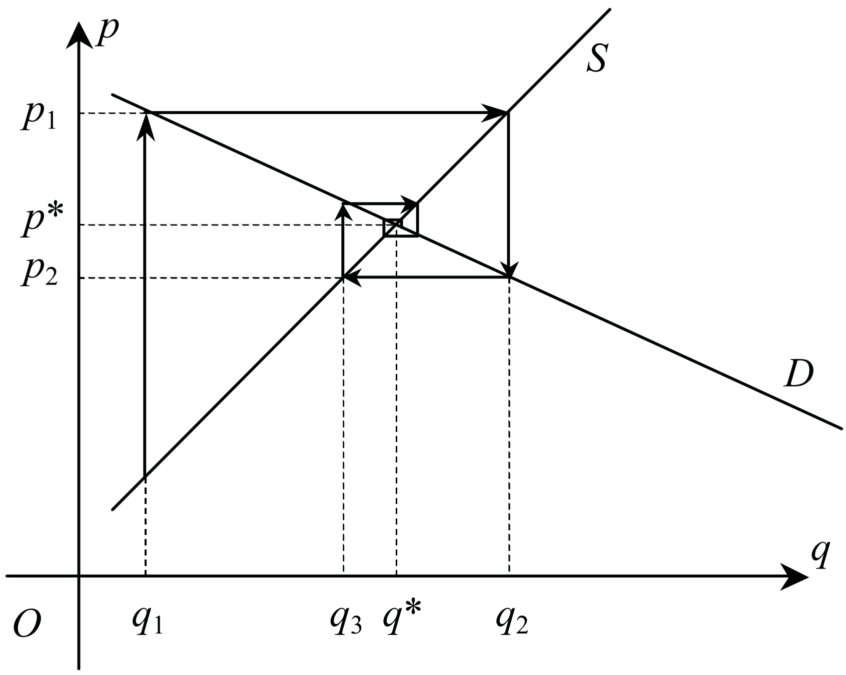

Рис. 1.16

Установление равновесной цены – одна из главных задач рынка. Рассмотрим простую модель поиска равновесной цены – так называемую *паутинную модель*. Она объясняет феномен регулярно повторяющихся циклов изменения объемов продажи и цен, например, сельскохозяйственных товаров.

Предположим, что решение о величине объема производства принимается в зависимости от цены товара в предыдущий период времени. Так, площадь, отводимую под сельскохозяйственную культуру, выбирают в зависимости от ее цены, сложившейся в предыдущем году.

Рассмотрим ситуацию, изображенную на рис. 1.16.

Пусть в первый торговый период предложение q_1 мало по сравнению с равновесным предложением q^*. Тогда в этом периоде из-за недостатка товара на него установится высокая цена p_1. Учитывая это обстоятельство, производители расширят выпуск и во втором периоде предложат большее количество товара q_2. Поскольку $q_2 > q^*$, во втором периоде будет наблюдаться избыток предложения, что приведет к снижению цены до p_2. Падение цены вызовет снижение производства, и в третьем торговом периоде предложение q_3, так же как и в первом, станет меньше равновесного, поэтому дальнейшая динамика предложения и спроса будет бесконечно повторять описанный выше цикл. При этом амплитуда колебаний цен будет постепенно уменьшаться, стремясь к 0, т.е. "паутина скручивается".

Такой характер динамики соответствует случаю, когда линия D спроса в каждой точке круче линии предложения S относительно оси цен.

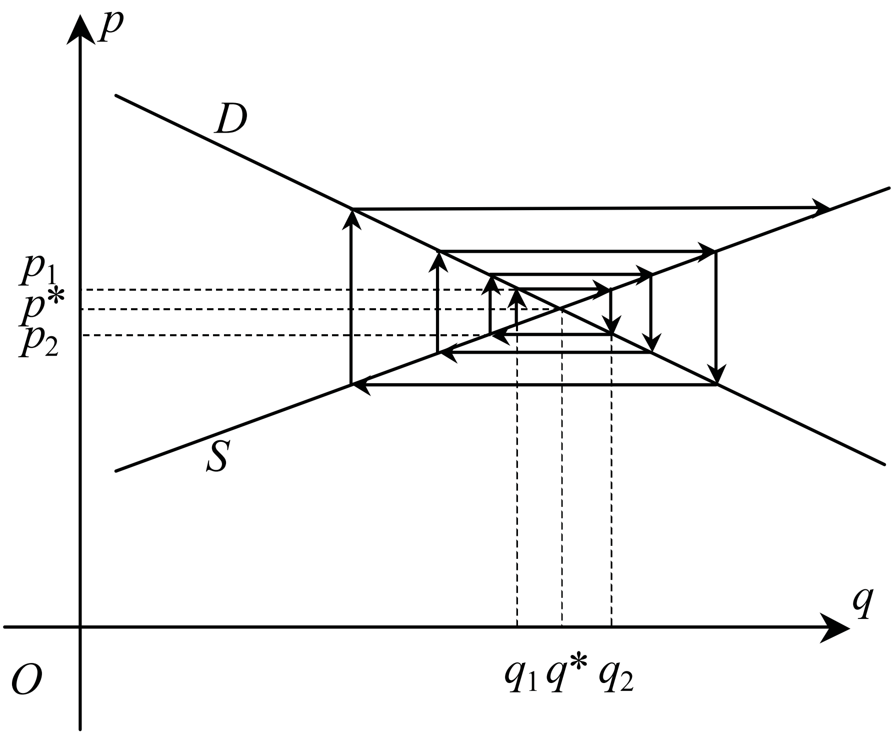

Рис. 1.17

Впрочем, "паутина" может и "раскручиваться", если линия спроса более пологая, чем линия предложения. Соответствующая траектория колебаний спроса и предложения отображена на рис. 1.17.

§ 1.13. Функции нескольких переменных

1. Определение функции нескольких переменных. Линии и поверхности уровня

До сих пор рассматривались числовые функции $y = f(x)$ одной переменной x. Областью определения такой функции являлось множество $X \subset \mathbb{R}$. Числовая *функция n переменных* характеризуется тем, что областью ее определения является подмножество X *пространства* $\mathbb{R}^n$, $n > 1$. В этом случае значение аргумента x представляет собой точку $(x_1, x_2,\ldots, x_n)$ из $\mathbb{R}^n$; соответственно пишут

$$y = f(x_1, x_2,\ldots, x_n) \tag{1.17}$$

(переменная y является функцией от n переменных $x_1, x_2,\ldots, x_n$, где $(x_1, x_2,\ldots, x_n) \in X$).

Для функций *двух* переменных ($n = 2$) вместо x_1, x_2, y пишут обычно x, y, z, тогда (1.17) принимает вид

$$z = f(x, y); \tag{1.18}$$

точка (x, y) пробегает область определения $X \subset \mathbb{R}^2$ функции f.

Линией уровня функции (1.18) называют линию

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

на координатной плоскости, в точках которой функция f принимает постоянное значение C.

Например, для функции $z = \ln x + \ln y$ ($x > 0$, $y > 0$) любая из линий уровня определяется условиями $x > 0$, $y > 0$, $xy = C$ (естественно, $C > 0$) и представляет собой ветвь гиперболы $xy = C$, расположенную в первой четверти.

При $n > 2$ следует говорить не о линиях, а о *множествах уровня.* Множество уровня имеет уравнение

$$f(x_1, x_2,\ldots, x_n) = C$$

и истолковывается как "поверхность" в $\mathbb{R}^n$.

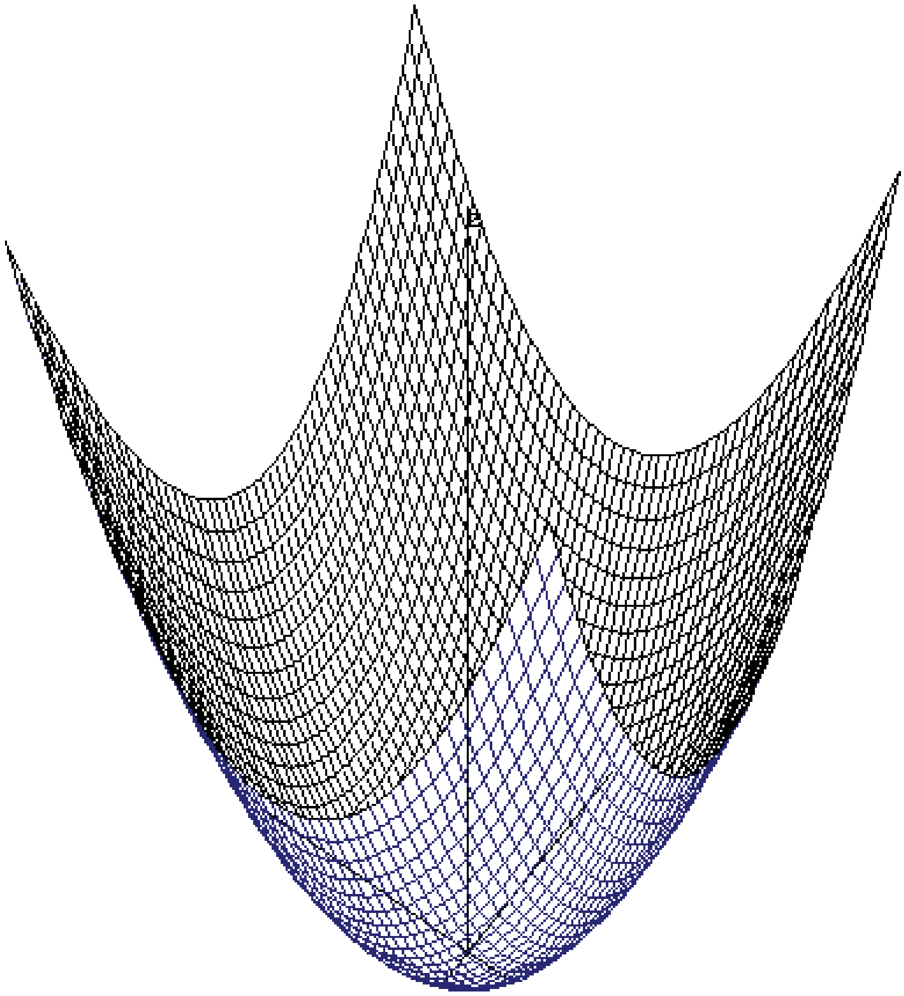

Рис. 1.18

Например, поверхность уровня функции

$$u = x^2 + y^2 + z^2$$

есть поверхность в $\mathbb{R}^3$; она определяется уравнением

$$x^2 + y^2 + z^2 = C,$$

где $C \geq 0$, и представляет собой сферу радиуса $\sqrt{C}$ с центром в начале координат. Для функции

$$u = \frac{1 + x^2 + y^2}{z} \quad (z \neq 0)$$

поверхность уровня определяется уравнением

$$1 + x^2 + y^2 = Cz$$

и представляет собой поверхность вращения с осью z (сечения плоскостями $z = \mathrm{const}$ – окружности). Одна из таких поверхностей уровня $C = 1$ показана на рис. 1.18.

2. Элементарные функции нескольких переменных

Практические приложения математического анализа, связанные с функциями нескольких переменных, обеспечиваются в основном *элементарными функциями*.

Определение. *Функция $f(x_1, \ldots, x_n)$, заданная некоторым аналитическим выражением, называется* ***элементарной***, *если в ее запись входят только*: 1) *символы самих переменных* $x_1, \ldots, x_n$; 2) *знаки арифметических операций* $+, -, \cdot, :$; 3) *скобки*; 4) *символы основных элементарных функций одной переменной*:

$(\ldots)^a$ – *символ степенной функции*,

$a^{(\ldots)}$ – *символ показательной функции*,

$\log_a(\ldots)$ – *символ логарифмической функции*,

$\sin(\ldots)$, $\cos(\ldots)$, $\mathrm{tg}(\ldots)$ – *символы тригонометрических функций*,

$\arcsin(\ldots)$, $\arccos(\ldots)$, $\mathrm{arctg}(\ldots)$ – *символы обратных тригонометрических функций*.

Областью определения элементарной функции $f(x_1, \ldots, x_n)$ можно считать множество всех точек $(x_1, \ldots, x_n) \in \mathbb{R}^n$, для которых выражение, задающее функцию, имеет смысл (ОДЗ).

Например, функция

$$y = 2^{\sqrt{1-\sin^2 x_1}} + \log_2(x_2 x_3) + \frac{1}{(x_1 - 1)(x_2 - 1)}$$

является элементарной функцией трех переменных x_1, x_2, x_3. Ее область определения задается в пространстве $\mathbb{R}^3$ условиями

$$x_2 x_3 > 0, \quad x_1 \neq 1, \quad x_2 \neq 1.$$

§ 1.14. Сходимость точек в $\mathbb{R}^n$. Открытые и замкнутые множества. Предел и непрерывность для функций нескольких переменных

1. Расстояние между точками в $\mathbb{R}^n$

Напомним, что в пространстве $\mathbb{R}^1$ (т.е. на числовой прямой) расстояние между точками x_1 и x_2 равно $|x_1 - x_2|$, в пространстве $\mathbb{R}^2$ для расстояния между точками $P(x_1, y_1)$ и $Q(x_2, y_2)$ справедлива формула

$$\rho(P,Q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \qquad (1.19)$$

а в пространстве $\mathbb{R}^3$ – формула

$$\rho(P,Q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}, \qquad (1.20)$$

где $P = P(x_1, y_1, z_1)$, $Q = Q(x_2, y_2, z_2)$.

В пространстве $\mathbb{R}^n$, где $n > 3$, о расстоянии можно говорить лишь в условном смысле, так как точки в $\mathbb{R}^n$ не имеют непосредственного геометрического истолкования. Аналогично формулам (1.19) и (1.20) определяем расстояние в $\mathbb{R}^n$, где $n > 3$, формулой

$$\rho(p,q) = |p - q| = \sqrt{(x_1' - x_1'')^2 + \ldots + (x_n' - x_n'')^2}, \qquad (1.21)$$

где $p = (x_1', x_2', \ldots, x_n')$, $q = (x_1'', x_2'', \ldots, x_n'')$ – две произвольные точки из $\mathbb{R}^n$.

Введенное таким путем расстояние в $\mathbb{R}^n$ обладает следующими свойствами:

1) $\rho(p, q) > 0$, если $p \neq q$, и $\rho(p, p) = 0$;

2) $\rho(p, q) = \rho(q, p)$;

3) $\rho(p, q) + \rho(q, r) \geq \rho(p, r)$, $\qquad (1.22)$

каковы бы ни были точки p, q, r.

Свойства 1 и 2 очевидным образом следуют из определения расстояния. Свойство 3 носит название "неравенство треугольника". В случае пространств $\mathbb{R}^2$ и $\mathbb{R}^3$ оно имеет очевидный смысл, так как выражает тот факт, что сумма двух сторон треугольника всегда больше либо равна третьей стороне.

Доказательство неравенства 3 (1.22) дается в приложении 1 к данной главе.

При рассмотрении функций в $\mathbb{R}^n$ широко используется геометрический язык, хотя буквальное понимание геометрических терминов возможно только при $n = 2$ и $n = 3$.

Определение. *Пусть p_0 – точка в $\mathbb{R}^n$ и ε – положительное число.* ***Открытым шаром,*** *или просто* ***шаром*** *радиуса ε с центром p_0, называется множество всех точек, расстояние которых от p_0 меньше ε:*

$$\{p \in \mathbb{R}^n \mid \rho(p_0, p) < \varepsilon\}. \tag{1.23}$$

Шар радиуса ε с центром p_0 обозначается $B(p_0, \varepsilon)$ или $U_\varepsilon(p_0)$. Множество $U_\varepsilon(p_0)$ называют ***ε-окрестностью*** точки p_0. Разумеется, при $n = 2$ вместо слова "шар" следует говорить "круг" либо использовать нейтральный термин "ε-окрестность".

Определение. *Множество $X \subset \mathbb{R}^n$ называется* ***ограниченным****, если оно целиком содержится в некотором шаре.*

Нетрудно показать, что ограниченность множества X означает, что существует такое число $C > 0$, что координаты любой точки $p = (x_1, x_2,\ldots, x_n)$ из X по модулю не превосходят C:

$$|x_1| \le C, |x_2| \le C, \ldots, |x_n| \le C.$$

2. Сходимость точек в $\mathbb{R}^n$

Определение 1. *Пусть $\{p_n\}$ – последовательность точек в $\mathbb{R}^n$. Мы говорим, что эта последовательность* ***сходится к точке*** *p_0, если числовая последовательность $\{\rho(p_n, p_0)\}$ имеет предел 0.*

Иначе говоря, последовательность $\{p_n\}$ сходится к p_0, если расстояние $\rho(p_n, p_0)$ неограниченно уменьшается с ростом n.

Можно дать и другое определение сходящейся последовательности – не через расстояние между точками, а через координаты точек. Для сокращения записей дадим это определение для $n = 2$.

Определение 2. *Пусть $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$, … – последовательность точек в $\mathbb{R}^2$. Мы скажем, что эта последовательность* ***сходится*** *к точке $p_0 = (x_0, y_0)$, если числовая последовательность $x_1, x_2, \ldots$ сходится к числу x_0, а числовая последовательность $y_1, y_2, \ldots$ – к числу y_0.*

Чтобы установить эквивалентность этих определений, заметим сначала, что справедливы следующие неравенства: если $p = (x, y)$ и $p' = (x', y')$, то

$$\rho(p, p') \le |x - x'| + |y - y'|, \tag{1.24}$$

$$|x - x'| \le \rho(p, p'),\ |y - y'| \le \rho(p, p'). \tag{1.25}$$

Справедливость каждого из этих неравенств устанавливается возведением левой и правой частей в квадрат.

Пусть теперь последовательность $\{p_n\}$ сходится к p_0 в смысле определения 1. Это означает, что последовательность $\{\rho(p_n, p_0)\}$ стремится к нулю. Записав для точек p_n, p_0 неравенства (1.25), получим

$$0 \le |x_n - x_0| \le \rho(p_n, p_0), \quad 0 \le |y_n - y_0| \le \rho(p_n, p_0),$$

откуда в силу известных свойств пределов вытекает, что каждая из последовательностей $\{|x_n - x_0|\}$ и $\{|y_n - y_0|\}$ стремится к нулю. Следовательно, $x_n \to x_0$ и $y_n \to y_0$, т.е. последовательность $\{p_n\}$ сходится к p_0 в смысле определения 2. Обратно, пусть $\{p_n\}$ сходится к p_0 в смысле определения 2. На основании неравенства (1.25), примененного к точкам p_n, p_0, получим

$$\rho(p_n, p_0) \le |x_n - x_0| + |y_n - y_0|. \tag{1.26}$$

Каждая из последовательностей $\{|x_n - x_0|\}$, $\{|y_n - y_0|\}$ по условию стремится к 0; тем самым последовательность $\{|x_n - x_0| + |y_n - y_0|\}$ также стремится к 0. Отсюда и из (1.26) вытекает, что последовательность $\{\rho(p_n, p_0)\}$ стремится к 0, т.е. что последовательность $\{p_n\}$ сходится к точке p_0 в смысле определения 1.

Установим важную для дальнейшего лемму о сходимости. Предварительно введем такое определение.

Пусть

$$p_1, p_2, \ldots \tag{1.27}$$

– последовательность каких-то элементов, а n_1, n_2, … – возрастающая последовательность натуральных чисел; тогда последовательность $p_{n_1}, p_{n_2}, \ldots$ называется *подпоследовательностью* последовательности (1.27).

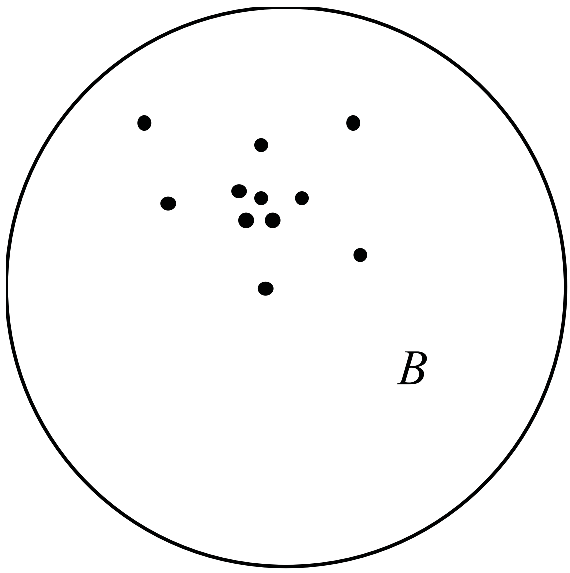

Рис. 1.19

Проще говоря, подпоследовательность – это любая бесконечная часть данной последовательности.

Лемма. *Всякая ограниченная последовательность точек в пространстве $\mathbb{R}^n$ содержит подпоследовательность, сходящуюся к некоторой точке из $\mathbb{R}^n$.*

Заметим, что с геометрической точки зрения содержание леммы в достаточной мере очевидно. Если последовательность точек на плоскости (в $\mathbb{R}^2$) ограничена, т.е. заключена внутри некоторого круга B, то в нем ввиду бесконечности этой последовательности обязательно должны найтись места "сгущения" данной последовательности, т.е. должны существовать подпоследовательности, сходящиеся к некоторым точкам круга (рис. 1.19).

Доказательство леммы дается в приложении 2 к данной главе.

3. Открытые и замкнутые множества в $\mathbb{R}^n$

Определение. *Пусть X – множество в пространстве* $\mathbb{R}^n$. *Точка p называется*:

внутренней *точкой множества X, если она содержится в X вместе с некоторой своей ε-окрестностью*;

внешней *точкой по отношению к X, если она является внутренней для дополнения X в* $\mathbb{R}^n$;

граничной *точкой для X, если она не является ни внутренней, ни внешней точкой для X, иначе говоря, если любая ее окрестность содержит как точки, принадлежащие X, так и точки, не принадлежащие X* (рис. 1.20).

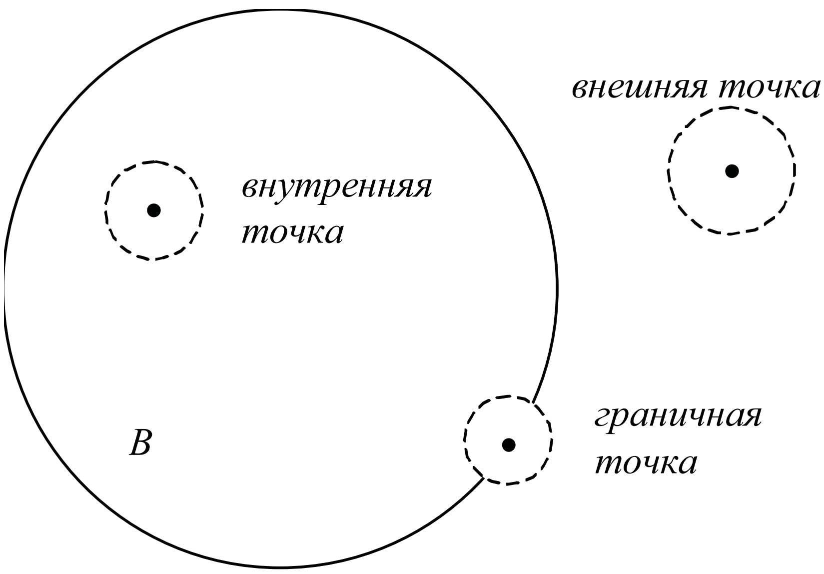

Рис. 1.20

Из данного определения следует, что любая точка $\mathbb{R}^n$. является либо внутренней точкой для множества *X*, либо внешней, либо граничной по отношению к *X*. При этом любая внутренняя точка принадлежит множеству X, любая внешняя точка не принадлежит X; что же касается граничных точек, то они могут как принадлежать, так и не принадлежать X.

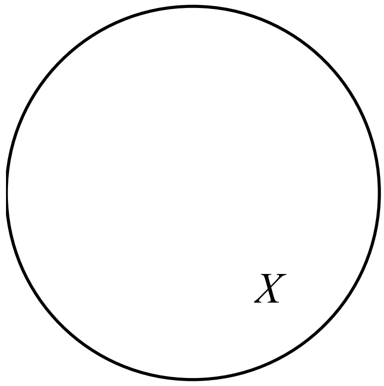

Рис. 1.21

На рис. 1.21 множество X представляет собой круг на плоскости (в $\mathbb{R}^2$) вместе с граничной окружностью; граничными точками множества X являются точки этой окружности. Если в качестве X взять круг без граничной окружности, множество граничных точек останется тем же.

Определение. *Множество X называется* ***открытым****, если все его точки внутренние. Множество X называется* ***замкнутым****, если оно содержит все свои граничные точки.*

Упомянутый выше круг X (см. рис. 1.21) без граничной окружности является открытым множеством; тот же самый круг вместе с граничной окружностью – замкнутое множество.

4. Предельные точки множества. Изолированные точки

Примем следующее определение.

Определение. *Пусть X – множество в $\mathbb{R}^n$. Точка p_0 называется* ***предельной*** *для X, если в любой ε-окрестности точки p_0 имеются точки множества X, отличные от p_0.*

При этом сама точка p_0 может как принадлежать, так и не принадлежать множеству X.

Очевидно, любая внутренняя точка множества X является предельной для X. Столь же очевидно, что других предельных точек для X, кроме внутренних и граничных, не существует.

Определение. *Точка* $p_0 \in X$ *называется* ***изолированной*** *точкой множества X, если у нее существует ε-окрестность, в которой никаких других точек из X, кроме* p_0, *нет.*

Ясно, что любая точка множества X является либо изолированной, либо предельной для X.

5. Предел и непрерывность функций нескольких переменных

Определение. *Пусть на множестве* $X \subset \mathbb{R}^n$ *задана функция* $f(p)$ *и пусть* p_0 *– предельная точка для X. Число a называется* ***пределом функции*** *f в точке* p_0, *если для любой сходящейся к* p_0 *последовательности* $\{p_n\}$, *где все* $p_n \neq p_0$, *соответствующая числовая последовательность* $\{f(p_n)\}$ *сходится к числу a.* Запись:

$$\lim_{p \to p_0} f(p) = a,$$

или в координатной форме:

$$\lim_{x_1 \to x_1^0, \ldots, x_n \to x_n^0} f(x_1, \ldots, x_n) = a.$$

Определение. *Функция* $f(p)$, *определенная на множестве* $X \subset \mathbb{R}^n$, *называется* ***непрерывной в точке*** $p_0 \in X$, *если*

$$\lim_{p \to p_0} f(p) = f(p_0),$$

или же, если p_0 *– изолированная точка множества X.*

Все свойства пределов, установленные ранее для функций одной переменной, остаются справедливыми для функций n переменных (доказательства повторяются дословно). Напомним эти свойства.

1. *Если* $\lim\limits_{p \to p_0} f(p) = a$ *и* $\lim\limits_{p \to p_0} g(p) = b$, *то*

$$\lim_{p \to p_0} (f(p) + g(p)) = a + b, \quad \lim_{p \to p_0} f(p) \cdot g(p) = a \cdot b, \quad \lim_{p \to p_0} \frac{f(p)}{g(p)} = \frac{a}{b}$$

(*последнее при* $b \neq 0$).

2. *Если существует* $\lim\limits_{p \to p_0} f(p)$, *то в некоторой ε-окрестности точки* p_0 *функция* $f(p)$ *ограничена.*

3. *Если в некоторой окрестности точки* p_0 *имеем* $f(p) \geq g(p)$ *и существуют* $\lim\limits_{p\to p_0} f(p) = a,\ \lim\limits_{p\to p_0} g(p) = b$, *то* $a \geq b$.

4. *Если в некоторой окрестности точки* p_0 *имеем*

$$f(p) \geq g(p) \geq h(p),$$

причем пределы $\lim\limits_{p\to p_0} f(p)$ *и* $\lim\limits_{p\to p_0} h(p)$ *существуют и равны одному и тому же числу* a, *то и* $\lim\limits_{p\to p_0} g(p) = a$.

Остается справедливой и теорема о непрерывности суммы, разности, произведения и частного двух непрерывных функций. В частности, для случая двух переменных отсюда следует, что любая функция вида

$$\frac{P(x, y)}{Q(x, y)},$$

где $P(x, y)$ и $Q(x, y)$ – два многочлена от x и y, непрерывна в любой точке (x, y), где $Q(x, y) \neq 0$.

Сохраняет силу и теорема о постоянстве знака непрерывной функции: если $f(p_0) \neq 0$, то в некоторой окрестности точки p_0 функция $f(p)$ имеет тот же знак, что и в точке p_0.

В заключение данного параграфа введем такое определение.

Определение. *Функция* $f(p)$, *определенная на множестве* $X \subset \mathbb{R}^n$, *называется* ***непрерывной на этом множестве***, *если она непрерывна в каждой точке множества* X.

Возвращаясь к введенному в § 1.13 классу элементарных функций от n переменных, отметим, что справедлива следующая теорема.

Теорема 1.10. *Любая элементарная функция непрерывна на всей области ее определения.*

Для функции одной переменной мы доказали такую теорему в § 1.11, сославшись на непрерывность основных элементарных функций. Для функций n переменных доказательство по существу остается тем же.

§ 1.15. Свойства непрерывных функций на ограниченных замкнутых множествах

При изучении числовых функций одной переменной x мы познакомились с рядом теорем о свойствах функций, непрерывных на отрезке. Наиболее важной среди них является теорема о том, что функция, непрерывная на отрезке, достигает в некоторой точке своего наибольшего и в некоторой точке – своего наименьшего значения. Обобщается ли эта теорема на функции нескольких переменных? И если да, то каким образом? Докажем в этом направлении следующие две теоремы.

Теорема 1.11. *Если непрерывная функция f от n переменных задана на ограниченном и замкнутом множестве $X \subset \mathbb{R}^n$, то она ограничена на этом множестве.*

Теорема 1.12. *Если непрерывная функция f от n переменных задана на ограниченном и замкнутом множестве $X \subset \mathbb{R}^n$, то существуют точка $p_0 \in X$, в которой f принимает свое наименьшее значение, и точка $q_0 \in X$, в которой f принимает свое наибольшее значение на X.*

Доказательство теоремы 1.11. Будем рассуждать от противного. Пусть функция f не ограничена на X. Тогда существует такая точка $p_1 \in X$, что $|f(p_1)| > 1$, затем такая точка $p_2 \in X$, что $|f(p_2)| > 2$ и т.д. Последовательность $p_1, p_2,\ldots$ ограничена (вслед за X), поэтому на основании леммы из § 1.14 существует подпоследовательность $p_{n_1}, p_{n_2},\ldots$, сходящаяся к некоторой точке p_0. В силу замкнутости множества X имеем $p_0 \in X$. Ввиду непрерывности функции f в точке p_0 числовая последовательность $f(p_{n_1}), f(p_{n_2}),\ldots$ должна сходиться к числу $f(p_0)$. Но это противоречит тому, что указанная последовательность неограничена: ведь по построению $|f(p_{n_i})| > n_i$ для любого номера i. Полученное противоречие доказывает теорему 1.11.

Доказательство теоремы 1.12 дословно повторяет доказательство соответствующей теоремы для функций одной переменной (см. свойство 4 непрерывных функций в § 1.10), и мы не будем приводить его повторно.

С помощью теоремы 1.12 удается в ряде случаев решить вопрос о достижении наименьшего и наибольшего значений для функции, заданной на *неограниченном* множестве. Пусть, например, функция $f(x, y)$ определена и непрерывна на всей плоскости $\mathbb{R}^2$. *Если можно указать такую область D, ограниченную и замкнутую, и такое число α, что значение f в любой точке, лежащей вне D, больше α, в то время как значение f хотя бы в одной точке области D не больше α, то наименьшее значение f в области D будет одновременно и наименьшим значением f на всей плоскости.* Действительно, пусть $f(x_0, y_0) = \min f$ в D, где $(x_0, y_0) \in D$. Так как по предположению в области D имеются точки, в которых $f \le \alpha$, то и $f(x_0, y_0) \le \alpha$. В то же время вне D имеем $f > \alpha$. Следовательно, $f(x_0, y_0) = \min f$ во всей плоскости.

Описанная ситуация имеет место, например, если $f(p) \to \infty$, когда точка p неограниченно удаляется от начала координат. Скажем, если

$$f = x^2 + y^2 - 2\sin x,$$

то ввиду $|\sin x| \le 1$ имеем, очевидно, $f \to \infty$ при $x^2 + y^2 \to \infty$. Взяв в качестве D круг радиусом 2 с центром в начале координат, получим всюду вне D $f(x, y) > 2$, в то же время $f(0, 0) = 0 < 2$. Значит, наименьшее значение f в круге D будет одновременно минимумом f во всей плоскости. Заметим, что методами дифференциального исчисления для функций нескольких переменных легко найти точку, в которой $f = f_{\min}$. Это точка $(x_0, 0)$, где x_0 есть корень (единственный) уравнения $x = \cos x$.

Аналогично, если можно указать область D (ограниченную и замкнутую) и число α так, что всюду вне D имеем $f < \alpha$, в то время как в некоторой точке из D имеем $f \ge \alpha$, то *наибольшее значение f в области D будет одновременно максимумом f во всей плоскости*. В частности, так будет, если $f \to 0$ при $x^2 + y^2 \to \infty$.

В качестве примера покажем, что среди прямоугольных параллелепипедов, имеющих заданную поверхность $2S$, существует параллелепипед наибольшего объема.

Обозначив через x, y, z стороны параллелепипеда, выходящие из одной вершины, будем иметь $xy + yz + xz = S$, откуда

$$z = \frac{S - xy}{x + y}.$$

Следовательно, объем параллелепипеда

$$V = xyz = \frac{xy\left(S - xy\right)}{x + y}.$$

По смыслу задачи функция $V(x, y)$ рассматривается в бесконечной области $D = \{x > 0, y > 0, xy < S\}$ плоскости Oxy. Эта область изображена на рис. 1.22.

Доопределим функцию V соглашением: $f = 0$ на границе области D; в результате получим функцию f, непрерывную (проверьте!) в каждой точке *замкнутой* области D.

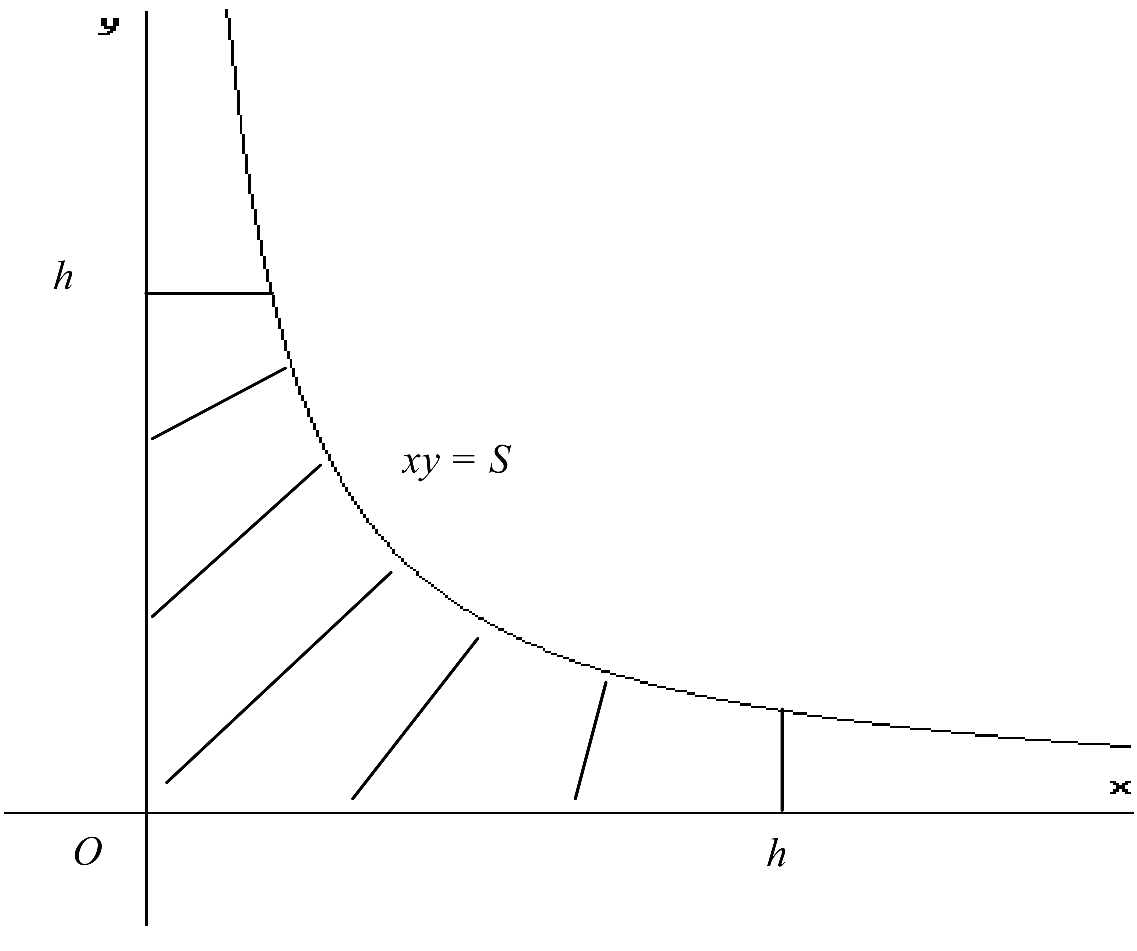

Рис. 1.22

На том же рис. 1.22 с помощью штриховки (наклонной) показана ограниченная область D_h: $x \le h$, $y \le h$, являющаяся частью D.

Поскольку справа от D_h мы имеем $y \le S/h,$ то можем записать (там же)

$$V = \frac{y}{1+y/x}(S - xy) \le \frac{S}{h} \cdot S = \frac{S^2}{h} \to 0 \text{ при } h \to \infty;$$

то же самое будет и сверху от D_h. Это показывает, что при достаточно большом h максимум V в области D_h будет одновременно и максимумом V во всей области D. Остается добавить, что ввиду ограниченности и замкнутости области D_h и непрерывности функции V максимум V в области D_h обязательно существует.

§ 1.16. Множества, заданные с помощью неравенств

Рассмотрим в пространстве $\mathbb{R}^n$ множество всех точек $(x_1, \ldots, x_n)$, удовлетворяющих неравенству

$$f(x_1, \ldots, x_n) \ge 0, \tag{1.28}$$

где $f(x_1, \ldots, x_n)$ – некоторая функция, определенная и непрерывная на всем $\mathbb{R}^n$. Обозначим это множество через X и покажем, что *множество X – замкнутое*.

Пусть $p_0 = (x_1^0, \ldots, x_n^0)$ – предельная точка множества X. Мы должны доказать, что $p_0 \in X$, т.е. что $f(x_1^0, \ldots, x_n^0) \ge 0.$ По условию имеем $p_0 = \lim\limits_{n\to\infty} p_n,$ где $p_1, p_2, \ldots$ – некоторая последовательность точек из X. Но тогда в силу непрерывности функции f должно быть $f(p_0) = \lim\limits_{n\to\infty} f(p_n)$. Учитывая, что $f(p_n) \ge 0$ для всех n (ведь $p_n \in X$), находим, что $f(p_0) \ge 0$, т.е. $p_0 \in X$, что и требовалось доказать.

Наше рассуждение остается в силе, если множество X задано не одним неравенством (1.28), а *системой* из нескольких неравенств

$$\begin{cases} f_1(x_1, \ldots, x_n) \ge 0, \\ \quad \ldots\ldots\ldots\ldots \\ f_s(x_1, \ldots, x_n) \ge 0, \end{cases} \tag{1.29}$$

где $f_1, \ldots, f_s$ – непрерывные функции на $\mathbb{R}^n$.

Таким образом, справедливо следующее предложение.

Множество $X \subset \mathbb{R}^n$, *заданное с помощью системы неравенств* (1.29), *где все функции* $f_1, \ldots, f_s$ *непрерывны*, *замкнуто в* $\mathbb{R}^n$.

Приложения к главе 1

Приложение 1

Докажем "неравенство треугольника" в $\mathbb{R}^n$ (см. (1.22)):

$$\rho(p, q) + \rho(q, r) \geq \rho(p, r).$$

Проверим сначала такое неравенство:

$$\sqrt{\sum_{i=1}^{n}(a_i + b_i)^2} \leq \sqrt{\sum_{i=1}^{n} a_i^2} + \sqrt{\sum_{i=1}^{n} b_i^2}, \tag{1.30}$$

где $a_1, \ldots, a_n, b_1, \ldots, b_n$ – какие угодно числа.

Взяв любое число x, запишем

$$\sum_{i=1}^{n}(a_i x + b)^2 = x^2\sum_{i=1}^{n} a_i^2 + 2x\sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} b_i^2 = \alpha x^2 + 2\beta x + \gamma, \tag{1.31}$$

где α, β, γ обозначают соответственно $\sum a_i^2, \sum a_i b_i, \sum b_i^2$. Очевидно, $\alpha \geq 0$, $\gamma \geq 0$. Квадратный трехчлен $\alpha x^2 + 2\beta x + \gamma$, как показывает левая часть равенства (1.31), неотрицателен при любом значении x. Следовательно, его дискриминант $\beta^2 - \alpha\gamma \leq 0$, откуда имеем $\beta^2 \leq$ $\leq \alpha\gamma$, или

$$\sum a_i b_i \leq \sum a_i^2 \cdot \sum b_i^2. \tag{1.32}$$

Но если возвести в квадрат обе части (1.30) и сократить слева и справа равные слагаемые, то как раз получим неравенство (1.32).

Опираясь на неравенство (1.30), докажем теперь "неравенство треугольника". Если в (1.30) положим

$$a_i = y_i - x_i,\ b_i = z_i - y_i, \text{ так что } a_i + b_i = z_i - x_i,$$

то придем к

$$\sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} + \sqrt{\sum_{i=1}^{n}(z_i - y_i)^2} \geq \sqrt{\sum_{i=1}^{n}(z_i - x_i)^2},$$

т.е. к "неравенству треугольника" для трех точек p, q, r в $\mathbb{R}^n$.

Приложение 2

Докажем лемму из п. 2 § 1.14.

Лемма. *Всякая ограниченная последовательность точек в пространстве $\mathbb{R}^n$ содержит подпоследовательность, сходящуюся к некоторой точке из $\mathbb{R}^n$.*

Доказательство. Начнем со случая $n = 1$. Данную нам ограниченную последовательность точек из $\mathbb{R}^n$ (т.е. чисел) обозначим

$$x_1, x_2, \ldots \tag{1.33}$$

По условию все числа x_i принадлежат некоторому отрезку $[a, b] = \Delta$. Разделим отрезок Δ пополам. Хотя бы одна из двух половин содержит подпоследовательность последовательности (1.33) (поскольку (1.33) – *бесконечная* последовательность). Обозначим эту половину Δ_1, а соответствующую подпоследовательность, содержащуюся в (1.33), через Π_1. Разобьем отрезок Δ_1 пополам и выберем ту из половин, которая содержит подпоследовательность последовательности Π_1; обозначим выбранную половину Δ_2, а соответствующую подпоследовательность – Π_2. Этот процесс можно продолжать до бесконечности. Получаемые при этом отрезки Δ_1, Δ_2, … образуют стягивающуюся систему, так как каждый последующий отрезок есть половина предыдущего. По известной теореме должна существовать точка x_0, принадлежащая всем этим отрезкам.

Пусть:

x_{n_1} – произвольная точка из Π_1;

x_{n_2} – произвольная точка из Π_2, причем такая, что $n_2 > n_1$ (существование такой точки вытекает из того, что подпоследовательность Π_1 бесконечная);

x_{n_3} – произвольная точка из Π_3, причем такая, что $n_3 > n_2$, и так далее. Поскольку $x_{n_1} \in \Delta_1, x_{n_2} \in \Delta_2, \ldots$, а отрезки Δ_1, Δ_2, … стягиваются к точке x_0, то последовательность $x_{n_1}, x_{n_2}, \ldots$ и есть искомая подпоследовательность в (1.33). Итак, для случая $n = 1$ лемма доказана.

Пусть теперь $n = 2$. Данную нам последовательность точек из $\mathbb{R}^n$ обозначим так:

$$p_1 = (x_1, y_1), p_2 = (x_2, y_2), \ldots \tag{1.34}$$

По условию эта последовательность ограничена, т.е. содержится в некотором круге. Тем самым она содержится и в некотором прямоугольнике вида $a \le x \le b$, $c \le y \le d$.

Рассмотрим последовательность x_1, x_2, … (первые координаты точек p_1, p_2, …). Она является ограниченной, так как вся содержится в $[a, b]$. По доказанному найдется подпоследовательность $x_{n_1}, x_{n_2}, \ldots$, сходящаяся к некоторому числу x_0. Рассмотрим тогда вторые координаты точек $p_{n_1}, p_{n_2}, \ldots$. Они образуют последовательность $y_{n_1}, y_{n_2}, \ldots$, которая также является ограниченной (вся содержится в $[c, d]$). Значит, существует подпоследовательность этой последовательности $y_{m_1}, y_{m_2}, \ldots$, которая сходится к некоторому числу y_0. Последовательность точек $p_{m_1}, p_{m_2}, \ldots$ является подпоследовательностью исходной последовательности (1.32). Первые координаты этих точек по построению сходятся к x_0, а вторые – к y_0. Значит, эта последовательность сходится к точке $p_0 = (x_0, y_0)$. Для случая $n = 2$ лемма доказана.

Далее таким же путем можно рассмотреть случаи $n = 3$, затем $n = 4$ и т.д.

Приложение 3

Докажем теорему о равномерной непрерывности (свойство 6 непрерывных функций из § 1.10).

Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$. Предположим, рассуждая от противного, что для некоторого $\varepsilon > 0$ не существует такого $\delta > 0$, что из $|x_1 - x_2| < \delta$ следует $|f(x_1) - f(x_2)| < \varepsilon$. Это означает, что каково бы ни было число $\delta > 0$, обязательно найдется пара точек x_1 и x_2, таких что $|x_1 - x_2| < \delta$, но $|f(x_1) - f(x_2)| \geq \varepsilon$. В частности, при любом натуральном n найдутся такие точки x_n и x'_n из $[a, b]$, что $|x_n - x'_n| < 1/n$, но $|f(x_n) - f(x'_n)| \geq \varepsilon$. Из ограниченной последовательности $\{x_n\}$, согласно лемме из приложения 2, можно выбрать сходящуюся подпоследовательность. Чтобы не усложнять обозначения, будем считать, что сама последовательность $\{x_n\}$ и является такой подпоследовательностью, т.е. что существует $\lim x_n = x_0$. Теперь из ограниченной последовательности $\{x'_n\}$ выберем сходящуюся подпоследовательность $\{x'_{n_i}\}$ и только такими парами x_{n_i}, x'_{n_i} и ограничимся. Короче говоря, будем считать, что существуют оба предела $\lim x_n = x_0$ и $\lim x'_n = x'$. Из того, что $|x_n - x'_n| < \dfrac{1}{n}$ для любого n, следует $x_0 = x_0'$. Таким образом, $\lim x_n = x_0$ и $\lim x'_n = x_0$. Но тогда из $|f(x_n) - f(x'_n)| \geq \varepsilon$ ввиду непрерывности функции $f(x)$ должно следовать $|f(x_0) - f(x_0)| \geq \varepsilon$, т.е. $0 \geq \varepsilon$, что невозможно.

Приложение 4

В данном приложении мы докажем критерий продуктивности матрицы Леонтьева из гл. 3 первой части данного учебника.

Теорема 1.13. *Если существует вектор* $\vec{x}^0 \geq 0$, *такой что выполняется неравенство* $\vec{x}^0 > A\vec{x}^0$, *то матрица Леонтьева A продуктивна.*

Доказательство. Пусть задан произвольный вектор спроса $\vec{y} \geq 0$. Так как по условию $\vec{z}^0 = \vec{x}^0 - A\vec{x}^0 > 0$, то найдется положительное число λ, такое что $\lambda\vec{z}^0 > \vec{y}$. Подставляя в это неравенство выражение для $\vec{z}^0$, получим

$$\lambda\vec{x}^0 - A(\lambda\vec{x}^0) > \vec{y}.$$

Итак, для вектора $\vec{x} = \lambda\vec{x}^0$ имеем

$$\vec{x} \ge A\vec{x} + \vec{y}. \tag{1.35}$$

Положим $\vec{x}^0 = \vec{y}$ и рассмотрим последовательность

$$\vec{x}^{i+1} = A\vec{x}^i + \vec{y} \qquad (i = 0, 1, 2, \ldots). \tag{1.36}$$

1. *Эта последовательность возрастающая*: $\vec{x}^{i+1} \ge \vec{x}^i$.
Доказательство. Рассуждаем по индукции:

$$\vec{x}^1 = A\vec{x}^0 + \vec{y} = A\vec{y} + \vec{y} \ge \vec{y} = \vec{x}^0,$$

$$\vec{x}^{i+1} = A\vec{x}^i + \vec{y} \ge A\vec{x}^{i-1} + \vec{y} = \vec{x}^i.$$

2. *Эта последовательность ограничена сверху вектором* $\vec{x}$.
Доказательство проведем также по индукции:

$$\vec{x} \ge A\vec{x} + \vec{y} \text{ (из (1.35))} \Rightarrow \vec{x} \ge \vec{y} = \vec{x}^0,$$

$$\vec{x} \ge A\vec{x} + \vec{y} \ge A\vec{x}^i + \vec{y} = \vec{x}^{i+1}$$

(из индуктивного предположения $\vec{x} \ge \vec{x}^i$ следует неравенство $A\vec{x} \ge A\vec{x}^i$). Отсюда вытекает существование предела последовательности $\lim\limits_{n\to\infty} \vec{x}^i = \tilde{x}$. Переходя к пределу в равенстве (1.36), получим $\tilde{x} = A\tilde{x} + \vec{y}$. Это показывает, что неотрицательный вектор $\tilde{x}$ является искомым решением уравнения Леонтьева

$$\vec{x} = A\vec{x} + \vec{y}$$

для заданного вектора конечного спроса $\vec{y}$. Теорема доказана полностью.

ГЛАВА 2

ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ

§ 2.1. Производная функции в точке

1. Определение производной

Пусть функция $y = f(x)$ определена в некоторой окрестности точки x_0. Для любой точки x из этой окрестности *приращение* Δx определяется формулой $\Delta x = x - x_0$, откуда $x = x_0 + \Delta x$. *Приращением функции* $y = f(x)$ *в точке* x_0 называется разность

$$\Delta y = f(x) - f(x_0) = f(x_0+\Delta x) - f(x_0).$$

Определение. ***Производной*** *от функции* $y = f(x)$ *в точке* x_0 *называется предел отношения* $\dfrac{\Delta y}{\Delta x}$, *когда* $\Delta x \to 0$ (*при условии, что предел отношения* $\dfrac{\Delta y}{\Delta x}$ *существует*).

Производная функции $y = f(x)$ в точке x_0 обозначается $y'(x_0)$ или $f'(x_0)$. Если необходимо указать, какая переменная является независимой (в данном случае x – независимая переменная), то используются обозначения:

$$y'_x, \quad \frac{dy}{dx}, \quad \left.\frac{dy}{dx}\right|_{x=x_0}.$$

Например, выражение $\left.\dfrac{d}{dx}\left(x+\dfrac{1}{x}\right)\right|_{x=2}$ следует понимать как производную функции $y = x + \dfrac{1}{x}$ в точке $x_0 = 2$.

Определение производной можно записать в виде формулы

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \qquad (2.1)$$

Конечно, предел (2.1) может не существовать. В этом случае говорят, что функция $f(x)$ не имеет производной в точке x_0. Если предел (2.1) равен $+\infty$ или $-\infty$, то говорят, что функция $f(x)$ имеет в точке x_0 *бесконечную* производную (равную $+\infty$ или $-\infty$ соответственно).

Пример 2.1. Рассмотрим функцию $f(x) = C = \text{const}$. Для любой точки x_0 приращение $\Delta y = 0$. Поэтому

$$C' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = 0.$$

Пример 2.2. Пусть $f(x) = x$. Для любой точки x_0 приращение функции $\Delta y = \Delta x$. Поэтому

$$x' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = 1.$$

Пример 2.3. Пусть $f(x) = \sqrt[3]{x}$. Для любой точки $x_0 \neq 0$ находим производную

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\sqrt[3]{x_0 + \Delta x} - \sqrt[3]{x_0}}{\Delta x} = \lim_{x \to x_0} \frac{\sqrt[3]{x} - \sqrt[3]{x_0}}{x - x_0} =$$

$$= \lim_{x \to x_0} \frac{1}{\sqrt[3]{x^2} + \sqrt[3]{x x_0} + \sqrt[3]{x_0^2}} = \frac{1}{3} x_0^{-\frac{2}{3}}.$$

Если же $x_0 = 0$, то

$$f'(0) = \lim_{\Delta x \to 0} \frac{\sqrt[3]{\Delta x}}{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{\sqrt[3]{(\Delta x)^2}} = +\infty.$$

2. Физический смысл производной

Для произвольной функции $f(x)$ рассмотрим динамическую систему: точка P движется по координатной прямой Oy таким образом, что в каждый момент времени x ее координата $y = f(x)$. Пусть A – точка координатной прямой Oy, в которой точка P находилась в момент времени x_0, B – точка, в которой P находилась в момент $x_0 + \Delta x > x_0$. Если точка P двигалась по направлению из $-\infty$ в $+\infty$, то $f(x_0 + \Delta x) > f(x_0)$ и расстояние от A до B равно

$$f(x_0+\Delta x) - f(x_0) = \Delta y.$$

Так как время, за которое точка P переместилась из A в B, равно Δx, то средняя скорость точки P на отрезке AB равна $v_{\text{ср.}} = \Delta y / \Delta x$. Мгновенная скорость v (или просто скорость) точки P в момент времени x_0 – это предел

$$v = \lim_{\Delta x \to 0} v_{\text{ср.}} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = f'(x_0).$$

Поэтому $f'(x_0)$ – это мгновенная скорость точки P.

Если же точка двигалась по направлению из $+\infty$ в $-\infty$, то $f(x_0 + \Delta x) < f(x_0)$, и расстояние от A до B равно $f(x_0) - f(x_0 + \Delta x) = -\Delta y$. В этом случае мгновенная скорость $v = \lim\limits_{\Delta x \to 0} \dfrac{-\Delta y}{\Delta x} = -f'(x_0)$. Откуда получаем $f'(x_0) = -v$. Таким образом, физический смысл производной в том, что $f'(x_0)$ – это мгновенная скорость (с учетом направления). В самых различных задачах (в том числе и экономических) производная функции $y = f(x)$ интерпретируется как скорость изменения величины y относительно x.

3. Геометрический смысл производной

Рассмотрим определение производной с геометрической точки зрения. Пусть Г – график функции $y = f(x)$; $A(x_0, f(x_0))$ и $B(x_0 + \Delta x, f(x_0 + \Delta x))$ – две точки на Г (рис. 2.1).

Прямая AB называется *секущей.* Будем считать, что $f(x)$ непрерывная функция. Тогда $f(x_0 + \Delta x) \to f(x_0)$ при $\Delta x \to 0$, т.е. точка B стремится к A, когда $\Delta x \to 0$.

Пусть $\varphi \in (-\pi/2, \pi/2)$ – угол наклона секущей AB относительно оси Ox. Если существует предел $\lim\limits_{\Delta x \to 0} \varphi = \varphi_0$, то прямая, проходящая через точку A и образующая с осью Ox угол φ_0, называется *касательной* к графику Г в точке A. Таким образом, касательная – это предельное положение, к которому стремится секущая AB, когда $\Delta x \to 0$.

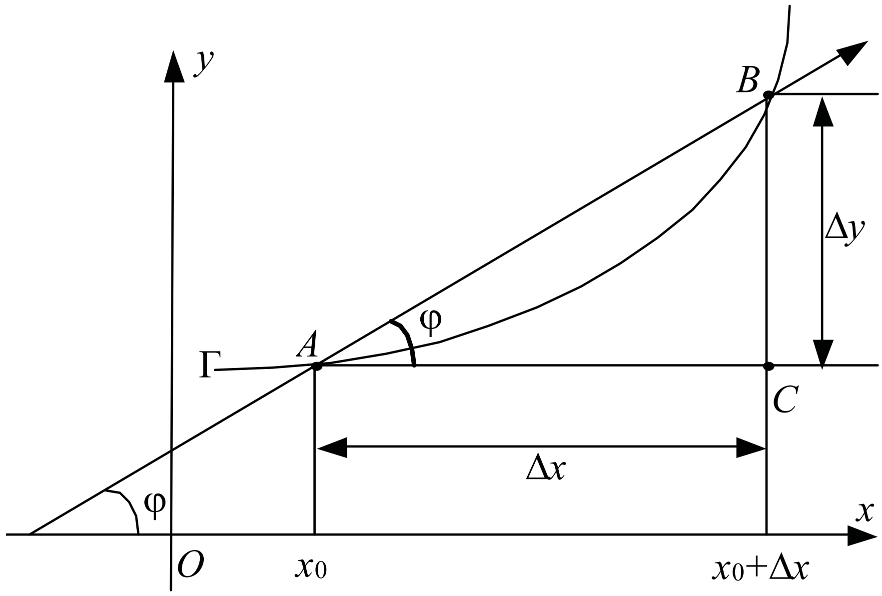

Рис. 2.1

Отметим, что вертикальная касательная может получиться в двух случаях:

$$\varphi_0 = \frac{\pi}{2} \text{ и } \varphi_0 = -\frac{\pi}{2}.$$

Пусть $C(x_0 + \Delta x, f(x_0))$ – точка, дополняющая отрезок AB до прямоугольного треугольника $ABC.$ Так как сторона AC этого треугольника параллельна оси Ox, имеем

$$\operatorname{tg}\varphi = \frac{\Delta y}{\Delta x}. \tag{2.2}$$

Переходя к пределу в левой и правой частях равенства (2.2) при $\Delta x \to 0$, получим

$$\operatorname{tg}\varphi_0 = f'(x_0)\ .$$

Поэтому геометрический смысл производной состоит в том, что $f'(x_0)$ – это тангенс угла наклона касательной к графику $y = f(x)$ в точке $(x_0, f(x_0))$.

4. Уравнение касательной

Найдем уравнение касательной к графику Γ функции $y = f(x)$ в точке $A(x_0, f(x_0))$. Будем искать это уравнение в виде $y = kx + b$. Так как $A \in \Gamma$, должно выполняться равенство $f(x_0) = kx_0 + b$, откуда $b = f(x_0) - kx_0$. Следовательно, касательная задается уравнением

$$y = kx + f(x_0) - kx_0 = f(x_0) + k(x - x_0).$$

Поскольку угловой коэффициент касательной $k = f'(x_0)$, уравнение касательной имеет вид

$$y = f(x_0) + f'(x_0)(x - x_0). \qquad (2.3)$$

П р и м е р 2.4. Найти уравнение касательной к графику функции $y = \sqrt[3]{x}$ в точке (1; 1).

Р е ш е н и е. Так как $\left(\sqrt[3]{x}\right)'_{x=1} = \left(\frac{1}{3}x^{-\frac{2}{3}}\right)_{x=1} = \frac{1}{3}$ (см. пример 2.3), уравнение касательной

$$y = 1 + \frac{1}{3}(x - 1) = \frac{2}{3} + \frac{1}{3}x.$$

5. Односторонние производные

Если в формуле (2.1) обычный предел заменить на односторонний, то получается определение односторонней производной. *Правой производной* от функции $f(x)$ в точке x_0 называется предел

$$f'_{+}(x_0) = \lim_{\Delta x \to 0+0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Левой производной от функции $f(x)$ в точке x_0 называется предел

$$f'_{-}(x_0) = \lim_{\Delta x \to 0-0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Из свойств пределов функций следует, что в случае, когда функция $f(x)$ имеет правую и левую производные в точке x_0 и они равны, производная $f'(x_0)$ существует и $f'(x_0) = f'_{+}(x_0) = f'_{-}(x_0)$. Если же $f'_{+}(x_0) \neq f'_{-}(x_0)$, то функция $f(x)$ не имеет производной в точке x_0.

Пример 2.5. Показать, что функция $f(x) = |x|$ не имеет производной в нуле.

Решение. Имеем:

$$f'_{+}(0) = \lim_{\Delta x \to 0+0} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to 0+0} \frac{\Delta x}{\Delta x} = 1,$$

$$f'_{-}(0) = \lim_{\Delta x \to 0-0} \frac{|\Delta x|}{\Delta x} = \lim_{\Delta x \to 0-0} \frac{-\Delta x}{\Delta x} = -1.$$

Так как $f'_{+}(0) \neq f'_{-}(0)$, функция $|x|$ не имеет производной в нуле.

Вообще, для функции $y=|f(x)|$ точки, где $f = 0$, будут, как правило, точками излома графика, и в них y' не существует.

§ 2.2. Дифференцируемость и непрерывность

Определение. *Функция $y=f(x)$ называется* ***дифференцируемой*** *в точке x_0, если ее приращение в точке x_0 можно представить в виде*

$$\Delta y = A\Delta x + \alpha(\Delta x)\Delta x, \tag{2.4}$$

где A – некоторое число, $\alpha(\Delta x)$ – функция от Δx, являющаяся бесконечно малой при $\Delta x \to 0$.

Следующая теорема устанавливает связь между дифференцируемостью функции в точке и существованием производной в этой же точке.

Теорема 2.1. *Для того чтобы функция $f(x)$ была дифференцируемой в точке x_0, необходимо и достаточно, чтобы она имела в этой точке конечную производную.*

Доказательство. *Необходимость.* Пусть функция $y = f(x)$ дифференцируема в точке x_0. Тогда ее приращение Δy можно представить в виде (2.4). Имеем

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\frac{A\Delta x + \alpha(\Delta x)\Delta x}{\Delta x} = \lim_{\Delta x\to 0}\left(A + \alpha(\Delta x)\right) = A.$$

Следовательно, производная $f'(x_0)$ существует и $f'(x_0) = A$.

Достаточность. Пусть существует конечная производная $f'(x_0) = A$. Тогда, по определению производной,

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = A.$$

Положим $\alpha(0) = 0$ и $\alpha(\Delta x) = \dfrac{\Delta y}{\Delta x} - A,$ если $\Delta x \neq 0$. Определенная так функция $\alpha(\Delta x)$ является бесконечно малой при $\Delta x \to 0$. Действительно,

$$\lim_{\Delta x\to 0}\alpha(\Delta x) = \lim_{\Delta x\to 0}\left(\frac{\Delta y}{\Delta x} - A\right) = A - A = 0.$$

Кроме того, $\Delta y = A\Delta x + \alpha(\Delta x)\Delta x$. Тем самым доказано, что функция $f(x)$ дифференцируема в точке x_0.

Замечание. Если функция $y = f(x)$ дифференцируема в точке x_0, то из равенства (2.4) следует, что $\Delta y \to 0$, когда $\Delta x \to 0$, т.е. *функция* $f(x)$ *непрерывна* в x_0. Обратное утверждение неверно: функция может быть непрерывной в некоторой точке x_0, но не быть дифференцируемой в x_0.

В качестве примера рассмотрим функцию $y = |x|$. Она непрерывна в каждой точке. С другой стороны, в точке $x_0 = 0$ функция $y = |x|$ не имеет производной (см. пример 2.5) и по теореме 2.1 не является дифференцируемой.

Пусть функция $y = f(x)$ дифференцируема в точке x_0. Тогда Δy можно представить в виде (2.4). Как следует из доказательства теоремы 2.1, коэффициент A в формуле (2.4) совпадает с производной $f'(x_0)$. Запишем формулу (2.4) следующим образом:

$$f(x) = f(x_0) + f'(x_0)\Delta x + \alpha(\Delta x)\Delta x. \qquad (2.5)$$

Далее, уравнение касательной в точке x_0 (см. формулу (2.3)) эквивалентно уравнению

$$y = f(x_0) + f'(x_0)\Delta x. \qquad (2.6)$$

Сравнивая формулы (2.5) и (2.6), видим, что расстояние от точки $P(x, f(x))$ на графике функции до точки $Q(x, f(x_0) + f'(x_0)\Delta x)$ на касательной (рис. 2.2)

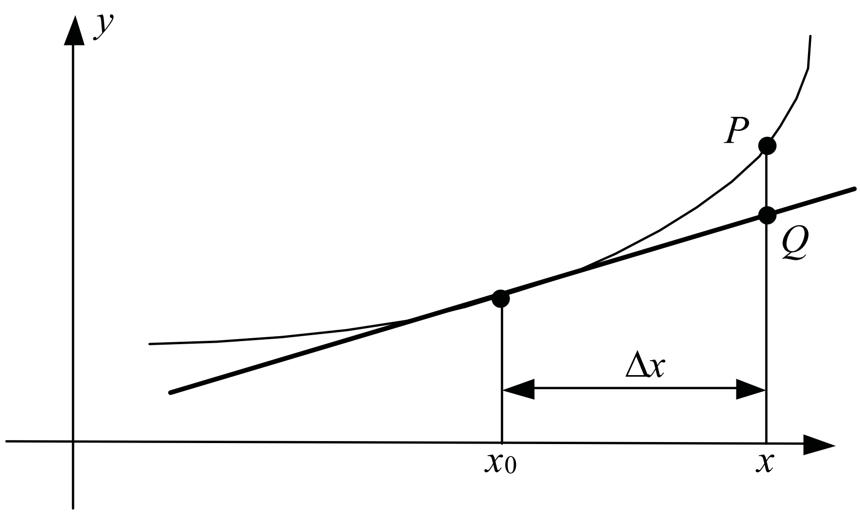

Рис. 2.2

равно $\alpha(\Delta x)\Delta x$, т.е. является бесконечно малой более высокого порядка, чем Δx, когда $\Delta x \to 0$.

Вывод: *геометрический смысл дифференцируемости функции $f(x)$ в точке x_0 состоит в том, что расстояние от точки на ее графике до соответствующей точки на касательной стремится к нулю "быстрее", чем Δx.*

§ 2.3. Правила дифференцирования

Процедура вычисления производной $f'(x_0)$ называется *дифференцированием* функции $f(x)$ в точке x_0. Теоремы этого параграфа устанавливают правила дифференцирования, которые сводят вычисление производных одних функций к вычислению производных других (более простых) функций. В § 2.4 мы увидим, что правила дифференцирования позволяют эффективно находить производные элементарных функций.

Теорема 2.2. *Если функции $u = f(x)$, $v = g(x)$ дифференцируемы в точке x_0, то сумма, разность, произведение и частное этих функций (частное при условии $g(x_0) \neq 0$) также дифференцируемы в точке x_0, и выполняются следующие формулы:*

$$\frac{d}{dx}\big(f(x) \pm g(x)\big)\Big|_{x=x_0} = f'(x_0) \pm g'(x_0),$$

$$\frac{d}{dx}\big(f(x)g(x)\big)\Big|_{x=x_0} = f'(x_0)g(x_0) + f(x_0)g'(x_0),$$

$$\frac{d}{dx}\big(f(x)/g(x)\big)\Big|_{x=x_0} = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{\big(g(x_0)\big)^2},$$

или, коротко:

$$(u \pm v)' = u' \pm v',$$

$$(uv)' = u'v + uv',$$

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}.$$

Д о к а з а т е л ь с т в о. Пусть приращения функций u, v, $u \pm v$, uv и $\dfrac{u}{v}$ вычисляются только в точке x_0, так что

$$\Delta u = f(x_0 + \Delta x) - f(x_0),\ \Delta v = g(x_0 + \Delta x) - g(x_0),$$
$$\Delta(u + v) = f(x_0 + \Delta x) + g(x_0 + \Delta x) - (f(x_0) + g(x_0))$$

и т.д. Нетрудно видеть, что приращение суммы $\Delta(u+v)$ равно сумме приращений $\Delta u + \Delta v$. Действительно,

$$\Delta(u+v) = (u+\Delta u) + (v+\Delta v) - (u+v) = \Delta u + \Delta v.$$

Аналогично

$$\Delta(u-v) = (u+\Delta u) - (v+\Delta v) - (u-v) = \Delta u - \Delta v.$$

Поэтому

$$(u \pm v)' = \lim_{\Delta x \to 0} \frac{\Delta(u \pm v)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} \pm \lim_{\Delta x \to 0} \frac{\Delta v}{\Delta x} = u' \pm v'.$$

Найдем приращение произведения:

$$\Delta(uv) = (u + \Delta u)(v + \Delta v) - uv = u\Delta v + v\Delta u + \Delta u\Delta v.$$

Из дифференцируемости функций u и v в точке x_0 следует их непрерывность в x_0 (см. § 2.2), поэтому $\Delta u \to 0$ и $\Delta v \to 0$, когда $\Delta x \to 0$. Следовательно,

$$(uv)' = \lim_{\Delta x \to 0} \frac{\Delta(uv)}{\Delta x} = \lim_{\Delta x \to 0} \frac{u\Delta v + v\Delta u + \Delta u\Delta v}{\Delta x} = u \lim_{\Delta x \to 0} \frac{\Delta v}{\Delta u} + v \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} +$$
$$+ \lim_{\Delta x \to 0} \Delta u \lim_{\Delta x \to 0} \frac{\Delta v}{\Delta x} = uv' + vu' + 0v' = u'v + uv'.$$

Найдем приращение частного:

$$\Delta\left(\frac{u}{v}\right)=\frac{u+\Delta u}{v+\Delta v}-\frac{u}{v}=\frac{v(u+\Delta u)-u(v+\Delta v)}{(v+\Delta v)v}=\frac{v\Delta u-u\Delta v}{(v+\Delta v)v}.$$

Отсюда получаем

$$\left(\frac{u}{v}\right)'=\lim_{\Delta x\to 0}\frac{\Delta\left(\frac{u}{v}\right)}{\Delta x}=\lim_{\Delta x\to 0}\frac{v\Delta u-u\Delta v}{(v+\Delta v)v\Delta x}=\lim_{\Delta x\to 0}\frac{v\frac{\Delta u}{\Delta x}-u\frac{\Delta v}{\Delta x}}{(v+\Delta v)v}=\frac{u'v-uv'}{v^2}.$$

Пример 2.6. Пусть C = const. Так как $x'=1$ и $C'=0$ (см. примеры 2.1 и 2.2), то $(x+C)'=1$.

Пример 2.7. $(x^2)'=(xx)'=x'x+xx'=2x.$

Пример 2.8. $\left(\frac{1}{x}\right)'=\frac{1'x-1x'}{x^2}=\frac{0x-1\cdot 1}{x^2}=-\frac{1}{x^2}.$

Теорема 2.3. *Если функция $y=f(x)$ имеет обратную функцию $x=g(y)$ и в точке x производная $f'(x_0)\neq 0$, то обратная функция $g(y)$ дифференцируема в точке $y_0=f(x_0)$ и*

$$g'(y_0)=\frac{1}{f(x_0)},$$

или

$$x'_y=\frac{1}{y'_x}.$$

Доказательство. Положим $a=f'(x_0)$. Тогда из дифференцируемости $f(x)$ в x_0 следует, что приращение $\Delta y=f(x_0+\Delta x)-f(x_0)$ можно представить в виде

$$\Delta y = a\Delta x + \alpha\Delta x = (a+\alpha)\Delta x,$$

где $\alpha = \alpha(\Delta x) \to 0$ при $\Delta x \to 0$. Так как $a \neq 0$, то отсюда следует, что $\Delta x \to 0$, когда $\Delta y \to 0$. Имеем

$$g'(y_0) = \lim_{\Delta y \to 0} \frac{g(y + \Delta y) - g(y_0)}{\Delta y} = \lim_{\Delta y \to 0} \frac{\Delta x}{\Delta y} = \lim_{\Delta x \to 0} \left(\frac{\Delta y}{\Delta x} \right)^{-1} = \frac{1}{f'(x_0)}.$$

Пример 2.9. Рассмотрим функцию $y = x^2$ на промежутке $[0; +\infty)$. Для нее существует обратная функция $x = \sqrt{y}$. В точке $x_0 \neq 0$ производная $\frac{dy}{dx} = 2x_0 \neq 0$. Следовательно, производная обратной функции $x = \sqrt{y}$ в точке $y_0 = x_0^2$ будет $\frac{dx}{dy} = \frac{1}{2x_0}$. Учитывая, что $x_0 = \sqrt{y_0}$, получим формулу

$$\frac{d}{dy}\left(\sqrt{y}\right)\Big|_{y=y_0} = \frac{1}{2\sqrt{y_0}}.$$

В частности, если $y_0 = 1$, то $\frac{d}{dy}\left(\sqrt{y}\right)\Big|_{y=1} = \frac{1}{2}$.

Теорема 2.4. *Если функция $y = f(x)$ дифференцируема в точке x_0, а функция $x = g(t)$ дифференцируема в точке t_0 и $g(t_0) = x_0$, то сложная функция $y = f(g(t))$ также дифференцируема в t_0 и выполняется следующее равенство*:

$$\frac{d}{dt} f\left(g(t)\right)\Big|_{t=t_0} = f'(x_0) \cdot g'(t_0),$$

или

$$y'_t = y'_x \cdot x'_t.$$

Доказательство. Функция $y = f(x)$ дифференцируема в точке x_0, поэтому ее приращение можно представить как

$$\Delta y = f'(x_0)\Delta x + \alpha(\Delta x)\Delta x,$$

где $\Delta x \to 0$ при $\Delta t \to 0$, поскольку функция $g(t)$ непрерывна (следствие дифференцируемости) в точке t_0. Так как $\alpha(\Delta x) \to 0$ при $\Delta x \to 0$, то $\alpha(\Delta x) \to 0$ и при $\Delta t \to 0$. Поэтому

$$\frac{d}{dt}f(g(t))\Big|_{t=t_0} = \lim_{\Delta t\to 0}\left(f'(x_0)\frac{\Delta x}{\Delta t} + \alpha(\Delta x)\frac{\Delta x}{\Delta t}\right) = f'(x_0)g'(t_0) + 0\cdot g'(t_0) =$$
$$= f'(x_0)g'(t_0).$$

Пример 2.10. Рассмотрим функции $y=\sqrt{x}$ и $x = t^2 + 1$. Для любого $t_0 \in \mathbb{R}$ выполняется неравенство $t_0^2 + 1 > 0$. Значит, функция $y=\sqrt{x}$ дифференцируема (пример 2.9) в каждой точке x_0 вида $x_0 = t_0^2 + 1$. Применяя теорему 2.4, получим

$$\frac{d}{dt}\sqrt{t^2+1}\Big|_{t=t_0} = \frac{d}{dx}\sqrt{x}\Big|_{x=t_0^2+1}\cdot\frac{d}{dt}\left(t^2+1\right)\Big|_{t=t_0} = \frac{1}{2\sqrt{t_0^2+1}}\cdot 2t_0 = \frac{t_0}{\sqrt{t_0^2+1}}.$$

Теперь мы можем изменить обозначения. После замены t_0 на x запись формулы становится более компактной:

$$\left(\sqrt{x^2+1}\right)' = \frac{x}{\sqrt{x^2+1}}.$$

§ 2.4. Производные элементарных функций

Пусть $f(x)$ – некоторая функция. Если x – это обозначение фиксированного числа, то запись $f'(x)$ – это тоже обозначение фиксированного числа. Если же x – это обозначение переменной, то запись $f'(x)$ – это обозначение *производной функции*. *Областью определения* производной функции $f'(x)$ является множество всех точек x_0, в которых функция $f(x)$ имеет конечную производную. Значением производной функции $f'(x)$ в точке x_0 является производная функции $f(x)$ в этой точке.

В этом параграфе мы докажем, что производные основных элементарных функций являются элементарными функциями. Отсюда и из теорем 2.2 и 2.4 следует, что *производная любой элементарной функции также является элементарной функцией*.

1. Производная логарифмической функции

Для функции $y = \log_a x$ имеем

$$\frac{\Delta y}{\Delta x} = \frac{\log_a (x + \Delta x) - \log_a x}{\Delta x} = \frac{\log_a \left(1 + \frac{\Delta x}{x}\right)}{\Delta x} = \frac{1}{x} \frac{\log_a \left(1 + \frac{\Delta x}{x}\right)}{\frac{\Delta x}{x}} =$$

$$= \frac{1}{x} \frac{\log_a (1 + t)}{t} = \frac{1}{x} \log_a (1 + t)^{\frac{1}{t}},$$

где $t = \frac{\Delta x}{x}$. Используя непрерывность функции $\log_a x$ в точке $x = e$ и первый замечательный предел, найдем производную логарифмической функции

$$\left(\log_a x\right)' = \frac{1}{x} \log_a \left(\lim_{t \to 0} (1 + t)^{\frac{1}{t}} \right) = \frac{1}{x} \log_a e = \frac{1}{x \ln a}.$$

Для $a = e$ получаем, в частности, формулу

$$\left(\ln x\right)' = \frac{1}{x}. \tag{2.7}$$

Для функции $\ln|x|$ также имеем

$$\left(\ln |x|\right)' = \frac{1}{x}. \tag{2.8}$$

Действительно, если $x > 0$, то $|x| = x$ и $(\ln |x|)' = (\ln x)' = \frac{1}{x}$. Если $x < 0$, то $|x| = -x$ и по теореме 2.4

$$(\ln |x|)' = (\ln(-x))' = \frac{1}{-x}(-x)' = \frac{1}{x}.$$

2. Производная показательной функции

Функция $y = a^x$ является обратной для функции $x = \log_a y$. По теореме 2.3

$$y'_x = \frac{1}{x'_y} = \frac{1}{\dfrac{1}{y\ln a}} = y\ln a.$$

Поскольку $y = a^x$, получаем $(a^x)' = a^x \ln a$. В частности, для $a = e$ имеем

$$(e^x)' = e^x. \tag{2.9}$$

3. Производная степенной функции

Функция $y = x^a$ при $x > 0$ может быть представлена в виде $x^a = e^{a\ln x}$. Используя теорему 2.4 и формулы (2.7), (2.9), найдем

$$\left(x^a\right)' = \left(e^{a\ln x}\right)' = e^{a\ln x}(a\ln x)' = x^a \cdot \frac{a}{x} = ax^{a-1}.$$

Если $x < 0$, то функцию можно представить иначе:

$$x^a = (-1)^a(-x)^a = (-1)^a e^{a\ln(-x)}.$$

Далее,

$$(x^a)' = (-1)^a e^{a\ln(-x)}[a\ln(-x)]' = (-1)^a(-x)^a \frac{a}{-x}(-1) = (-1)^{a-1}(-x)^{a-1}a = ax^{a-1}.$$

Пусть теперь $x = 0$. Выражение $a\cdot 0^{a-1}$ определено только тогда, когда $a > 1$. В этом случае

$$\left(x^a\right)'_{x=0} = \lim_{\Delta x\to 0}\frac{(\Delta x)^a}{\Delta x} = \lim_{\Delta x\to 0}(\Delta x)^{a-1} = 0.$$

Следовательно, $(x^a)'_{x=0} = (ax^{a-1})_{x=0} = 0$. Рассмотрев случаи $x > 0$, $x < 0$ и $x = 0$, мы пришли к следующему выводу: *производная степенной функции $y = x^a$ может быть найдена по формуле*

$$(x^a)' = ax^{a-1} \tag{2.10}$$

для любых a и x, для которых имеет смысл правая часть этой формулы.

Используя формулу (2.10), можно получить, например, такие равенства: $\left(\sqrt[3]{x^4}\right)' = \left(x^{4/3}\right)' = \frac{4}{3}\sqrt[3]{x}$, $\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$ и т.д.

4. Производные тригонометрических функций

С помощью формулы $\sin a - \sin b = 2\sin\frac{a-b}{2}\cos\frac{a+b}{2}$, первого замечательного предела и непрерывности функции $\cos x$ найдем

$$(\sin x)' = \lim_{\Delta x\to 0}\frac{\sin(x+\Delta x)-\sin x}{\Delta x} = \lim_{\Delta x\to 0}\frac{2\sin\frac{\Delta x}{2}\cos\left(x+\frac{\Delta x}{2}\right)}{\Delta x} =$$

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

Используя тригонометрическое тождество $\cos x = \sin\left(x+\frac{\pi}{2}\right)$ и правило дифференцирования сложной функции (теорему 2.4), получим

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

Далее, $(\operatorname{tg} x)' = \left(\dfrac{\sin x}{\cos x}\right)' = \dfrac{(\sin x)'\cos x - \sin x(\cos x)'}{\cos^2 x} = \dfrac{\cos^2 x + \sin^2 x}{\cos^2 x} =$

$= \dfrac{1}{\cos^2 x}$.

Итак, мы нашли производные функций $\sin x$, $\cos x$ и $\operatorname{tg} x$:

$$(\sin x)' = \cos x, \tag{2.11}$$

$$(\cos x)' = -\sin x, \tag{2.12}$$

$$(\operatorname{tg} x)' = \frac{1}{\cos^2 x}. \tag{2.13}$$

5. Производные обратных тригонометрических функций

Функция $y = \arcsin x$ является обратной для функции $x = \sin y$. Следовательно,

$$(\arcsin x)'_x = \frac{1}{(\sin y)'_y} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}}.$$

Аналогично функция $y = \arccos x$ является обратной для функции $x = \cos y$. Поэтому

$$(\arccos x)'_x = \frac{1}{(\cos y)'_y} = \frac{1}{-\sin y} = -\frac{1}{\sqrt{1-\cos^2 y}} = -\frac{1}{\sqrt{1-x^2}}.$$

Функция $y = \operatorname{arctg} x$ – обратная функция для функции $x = \operatorname{tg} y$. Откуда

$$(\operatorname{arctg} x)'_x = \frac{1}{(\operatorname{tg} y)'_y} = \frac{1}{\dfrac{1}{\cos^2 y}} = \frac{1}{1+\operatorname{tg}^2 y} = \frac{1}{1+x^2}.$$

В итоге имеем следующие формулы:

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}, \tag{2.14}$$

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}, \tag{2.15}$$

$$(\operatorname{arctg} x)' = \frac{1}{1+x^2}. \tag{2.16}$$

Пример 2.11. Рассмотрим функцию (многочлен от x)

$$P(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n.$$

Имеем $P'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \ldots + n a_n x^{n-1}$. Следовательно, функция $P(x)$ дифференцируема при любом $x \in \mathbb{R}$.

Пример 2.12. Найдем производную рациональной функции

$$f(x) = \frac{P(x)}{Q(x)} = \frac{a_0 + \ldots + a_n x^n}{b_0 + \ldots + b_m x^m}.$$

Имеем

$$f'(x) = \frac{P'(x)Q(x) - P(x)Q'(x)}{Q^2(x)}.$$

Функция $Q^2(x)$ обращается в нуль в тех же точках, что и функция $Q(x)$. Поэтому рациональная функция дифференцируема в каждой точке, в которой она определена.

Необходимо отметить, что правила дифференцирования (теоремы 2.2, 2.3 и 2.4) содержат лишь *достаточные* условия дифференцируемости. Может случиться так, что функция $f(x)$ дифференцируема в некоторой точке x_0, тогда как правила дифференцирования, примененные для вычисления $f'(x)$, приводят к неопределенному результату.

Пример 2.13. Пусть $f(x)=\sqrt{\sin x^4}$. Применяя правило дифференцирования сложной функции, получим

$$f'(x)=\frac{1}{2\sqrt{\sin x^4}}(\sin x^4)'=\frac{\cos x^4}{2\sqrt{\sin x^4}}\left(x^4\right)'=\frac{2x^3\cos x^4}{\sqrt{\sin x^4}}.$$

В точке $x_0 = 0$ функция $\dfrac{2x^3\cos x^4}{\sqrt{\sin x^4}}$ не определена. Однако из этого вовсе не следует, что $f'(0)$ не существует. Используя определение производной, убеждаемся в том, что

$$f'(0)=\lim_{\Delta x\to 0}\frac{f(\Delta x)-f(0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{\sqrt{\sin(\Delta x)^4}}{\Delta x}=$$

$$=\lim_{\Delta x\to 0}\Delta x\lim_{\Delta x\to 0}\sqrt{\frac{\sin(\Delta x)^4}{(\Delta x)^4}}=0\cdot 1=0.$$

§ 2.5. Дифференциал и приближенные вычисления

Определение. ***Дифференциалом*** *функции $f(x)$ в точке x_0 называется линейная функция приращения Δx вида $f'(x_0)\Delta x$.*

Дифференциал функции $y = f(x)$ обозначается dy или $df(x_0)$. Главное назначение дифференциала состоит в том, чтобы заменить приращение Δy, совершив при этом по возможности меньшую ошибку.

Напомним, что наличие конечной производной $f'(x_0)$ влечет дифференцируемость (см. § 2.2) функции $f(x)$ в точке x_0, т.е. возможность представить приращение функции Δy в виде

$$\Delta y=f'(x_0)\Delta x+\alpha(\Delta x)\Delta x, \qquad (2.17)$$

где $\alpha(\Delta x)\to 0$ при $\Delta x\to 0$. Из формулы (2.17) следует, что ошибка в приближенном равенстве

$$\Delta y\approx dy \qquad (2.18)$$

(равная $\alpha(\Delta x)\Delta x$) является бесконечно малой более высокого порядка, чем Δx, когда $\Delta x \to 0$. Равенство (2.18) часто используют в приближенных вычислениях.

Пример 2.14. Пусть r – радиус круга, $S = \pi \cdot r^2$ – его площадь. Используя (2.18), получим $\Delta S \approx dS = 2\pi r\Delta r$, т.е. приращение площади круга приближенно равняется произведению длины его окружности на приращение радиуса.

Формулу (2.18) можно записать в виде равенства

$$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x, \qquad (2.19)$$

которое также используется в приближенных вычислениях.

Если $x_0 = 0$, то $\Delta x = x$ и формула (2.19) приобретает вид:

$$f(x) \approx f(0) + f'(0)x. \qquad (2.20)$$

Используя равенство (2.20), следует помнить, что величина $|x|$ (так же, как и $|\Delta x|$ в (2.18), (2.19)) должна быть достаточно малой.

Пример 2.15. Пусть $f(x) = \sqrt{1+x}$. Так как $f'(0) = \dfrac{1}{2}$, то при достаточно малых x имеем $\sqrt{1+x} \approx 1 + \dfrac{1}{2}x$.

Если вместо $f(x) = \sqrt{1+x}$ взять функции $(1+x)^\alpha$, e^x, $\ln(1+x)$ или $\sin x$, то получатся (проверьте самостоятельно!) следующие приближенные равенства:

$$(1+x)^\alpha \approx 1 + \alpha x, \qquad (2.21)$$

$$e^x \approx 1 + x, \qquad (2.22)$$

$$\ln(1+x) \approx x, \qquad (2.23)$$

$$\sin x \approx x, \qquad (2.24)$$

в которых $|x| << 1$.

Пример 2.16. Пусть r – ставка банковского процента (проценты за год). Найдем количество лет T, в течение которых первоначальная сумма вклада увеличится в два раза. Так как за 1 год вклад увеличивается в $\left(1+\frac{r}{100}\right)$ раз, то за T лет вклад увеличится в $\left(1+\frac{r}{100}\right)^T$ раз. Фактически нам необходимо решить уравнение $\left(1+\frac{r}{100}\right)^T = 2$. Логарифмируя это уравнение, получим

$$T\ln\left(1+\frac{r}{100}\right)=\ln 2,$$

откуда $T=\dfrac{\ln 2}{\ln\left(1+\frac{r}{100}\right)}$. Используя приближенное равенство (2.23), найдем $T\approx\dfrac{100\ln 2}{r}$. Так как $\ln 2\approx 0{,}7$, то время удвоения вклада будет (“правило 70”)

$$T\approx\frac{70}{r}.$$

Если, например, процентная ставка – 10 % годовых, то время удвоения вклада составит приблизительно 7 лет.

Итак, мы убедились, что дифференциал можно использовать в приближенных вычислениях. Сделаем теперь несколько замечаний, касающихся вычисления самого дифференциала.

1. Дифференциал константы равен нулю:

$$dC=C'\Delta x=0.$$

2. Дифференциал независимой переменной x равен ее приращению:

$$dx=x'\Delta x=1\cdot\Delta x=\Delta x.$$

3. Правила дифференцирования суммы, разности, произведения и частного функций (теорема 2.2) влекут соответствующие правила вычисления дифференциала:

$$d(u \pm v) = du \pm dv, \tag{2.25}$$

$$d(uv) = (du)v + u(dv), \tag{2.26}$$

$$d\left(\frac{u}{v}\right) = \frac{(du)v - u(dv)}{v^2}. \tag{2.27}$$

Проверим, например, равенство (2.26). Действительно,

$$d(uv) = (uv)'dx = (u'v + uv')dx = (du)v + u(dv).$$

Пример 2.17. Найти $d\left(1 + \dfrac{\sin x}{x}\right)$.

Решение.

$$d\left(1 + \frac{\sin x}{x}\right) = d(1) + d\left(\frac{\sin x}{x}\right) = 0 + \frac{(d\sin x)x - \sin x dx}{x^2} =$$

$$= \frac{\cos x \cdot \Delta x \cdot x - \sin x \cdot \Delta x}{x^2} = \frac{x\cos x - \sin x}{x^2}\Delta x.$$

§ 2.6. Предельные величины в экономике

Теоретический анализ разнообразных явлений экономики использует ряд *предельных величин*. Перечислим лишь некоторые: предельные издержки, предельный доход, предельная производительность, предельная полезность, предельная склонность к потреблению и т.д. Все эти величины самым тесным образом связаны с понятием производной. В качестве характерного примера рассмотрим предельные издержки.

Пусть q – количество произведенной продукции, $C(q)$ – соответствующие данному выпуску издержки. Предельные издержки обозначаются MC и определяются как дополнительные издержки, связанные с производством еще одной единицы продукции. Другими словами, $MC = C(q + \Delta q) - C(q)$, где $\Delta q = 1$. Используя равенство $\Delta C \approx dC$, получим $MC = \Delta C \approx dC = C'(q) \cdot \Delta q = C'(q)$. Таким образом, данное выше определение MC, по существу не противоречит другому распространенному определению, согласно которому $MC = C'(q)$.

Пример 2.18. Пусть $C(q) = 1500q - 2q^2 + 0{,}002q^3$. Тогда дополнительные издержки, связанные с увеличением выпуска от q до $q+1$, составят $\Delta C = C(q + 1) - C(q)$, что приближенно равно $C'(q) = 1500 - 4q + 0{,}006q^2$. В табл. 2.1 даны значения ΔC и $C'(q)$ в точках $q = 100, 200, ..., 1000$.

Таблица 2.1

Q	C′	ΔC
100	1160	1158,6
200	940	939,2
300	840	839,8
400	860	860,4
500	1000	1001,0
600	1260	1261,6
700	1640	1642,2
800	2140	2142,8
900	2760	2763,4
1000	3500	3504,0

Общая схема введения предельных величин такова: пусть величина Y является функцией от величины X, тогда предельная величина MY (по X) определяется как отношение $\dfrac{\Delta Y}{\Delta X}$. Приращение ΔX в различных случаях задается по-разному. В одних случаях ΔX – это наиболее естественная единица измерения величины X, в других случаях ΔX – это разность между соседними значениями X в таблице, задающей функцию Y от X. В теоретических вопросах, однако, более удобным является определение MY, основанное на равенстве

$$MY = Y'_X .$$

Конечно, при таком определении приходится дополнительно предполагать, что Y является дифференцируемой функцией от X.

§ 2.7. Логарифмическая производная

Определение. ***Логарифмической производной*** *положительной функции $y = f(x)$ называется производная $(\ln y)'_x$.*

Так как $(\ln x)' = \dfrac{1}{x}$, то по правилу дифференцирования сложной функции получим следующее соотношение для логарифмической производной:

$$(\ln y)' = \frac{y'}{y}. \qquad (2.28)$$

Если производную y' рассматривать как скорость изменения функции y, то величину $\dfrac{y'}{y}$ естественно считать ее относительной скоростью изменения. Из формулы (2.28) становится ясным, почему логарифмическую производную $(\ln y)'$ называют также *относительной скоростью изменения функции y, или ее темпом роста*.

Для не обращающейся в нуль функции $y = f(x)$ логарифмическую производную определяют как $(\ln|y|)'$. Из (2.8) следует, что формулу (2.28) можно записать в более общем виде:

$$(\ln|y|)' = \frac{y'}{y}. \qquad (2.29)$$

С помощью логарифмической производной удобно вычислять обычную производную в тех случаях, когда логарифмирование упрощает вид функции. Такое вычисление основано на формуле

$$y' = y(\ln|y|)', \tag{2.30}$$

полученной из формулы (2.29) умножением на y.

Используя формулу (2.30), найдем производную функции вида $y = u^v$, где $u = u(x)$, $v = v(x)$ – дифференцируемые функции, производные которых известны:

$$y' = y(\ln y)' = u^v(v \ln u)' = u^v(v' \ln u + v(\ln u)') = u^v\left(v' \ln u + v\frac{u'}{u}\right).$$

Пример 2.19. Найти производную функции $y = \sqrt[3]{x}|x|^x$.

Решение. Применяя тождество (2.30), находим

$$y' = y(\ln|y|)' = y\left(\ln|x|^{x+\frac{1}{3}}\right)' = y\left[\left(x + \frac{1}{3}\right)\ln|x|\right]' =$$

$$= y\left[\ln|x| + \left(x + \frac{1}{3}\right)\frac{1}{x}\right] = \sqrt[3]{x}|x|^x\left(\ln|x| + 1 + \frac{1}{3x}\right).$$

Пусть $K = K(t)$ – приближенная величина вклада в момент времени t. Можно ли определить (приближенно) ставку банковского процента r по функции $K(t)$? Если проценты начисляются один раз за период времени Δt, то проценты за период составят $Kr\Delta t$ (мы считаем, что r – номинальная ставка в за год, Δt – доля года). Так как приращение вклада и проценты по вкладу – одно и то же, то $\Delta K = Kr\Delta t$. Отсюда

$$r = \frac{\Delta K}{K\Delta t}. \tag{2.31}$$

Предположим, что функция $K(t)$ имеет производную $K'(t)$. Тогда мы можем заменить в равенстве (2.31) приращение ΔK на дифференциал $dK = K' \Delta t$, в результате получим

$$r \approx \frac{K'\Delta t}{K\Delta t} = \frac{K'}{K} = (\ln K)'. \tag{2.32}$$

Вывод: *ставка банковского процента r совпадает с логарифмической производной от величины вклада.*

Пример 2.20. Пусть $K(t) = K_0(t+1)^{1,5}$, где t – число лет от открытия вклада, K_0 – величина вклада в начальный момент времени $t = 0$. Тогда мы можем определить, как изменялась ставка процента $r = r(t)$. Действительно,

$$r \approx \left(\ln K_0(t+1)^{1,5}\right)' = \left[\ln K_0 + 1{,}5\ln(t+1)\right]' = \frac{1{,}5}{t+1},$$

или, в процентах, $r \approx (t+1)^{-1}150\%$. Так, через два года после открытия вклада ставка была $r \approx 50\%$ годовых, через 5 лет ставка уменьшилась до 25% годовых и т.д. Отметим, что абсолютная скорость роста вклада при этом не убывала, а возрастала, поскольку $K' = 1{,}5K_0\sqrt{t+1}$.

Разобранные примеры – далеко не единственная область применения логарифмической производной. С ее помощью можно получить мгновенную оценку доходности какого-либо актива (если, конечно, известна его стоимость как функция от времени). Пусть, например, в обращение сроком на 1 год выпущен вексель. Предположим, что его рыночная стоимость меняется линейным образом от 75% на момент выпуска до 100% на момент погашения, т.е. $B(t) = \frac{t+3}{4}b_1$, где $0 \le t < 1$ – время в долях года, b_1 – конечная стоимость векселя. За первую неделю после выпуска относительный прирост его стоимости будет приблизительно

$$\frac{B\left(\frac{1}{52}\right) - B(0)}{B(0)} = \frac{\left(\frac{1}{52}+3\right)b_1 - 3b_1}{3b_1} = \frac{1}{156} = 0{,}641\%.$$

Таким образом, за первую неделю доходность векселя составит $52 \cdot 0{,}641 = 33{,}3\%$ годовых. За последнюю неделю относительный прирост стоимости будет

$$\frac{B(1) - B\left(\frac{51}{52}\right)}{B\left(\frac{51}{52}\right)} = \frac{(1+3) - \left(\frac{51}{52} + 3\right)}{\left(\frac{51}{52} + 3\right)} = \frac{1}{207} = 0{,}483\%,$$

что составляет $52 \cdot 0{,}483 = 25{,}1\%$ годовых. Нетрудно убедиться, что полученные результаты незначительно отличаются от мгновенных темпов роста стоимости векселя. Действительно,

$$\left(\ln B(t)\right)' = \left[\ln(t+3) + \ln\left(\frac{b_1}{4}\right)\right]' = \frac{1}{t+3}.$$

Таким образом, мгновенная доходность в момент $t = 0$ равна $\frac{1}{3}$, что составляет 33,3% годовых, и в момент $t = 1$ равна 1/4, что составляет 25% годовых.

Предположим, что преобладающая ставка процента r на денежном рынке равна $\frac{2}{7}$ ($\approx$ 29% годовых). Решая неравенство $\frac{1}{t+3} > \frac{2}{7}$, найдем, что при $t < \frac{1}{2}$, т.е. в первой половине года, вексель обеспечивает доходность больше r. Таким образом, владельцу векселя выгодно продать его через 6 месяцев после его выпуска. Возникает естественный вопрос: что побуждает покупателя приобрести вексель, доходность которого упала ниже r? Существует несколько мотивов покупки. Возможно, например, что покупатель ожидает в ближайшем будущем падение ставки r ниже доходности векселя. Другой мотив – покупатель стремится объявить банкротом эмитента векселя.

Рассмотренный пример с векселем является частным случаем более общей модели. Пусть $A(t)$ – стоимость некоторого актива A в момент времени t, r – доходность от вложения денег в другие активы. Считаем для простоты, что r не зависит от времени. Когда выгодно покупать или продавать актив A? Для ответа на данный вопрос найдем интервал времени, в течение которого мгновенная доходность актива A будет больше r. Так как мгновенная доходность A совпадает с темпом роста его стоимости, искомый интервал времени задается неравенством

$$(\ln A(t))' > r. \tag{2.33}$$

Если неравенство (2.33) задает интервал (t_1, t_2), то актив A следует купить в момент t_1 и продать в момент t_2. Множество решений неравенства (2.33) может иметь и более сложную структуру. Если, например, это множество является объединением двух интервалов $(-\infty, t_1) \cup (t_2, +\infty)$, то актив A выгодно продать в момент t_1 и снова купить в момент t_2.

Пример 2.21. Пусть $r = 10\%$ годовых, $A(t) = Ce^{\operatorname{arctg} t}, C - const.$ В какой момент времени выгоднее всего купить (продать) актив A?

Решение. Для функции $A(t) = Ce^{\operatorname{arctg} t}$ неравенство (2.33) преобразуется к виду

$$\frac{1}{1+t^2} > 0{,}1.$$

Отсюда $t^2 < 9$. Поэтому доходность актива более 0,1 на интервале (–3; 3). Следовательно, актив выгодно купить в момент $t = -3$ и продать в момент $t = 3$.

§ 2.8. Эластичность и ее свойства

Понятие эластичности было введено Аланом Маршаллом в связи с анализом функций спроса. По существу, это понятие является чисто математическим и может применяться при анализе любых дифференцируемых функций.

Определение. ***Эластичностью*** *функции $y = f(x)$ в точке x_0 называется следующий предел*:

$$E_{yx}(x_0) = \lim_{\Delta x \to 0}\left(\frac{\Delta y}{y} : \frac{\Delta x}{x}\right). \tag{2.34}$$

Говорят также, что $E_{yx}(x_0)$ – ***коэффициент эластичности*** *y по x.*

Если из контекста ясно, в какой точке определяется эластичность и какая переменная является независимой, то в обозначении эластичности могут опускаться отдельные символы. Мы будем часто использовать сокращенные обозначения E_y и E_{yx}.

Из определения эластичности вытекает, что при достаточно малых Δx выполняется приближенное равенство

$$\frac{\Delta y}{y} : \frac{\Delta x}{x} \approx E_y,$$

которое можно записать в виде

$$\frac{\Delta y}{y} \approx E_y \frac{\Delta x}{x}. \tag{2.35}$$

Вывод: *эластичность E_y – это коэффициент пропорциональности между относительными изменениями величин y и x.*

Если, например, x увеличится на один процент, то y увеличится (приближенно) на E_y процентов.

Далее потребуется несколько формул, выражающих эластичность через производные функции. Заметим, во-первых, что

$$E_{yx}(x_0) = \lim_{\Delta x \to 0} \left(\frac{f(x_0 + \Delta x) - f(x_0)}{f(x_0)} : \frac{\Delta x}{x_0} \right) =$$

$$= \frac{x_0}{f(x_0)} \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{x_0}{f(x_0)} f'(x_0),$$

или

$$E_y = \frac{x}{y} y'. \tag{2.36}$$

Таким образом, если $x_0 \neq 0$ и $f(x_0) \neq 0$, то для существования конечного предела (2.34) в точке x_0 необходимо и достаточно, чтобы существовала конечная производная $f'(x_0)$. Далее, представим отношение $\dfrac{y'}{y}$ как логарифмическую производную. Соответственно, формула (2.36) запишется тогда в виде

$$E_y = x(\ln y)'. \tag{2.37}$$

Наконец, заметим, что $x = \dfrac{1}{(1/x)} = \dfrac{1}{(\ln x)'}$, поэтому эластичность E_y совпадает с отношением логарифмических производных

$$E_y = \frac{(\ln y)'}{(\ln x)'}. \tag{2.38}$$

Рассмотрим теперь ряд свойств эластичности[1].

Свойство 1. *Эластичность в точке x_0 суммы $y = y_1 + ... + y_n$ положительных функций $y_i = f_i(x)$ (i=1, 2, ..., n) удовлетворяет соотношению*

$$E_{min} \le E_y \le E_{max}, \tag{2.39}$$

где E_{min} (E_{max}) – это минимальная (максимальная) эластичность в точке x_0 функций y_i.

Доказательство. Воспользуемся формулой (2.36):

$$E_y = \frac{xy'}{y} = \frac{xy_1' + ... + xy_n'}{y} = \frac{y_1E_1 + ... + y_nE_n}{y},$$

где E_i – эластичность y_i в точке x_0.

[1] К перечисленным ниже можно добавить еще такое свойство: эластичность (2.34) является безразмерной величиной и не зависит от выбора единиц измерения величин y и x. Это утверждение следует из безразмерности и независимости от выбора единиц измерения отношений $\dfrac{\Delta y}{y}$ и $\dfrac{\Delta x}{x}$.

Пусть $\alpha_i = \dfrac{y_i}{y}$ $(i = 1, ...,n)$ – доля y_i в сумме y. Тогда

$$E_y = \alpha_1 E_1 + ... + \alpha_n E_n. \tag{2.40}$$

Так как $y_i > 0$, все $\alpha_i>0$ и $\alpha_1 + ... + \alpha_n = 1$. Из определения E_{min} следует, что $E_i \geq E_{min}$. Используя (2.40), получим

$$E_y \geq \alpha_1 E_{min} + ... + \alpha_n E_{min} = E_{min}.$$

Далее, $E_i \leq E_{max}$, поэтому

$$E_y \leq \alpha_1 E_{max} + ... + \alpha_n E_{max} = E_{max}.$$

Свойство 2. *Эластичность произведения функций $u = u(x)$ и $v = v(x)$ в точке равна сумме эластичностей функций u и v в той же точке*:

$$E_{u\cdot v} = E_u + E_v. \tag{2.41}$$

Доказательство. Используя формулу (2.37) и свойства логарифмической функции, получим

$$E_{u\cdot v} = x(\ln uv)' = x(\ln u + \ln v)' = x(\ln u)' + x(\ln v)' = E_u + E_v.$$

Свойство 3. *Эластичность частного функций $u = u(x)$ и $v = v(x)$ в точке x_0 $(v(x_0) \neq 0)$ равна разности эластичностей функций* u и v *в той же точке*:

$$E_{u/v} = E_u - E_v. \tag{2.42}$$

Доказательство. Снова используем формулу (2.37) и свойства логарифмической функции:

$$E_{u/v} = x\left(\ln\frac{u}{v}\right)' = x(\ln u - \ln v)' = x(\ln u)' - x(\ln v)' = E_u - E_v.$$

Свойство 4. *Для функций $y = f(x)$ и $x = g(t)$ эластичность y по t в точке t_0 удовлетворяет следующему равенству*:

$$E_{yt}(t_0) = E_{yx}(g(t_0))E_{xt}(t_0). \tag{2.43}$$

Д о к а з а т е л ь с т в о. Применяя формулу (2.38) и правило дифференцирования сложной функции, получим

$$E_{yt}(t_0)=\frac{(\ln y)'_t}{(\ln t)'_t}=\frac{(\ln y)'_t}{(\ln x)'_t}\frac{(\ln x)'_t}{(\ln t)'_t}=\frac{(\ln y)'_x x'_t}{(\ln x)'_x x'_t}E_{xt}(t_0)=E_{yx}(g(t_0))E_{xt}(t_0).$$

С в о й с т в о 5. *Для функции* $y=f(x)$ *эластичность обратной функции* $x=g(y)$ *в точке* y_0 *удовлетворяет соотношению*

$$E_{xy}(y_0)=E_{yx}^{-1}(g(y_0)). \tag{2.44}$$

Д о к а з а т е л ь с т в о. Поскольку $g(y)$ – обратная функция, то выполняется тождество

$$f(g(y))=y.$$

Применяя формулу (2.43), получим

$$E_{yx}(g(y_0))E_{xy}(y_0)=E_{yy}(y_0).$$

Так как $E_{yy}(y_0)=\lim\limits_{\Delta y\to 0}\left(\dfrac{\Delta y}{y}:\dfrac{\Delta y}{y}\right)=1,$ то

$$E_{yx}(g(y_0))E_{xy}(y_0)=1,$$

откуда следует равенство (2.44).

П р и м е р 2.22. Пусть $y=C=\text{const}$. Используя (2.36) , получим

$$E_C=x\frac{C'}{C}=0.$$

П р и м е р 2.23. Найти эластичность функции $y=x+C$, где C – постоянная.

Р е ш е н и е. По формуле (2.36) получаем $E_y=x\dfrac{(x+C)'}{x+C}=\dfrac{x}{x+C}$.

Пример 2.24. Найти эластичность функции $y = x^a$.

Решение. Применяя формулу (2.37), находим эластичность

$$E_y = x\left(\ln x^a\right)' = x(a\ln x)' = xa\frac{1}{x} = a.$$

1. Геометрический смысл эластичности

Напомним геометрический смысл производной: $f'(x_0)$ – это тангенс угла наклона касательной к графику функции $y = f(x)$ в точке $C(x_0,\ y_0)$, $y_0 = f(x_0)$. Геометрический смысл эластичности функции $f(x)$ в точке x_0 связан с разбиением данной касательной на отрезки точками A, B и C , где $A(x_a,\ 0)$ – точка пересечения касательной с осью Ox, $B(0, y_b)$ – точка пересечения касательной с осью Oy (рис. 2.3).

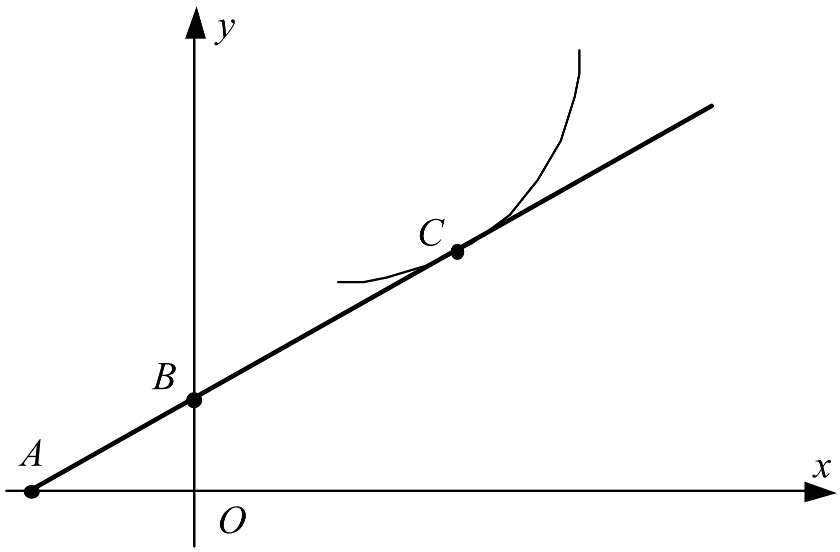

Рис. 2.3

Если эластичность y по x положительна, то она совпадает с отношением длин отрезков BC и AC:

$$E_{yx}(x_0) = \frac{BC}{AC}\ . \tag{2.45}$$

Если же эластичность y по x в точке x_0 отрицательна (рис. 2.4), то выполняется следующее соотношение:

$$E_{yx}(x_0) = -\frac{BC}{AC}. \tag{2.46}$$

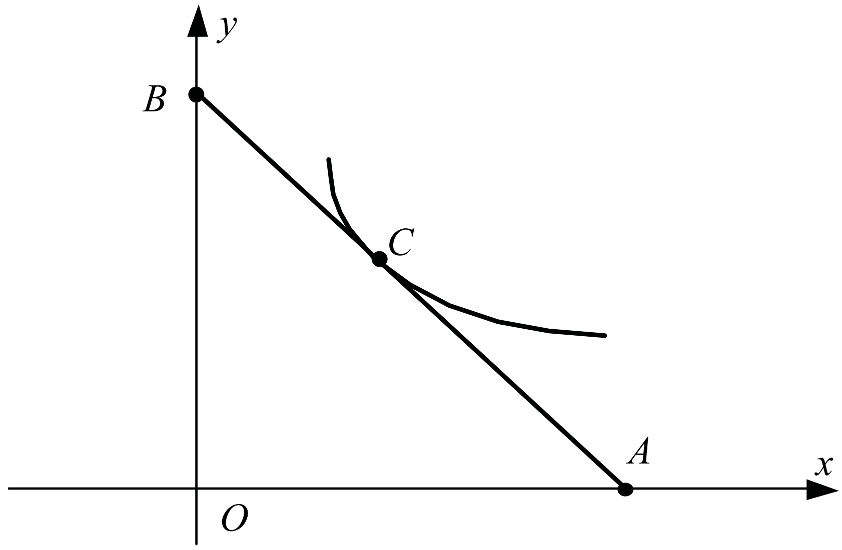

Рис. 2.4

С помощью векторов формулы (2.45) и (2.46) можно записать в виде одной формулы:

$$\overrightarrow{BC} = E_{yx}(x_0)\overrightarrow{AC}. \tag{2.47}$$

Докажем равенство (2.47). Уравнение касательной имеет вид

$$y = y_0 + f'(x_0)(x - x_0). \tag{2.48}$$

Подставив в уравнение (2.48) координаты точки $B(0, y_b)$, получим

$$y_b = y_0 - f'(x_0)x_0,$$

Если же подставить в уравнение (2.48) координаты точки $A(x_a, 0)$, то получим равенство

$$0 = y_0 + f'(x_0)(x_a - x_0),$$

из которого находим

$$x_a = x_0 - \frac{y_0}{f'(x_0)}.$$

Так как

$$\overrightarrow{AC} = (x_0 - x_a, y_0) = \left(\frac{y_0}{f'(x_0)}, y_0\right)$$

и

$$\overrightarrow{BC} = (x_0, y_0 - y_b) = (x_0, f'(x_0)x_0),$$

то

$$\overrightarrow{BC} = \frac{f'(x_0)x_0}{y_0}\overrightarrow{AC} = E_{yx}(x_0)\overrightarrow{AC}.$$

2. Ценовая эластичность спроса

Пусть $D = D(p)$ – спрос (в натуральных единицах) на некоторый товар при цене p. Так как при увеличении цены спрос уменьшается, то эластичность спроса $E_D < 0$. Спрос называется *эластичным*, если $|E_D| > 1$, и *неэластичным*, если $|E_D| < 1$. Термин *совершенно неэластичный* спрос означает крайний случай, когда изменение цены не приводит ни к какому изменению спроса. В этом случае $E_D = 0$. В другом крайнем случае, когда самое малое снижение цены побуждает покупателя увеличивать покупки от нуля до предела своих возможностей, говорят, что спрос является *совершенно эластичным*. Можно считать, что для совершенно эластичного спроса $|E_D| = \infty$.

Если спрос со стороны отдельных покупателей или групп покупателей является эластичным (неэластичным), то и суммарный спрос также является эластичным (неэластичным). Это утверждение следует из первого свойства эластичности. Действительно, пусть спрос $D_i(p)$ каждого покупателя ($i = 1, 2, \dots, n$) является эластичным. Тогда для эластичности E_{D_i} спроса D_i имеем $E_{D_i} < -1$.

Пусть $E_{max} = \max\{E_{D_1}, E_{D_2}, \dots, E_{D_n}\}$. Поскольку $E_{D_i} < -1$ для всех $i = 1, \dots, n$, то и $E_{max} < -1$. Применяя формулу (2.40), получим для эластичности суммарного спроса неравенство $E_D \le E_{max} < -1$. Следовательно, суммарный спрос также является эластичным. Для неэластичного спроса доказательство аналогично.

Если продавцы обладают достаточными запасами товара, то $D = D(p)$ – это не только количество спрашиваемого товара, но и одновременно количество проданного товара. В этом случае общая выручка всех продавцов $R = pD$. Находим эластичность выручки по цене:

$$E_R = \frac{R'}{R} p = \frac{D + pD'}{pD} p = 1 + \frac{D'}{D} p = 1 + E_D \, .$$

Следовательно, при эластичном спросе $E_R < 0$, а при неэластичном спросе $E_R > 0$.

Вывод: *Если спрос эластичен, то изменение цены вызывает изменение общей выручки в противоположном направлении. Если же спрос неэластичен, то изменение общей выручки происходит в том же направлении, что и изменение цены.*

Проводя графический анализ эластичности спроса, следует помнить, что в экономической теории принято ось цен изображать вертикально, несмотря на то, что спрос D рассматривается при этом как функция цены p. С учетом этого замечания, используя формулу (2.46) (или (2.47)), приходим к следующему.

Вывод: *эластичность (неэластичность) кривой спроса в точке C означает, что отрезок касательной, проведенной к этой кривой в точке C, с концами A и B на осях координат (рис. 2.5), делится точкой C так, что отрезок AC короче (соответственно длиннее) отрезка BC.*

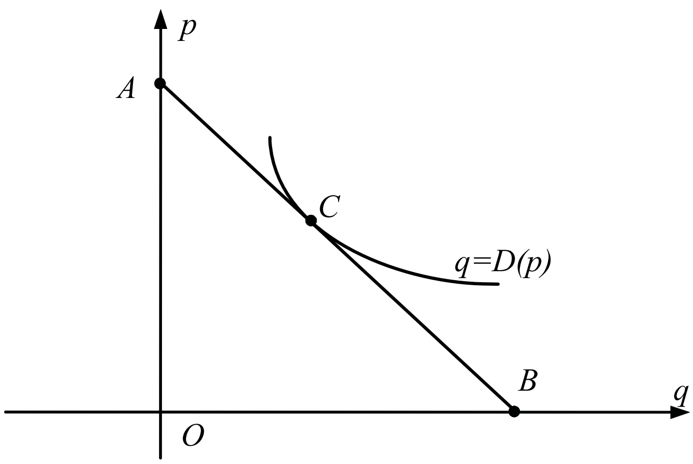

Рис. 2.5

Для линейной функции спроса $D(p) = kp + b(k < 0)$ касательная к графику совпадает с прямой спроса. В этом случае точки C, в которых спрос эластичен (неэластичен), расположены выше (соответственно ниже) середины M отрезка AB (рис. 2.6).

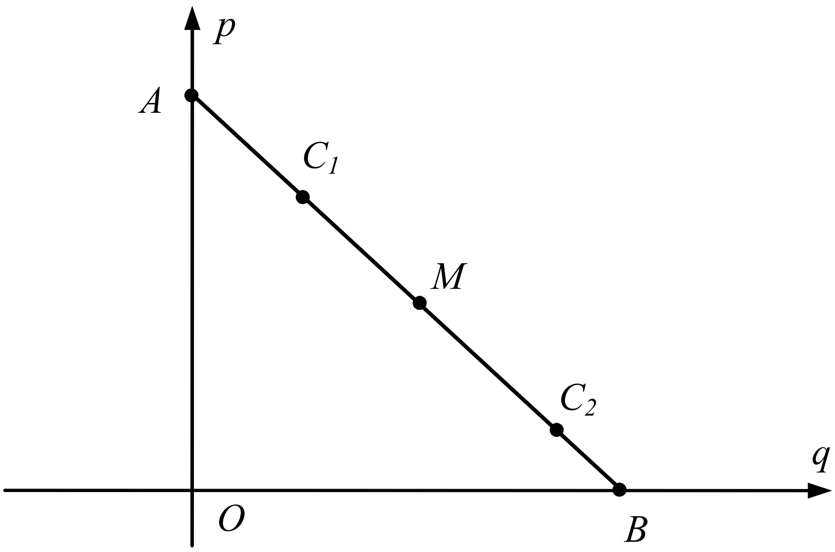

Рис. 2.6.

В точке C_1 спрос эластичен, а в точке C_2 спрос неэластичен.

§ 2.9. Распределение налогового бремени

Пусть p – цена товара на некотором рынке, $S(p)$ – его предложение при цене p, $D(p)$ – спрос. Равновесная цена p_0 определяется уравнением

$$S(p_0) = D(p_0).$$

Предположим, что вводится дополнительный налог с производителей в размере t с каждой единицы товара. Так как зависимость предложения от цены определяется прибылью, а при цене p и налоге t прибыль такая же, что и при цене $p - t$ и отсутствии дополнительного налога, то $S_t(p) = S(p - t)$, где $S_t(p)$ – функция предложения после введения налога. Таким образом, кривая предложения после введения налога сдвигается на t вверх (рис. 2.7).

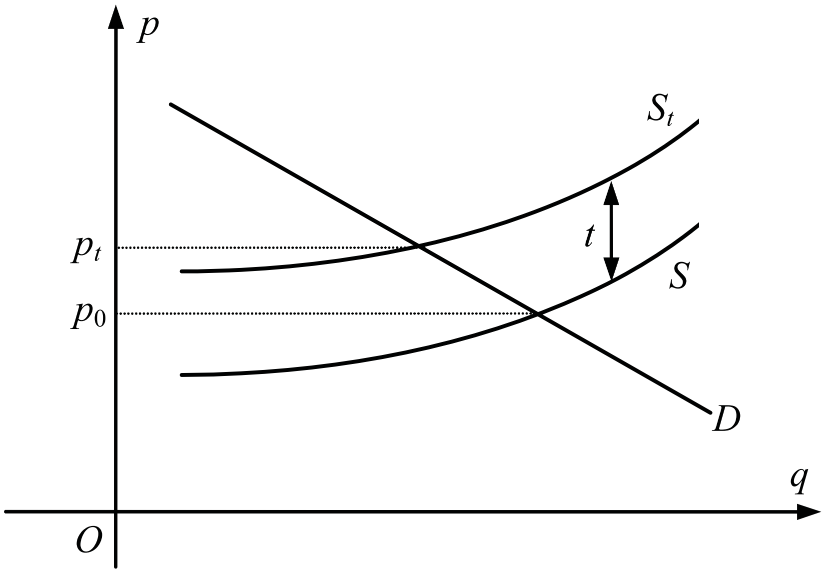

Рис. 2.7

Пусть p_t – новая равновесная цена. Равенство спроса и предложения при цене p_t выражается уравнением

$$S_t(p_t) = D(p_t),$$

которое эквивалентно уравнению

$$S(p_t - t) = D(p_t). \qquad (2.49)$$

Заменяя приращения функций $S(p)$ и $D(p)$ в точке p_0 на их дифференциалы, получим (приближенные) равенства:

$$S(p_0 + \Delta p) = S(p_0) + S'(p_0)\Delta p, \qquad (2.50)$$

$$D(p_0 + \Delta p) = D(p_0) + D'(p_0)\Delta p, \qquad (2.51)$$

в которых $\Delta p = p_t - p_0$ есть изменение равновесной цены. С учетом (2.50), (2.51) равенство (2.49) приобретает вид

$$S(p_0) + S'(p_0)(\Delta p - t) = D(p_0) + D'(p_0)\Delta p.$$

Так как $S(p_0) = D(p_0)$, то последнее равенство эквивалентно равенству

$$S'(p_0)(\Delta p - t) = D'(p_0)\Delta p.$$

Отсюда

$$\Delta p = \frac{tS'(p_0)}{S'(p_0) - D'(p_0)}. \tag{2.52}$$

Итак, после введения дополнительного налога затраты потребителя на покупку 1 ед. товара увеличатся на величину Δp, которую можно (приближенно) рассчитать по формуле (2.52). Соответственно доход производителя (также на 1 ед. продукции) уменьшится на

$$t - \Delta p = -\frac{tD'(p_0)}{S'(p_0) - D'(p_0)}.$$

Следовательно, налоговое бремя распределяется между потребителями и производителями продукции в отношении

$$\Delta p : [t - \Delta p] = S'(p_0) : [-D'(p_0)].$$

Поскольку в точке p_0 спрос равен предложению,

$$S'(p_0) : [-D'(p_0)] = \frac{S'(p_0)p_0}{S(p_0)} : \frac{-D'(p_0)p_0}{D(p_0)} = E_S : [-E_D],$$

где E_D, E_S – коэффициенты эластичности спроса и предложения в точке p_0.

Пример 2.25. Пусть ценовая эластичность спроса равна (–3), ценовая эластичность предложения равна 2, а вводимый налог $t = 100$. Тогда цена после введения этого налога увеличится на

$\frac{2}{2+3} \cdot 100 = 40,$ а прибыль производителя от одной единицы продукции уменьшится на 100 – 40 =60.

§ 2.10. Теоремы о промежуточных значениях

Теорема 2.5 (*теорема Ферма*). *Пусть функция $f(x)$ определена на интервале (a, b) и в некоторой точке x_0 этого интервала принимает наибольшее или наименьшее значение. Тогда возможны только два случая*:

1) *производная* $f'(x_0)$ *не существует*,

2) $f'(x_0) = 0.$

Доказательство. Если $f'(x_0)$ не существует, то доказывать нечего. Если $f'(x_0)$ существует, то существуют и равны обе односторонние производные в точке x_0:

$$f'_+(x_0) = f'_-(x_0).$$

Предположим, для определенности, что x_0 – точка максимума. Тогда для любой точки $x = x_0 + \Delta x$ из интервала (a, b) выполняется неравенство $\Delta y = f(x) - f(x_0) \le 0$. Следовательно,

$$\frac{\Delta y}{\Delta x} \le 0, \text{ если } \Delta x > 0, \text{ и } \frac{\Delta y}{\Delta x} \ge 0, \text{ если } \Delta x < 0.$$

Переходя к пределам, получим:

$$f'_+(x_0) = \lim_{\Delta x \to 0+0} \frac{\Delta y}{\Delta x} \le 0,$$

$$f'_-(x_0) = \lim_{\Delta x \to 0-0} \frac{\Delta y}{\Delta x} \ge 0.$$

Поскольку $f'_+(x_0) = f'_-(x_0) = f'(x_0),$ имеем $f'(x_0) = 0.$

Геометрический смысл теоремы Ферма состоит в том, что касательная к графику функции $y = f(x)$ в точке $(x_0, f(x_0))$ параллельна оси Ox, если x_0 – точка максимума или минимума функции $f(x)$ на интервале (a, b) (рис. 2.8 *а*, *б*).

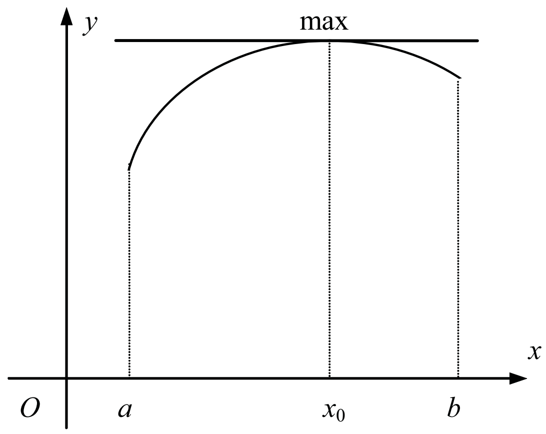

Рис. 2.8, *а*

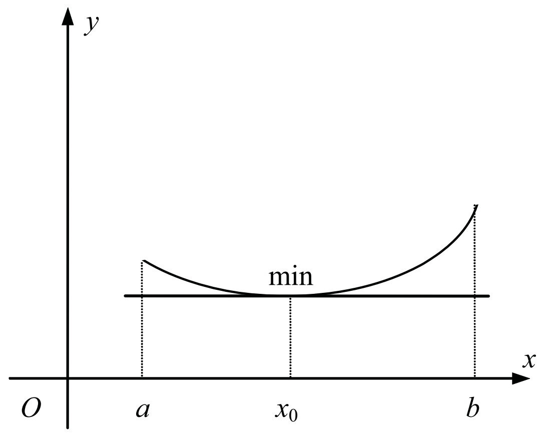

Рис. 2.8, *б*

Однако в точке максимума (минимума) x_0 производная $f'(x_0)$ может и не существовать (рис. 2.8 *в*, *г*).

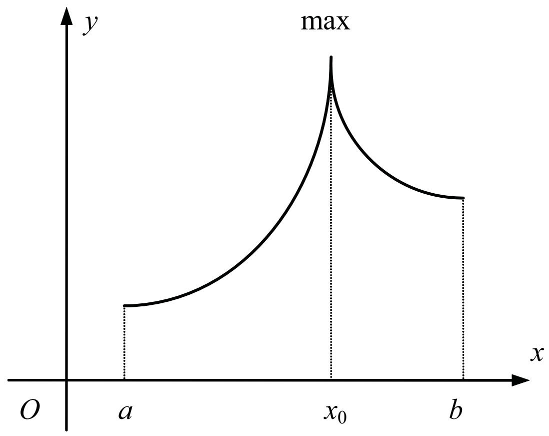

Рис. 2.8, *в*

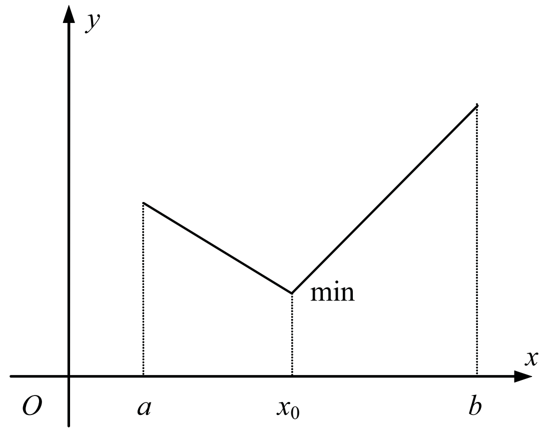

Рис. 2.8, *г*

Теорема 2.6 (*теорема Ролля*). *Пусть функция* $f(x)$:

1) *непрерывна на отрезке* $[a, b]$,

2) *дифференцируема на интервале* (a, b),

3) $f(a) = f(b)$.

Тогда существует точка $c \in (a, b)$, *в которой* $f'(c) = 0$.

Доказательство. Так как функция $f(x)$ непрерывна на отрезке $[a, b]$, то она принимает в некоторых точках x_{min} и x_{max} минимальное и максимальное значения: $m = f(x_{min})$, $M = f(x_{max})$. Если $m = M$, то $f(x) = m = \text{const}$ и в любой точке интервала (a, b) производная $f'(x)=0$. Поэтому мы можем считать, что $m \neq M$. Положим $c = x_{min}$, если $f(a) \neq m$, и $c = x_{max}$, если $f(a) = m$. При таком определении c имеем $c \neq a$. Поскольку $f(b) = f(a)$, то $f(b) \neq f(c)$, поэтому $c \neq b$. Итак, c – это точка максимума или минимума функции $f(x)$ и $c \in (a, b)$. По теореме Ферма $f'(c)=0$.

Теорема 2.7 (*теорема Лагранжа*). *Пусть функция $f(x)$ непрерывна на отрезке $[a, b]$ и дифференцируема на интервале (a, b). Тогда существует такая точка $c \in (a, b)$, что*

$$f(b)-f(a)=f'(c)(b-a). \tag{2.53}$$

Формула (2.53) называется *формулой Лагранжа*, или *формулой конечных приращений*.

Доказательство. Введем вспомогательную функцию

$$h(x)=(f(x)-f(a))(b-a)-(f(b)-f(a))(x-a).$$

Тогда:

1) $h(a) = h(b) = 0$;

2) $h(x)$ непрерывна (дифференцируема) в тех же точках, где непрерывна (дифференцируема) функция $f(x)$, т.е. $h(x)$ непрерывна на $[a, b]$ и дифференцируема в (a, b). По теореме Ролля существует точка $c \in (a, b)$, в которой $h'(c)=0$. Так как

$$h'(c)=f'(c)(b-a)-f(b)+f(a)=0,$$

то в точке c выполняется равенство (2.53).

Чтобы понять геометрический смысл теоремы Лагранжа, запишем формулу (2.53) в виде

$$\frac{f(b)-f(a)}{b-a}=f'(c). \tag{2.54}$$

Производная $f'(c)$ – это тангенс угла наклона касательной к графику функции $y=f(x)$ в точке $C(c, f(c))$, а отношение

$\dfrac{f(b)-f(a)}{b-a}$ – это тангенс угла наклона секущей, проходящей через точки $A(a, f(a))$ и $B(b, f(b))$ (рис. 2.9).

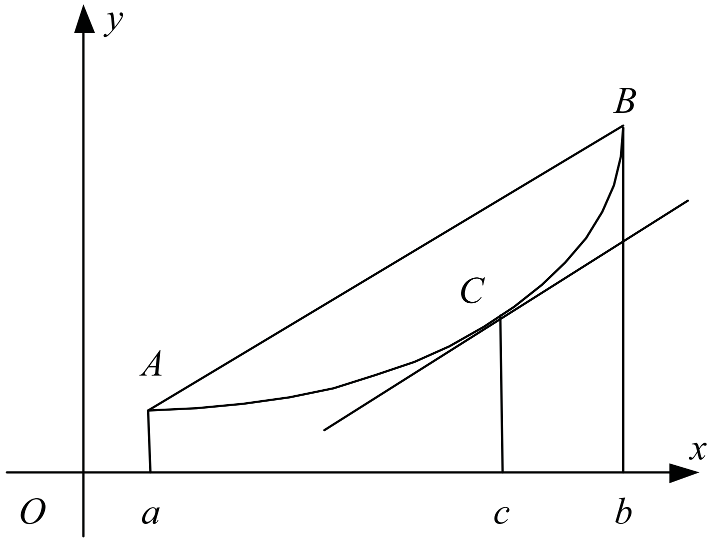

Рис. 2.9

Формула (2.54) означает, что на интервале (a, b) найдется точка c, для которой касательная к графику в точке $(c, f(c))$ параллельна секущей AB.

Теорема Лагранжа также допускает следующую экономическую интерпретацию. Пусть $f(x)$ – функция, выражающая зависимость издержек от объема производства x. Положим в (2.54) $a = 0$. Тогда левая часть (2.54) – средние (переменные) издержки, а правая часть – предельные издержки.

Вывод: *если функция издержек дифференцируема, то при любом объеме производства b средние переменные издержки совпадают с предельными издержками при некотором меньшем объеме производства c.*

Теорема 2.8 (*теорема Коши*). *Пусть функции $f(x)$ и $g(x)$ непрерывны на отрезке $[a, b]$ и дифференцируемы на интервале (a, b). Пусть, кроме того, $g'(x) \neq 0$ на (a, b). Тогда существует точка $c \in (a, b)$, такая, что*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}. \tag{2.55}$$

Д о к а з а т е л ь с т в о. Докажем сначала, что $g(b) \neq g(a)$. Действительно, если бы выполнялось равенство $g(b) = g(a)$, то по теореме Ролля нашлась бы точка $x_0 \in (a, b)$, в которой $g'(x_0) = 0$. А это противоречит условию теоремы.

Определим на $[a, b]$ вспомогательную функцию

$$h(x) = f(x) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Нетрудно убедиться в том, что $h(x)$ непрерывна на $[a, b]$, дифференцируема на (a, b) и $h(a) = h(b) = 0$. По теореме Ролля существует точка $c \in (a, b)$, такая, что $h'(c) = 0$. Поскольку

$$h'(c) = f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(c) = 0,$$

с учетом $g'(c) \neq 0$ получаем формулу (2.55).

§ 2.11. Правило Лопиталя

Т е о р е м а 2.9 (*п р а в и л о Л о п и т а л я*). *Пусть A – число, символ одностороннего предела* ($A = a \pm 0$) *или символ бесконечности* ($A = \pm\infty$). *Пусть функции $f(x)$ и $g(x)$ либо обе бесконечно малые, либо обе бесконечно большие при $x \to A$. Тогда если существует предел $\lim\limits_{x \to A} \frac{f'(x)}{g'(x)}$* (*конечный или бесконечный*), *то существует и предел $\lim\limits_{x \to A} \frac{f(x)}{g(x)}$, при этом выполняется равенство*

$$\lim_{x \to A} \frac{f(x)}{g(x)} = \lim_{x \to A} \frac{f'(x)}{g'(x)}.$$

Д о к а з а т е л ь с т в о этой теоремы дадим только в случае, когда $f(x)$, $g(x)$ – бесконечно малые функции и $A = a$ – число. Изменим, если это необходимо, определение функций $f(x)$ и $g(x)$ в точке

a так, чтобы значения этих функций в точке a были бы равны нулю: $f(x) = g(x) = 0$. Так как $\lim\limits_{x\to a} f(x) = 0 = f(a)$ и $\lim\limits_{x\to a} g(x) = 0 = g(a)$, то $f(x)$ и $g(x)$ непрерывны в точке a, и к этим функциям можно применить теорему Коши (см. § 2.10). Учитывая, что $f(a) = g(a) = 0$, получим

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c)}{g'(c)}$$

для некоторой точки c, расположенной между точками a и x. При $x \to a$ имеем $c \to a$ и, следовательно,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{c\to a} \frac{f'(c)}{g'(c)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Напомним, что предел $\lim\limits_{x\to A} \dfrac{f(x)}{g(x)}$ называется *неопределенностью* вида $\dfrac{0}{0}$ (или $\dfrac{\infty}{\infty}$), если $f(x) \to 0$, $g(x) \to 0$ (соответственно $|f(x)| \to +\infty, |g(x)| \to +\infty$), когда $x \to A$. Правило Лопиталя позволяет во многих случаях найти предел $\lim\limits_{x\to A} \dfrac{f(x)}{g(x)}$ вида $\dfrac{0}{0}, \dfrac{\infty}{\infty}$ или, как говорят, раскрыть неопределенность.

Пример 2.26. Найти $\lim\limits_{x\to\infty} \dfrac{\ln x}{\sqrt{x}}$.

Решение. Имеем неопределенность $\dfrac{\infty}{\infty}$. Применяя правило Лопиталя, получим

$$\lim_{x\to\infty} \frac{\ln x}{\sqrt{x}} = \lim_{x\to\infty} \frac{(\ln x)'}{\left(\sqrt{x}\right)'} = \lim_{x\to\infty} \frac{\dfrac{1}{x}}{\dfrac{1}{2\sqrt{x}}} = \lim_{x\to\infty} \frac{2\sqrt{x}}{x} = 0.$$

Пример 2.27. Найти $\lim\limits_{x\to\infty}\dfrac{x^n}{e^x}$, n – натуральное число.

Решение. Имеем неопределенность $\dfrac{\infty}{\infty}$. Применяя правило Лопиталя, найдем

$$\lim_{x\to\infty}\frac{x^n}{e^x}=\lim_{x\to\infty}\frac{\left(x^n\right)'}{\left(e^x\right)'}=\lim_{x\to\infty}\frac{nx^{n-1}}{e^x}.$$

Снова получили неопределенность $\dfrac{\infty}{\infty}$ (если $n-1>0$). Применим правило Лопиталя повторно:

$$\lim_{x\to\infty}\frac{nx^{n-1}}{e^x}=\lim_{x\to\infty}\frac{n(n-1)x^{n-2}}{e^x}.$$

Если $n-2>0$, то $x^{n-2}\to\infty$ и правило Лопиталя можно применить еще раз и т.д. В результате получаем цепочку равенств

$$\lim_{x\to\infty}\frac{x^n}{e^x}=\lim_{x\to\infty}\frac{nx^{n-1}}{e^x}=\lim_{x\to\infty}\frac{n(n-1)x^{n-2}}{e^x}=\ldots=\lim_{x\to\infty}\frac{n!}{e^x}=0.$$

Пример 2.28. Найти $\lim\limits_{x\to\infty}\left(e^x-x^n\right)$.

Решение. Имеем неопределенность $\infty-\infty$. После тождественного преобразования $e^x-x^n=e^x\left(1-\dfrac{x^n}{e^x}\right)$ получим

$$\lim_{x\to\infty}\left(e^x-x^n\right)=\lim_{x\to\infty}e^x\cdot\left(1-\lim_{x\to\infty}\frac{x^n}{e^x}\right)=+\infty\cdot(1-0)=+\infty,$$

где мы использовали равенство $\lim\limits_{x\to\infty}\dfrac{x^n}{e^x}=0$ (пример 2.27).

Пример 2.29. Найти $\lim\limits_{x\to\infty}\left(\sqrt{x^2+x}-\sqrt[3]{x^3+x^2}\right)$.

Решение. Имеем неопределенность $\infty-\infty$. Раскроем эту неопределенность, применяя правило Лопиталя после тождественного преобразования и замены переменной $y=\dfrac{1}{x}$:

$$\lim_{x\to\infty}\left(\sqrt{x^2+x}-\sqrt[3]{x^3+x^2}\right)=\lim_{x\to\infty}x\left(\sqrt{1+\frac{1}{x}}-\sqrt[3]{1+\frac{1}{x}}\right)=\lim_{y\to 0}\frac{\sqrt{1+y}-\sqrt[3]{1+y}}{y}=$$

$$=\lim_{y\to 0}\frac{\left(\sqrt{1+y}-\sqrt[3]{1+y}\right)'}{y'}=\lim_{y\to 0}\left[\frac{1}{2}(1+y)^{-\frac{1}{2}}-\frac{1}{3}(1+y)^{-\frac{2}{3}}\right]=\frac{1}{6}.$$

Пример 2.30. Найти $\lim\limits_{x\to 0+0}x\ln x$.

Решение. Имеем неопределенность вида $0\cdot\infty$. После преобразования $x\ln x=\dfrac{\ln x}{\left(\dfrac{1}{x}\right)}$ получаем неопределенность $\lim\limits_{x\to 0+0}\dfrac{\ln x}{\left(\dfrac{1}{x}\right)}$ вида $\dfrac{\infty}{\infty}$, которая легко раскрывается по правилу Лопиталя:

$$\lim_{x\to 0+0}\frac{\ln x}{\left(\frac{1}{x}\right)}=\lim_{x\to 0+0}\frac{(\ln x)'}{\left(\frac{1}{x}\right)'}=\lim_{x\to 0+0}\frac{\left(\frac{1}{x}\right)}{\left(-\frac{1}{x^2}\right)}=\lim_{x\to 0+0}(-x)=0.$$

В ряде случаев по правилу Лопиталя удается раскрыть неопределенности вида 0^0, 1^∞ и ∞^0. Для этого следует воспользоваться тождеством $u(x)^{v(x)}=e^{v(x)\ln u(x)}$, которое приводит указанные неопределенности к виду $0\cdot\infty$.

Пример 2.31. Найти $\lim\limits_{x\to 0+0}x^x$.

Решение. Имеем

$$\lim_{x\to 0+0}x^x=\lim_{x\to 0+0}e^{x\ln x}=e^{\lim\limits_{x\to 0+0}x\ln x}=e^0=1.$$

Здесь мы воспользовались тем, что $\lim\limits_{x \to 0+0} x \ln x = 0$ (см. пример 2.30), и непрерывностью функции $y = e^x$.

§ 2.12. Цены, предельные издержки и объем производства

Пусть q – выпуск продукции (в натуральных единицах); $R(q)$ – выручка от продаж; $C(q)$ – издержки производства, связанные с выпуском q ед. продукции. Тогда прибыль

$$\pi(q) = R(q) - C(q).$$

Предположим, что выполняются следующие условия.

1. Функции $R(q)$, $C(q)$ определены на полуинтервале $[0, +\infty)$ и дифференцируемы при $q > 0$.

2. Максимум прибыли достигается в некоторой точке $q^* \neq 0$.

Замечание. В случае, когда максимум прибыли $\pi(q^*)$ положителен ($\pi(q^*) > 0$), условие $q^* \neq 0$ естественным образом выполняется, поскольку $\pi(0) \leq 0$ (нет выпуска – нет выручки, нет выручки – нет прибыли).

Пусть условия 1 и 2 выполнены. Тогда функция

$$\pi(q) = R(q) - C(q)$$

дифференцируема и имеет на интервале $(0, +\infty)$ максимум в точке $q^* \neq 0$. По теореме Ферма $\pi'(q^*) = 0$. Так как $\pi'(q) = R'(q) - C'(q)$, то в точке $q = q^*$ получаем равенство

$$R'(q^*) = C'(q^*). \tag{2.56}$$

В экономической теории равенство (2.56) объясняется как правило, согласно которому фирма, максимизирующая свою прибыль, устанавливает объем производства таким образом, чтобы предельная выручка была равна предельным издержкам.

В случае, когда объем производства q не влияет на цену продукции p, имеем $R(q) = pq$ и $R'(q) = p$. Следовательно, равенство (2.56) принимает вид

$$p = C'(q^*). \tag{2.57}$$

Пример 2.32. Найти объем производства, если $p = 15$, $C(q) = q^3 + 3q$.

Решение. Прибыль при производстве q единиц продукции будет $\pi(q) = 15q - q^3 - 3q = q(12 - q^2)$. Очевидно, что $\pi(q) \geq 0$, если $q \in \left[0; \sqrt{12}\right]$. Так как $\pi(q)$ – непрерывная функция, то на отрезке $\left[0; \sqrt{12}\right]$ в некоторой точке q^* она принимает наибольшее значение на этом отрезке. Поскольку $\pi(q) \leq 0$ при $q > \sqrt{12}$, то $\pi(q^*)$ – наибольшее значение при любом $q \geq 0$. Используя равенство (2.57), получим

$$15 = C'(q^*) = 3(q^*)^2 + 3,$$

откуда $q^* = 2$. Так как фирма стремится получить максимальную прибыль, то она будет производить 2 ед. продукции.

Фактически мы выяснили, что при цене $p = 15$ фирма предложит на продажу 2 ед. продукции. Ясно, что немного усложнив рассуждения, мы могли бы найти функцию предложения $S(p)$ данной фирмы (см. в связи с этим § 2.15, где этот вопрос подробно изучается).

Рассмотрим теперь более общий случай, когда цена продукции p является дифференцируемой функцией $p = p(q)$ от объема выпуска q. Имеем

$$R'(q) = \left(p(q)q\right)' = p'(q)q + p(q) =$$

$$= \frac{p'(q)}{p(q)} q \cdot p(q) + p(q) = p(q)\left(E_{pq}(q) + 1\right).$$

Равенство (2.56) запишем следующим образом

$$p(q^*)\left(E_{pq}(q^*) + 1\right) = C'(q^*).$$

Получаем уравнение для цены:

$$p(q^*) = \frac{C'(q^*)}{E_{pq}(q^*) + 1}\,. \qquad (2.58)$$

Так как $E_{pq}(q) \le 0$, то из равенства (2.58) следует, что цена $p(q^*)$ не ниже предельных издержек $C'(q^*)$. В действительности, если фирма занимает существенную долю рынка, то увеличение ее выпуска приводит к насыщению рынка и падению цены. В этом случае $E_{pq}(q^*) < 0$ и из (2.58) следует, что цена $p(q^*)$ *больше* предельных издержек $C'(q^*)$.

Предположим теперь, что фирма является монополией. В этом случае при цене p она будет производить столько единиц продукции, сколько требуется покупателям при цене p, т.е. $q = D(p)$, где $D(p)$ – функция спроса. Таким образом, функция $D(p)$ будет обратной функцией для функции $p(q)$. Из свойств эластичности следует, что

$$E_{pq}(q^*) = E_{Dp}^{-1}(p(q^*)).$$

Пусть $p^* = p(q^*)$ цена, соответствующая выпуску q^*. Оптимальный выпуск q^* можно представить как $q^* = D(p^*)$ – спрос при наилучшей (для монополии) цене p^*. Уравнение (2.58) приобретает вид

$$p^* = \frac{C'\big(D(p^*)\big)}{E_{Dp}^{-1}\big(p^*\big) + 1}\,. \qquad (2.59)$$

Пусть, например, $E_D = -1{,}2$. Тогда $\left(E_D^{-1} + 1\right)^{-1} = \left(-\frac{5}{6} + 1\right)^{-1} = 6$, т.е. цена монополиста p^* в 6 раз (!) выше его предельных издержек.

При неэластичном спросе $|E_D| < 1$ знаменатель дроби в правой части равенства (2.59) $E_D^{-1} + 1 < 0$ и формула (2.59) не имеет смысла. Это значит, что сделанные ранее предположения невыполнимы. Фактически при неэластичном спросе монополия, стремящаяся увеличить свою прибыль, будет снижать объем выпуска. При этом издержки будут снижаться, а цена и прибыль – увеличиваться.

В некоторый момент начнется массовый отказ (из-за отсутствия средств) потребителей от продукции данной монополии. Спрос снова станет эластичным.

Пример 2.33. Пусть $C(q)=\frac{1}{2}q^2$ – издержки фирмы-монополиста, $D(p) = 40 - 2p$ есть функция спроса. Найти зависимость цены p от количества произведенной продукции q.

Решение. Так как $q = D(p) = 40 - 2p$, то $\frac{1}{2}q = 20 - p$ и $p = 20 - \frac{1}{2}q$. Итак, для функции спроса $D(p)$ мы нашли обратную функцию $p(q) = 20 - \frac{1}{2}q$. Прибыль $\pi(q)$ имеет вид

$$\pi(q) = \left(20 - \frac{1}{2}q\right)q - \frac{1}{2}q^2 = 20q - q^2.$$

В точке q^* максимума прибыли выполняется равенство

$$\pi'(q^*) = 20 - 2q^* = 0.$$

Находим оптимальный (для монополии) объем производства $q^* = 10$. Соответствующая цена будет $p^* = p(q^*) = 20 - \frac{1}{2}q^* = 15.$ При этом предельные издержки $C'(q^*) = 10$. Таким образом, цена, наиболее выгодная для монополии, в полтора раза выше ее предельных издержек. Этот же результат можно получить и по формуле (2.59). Действительно,

$$E_{Dp}(p) = p\frac{D'}{D} = p\frac{-2}{40-2p} = -\frac{p}{20-p},$$

поэтому $E_{Dp}(p^*) = E_{Dp}(15) = -3.$ Следовательно,

$$\left(E_{Dp}^{-1}(p^*) + 1\right)^{-1} = \left(-\frac{1}{3} + 1\right)^{-1} = 1{,}5.$$

§ 2.13. Высшие производные

Пусть $f'(x)$ – производная функции $f(x)$. Функция $f'(x)$ называется также *первой производной*. Производная от $f'(x)$ называется *второй производной* функции $f(x)$ и обозначается $f''(x)$ или $f^{(2)}(x)$. *Третьей производной* функции $f(x)$ называется производная от $f''(x)$, она обозначается $f'''(x)$ или $f^{(3)}(x)$. Вообще, *n-й производной* от функции $f(x)$ называется производная от ее $(n-1)$-й производной: $f^{(n)}(x)=(f^{(n-1)}(x))'$. Говорят также, что $f^{(n)}(x)$ – это *производная порядка n* от функции $f(x)$.

Если x_0 – это фиксированная точка, то символ $f^{(n)}(x_0)$ обозначает производную n-го порядка от функции $f(x)$ в точке x_0. Для ее существования необходимо существование производной $f^{(n-1)}(x)$ не только в точке x_0, но и в некоторой окрестности этой точки.

1. Производная порядка *n* степенной функции

Пусть $y=x^{\alpha}$ – степенная функция с произвольным (не равным нулю) показателем α. Первая производная $y'=\alpha x^{\alpha-1}$. Если $\alpha-1\neq$ $\neq 0$, то вторая производная $y^{(2)}=\alpha(\alpha-1)x^{\alpha-2}$. Если $\alpha-2\neq 0$, то $y^{(3)}=\alpha(\alpha-1)(\alpha-2)x^{\alpha-3}$ и т.д. Таким образом, если α не является натуральным числом, то n-я производная имеет вид:

$$\left(x^{\alpha}\right)^{(n)}=\alpha(\alpha-1)...(\alpha-n+1)x^{\alpha-n}. \qquad (2.60)$$

Если же α – натуральное число, то формула (2.60) имеет смысл только для $n\leq\alpha$. Рассмотрим подробнее случай, когда α – натуральное число, а порядок производной $n=\alpha$. В этом случае формула (2.60) выглядит так:

$$\left(x^{n}\right)^{(n)}=n(n-1)...2\cdot 1.$$

Напомним, что произведение натуральных чисел от 1 до n называется *факториалом* числа n и обозначается $n!$. Таким образом, имеем $\left(x^n\right)^{(n)} = n!$. Для натурального α в случае $n > \alpha$, очевидно, n-я производная от x^α равна нулю.

Пример 2.34.

$$\begin{aligned}
\left(x^4\right)' &= 4x^3, \\
\left(x^4\right)'' &= 4\cdot 3x^2 = 12x^2, \\
\left(x^4\right)^{(3)} &= 4\cdot 3\cdot 2x = 24x, \\
\left(x^4\right)^{(4)} &= 4\cdot 3\cdot 2\cdot 1 = 4! = 24, \\
\left(x^4\right)^{(n)} &= 0,\ n > 4.
\end{aligned}$$

Пример 2.35.

$$\left(\sqrt{x}\right)' = \frac{1}{2}x^{-\frac{1}{2}}, \qquad \left(\sqrt{x}\right)^{(2)} = -\left(\frac{1}{2}\right)^2 x^{-\frac{3}{2}},$$

$$\left(\sqrt{x}\right)^{(3)} = \left(\frac{1}{2}\right)^3 3x^{-\frac{5}{2}}, \quad \left(\sqrt{x}\right)^{(4)} = -\left(\frac{1}{2}\right)^4 3\cdot 5\cdot x^{-\frac{7}{2}},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\left(\sqrt{x}\right)^{(n)} = -\left(-\frac{1}{2}\right)^n 3\cdot 5\cdot 7\cdot \ldots \cdot (2n-3)\cdot x^{-\frac{2n-1}{2}}.$$

2. Производная порядка *n* показательной функции $y = a^x$ $(0 < a \neq 1)$

Последовательно дифференцируя, находим

$$y' = a^x \ln a,\ y'' = a^x (\ln a)^2, \ldots,\ y^{(n)} = a^x (\ln a)^n.$$

В частности, если $y = e^x$, то для любого n имеем

$$\left(e^x\right)^{(n)} = e^x.$$

3. Производные порядка *n* функций sin *x*, cos *x*

Пусть $y=\sin x$. Последовательно дифференцируя, получим

$$y'=\cos x=\sin\left(x+\frac{\pi}{2}\right),\ y''=-\sin x=\sin\left(x+2\cdot\frac{\pi}{2}\right),\ \ldots,\ y^{(n)}=\sin\left(x+n\frac{\pi}{2}\right).$$

Аналогично выглядит формула для производной n-го порядка-функции $\cos x$:

$$(\cos x)^{(n)}=\cos\left(x+n\frac{\pi}{2}\right).$$

Пример 2.36.

$$(\cos x)^{(10)}=\cos\left(x+10\cdot\frac{\pi}{2}\right)=\cos(x+\pi)=-\cos x.$$

4. Производная порядка *n* функции *y* = ln*x*

Последовательно дифференцируя, находим производные

$$y'=\frac{1}{x},\ y''=-\frac{1}{x^2},\ y^{(3)}=\frac{2}{x^3},\ \ldots,\ y^{(n)}=(-1)^{n-1}\frac{(n-1)!}{x^n}.$$

Пример 2.37.

$$(\ln x)^{(7)}=\frac{6!}{x^7}=\frac{720}{x^7}.$$

§ 2.14. Применение производных к исследованию функций

1. Возрастание и убывание функции

Пусть $X\subset\mathbb{R}$ – промежуток числовой прямой, т.е. X – это либо отрезок, либо полуинтервал, либо интервал. Напомним, что функция $f(x)$ называется *возрастающей* (*убывающей*) на промежутке X, если для любых $x_1, x_2\in X$, $x_1<x_2$, верно неравенство

$$f(x_1)<f(x_2)\quad \bigl(f(x_1)>f(x_2)\bigr).$$

Если в последнем неравенстве допускается равенство, то функция называется *неубывающей* (*невозрастающей*). Так, например, функция, принимающая постоянное значение $f(x) = \text{const},$ является одновременно неубывающей и невозрастающей. По знаку производной $(\text{"0"}, \text{"+"}, \text{"–"})$ можно судить о характере изменения функции на промежутке (const, возрастает, убывает). В развернутом виде это утверждение дается ниже в теоремах 2.10, 2.11 и 2.12.

Теорема 2.10. *Пусть функция* $f(x)$ *непрерывна на промежутке X, а ее производная* $f'(x)$ *обращается тождественно в нуль внутри X. Тогда* $f(x) = const$ *на X.*

Доказательство. Пусть $x_1 < x_2$ – две точки из *X*. По теореме Лагранжа найдется точка $c \in (x_1, x_2).$ для которой

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$

Так как $x_1 < c < x_2$, то точка *c* является внутренней точкой промежутка *X*. Поэтому $f'(c) = 0$ и $f(x_2) = f(x_1).$ Следовательно, $f(x)$ принимает постоянное значение на *X*.

Следствие 1. *Если эластичность функции y = f(x) тождественно равна нулю, то y = C – const.*

Доказательство. Так как $E_y = x\dfrac{y'}{y} = 0$ влечет $y' = 0,$ то по теореме 2.10 получаем, что *y – const*.

Следствие 2. *Если эластичность* $E_{yx}(x) = \alpha$ *не зависит от x, то* $y = Cx^{\alpha},$ *где C – некоторая постоянная.*

Доказательство. Так как эластичность функции x^{α} равна $\alpha,$ то эластичность функции $z = \dfrac{y}{x^{\alpha}}$ будет $E_{zx} = \alpha - \alpha = 0.$ По следствию 1 имеем *z* = *C* = const. Отсюда следует, что $y = Cx^{\alpha}.$

Теорема 2.11. *Пусть функция f(x) непрерывна на промежутке X и дифференцируема во внутренних точках X. Тогда*:

а) е*сли производная* $f'(x) \geq 0$ *внутри X, то f(x) не убывает на X*;

б) *если производная* $f'(x)$ *неотрицательна внутри* X *и не обращается тождественно в нуль ни на каком интервале из промежутка* X, *то* $f(x)$ *возрастает на* X.

Доказательство. а) Пусть $x_1 < x_2$ – две точки из промежутка X. По теореме Лагранжа найдется точка $c \in (x_1, x_2)$, для которой

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$

Так как $x_1 < c < x_2$, то точка c является внутренней точкой промежутка X. Поэтому $f'(c) \geq 0$ и $f(x_2) \geq f(x_1)$. Таким образом, мы доказали, что функция $f(x)$ не убывает на промежутке X.

б) Из а) следует, что для любой точки $x \in (x_1, x_2)$ выполняются неравенства $f(x_1) \leq f(x)$ и $f(x) \leq f(x_2)$. Предположим, что $f(x_1) = f(x_2)$. Тогда для любой точки $x \in (x_1, x_2)$ имеем $f(x_1) = f(x) = f(x_2)$, т.е. $f(x) = \text{const}$ на интервале (x_1, x_2). Но в этом случае $f'(x) = 0$ на (x_1, x_2), что противоречит условию б) теоремы. Поэтому $f(x_1) < f(x_2)$, а $f(x)$ возрастает на X.

Следующая теорема сводится к предыдущей заменой $f(x)$ на $-f(x)$.

Теорема 2.12. *Пусть функция* $f(x)$ *непрерывна на промежутке* X *и дифференцируема во внутренних точках* X. *Тогда*:

а) *если производная* $f'(x) \leq 0$ *внутри* X, *то* $f(x)$ *не возрастает на* X.

б) *если производная* $f'(x)$ *не положительна внутри* X *и не обращается тождественно в нуль ни на каком интервале из промежутка* X, *то* $f(x)$ *убывает на* X.

Пример 2.38. Найти интервалы монотонности функции

$$y = x^2 - 6x + 10.$$

Решение. Имеем $y' = 2x - 6 > 0$ при $x > 3$ и $y' = 2x - 6 < 0$ при $x < 3$. Поэтому функция $y = x^2 - 6x + 10$ убывает на интервале $(-\infty; 3)$ и возрастает на интервале $(3; +\infty)$.

Пример 2.39. Найти интервалы монотонности функции

$$y = x - \sin x.$$

Решение. Находим производную $y' = 1 - \cos x$. Поскольку $\cos x \le 1$ при любом x, то $y' \ge 0$. Кроме того, производная $y' = 1 - \cos x$ обращается в нуль только в изолированных точках $x = 2\pi k$, где k –целое число. Поэтому y' не обращается тождественно в нуль ни на каком интервале (a, b). Из достаточного условия возрастания функции следует, что функция y возрастает на $\mathbb{R}$.

2. Экстремум функции

Определение. *Точка x_0 называется **точкой локального максимума** (соответственно **локального минимума**) функции $f(x)$, если для всех x из некоторой окрестности точки x_0 выполняется неравенство $f(x) \le f(x_0)$ (соответственно $f(x) \ge f(x_0)$)* (рис. 2.10).

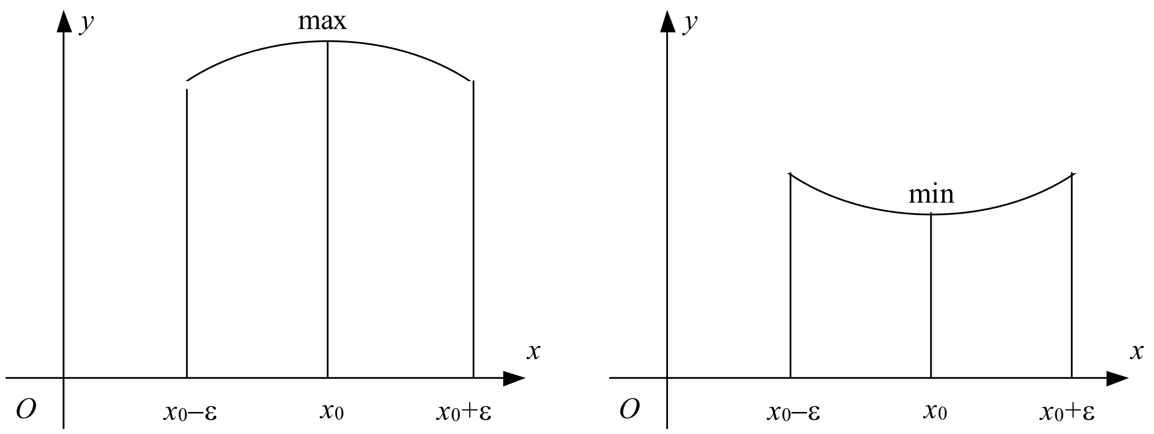

Рис. 2.10

Очевидно, что у функции может быть несколько локальных максимумов и несколько локальных минимумов, причем некоторый локальный максимум может оказаться меньше какого-то локального минимума (рис. 2.11).

Локальный максимум и локальный минимум объединяются общим названием *локальный экстремум*.

Теорема 2.13. *Для того чтобы дифференцируемая функция $f(x)$ имела в точке x_0 локальный экстремум, необходимо, чтобы в этой точке выполнялось равенство* $f'(x_0)=0$.

Доказательство. Поскольку x_0 – точка экстремума, то существует такой интервал $(x_0-\varepsilon, x_0+\varepsilon)$, на котором $f(x_0)$ – наибольшее или наименьшее значение. Тогда по теореме Ферма $f'(x_0)=0$.

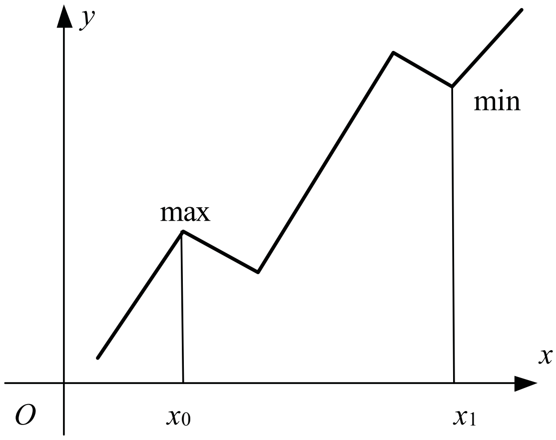

Рис. 2.11

Точки, в которых производная функции обращается в нуль, называются *стационарными*. Из теоремы 2.13 следует, что точка локального экстремума дифференцируемой функции является стационарной точкой. Обратное утверждение неверно. Рассмотрим, например, функцию $f(x)=x^3$. Эта функция возрастает на всей числовой прямой и, поэтому не имеет точек локального экстремума. В то же время точка $x_0=0$ является стационарной точкой, так как $f'(x_0)=3x_0^2=0$.

Теорема 2.14. *Пусть вторая производная дважды дифференцируемой функции $f(x)$ принимает на промежутке X только положительные (соответственно отрицательные) значения. Тогда если x_0 – стационарная точка функции $f(x)$, то $f(x_0)$ – наименьшее (соответственно наибольшее) значение функции $f(x)$ на всем промежутке X.*

Доказательство. Рассмотрим случай, когда вторая производная положительна. Тогда $(f'(x))'>0,$ и по достаточному условию возрастания функции получаем, что производная $f'(x)$ возрастает на промежутке X.

Пусть x_0 – стационарная точка, $f'(x_0) = 0$. Тогда слева от x_0 значения производной меньше нуля, а справа – больше нуля. Следовательно, слева от x_0 функция убывает, справа – возрастает. Поэтому $f(x_0)$ – наименьшее значение функции $f(x)$ на промежутке X. Аналогично доказывается, что в случае $f''(x) < 0$ значение $f(x_0)$ является наибольшим.

Пример 2.40. Найти наибольшее значение $f_{\max}$ функции

$$f(x) = \ln x - ax$$

для произвольного положительного значения параметра a.

Решение. Производная $f'(x) = x^{-1} - a$ обращается в нуль в точке $x_0 = a^{-1}$. Вторая производная $f''(x) = (x^{-1} - a)' = -x^{-2} < 0$. Поэтому значение в стационарной точке $f(x_0)$ является наибольшим значением функции на интервале $(0, +\infty)$. Таким образом, $f_{\max} = f(a^{-1}) = -\ln a - 1$.

Теорема 2.15. *Для того чтобы функция $f(x)$ имела локальный максимум (минимум) в стационарной точке x_0, достаточно, чтобы в окрестности этой точки существовала непрерывная вторая производная и выполнялось неравенство $f''(x_0) < 0$ ($f''(x_0) > 0$).*

Доказательство. Пусть $f''(x_0) < 0$. Так как $f''(x)$ непрерывна в точке x_0, то в некоторой ε-окрестности выполняется неравенство

$$f''(x) < 0, \quad x \in (x_0 - \varepsilon, x_0 + \varepsilon).$$

Следовательно, x_0 – точка локального максимума функции $f(x)$. Аналогично доказывается, что в случае $f''(x_0) > 0$ функция $f(x)$ имеет локальный минимум в стационарной точке x_0.

Для того чтобы установить, что некоторая точка x_0 является точкой локального экстремума, в тех случаях, когда условия теоремы 2.15 не выполняются, можно попытаться воспользоваться следующим критерием (очевидно вытекающим из достаточного условия возрастания (убывания) функции).

Теорема 2.16. *Если при переходе через точку x_0 производная дифференцируемой функции $f(x)$ меняет свой знак с плюса на минус, то точка x_0 – точка локального максимума функции $f(x)$, а если с минуса на плюс, то x_0 – точка локального минимума.*

Пример 2.41. Найти точки локального экстремума функции $y = x^2 - 2|x|$.

Решение. Если $x > 0$, то $y = x^2 - 2x$ и $y' = 2x - 2$. Если $x < 0$, то $y = x^2 + 2x$ и $y' = 2x + 2$. Решая уравнение $y' = 0$, находим стационарные точки $x_{12} = \pm 1$. При $x \neq 0$ вторая производная $y'' = (2x \pm 2)' = 2 > 0$. Согласно теореме 2.15 точки $x_1 = -1$, $x_2 = 1$ являются точками локального минимума. Осталось проверить точку $x = 0$. Слева от этой точки, на интервале (–1; 0), $y' > 0$; справа – на интервале (0;1), $y' < 0$. По теореме 2.16 точка $x = 0$ – точка локального максимума (рис. 2.12).

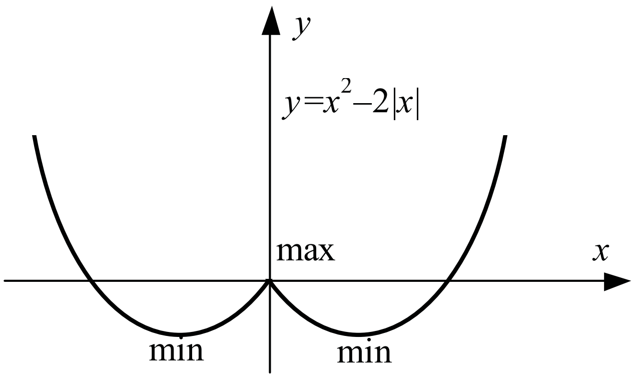

Рис. 2.12

3. Наибольшее и наименьшее значения функции на отрезке

Многие экономические задачи формулируются как задачи на нахождение максимального (минимального) значения функции на некотором множестве. Рассмотрим наиболее простой случай, когда требуется найти наибольшее (наименьшее) значение функции $f(x)$

на отрезке $[a, b]$. В главе 1 было доказано, что наибольшее (наименьшее) значение функции достигается в некоторой точке $x_0 \in [a, b]$. При этом возможны лишь следующие три случая: 1) $x_0 = a$, 2) $x_0 = b$, 3) $x_0 \in (a, b)$. Пусть $x_0 \in (a, b)$. Тогда x_0 – точка локального экстремума, и если существует $f'(x_0)$, то $f'(x_0) = 0$. Однако производная $f'(x_0)$ может и не существовать.

Определение. ***Критической точкой*** *функции $f(x)$ называется точка, в которой производная $f'(x)$ либо не существует, либо равна нулю.*

Из определения вытекает, что точка локального экстремума x_0 является критической точкой функции $f(x)$. Предположим, что критические точки функции $f(x)$ на интервале (a, b) образуют конечное множество $\{x_1, x_2, ..., x_n\}$. Из сказанного выше следует, что точка x_0, в которой функция принимает наибольшее (или наименьшее) значение, совпадает с одной из точек: a, b, x_1, ..., x_n. Поэтому для максимального значения $f_{\max}$ функции $f(x)$ на отрезке $[a, b]$ имеем равенство

$$f_{\max} = \max\{f(a), f(b), f(x_1), ..., f(x_n)\}.$$

Аналогично, для $f_{\min}$ выполняется равенство

$$f_{\min} = \min\{f(a), f(b), f(x_1), ..., f(x_n)\}.$$

Пример 2.42. Найти наибольшее и наименьшее значения функции $f(x) = x^2 e^x$ на отрезке [3; –3].

Решение. Имеем $f'(x) = 2xe^x + x^2 e^x = x(x+2)e^x$. Критические точки: $x_1 = 0$, $x_2 = -2$. Максимальное значение

$$f_{\max} = \max\{f(-3), f(3), f(0), f(-2)\} = \max\{9e^{-3}, 9e^3, 0, 4e^{-2}\} = 9e^3.$$

Минимальное значение $f_{\min} = \min\{9e^{-3}, 9e^3, 0, 4e^{-2}\} = 0.$

4. Оценка числа корней уравнения

Рассмотрим уравнение

$$f(x) = 0, \ x \in \mathbb{R},$$

где $f(x)$ – многочлен третьей степени вида $f(x) = x^3 + px + q$. Если коэффициент $p > 0$, то производная $f'(x) = 3x^2 + p > 0$, и функция $f(x)$ возрастает на $\mathbb{R}$. Следовательно, многочлен $f(x)$ имеет только один действительный корень. Если $p < 0$, то количество корней уравнения $f(x) = 0$ зависит от знака *дискриминанта* D, который определяется формулой[1]

$$D = -4p^3 - 27q^2.$$

Действительно, в случае $p < 0$ функция $f(x)$ имеет две стационарные точки $t_1 = -\sqrt{-\frac{p}{3}}$ и $t_2 = \sqrt{-\frac{p}{3}}$. Эти точки разбивают $\mathbb{R}$ на три промежутка таким образом, что $f(x)$ возрастает на интервалах $(-\infty, t_1)$, $(t_2, +\infty)$ и убывает на (t_1, t_2). Находим значения многочлена в стационарных точках:

$$f(t_1) = 2\sqrt{-\frac{p^3}{27}} + q, \quad f(t_2) = -2\sqrt{-\frac{p^3}{27}} + q.$$

Отсюда

$$f(t_1) \cdot f(t_2) = q^2 + \frac{4p^3}{27} = -\frac{D}{27}.$$

Таким образом, если $D > 0$, то числа $f(t_1)$ и $f(t_2)$ имеют противоположные знаки. Следовательно, на интервале (t_1, t_2) многочлен $f(x)$ имеет один корень. Кроме того, $f(x)$ убывает на (t_1, t_2), поэтому

[1] Эквивалентное определение дискриминанта: $D = (a - b)^2(b - c)^2(a - c)^2$, где a, b, c – действительные или комплексные корни многочлена $f(x)$.

$f(t_1) > 0$, а $f(t_2) < 0$. Так как $\lim\limits_{x\to-\infty} f(x) = -\infty$ и $f(t_1) > 0$, то $f(x)$ имеет еще один корень на интервале $(-\infty, t_1)$. Так как $\lim\limits_{x\to+\infty} f(x) = +\infty$ и $f(t_2) < 0$, то $f(x)$ имеет третий корень на интервале $(t_2, +\infty)$.

Вывод: *если дискриминант больше нуля, то многочлен имеет три действительных корня.*

Аналогично доказывается, что в случае $D < 0$ многочлен имеет только один действительный корень.

Покажем, как с помощью n-й производной $f^{(n)}(x)$ можно получить *оценку* числа корней уравнения $f(x) = 0$.

Лемма. *Если n раз дифференцируемая функция $f(x)$ обращается в нуль в k точках $x_1 < x_2 < \ldots < x_k$ на промежутке X, то ее n-я производная $f^{(n)}(x)$ обращается в нуль на промежутке X по меньшей мере в $(k - n)$ точках.*

Доказательство. Так как $f(x_1) = \ldots = f(x_k) = 0,$ то по теореме Ролля найдутся точки $t_1, t_2, \ldots, t_{k-1},$ такие что $t_i \in (x_i, x_{i+1})$ и

$$f'(t_i) = 0, \ (i = 1, \ldots, k-1).$$

Все точки $t_i \in X$, поэтому первая производная $f'(x)$ обращается в нуль на X по меньшей мере в $k - 1$ точках. Так как $f'(t_1) = \ldots = f'(t_{k-1}),$ то можно снова применить теорему Ролля. В результате ее применения получим точки $s_1, \ldots, s_{k-2}$, такие что $s_i \in (t_i, t_{i+1})$ и $f''(s_i) = 0 \ \ (i = 1, \ldots, k-2)$. Точки s_i принадлежат промежутку X, поэтому $f''(x)$ имеет на X по меньшей мере $k - 2$ точки, в которых она обращается в нуль.

Повторяя описанную выше процедуру n раз, получим $k - n$ точек, в которых n-я производная обращается в нуль.

Теорема 2.17. *Если уравнение $f^{(n)}(x) = 0$ имеет r корней на промежутке X, то уравнение $f(x) = 0$ имеет на промежутке X не более $n + r$ корней.*

Доказательство. Пусть k – число корней уравнения $f(x) = 0$ на X. По доказанной выше лемме имеем неравенство $r \geq k - n$. Отсюда $k \leq n + r$.

Пример 2.43. Доказать, что для любого многочлена $P(x)$ степени n уравнение $e^x + P(x) = 0$ имеет не более $n + 1$ корней.

Решение. Найдем $(n + 1)$-ю производную функции $f(x) = e^x + P(x)$. Учитывая, что $\left(e^x\right)^{(n+1)} = e^x$, а $(P(x))^{(n+1)} = 0$, получим $f^{(n+1)}(x) = e^x > 0$. Следовательно, уравнение $f^{(n+1)}(x) = 0$ не имеет корней на промежутке $(-\infty, +\infty)$. Поэтому уравнение $e^x + P(x) = 0$ имеет не более $n + 1$ корней.

§ 2.15. Функция предложения конкурентной фирмы

Пусть некоторая фирма A предлагает свою продукцию к продаже по цене p_A. Фирма A называется *конкурентной*, если существует такая цена p, что при цене $p_A \le p$ фирма может продать любое количество своей продукции, а при цене $p_A > p$ не может продать ничего. Таким образом, спрос на продукцию конкурентной фирмы имеет вид

$$D(p_A) = \begin{cases} +\infty, \text{ если } p_A \le p, \\ 0, \quad \text{если } p_A > p. \end{cases}$$

В данном случае кривая спроса является прямой, параллельной оси выпуска Oq (рис. 2.13), а сам спрос является совершенно эластичным. Чтобы какая-то реально существующая фирма могла считаться конкурентной, необходимо, чтобы максимальный объем ее выпуска составлял лишь незначительную долю от совокупного выпуска ее конкурентов.

Для конкурентной фирмы невыгодно продавать продукцию по цене $p_A < p$ и невозможно продавать по цене $p_A > p$. Поэтому мы будем считать, что $p_A = p$. Не имея возможности влиять на цену p, конкурентная фирма может увеличивать или уменьшать объем своего выпуска q так, чтобы получить максимальную прибыль $\pi(q)$. Пусть $C(q)$ – издержки, связанные с производством q ед. продукции. Тогда

$$\pi(q) = pq - C(q). \tag{2.61}$$

Далее считаем, что функция $C(q)$ определена и дифференцируема на промежутке $[0, +\infty)$. Таким образом, производная $C'(q)$ определена при $q \geq 0$ и отождествляется, как обычно, с предельными издержками $MC = C'(q)$. Рассмотрим подробнее функции издержек следующих двух типов.

Тип I. Предельные издержки возрастают на промежутке $[0, +\infty)$.

Тип II. Предельные издержки убывают на отрезке $[0, q_0]$ и возрастают на интервале $(q_0, +\infty)$.

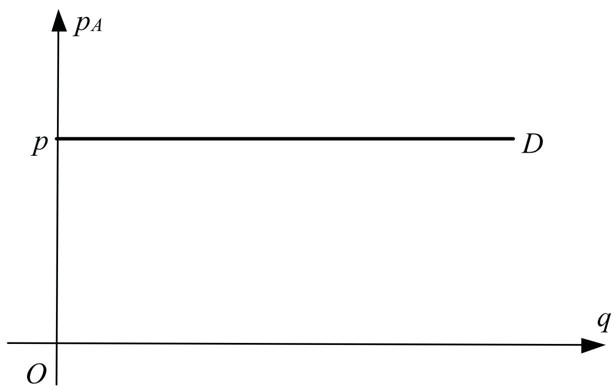

Рис. 2.13

В точке q^* локального максимума функции прибыли $\pi(q)$ выполняется равенство $\pi'(q^*) = p - C'(q^*) = 0$, из которого следует, что

$$p = C'(q^*). \tag{2.62}$$

Для функции издержек первого типа верно следующее утверждение:

если в точке q^* предельные издержки совпадают с ценой продукции (выполняется равенство (2.62)*), то $\pi(q^*)$ – наибольшее значение прибыли на промежутке* $[0, +\infty)$.

Действительно, функция $C'(q)$ возрастает на промежутке $[0, +\infty)$. Следовательно, $\pi'(q) = p - C'(q) > 0$, когда $q < q^*$, и

$\pi'(q) = p - C'(q) < 0,$ когда $q > q^*$. Поэтому прибыль $\pi(q)$ возрастает на отрезке $[0, q^*]$ и убывает на промежутке $[q^*, +\infty)$ (теоремы 2.11, 2.12). Таким образом, q^* – точка глобального максимума функции прибыли.

Как и любая коммерческая фирма, конкурентная фирма стремится максимизировать свою прибыль, поэтому при данной цене продукции p она устанавливает объем выпуска q равным q^*, где q^* – точка глобального максимума прибыли (2.61). Следовательно, равенство (2.62) влечет равенство $q^* = S(p)$, где $S(p)$ – функция предложения, т.е. $S(p)$ – обратная функция для функции $C'(q)$. С геометрической точки зрения данное утверждение означает, что кривая предложения (график $q = S(p)$) совпадает с кривой предельных издержек (графиком $p = C'(q)$) (рис. 2.14).

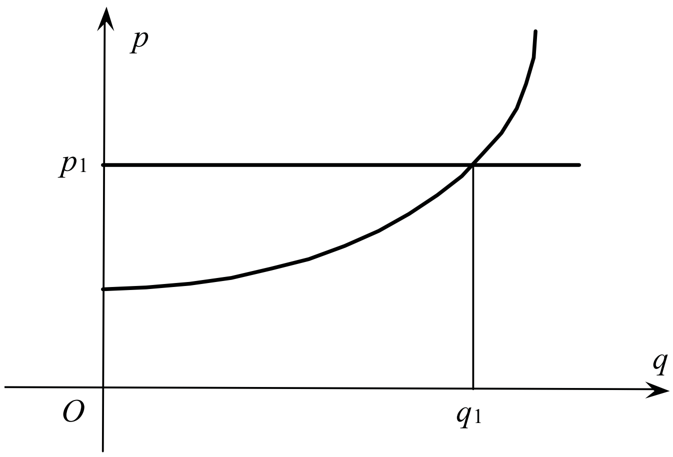

Рис. 2.14

Пример 2.44. Найти функцию предложения конкурентной фирмы $S(p)$, если ее функция издержек имеет вид

$$C(q) = q^2 + 6q + 5.$$

Решение. Находим предельные издержки $C'(q) = 2q + 6$. Решая уравнение $p = 2q + 6$ относительно q, получим функцию предложения

$$S(p) = \frac{1}{2}p - 3.$$

Замечание. Если в примере 2.44 цена $p = 10$, то предложение $S(10) = 2$. Соответствующие издержки будут $C(2) = 21$, прибыль $\pi(2) = 10 \cdot 2 - 21 = -1$. Зачем же фирма решила выпускать 2 ед. продукции, если при этом она имеет убытки в размере 1 ден. ед.? Не лучше ли было прекратить выпуск? Ответ на эти вопросы состоит в том, что при выпуске $q = 2$ фирма несет *минимальные* убытки. Если, например, она прекратит выпуск продукции $(q = 0)$, то убытки составят $-\pi(0) = C(0) = 5$ ден. ед. Очевидно, что убыточное производство возможно только на коротком интервале времени, поэтому найденная функция предложения является *функцией предложения в краткосрочном периоде*.

Перейдем теперь к изучению функции предложения конкурентной фирмы в случае функции издержек второго типа. Пусть q^* – точка глобального максимума функции $\pi(q)$ на промежутке $[0, +\infty)$, Если $q^* \neq 0$, то глобальный максимум будет также и локальным. Поэтому для точки $q^* \neq 0$ выполняется равенство (2.62). Следовательно, q^* – это либо 0, либо *решение* уравнения

$$p = C'(q). \tag{2.63}$$

Напомним, что в случае функции издержек второго типа предельные издержки $C'(q)$ убывают на отрезке $[0, q_0]$ и возрастают на интервале $(q_0, +\infty)$. Поэтому уравнение (2.63) может иметь не более одного решения q_1 на отрезке $[0, q_0]$ и не более одного решения q_2 на интервале $(q_0, +\infty)$. Производная прибыли $\pi'(q) = p - C'(q)$ меньше нуля на интервале $(0, q_1)$ и больше нуля на интервале (q_1, q_0). Поэтому прибыль $\pi(q)$ убывает на интервале $(0, q_1)$ и возрастает на интервале (q_1, q_0) (рис. 2.15). Следовательно, точка q_1 не может быть точкой локального максимума прибыли. Итак, мы доказали, что максимум прибыли (если он существует[1]) достигается либо в нуле, либо в точке q_2.

[1] Как правило, $\lim\limits_{q \to \infty} C'(q) = +\infty$. В этом случае легко доказать, что максимум прибыли достигается в некоторой точке.

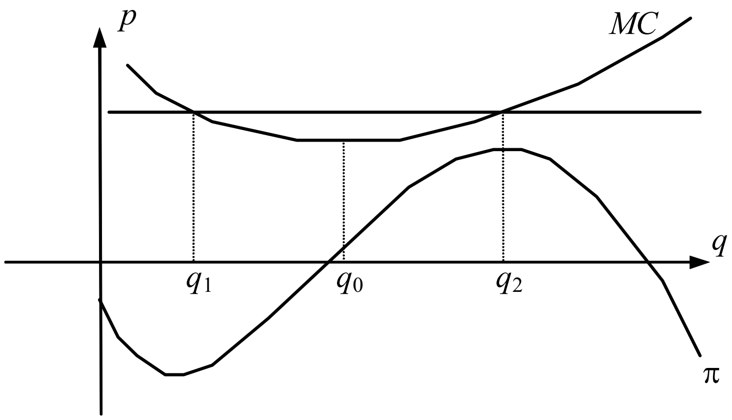

Рис. 2.15

Пусть $S(p)$ – функция предложения конкурентной фирмы. Так как $S(p)$ – точка глобального максимума функции прибыли $\pi(q)$, то либо $S(p) = 0$, либо $S(p) = q_2$. Для того чтобы сделать выбор между точками 0 и q_2 достаточно сравнить значения прибыли в этих точках:

$$S(p)=\begin{cases}0, & \text{если}\,\pi(0)>\pi(q_2);\\ q_2, & \text{если}\,\pi(q_2)>\pi(0).\end{cases}$$

Рассмотрим неравенство $\pi(q_2)>\pi(0)$. Имеем

$$\pi(q_2)>\pi(0)\Leftrightarrow pq_2-C(q_2)+C(0)>0\Leftrightarrow pq_2>C(q_2)-C(0).$$

Пусть $VC(q)=C(q)-C(0)$ – *переменные издержки*, связанные с производством q ед. продукции; $AVC(q)=\dfrac{VC(q)}{q}$ – средние переменные издержки. Тогда

$$\pi(q_2)>\pi(0)\Leftrightarrow pq_2>VC(q_2)\Leftrightarrow p>AVC(q_2).$$

В итоге получаем следующий вид функции предложения для функции издержек второго типа:

$$S(p)=\begin{cases}0, & \text{если } AVC(q_2)>p,\\ q_2, & \text{если } p>AVC(q_2),\end{cases} \tag{2.64}$$

где q_2 – единственное решение уравнения (2.63) на интервале $(q_0, +\infty)$.

С геометрической точки зрения формула (2.64) означает, что кривая предложения совпадает с участком кривой предельных издержек, расположенным выше кривой средних переменных издержек.

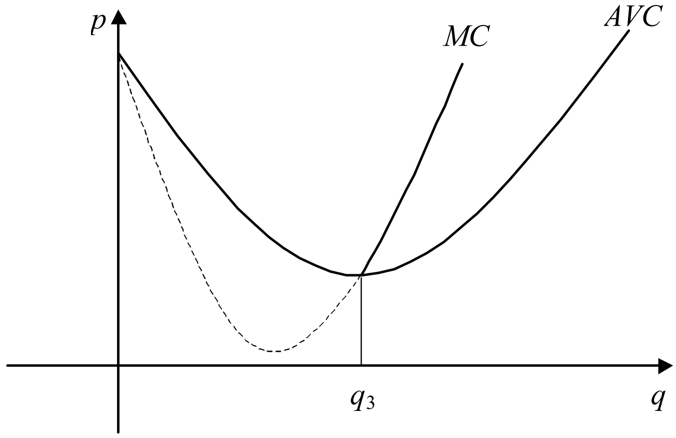

Рис. 2.16

Кривая предельных издержек пересекает кривую средних переменных издержек в точке минимума q_3 (рис. 2.16). Это не случайно, так как $AVC'(q_3)=0,$ а каждая такая точка является точкой пересечения кривых MC и $AVC.$ Действительно,

$$(AVC(q))' = \left(\frac{C(q)-C(0)}{q}\right)' = \frac{qC'(q)+C(0)-C(a)}{q^2} =$$

$$= \frac{C'(q)-AVC(q)}{q} = 0 \quad \Leftrightarrow \quad C'(q) = AVC(q).$$

П р и м е р 2.45. Найти функцию предложения конкурентной фирмы, если $C(q)=\frac{1}{3}q^3-2q^2+9q+20.$

Р е ш е н и е. Предельные издержки $C'(q)=q^2-4q+9,$ средние переменные издержки

$$AVC(q)=\frac{C(q)-C(0)}{q}=\frac{1}{3}q^2-2q+9.$$

Решая уравнение

$$p = C'(q) \Leftrightarrow q^2 - 4q + 9 - p = 0,$$

находим больший корень $q_2 = 2 + \sqrt{4 - (9 - p)} = 2 + \sqrt{p - 5}$. Неравенство $C'(q) > AVC(q)$ эквивалентно следующим неравенствам:

$$q^2 - 4q + 9 > \frac{1}{3}q^2 - 2q + 9 \Leftrightarrow \frac{2}{3}q^2 - 2q > 0 \Leftrightarrow q(q - 3) > 0.$$

Следовательно, кривая *MC* расположена выше кривой *AVC* на участке справа от точки $q = 3$. В точке $q = 3$ предельные издержки равны $C'(3) = 9 - 12 + 9 = 6$. В итоге имеем следующую функцию предложения:

$$S(p) = \begin{cases} 0, & \text{если } p < 6, \\ 2 + \sqrt{p - 5}, & \text{если } p > 6. \end{cases}$$

§ 2.16. Выпуклые функции

Пусть область определения $f(x)$ содержит промежуток X.

Определение. *Функция $f(x)$ называется **выпуклой** на промежутке X, если для любых точек a и b из X и любого $\alpha \in [0; 1]$ выполняется неравенство*

$$f(\alpha a + (1 - \alpha)b) \le \alpha f(a) + (1 - \alpha) f(b). \tag{2.65}$$

*Функция $f(x)$ называется **строго выпуклой** на промежутке X, если для любых двух различных точек a, $b \in X$ и любого $\alpha \in (0, 1)$ неравенство (2.65) выполняется как строгое неравенство. Функция $f(x)$ называется **вогнутой** (**строго вогнутой**) на промежутке X, если функция $-f(x)$ является выпуклой (строго выпуклой) на этом промежутке.*

Для того чтобы функция $f(x)$ была вогнутой на промежутке X, необходимо и достаточно, чтобы для любых $a, b \in X$ и $\alpha \in [0; 1]$ выполнялось неравенство

$$f(\alpha a + (1-\alpha)b \geq \alpha f(a) + (1-\alpha) f(b). \qquad (2.66)$$

Чтобы выяснить геометрический смысл неравенств (2.65) и (2.66), рассмотрим на координатной плоскости Oxy точки $A(a, f(a))$ и $B(b, f(b))$. Эти точки принадлежат графику функции $y = f(x)$ и соответствуют точкам a и b на оси Ox (рис. 2.17). Отрезок AB состоит из точек M с координатами

$$\begin{aligned} x_\alpha &= \alpha a + (1-\alpha)b, \\ y_\alpha &= \alpha f(a) + (1-\alpha) f(b), \end{aligned} \qquad (2.67)$$

где α принимает любое значение от 0 до 1.

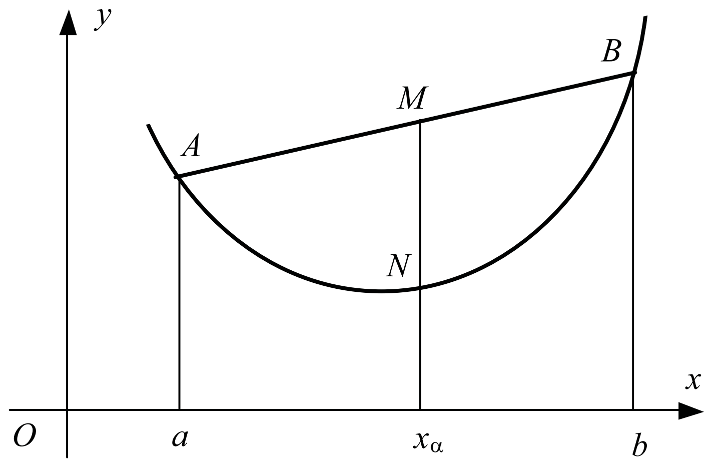

Рис. 2.17

Неравенство (2.65) означает, что $f(x_\alpha) \leq y_\alpha$. Следовательно, точка $N(x_\alpha, f(x_\alpha))$ графика функции расположена ниже, чем точка $M(x_\alpha, y_\alpha)$ отрезка AB. Поэтому про график выпуклой функции говорят, что он обращен выпуклостью вниз. Аналогично неравенство (2.66) означает, что $f(x_\alpha) \geq y_\alpha$, поэтому точка $N(x_\alpha, f(x_\alpha))$ расположена выше (рис. 2.18) соответствующей точки M отрезка AB. Говорят, что график вогнутой функции обращен выпуклостью вверх.

Пример 2.46. Исследуем на выпуклость линейную функцию

$$f(x) = px + q, \text{ где } p,\ q - const.$$

Имеем

$$\alpha f(a) + (1-\alpha) f(b) = \alpha pa + \alpha q + (1-\alpha) pb + (1-\alpha) q =$$
$$= p\big[\alpha a + (1-\alpha) b\big] + q = f(\alpha a + (1-\alpha) b).$$

Для данной функции $f(x)$ выполняются оба неравенства (2.65) и (2.66), поэтому линейная функция является одновременно и выпуклой, и вогнутой. В то же время линейная функция не является ни строго выпуклой, ни строго вогнутой, так как неравенства (2.65) и (2.66) выполняются только как равенства.

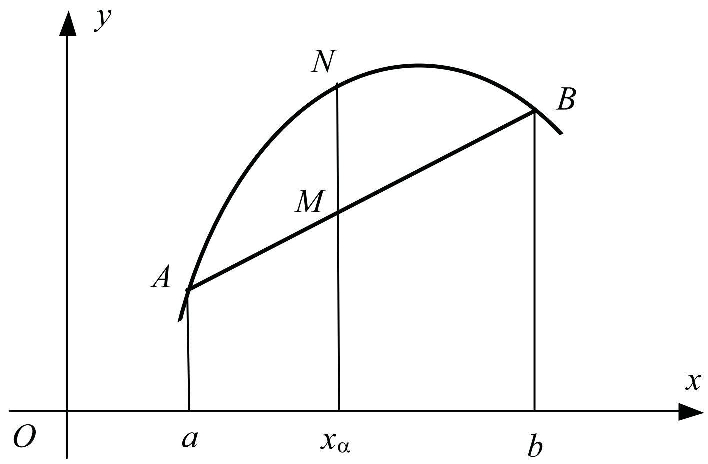

Рис. 2.18

С геометрической точки зрения одновременная выпуклость и вогнутость линейной функции означают, что отрезок, соединяющий две точки графика, совпадает с графиком (напомним, что график линейной функции – это прямая).

Теорема 2.18. *Пусть $f_1(x), \ldots, f_n(x)$ – выпуклые (соответственно вогнутые) функции на промежутке X. Тогда функция $f(x) = max\{f_1(x), \ldots, f_n(x)\}$ (соответственно $f(x) = min\{f_1(x), \ldots, f_n(x)\}$) также является выпуклой (соответственно вогнутой) на промежутке X.*

Доказательство. Рассмотрим случай максимума. Положим для простоты обозначений $n = 2$. Так как $f(x) = \max\{f_1(x), f_2(x)\}$, то для любых $x = a$ и $x = b$ из промежутка X выполняются неравенства $f_1(x) \le f(x)$ и $f_2(x) \le f(x)$. Из выпуклости функций $f_1(x)$ и $f_2(x)$ получаем следующие неравенства:

$$f_1(x_\alpha) \le \alpha f_1(a) + (1-\alpha) f_1(b) \le \alpha f(a) + (1-\alpha) f(b), \qquad (2.68)$$

$$f_2(x_\alpha) \le \alpha f_2(a) + (1-\alpha) f_2(b) \le \alpha f(a) + (1-\alpha) f(b), \qquad (2.69)$$

где $x_\alpha = \alpha a + (1-\alpha) b,\ \alpha \in [0,1]$. Из (2.68), (2.69) следует, что

$$f(x_\alpha) = \max\{f_1(x_\alpha), f_2(x_\alpha)\} \le \alpha f(a) + (1-\alpha) f(b),$$

т.е. $f(x)$ – выпуклая функция. Вогнутость функции $\min\{f_1(x), ..., f_n(x)\}$, где $f_i(x)$ – вогнутые функции, доказывается аналогично.

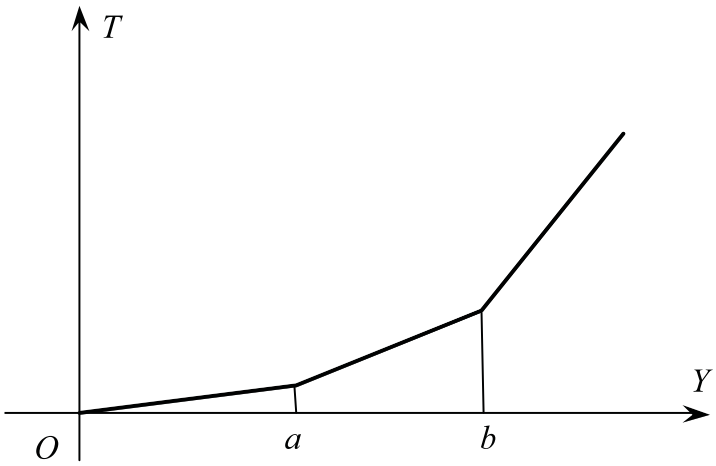

Рис. 2.19

Пример 2.47. Пусть $f(x) = ax + b$, где a и b – const. Тогда $|ax + b| = \max\{f(x), -f(x)\}$ – вогнутая функция на промежутке $(-\infty, +\infty)$, поскольку $f(x)$ и $-f(x)$ есть линейные функции.

Пример 2.48. Пусть сумма подоходного налога T определяется следующим образом:

1) если доход $Y \le a$, то T составляет 10% от Y;

2) если доход $Y > a$, то часть дохода, превышающая a, облагается налогом по ставке 20%;

3) если доход $Y > b$ (где $b > a$), то часть дохода, превышающая b, облагается 30%-ным налогом.

Функцию $T = T(Y)$ можно представить как максимум трех линейных функций $T(Y) = \max\{f_1(Y), f_2(Y), f_3(Y)\}$, где

$$f_1(Y) = 0{,}1 \cdot Y,$$
$$f_2(Y) = f_1(a) + 0{,}2(Y - a),$$
$$f_3(Y) = f_2(b) + 0{,}3(Y - b).$$

Таким образом, $T(Y)$ – выпуклая функция на промежутке $[0, +\infty)$ (график $T = T(Y)$ на рис. 2.19).

Теорема 2.19. *Сумма любого числа выпуклых (вогнутых) на промежутке X функций снова является выпуклой (вогнутой) функцией на промежутке X. Если при этом хотя бы одна из суммируемых функций является строго выпуклой (строго вогнутой), то и вся сумма также является строго выпуклой (строго вогнутой).*

Доказательство. Предположим, ради простоты обозначений, что суммируются только две функции $f(x)$ и $g(x)$. Рассмотрим случай, когда $f(x)$ и $g(x)$ – выпуклые на X функции. Пусть $a, b \in X$, $\alpha \in [0, 1]$. Запишем неравенство (2.65) для f и g:

$$f(\alpha a + (1-\alpha)b) \le \alpha f(a) + (1-\alpha) f(b), \tag{2.70}$$

$$g(\alpha a + (1-\alpha)b) \le \alpha g(a) + (1-\alpha) g(b). \tag{2.71}$$

Пусть $h(x) = f(x) + g(x)$ $(x \in X)$ – сумма функций f и g. Складывая неравенства (2.70) и (2.71), получим неравенство

$$h(\alpha a + (1-\alpha)b) \le \alpha h(a) + (1-\alpha) h(b), \tag{2.72}$$

означающее выпуклость функции $h(x)$. Если хотя бы одно из неравенств (2.70) или (2.71) является строгим, то их сумма – неравенство (2.72) также является строгим, что доказывает строгую выпуклость $h(x)$, когда хотя бы одна из выпуклых функций $f(x)$ или $g(x)$ является строго выпуклой. В случае, когда $f(x)$ и $g(x)$ – вогнутые функции, теорема доказывается аналогично.

Пример 2.49. Функция $y = |x - 1| + |x - 2| + |x - 3|$ есть выпуклая на множестве $\mathbb{R}$.

Теорема 2.20. *Пусть X – промежуток, X_0 – множество его внутренних точек; $f(x)$ – функция, непрерывная на X и дифференцируемая на X_0. Тогда:*

а) *если производная $f'(x)$ возрастает (убывает) на X_0, то $f(x)$ – строго выпуклая (строго вогнутая) функция на промежутке X;*

б) *если производная $f'(x)$ не убывает (не возрастает) на X_0, то $f(x)$ – выпуклая (вогнутая) функция на промежутке X.*

Доказательство. а) Предположим, что $f'(x)$ возрастает на X_0. Пусть $a < b$ – точки из промежутка X. Докажем, что для любого $\alpha \in (0, 1)$ выполняется неравенство

$$f(c) < \alpha f(a) + \beta f(b),$$

где $\beta = 1 - \alpha$, $c = \alpha a + \beta b$. По теореме Лагранжа найдутся такие точки $a_1 \in (a, c)$ и $b_1 \in (c, b)$, что

$$f(a) - f(c) = f'(a_1)(a - c), \tag{2.73}$$

$$f(b) - f(c) = f'(b_1)(b - c). \tag{2.74}$$

Поскольку $f'(x)$ возрастает на X_0, а a_1 и b_1 – внутренние точки промежутка X, имеем $f'(b_1) > f'(a_1)$. С учетом (2.74) получаем

$$f(b) - f(c) > f'(a_1)(b - c). \tag{2.75}$$

Складывая равенство (2.73) с неравенством (2.75), предварительно умножив (2.73) на α и (2.75) на β, получим

$$\alpha f(a) + \beta f(b) - f(c) > f'(a_1)[\alpha(a - c) + \beta(b - c)] = f'(a_1)[\alpha a + \beta b - c] = 0.$$

Отсюда следует, что $f(c) < \alpha f(a) + \beta f(b)$.

Итак, $f(x)$ – выпуклая функция, если $f'(x)$ возрастает. Предположим теперь, что производная $f'(x)$ убывает на X_0. Тогда производная функции $g(x) = -f(x)$ возрастает на X_0 и, как мы только что доказали, $g(x)$ строго выпукла на X. Следовательно, функция $f(x)$ является строго вогнутой на X.

б) Доказательство повторяет доказательство а) с заменой строгих неравенств на нестрогие.

Пусть функция $f(x)$ дважды дифференцируема. Для определения возрастания (убывания) ее производной $f'(x)$ достаточно выяснить знак производной от производной, т.е. знак $f''(x)$. Таким образом, комбинируя теорему 2.20 с достаточными условиями возрастания или убывания функции (теоремы 2.11, 2.12), получаем следующую теорему.

Теорема 2.21. *Пусть функция $f(x)$ непрерывна на промежутке X, а ее вторая производная $f''(x)$ определена на X_0 – множестве всех внутренних точек промежутка X. Тогда*:

а) *если $f''(x)$ положительна (отрицательна) на X_0, то $f(x)$ строго выпукла (строго вогнута) на X*;

б) *если $f''(x)$ неотрицательна (неположительна) на X_0, то $f(x)$ выпукла (вогнута) на промежутке X.*

Пример 2.50. Доказать строгую вогнутость функции $f(x)=\sqrt{x}$ на промежутке $[0, +\infty)$.

Решение. Функция $\sqrt{x}$ непрерывна на промежутке $[0, +\infty)$ и дважды дифференцируема во всех внутренних точках этого промежутка: $\left(\sqrt{x}\right)''=-\frac{1}{4}x^{-\frac{3}{2}}$. Так как $-\frac{1}{4}x^{-\frac{3}{2}}<0$ при $x>0$, то $\sqrt{x}$ есть строго вогнутая функция на промежутке $[0, +\infty)$.

Определение. *Функция $f(x)$ называется* ***выпуклой, строго выпуклой, вогнутой*** *или* ***строго вогнутой****, если она является таковой на всей области определения (которая в таком случае должна быть промежутком).*

Пример 2.51. Проверить строгую вогнутость функции $y = \ln x$.

Решение. Так как $(\ln x)''=(x^{-1})'=-x^{-2}<0$, если $x>0$, то $(\ln x)''<0$ во всей области определения функции $\ln x$. Следовательно, $y = \ln x$ – строго вогнутая функция.

Пример 2.52. Проверить строгую выпуклость функции

$$f(x)=\frac{1}{2}x^2+\cos x.$$

Р е ш е н и е. Находим вторую производную:

$$f''(x) = (x - \sin x)' = 1 - \cos x.$$

Так как $f''(x) \geq 0$ при любом x и $f''(x) = 0$ только в точках $x = 2\pi k$ (k – целое число), то производная $f'(x)$ возрастает на $\mathbb{R}$. Следовательно, функция $f(x) = \frac{1}{2}x^2 + \cos x$ строго выпукла на $\mathbb{R}$.

При построении графиков функций бывает полезно определить точки, отделяющие промежутки, на которых функция $f(x)$ выпукла, от промежутков, на которых $f(x)$ вогнута.

Такие точки называются точками *перегиба* функции $f(x)$, если в них существует производная $f'(x)$. Говорят также, что в точках перегиба функция (или ее график) меняет *направление выпуклости*.

Предположим, что вторая производная $f''(x)$ определена и непрерывна в окрестности точки x_0. Рассмотрим два случая: 1) $f''(x_0) > 0$; 2) $f''(x_0) < 0$.

В случае 1) найдется $\varepsilon > 0$, такое, что $f''(x)$ положительна на интервале $(x_0 - \varepsilon, x_0 + \varepsilon)$. По теореме 2.21 функция $f(x)$ строго выпукла на этом интервале. Следовательно, точка x_0 не может быть точкой перегиба $f(x)$ в случае 1).

В случае 2) найдется $\varepsilon > 0$ такое, что $f''(x)$ отрицательна на интервале $(x_0 - \varepsilon, x_0 + \varepsilon)$. По теореме 2.21 функция $f(x)$ строго вогнута на этом интервале. Поэтому точка x_0 не может быть точкой перегиба $f(x)$ и в случае 2).

Итак, если $f''(x) \neq 0$, то x_0 не является точкой перегиба $f(x)$.

В ы в о д: *для того чтобы точка x_0 была точкой перегиба функции $f(x)$, необходимо, чтобы выполнялось равенство*

$$f''(x_0) = 0.$$

П р и м е р 2.53. Найти промежутки выпуклости (вогнутости) и точки перегиба функции $y = e^{-\frac{x^2}{2}}$.

Р е ш е н и е. Находим вторую производную

$$y'' = \left(-xe^{-\frac{x^2}{2}} \right)' = -e^{-\frac{x^2}{2}} + x^2 e^{-\frac{x^2}{2}} = \left(x^2 - 1\right)e^{-\frac{x^2}{2}}.$$

Функция y'' обращается в нуль в точках –1 и 1. Эти точки разбивают $\mathbb{R}$ на три промежутка: (–∞, –1], [–1, 1], [1, +∞). На каждом из этих промежутков функция непрерывна, а ее вторая производная y'' имеет определенный знак внутри промежутка. Следовательно, функция y строго выпукла на (–∞, –1] и [1, +∞) ($y'' > 0$) и строго вогнута на отрезке [–1, 1] ($y'' < 0$). Точки –1 и 1 отделяют промежутки с различными направлениями выпуклости и поэтому являются точками перегиба.

Т е о р е м а 2.22. *Если функция $f(x)$ выпукла на промежутке X и дифференцируема в точке $a \in X$, то для любого $x \in X$ выполняется неравенство*

$$f(x) \geq f(a) + f'(a)(x-a). \tag{2.76}$$

Д о к а з а т е л ь с т в о. Поскольку $f(x)$ выпукла на X, то для любого $\alpha \in [0, 1]$ имеем неравенство

$$f\big(\alpha x + (1-\alpha)a\big) \leq \alpha f(x) + (1-\alpha) f(a),$$

которое, как нетрудно видеть, эквивалентно неравенству

$$\frac{f\big(a + \alpha(x-a)\big) - f(a)}{\alpha} \leq f(x) - f(a). \tag{2.77}$$

Положим $\Delta x = \alpha(x-a)$ и найдем предел левой части (2.77):

$$\lim_{\alpha \to 0} \frac{f\big(a+\alpha(x-a)\big) - f(a)}{\alpha} = (x-a) \lim_{\alpha \to 0} \frac{f\big(a+\alpha(x-a)\big) - f(a)}{(x-a)\alpha} =$$

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

Используя неравенство (2.77) и известное свойство пределов, получим неравенство $(x-a)f'(a) \le f(x)-f(a)$, из которого следует неравенство (2.76).

Напомним, что уравнение касательной к графику функции $y = f(x)$ в точке с абсциссой a имеет вид $y = f(a) + f'(a)(x-a)$, поэтому геометрический смысл теоремы 2.21 состоит в том, что график выпуклой функции расположен выше касательной к этому графику (рис. 2.20, *а*). Аналогично теореме 2.22 доказывается следующая теорема.

Теорема 2.23. *Если функция $f(x)$ вогнута на промежутке X и дифференцируема в точке $a \in X$, то для любого $x \in X$ выполняется неравенство*

$$f(x) \le f(a) + f'(a)(x-a). \tag{2.78}$$

Геометрический смысл теоремы 2.22 состоит в том, что график вогнутой функции расположен ниже касательной к этому графику (рис. 2.20, *б*).

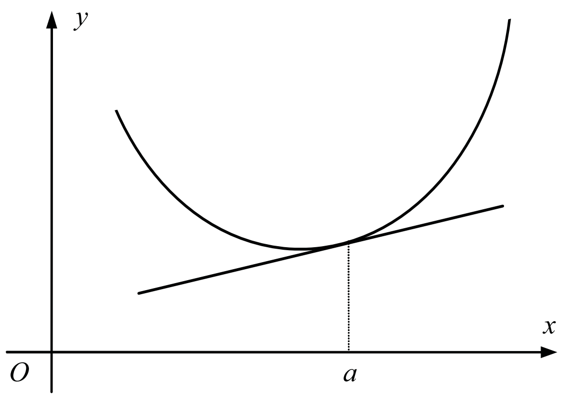

Рис. 2.20, *а*

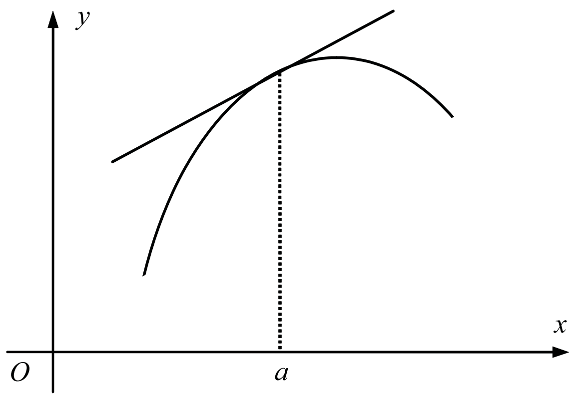

Рис. 2.20, *б*

Теорема 2.24. *Пусть $f(x)$ – выпуклая (вогнутая) функция на промежутке X, $a \in X$ – стационарная точка, т.е. $f'(a) = 0$. Тогда $f(a)$ – наименьшее (наибольшее) значение функции $f(x)$ на промежутке X.*

Доказательство. Для любой точки $x \in X$ из неравенства (2.76) (соответственно, (2.78)) получаем неравенство $f(x) \ge f(a)$ ($f(x) \le f(a)$), которое означает, что $f(a)$ – наименьшее (наибольшее) значение.

Пример 2.54. Проверить, что точка 0 является точкой глобального минимума функции $f(x) = |x-1| + e^x$.

Решение. Выше было доказано, что $|x - 1|$ есть выпуклая функция (пример 2.47). Функция e^x – также выпуклая функция, так как $\left(e^x\right)'' = e^x > 0.$ Следовательно, их сумма – выпуклая функция. В малой окрестности нуля имеем равенство $|x - 1| = 1 - x$, поэтому $|x-1|'_{x=0} = -1.$ Далее, $f'(0) = -1 + e^0 = 0,$ т.е. 0 – стационарная точка функции $f(x)$, выпуклой на $\mathbb{R}$.

Теорема 2.25. *Пусть $f(x)$ – строго выпуклая (строго вогнутая) функция на промежутке X. Тогда функция $f(x)$ имеет не более одной точки глобального минимума (максимума) на X.*

Доказательство. Пусть $f(x)$ – строго выпуклая функция. Предположим, что существуют две различные точки a и b глобального минимума $f(x)$ на X. Тогда для любого $\alpha \in (0, 1)$ имеем неравенство

$$f(c) < \alpha f(a) + (1-\alpha) f(b) = m, \tag{2.79}$$

где $c = \alpha a + (1-\alpha)b,\ \ m = f(a) = f(b)$ – наименьшее значение функции $f(x)$ на промежутке X. Точка $c \in [a, b]$, а поскольку X – промежуток, то $c \in X$. Неравенство (2.79) противоречит тому, что m – наименьшее значение на X. Следовательно, мы доказали, что строго выпуклая функция имеет не более одной точки глобального минимума на промежутке X. Аналогично доказывается, что строго вогнутая функция имеет не более одной точки глобального максимума.

Пример 2.55. Показать, что функция $y = e^x + e^{ax} + e^{bx} + e^{cx}$ имеет не более одной точки глобального минимума при любых фиксированных a, b и c.

Решение. Вторая производная $y'' = e^x + a^2 e^{ax} + b^2 e^{bx} + c^2 e^{cx}$ строго больше нуля в любой точке $x \in X$. Поэтому y – строго выпуклая функция в $\mathbb{R}$.

Замечание. Функция e^x является строго выпуклой, поэтому функция $f(x)$ из примера 2.54 также является строго выпуклой (по теореме 2.19), а точка 0 является единственной точкой глобального минимума функции $f(x)$ (в силу теоремы 2.25).

§ 2.17. Неравенство Йенсена и средние величины

Напомним, что множество M в $\mathbb{R}^2$ называется *выпуклым*, если для любых точек $A, B \in M$ отрезок AB целиком содержится в M. В предыдущем параграфе было доказано, что отрезок, соединяющий точки A и B графика выпуклой функции, расположен выше графика. Очевидно (рис. 2.21), что и отрезок $A'B'$, концы которого расположены выше графика выпуклой функции, также расположен выше графика.

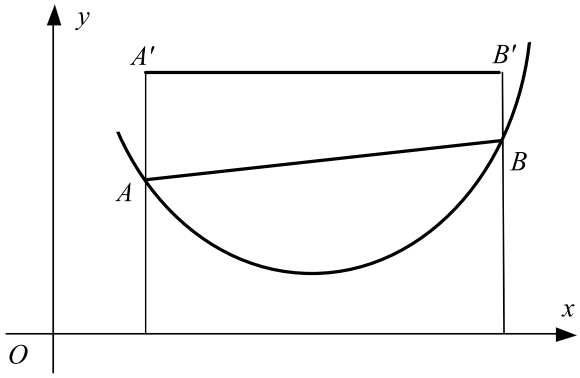

Рис. 2.21

Определение. ***Надграфиком*** *функции* $f(x)$ *называется множество* H_f *всех точек* (x_0, y_0), *таких, что функция* $f(x)$ *определена в точке* x_0 *и* $y_0 \geq f(x_0)$.

Фактически мы доказали следующую теорему.

Теорема 2.26. *Надграфик выпуклой функции является выпуклым множеством.*

Напомним, что выпуклой оболочкой точек $A_1, ..., A_n$ называется множество всех точек вида $\alpha_1 A_1 + ... + \alpha_n A_n$, где $\alpha_1, ..., \alpha_n$ – неотрицательные числа, сумма которых равна 1. Было доказано, что в случае, когда все точки A_i $(i = 1, ..., n)$ принадлежат выпуклому множеству M, выпуклая оболочка точек A_i также содержится в M.

Вывод: *выпуклая оболочка любых точек* $A_1, ..., A_n$ *из надграфика* H_f *выпуклой функции* $f(x)$ *содержится в* H_f.

В частности, это верно и для точек $A_i(x_i, f(x_i))$, расположенных на графике функции $y = f(x)$. Пусть $C(x_c, y_c)$ – произвольная точка из выпуклой оболочки точек A_i. Тогда ее координаты равны

$$x_c = \alpha_1 x_1 + ... + \alpha_n x_n,$$
$$y_c = \alpha_1 f(x_1) + ... + \alpha_n f(x_n).$$

Условие $C \in H_f$ означает, что $y_c \geq f(x_c)$, или

$$f(\alpha_1 x_1 + \ldots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \ldots + \alpha_n f(x_n). \tag{2.80}$$

Выражение (2.80) называется *неравенством Йенсена*. Это неравенство выполняется для любых неотрицательных чисел α_i с единичной суммой и для любых точек x_i из промежутка X, на котором функция $f(x)$ является выпуклой.

Положив в неравенстве (2.80) $\alpha_1 = \alpha_2 = \ldots = \alpha_n = \frac{1}{n}$, получим следующее неравенство:

$$f(\bar{x}) \leq \frac{f(x_1) + \ldots + f(x_n)}{n}, \tag{2.81}$$

где $\bar{x} = \frac{x_1 + \ldots + x_n}{n}$ – *среднее арифметическое* чисел x_1, ..., x_n. С помощью неравенства (2.81) нетрудно получить неравенства, сравнивающие среднее арифметическое с такими средними величинами как среднее геометрическое, среднее гармоническое и среднее квадратичное.

Средним геометрическим положительных чисел x_i называется число $\sqrt[n]{x_1 \ldots x_n}$. Функция $-\ln x$ является выпуклой, так как

$$(-\ln x)'' = \frac{1}{x^2} > 0$$

при $x > 0$. Используя неравенство (2.81) для функции $f(x) = -\ln x$, получаем неравенство

$$-\ln \bar{x} \leq -\frac{\ln x_1 + \ldots + \ln x_n}{n},$$

из которого следует, что

$$\sqrt[n]{x_1 \ldots x_n} = e^{\frac{1}{n}(\ln x_1 + \ldots + \ln x_n)} \leq e^{\ln \bar{x}} = \bar{x}.$$

Следовательно,

{среднее геометрическое} ≤ {среднее арифметическое}. (2.82)

Средним гармоническим положительных чисел x_i называется число

$$\frac{n}{x_1^{-1} + \ldots + x_n^{-1}}.$$

Вторая производная функции x^{-1} положительна $(x^{-1})'' = 2x^{-3} > 0$, если $x > 0$, поэтому функция x^{-1} выпукла на интервале (0, +∞). Из неравенства (2.81) для функции $f(x) = x^{-1}$ получаем неравенство $(\bar{x})^{-1} \le \frac{x_1^{-1} + \ldots + x_n^{-1}}{n}$, или $\bar{x} \ge \frac{n}{x_1^{-1} + \ldots + x_n^{-1}}$.

Таким образом,

{среднее арифметическое} ≥ {среднее гармоническое}. (2.83)

Средним квадратичным произвольных чисел x_i называется число

$$\sqrt{\frac{x_1^2 + \ldots + x_n^2}{n}}.$$

Так как $(x^2)'' = 2 > 0$ при любом x, то x^2 – выпуклая функция. Из неравенства (2.81) для функции $f(x) = x^2$ получаем неравенство

$$(\bar{x})^2 \le \frac{x_1^2 + \ldots + x_1^2}{n},$$

из которого следует, что

{среднее арифметическое} ≤ {среднее квадратичное}. (2.84)

С помощью средних величин можно сжимать имеющуюся информацию. Чаще всего для этого используют среднее арифметическое. Пусть, например, в народном хозяйстве занято 10 млн человек и по каждому работнику имеется информация о его доходах. При принятии экономических решений (введение налогов, изменение таможенных тарифов и т.д.) необходимо учитывать влияние этих решений на доходы населения. Однако непосредственно работать с массивом из 10 млн чисел и неудобно, и бесполезно, поскольку изменения в доходах каждого отдельного лица могут быть не связаны с принятыми решениями.

Естественно поэтому попытаться как-то сжать исходную информацию. Можно, например, рассчитать среднюю зарплату по отраслям народного хозяйства или по регионам. В любом случае работать с несколькими десятками показателей проще и зависимость усредненных данных от случайных факторов меньше.

Как уже отмечалось, для определения среднего значения чаще всего вычисляется среднее арифметическое. Тем не менее, в некоторых случаях целесообразно использовать другие средние величины. Пусть, например, имеются данные об индексах инфляции k_i по каждому из n лет ($i = 1, ..., n$). Так как k_i – это отношение уровня цен на конец i-го года к уровню цен на начало года, то за все n лет уровень цен увеличится в $k_1 k_2 ... k_n$ раз. Поэтому для определения среднего годового индекса цен лучше использовать среднее геометрическое чисел $k_1, k_2, \ldots, k_n$. Неравенство (2.82) позволяет сделать вывод о том, что средний индекс цен, полученный как среднее арифметическое, является завышенной оценкой “истинного” индекса цен, основанного на среднем геометрическом.

Пример 2.56. Пусть в течение 1, 2 и 3-го года цены увеличивались на 20%, а в течение 4 и 5-го снижались на 30%. Среднее годовое изменение уровня цен за 5 лет, полученное с помощью среднего арифметического, составит $\frac{1}{5}(20 + 20 + 20 - 30 - 30) = 0\%$, тогда как среднее геометрическое изменение цен будет

$$\left(\sqrt[5]{1{,}2 \cdot 1{,}2 \cdot 1{,}2 \cdot 0{,}7 \cdot 0{,}7} - 1\right) \cdot 100\% = -3{,}3\%$$

за год. Поскольку за 5 лет уровень цен действительно понизился, то данный пример подтверждает целесообразность применения среднего геометрического при определении среднего индекса цен за ряд последовательных лет.

Рассмотрим теперь пример, в котором “правильным” средним является среднее гармоническое. Пусть в обращении имеется n наличных рублей. Пусть $t_i\left(i=1,...,n\right)$ – среднее время, в течение которого рубль находился в собственности одного лица. Среднее время, в течение которого каждый рубль принадлежит одному лицу, можно определить как среднее арифметическое чисел $t_1,...,t_n$. Можно, однако, поступить по-другому: вычислить число оборотов каждого рубля за год по формуле

$$k_i=\frac{1}{t_i},$$

затем найти среднее арифметическое числа оборотов

$$\overline{k}=\frac{k_1+...+k_n}{n}$$

и, наконец, получить среднее время по формуле

$$t_{cp}=\frac{1}{\overline{k}}=\frac{n}{\frac{1}{t_1}+...+\frac{1}{t_n}}. \tag{2.85}$$

Пусть V – суммарный объем за год всех операций с участием наличных рублей. Среднее время, рассчитанное по формуле (2.85), является более предпочтительным, поскольку позволяет установить простую связь между массой наличных денег n и объемом операций V. Действительно, $V=k_1+...+k_n$, поэтому $n=t_{cp.}\cdot V$. Из неравенства (2.83) следует, что среднее время, рассчитанное по формуле

$$t_{cp.}=\frac{t_1+...+t_n}{n},$$

является завышенной оценкой среднего времени, рассчитанного по формуле (2.85).

Пример использования среднего квадратичного будет приведен в следующем параграфе.

§ 2.18. Формула Тейлора

Пусть функция $f(x)$ имеет n производных в точке x_0. Многочлен

$$T(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0)^1 + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

называется *n-м многочленом Тейлора* функции $f(x)$ в точке x_0. Очевидно, что в точке x_0 выполняется равенство $T(x_0) = f(x_0)$. Найдем первую производную многочлена Тейлора

$$T'(x) = f'(x_0) + \frac{f''(x_0)}{1!}(x - x_0)^1 + \frac{f'''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{(n-1)!}(x - x_0)^{n-1}.$$

Из полученной формулы следует, что $T'(x_0) = f'(x_0)$. Вторая производная многочлена Тейлора имеет вид

$$T''(x) = f''(x_0) + \frac{f'''(x_0)}{1!}(x - x_0)^1 + \ldots + \frac{f^{(n)}(x_0)}{(n-2)!}(x - x_0)^{n-2},$$

для нее также $T''(x_0) = f''(x_0)$. Аналогично $T'''(x_0) = f'''(x_0)$ и т.д. Таким образом, для любого k от 1 до n выполняется равенство

$$T^{(k)}(x_0) = f^{(k)}(x_0). \tag{2.86}$$

Теорема 2.27. *Пусть функция $f(x)$ имеет $(n + 1)$ производных в ε-окрестности точки x_0. Тогда для любой точки x из этой окрестности найдется точка c, расположенная между точками x и x_0, для которой выполняется следующая формула*

$$f(x) = T(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}, \tag{2.87}$$

где $T(x)$ – n-й многочлен Тейлора функции $f(x)$ в точке x_0.

Доказательство. Пусть $r(x) = f(x) - T(x)$. С учетом того, что $T(x_0) = f(x_0)$, имеем $r(x_0) = 0$. Определим еще одну функцию:

$$\varphi(x) = (x - x_0)^{n+1}.$$

Эта функция также, очевидно, обращается в нуль в точке x_0. Благодаря этому можем применить теорему Коши для преобразования частного функций $r(x)$ и $\varphi(x)$:

$$\frac{r(x)}{\varphi(x)}=\frac{r(x)-r(x_0)}{\varphi(x)-\varphi(x_0)}=\frac{r'(x_1)}{\varphi'(x_1)},$$

где x_1 – некоторая точка, расположенная между точками x_0 и x. Из формулы (2.86) следует, что $r'(x_0)=0$.

Кроме того, $\varphi'(x)=(n+1)(x-x_0)^n$, откуда $\varphi'(x_0)=0$. Применяя теорему Коши еще раз, получим

$$\frac{r'(x_1)}{\varphi'(x_1)}=\frac{r'(x_1)-r'(x_0)}{\varphi'(x_1)-\varphi'(x_0)}=\frac{r''(x_2)}{\varphi''(x_2)},$$

где точка x_2 расположена между x_0 и x_1. Применяя теорему Коши аналогичным образом $(n+1)$ раз, получим цепочку равенств

$$\frac{r(x)}{\varphi(x)}=\frac{r'(x_1)}{\varphi'(x_1)}=\frac{r''(x_2)}{\varphi''(x_2)}=\ldots=\frac{r^{(n+1)}(x_{n+1})}{\varphi^{(n+1)}(x_{n+1})},$$

где каждая точка x_{k+1} расположена между x_0 и x_k ($k = 1, ..., n$).

Так как $\varphi^{(n+1)}(x_{n+1})=(n+1)!$, а также

$$r^{(n+1)}(x_{n+1})=f^{(n+1)}(x_{n+1})-T^{(n+1)}(x_{n+1})=f^{(n+1)}(x_{n+1}),$$

то

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

Положив $c = x_{n+1}$ и учитывая, что $f(x) = T(x) + r(x)$, получаем равенство (2.87).

Формула (2.87) называется *формулой Тейлора*, а выражение

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

– *остаточным членом в форме Лагранжа*.

Положив $x_0 = 0$ в (2.87), получим *формулу Маклорена*:

$$f(x) = f(0) + \frac{f'(0)}{1!}x^1 + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1},$$

где c – точка, расположенная между 0 и x.

Положив $\Delta x = x - x_0$, $x = x_0 + \Delta x$, запишем формулу Тейлора (2.87) в другом виде:

$$f(x_0 + \Delta x) - f(x_0) = \frac{f'(x_0)}{1!}\Delta x + \frac{f''(x_0)}{2!}(\Delta x)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(\Delta x)^n +$$

$$+ \frac{f^{(n+1)}(c)}{(n+1)!}(\Delta x)^{n+1}.$$

Для функции $y = f(x)$ в случае $n = 1$ получаем

$$\Delta y = dy + \frac{f''(c)}{2}(\Delta x)^2. \qquad (2.88)$$

Формула (2.88) позволяет оценивать ошибку в приближенном равенстве $\Delta y \approx dy$, которое мы неоднократно использовали в предыдущих параграфах. Действительно, если на промежутке X вторая производная $f''(x)$ не превосходит по модулю некоторое число M, то и абсолютная величина ошибки в приближенном равенстве $\Delta y \approx dy$ для любых $x, x_0 \in X$ не превосходит $M(\Delta x)^2/2$. В § 2.5 с помощью приближенного равенства $\Delta y \approx dy$ были получены такие приближенные равенства, как $\sin x \approx x$, $e^x \approx 1 + x$ и др. Теперь мы можем оценить их погрешность.

Пример 2.57. Показать, что ошибка в приближенном равенстве $e^x \approx 1 + x$ не превосходит $ex^2/2$ для $x \in (-\infty, 1]$.

Решение. Воспользуемся непосредственно формулой Тейлора (2.87). Первый многочлен Тейлора для функции $f(x) = e^x$ в нуле будет $T(x) = f(0) + f'(0)x = 1 + x$. Поэтому ошибка в приближенном равенстве $e^x \approx 1 + x$ совпадает с остаточным членом $\dfrac{f''(c)}{2}x^2$. Вторая производная $f''(x)=e^x$ возрастает на $\mathbb{R}$, поэтому ее максимальное значение на промежутке $(-\infty, 1]$ равно e. Следовательно,

$$0<\frac{f''(c)}{2}x^2\le\frac{e}{2}x^2,\ \text{когда}\ x\in(-\infty,1].$$

Формула Тейлора не только позволяет оценить ошибки в известных нам ранее приближенных равенствах, но и получить приближенные равенства нового типа. Предположим, что $(n + 1)$-я производная функции $f(x)$ ограничена в окрестности точки x_0. Пусть $T_n(x)$ – n-й многочлен Тейлора функции $f(x)$ в точке x_0, $r_n(x)=\dfrac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}$ – остаточный член в форме Лагранжа. Тогда $r_n(x)$ является бесконечно малой более высокого порядка, чем $(x - x_0)^n$ при $x\to x_0$. Действительно,

$$\lim_{x\to x_0}\frac{r_n(x)}{(x-x_0)^n}=\lim_{x\to x_0}\frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)=0$$

в силу ограниченности $f^{(n+1)}(c)$ в окрестности x_0. Следовательно, ошибка в приближенном равенстве

$$f(x)\approx T_n(x) \tag{2.89}$$

также является бесконечно малой более высокого порядка, чем $(x - x_0)^n$, когда $x\to x_0$. Используя равенство (2.89), можно получить, например, следующие формулы (при $x\to 0$):

$$(1+x)^\alpha\approx 1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\ldots+\frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}x^n,$$

$$e^x\approx 1+\frac{x}{1!}+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!},$$

$$\ln(1+x)\approx x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots+(-1)^{n+1}\frac{x^n}{n},$$

$$\sin x\approx x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots+\frac{(-1)^k x^{2k+1}}{(2k+1)!},$$

$$\cos x\approx 1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots+\frac{(-1)^k x^{2k}}{(2k)!}.$$

Нетрудно видеть, что в каждой из приведенных выше формул ошибка является бесконечно малой более высокого порядка относительно x^n.

Пример 2.58. Показать, что ошибка в равенстве

$$\ln(1+x) \approx x - \frac{x^2}{2} \tag{2.90}$$

не превосходит $x^3/3$ для любого $x > 0$.

Решение. Найдем второй многочлен Тейлора $T_2(x)$ функции $\ln(1 + x)$ в нуле. Имеем

$$(\ln(1+x))' = \frac{1}{x+1}, \quad (\ln(1+x))'' = -\frac{1}{(x+1)^2},$$
$$T_2(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 = x - \frac{1}{2}x^2.$$

Далее, $(\ln(1+x))''' = 2(1+x)^{-3}$, поэтому остаточный член

$$r_3(x) = \frac{f'''(c)}{6}x^3 = \frac{x^3}{3(1+c)^3}.$$

Точка c расположена между $x_0 = 0$ и $x > 0$. Поэтому $c > 0$, $(1+c)^3 > 1$. Следовательно, ошибка в равенстве (2.90) равна $\dfrac{x^3}{3(1+c)^3}$ и не превосходит $\dfrac{x^3}{3}$ для любого $x > 0$.

Формула Тейлора, точнее, равенство $f(x) \approx T_2(x)$, может применяться в экономической статистике в следующей ситуации. Предположим, что для чисел $x_1, x_2, ..., x_n$ известно среднее арифметическое

$$a = \frac{x_1 + ... + x_n}{n}$$

и *среднее квадратичное отклонение*

$$\sigma = \sqrt{\frac{(x_1 - a)^2 + ... + (x_n - a)^2}{n}}.$$

Как найти среднее арифметическое вида

$$\bar{y} = \frac{f(x_1) + \ldots + f(x_n)}{n},$$

если числа $x_1, x_2, \ldots, x_n$ неизвестны, но известен отрезок, в котором они все содержатся? Конечно, точное значение $\bar{y}$ определить невозможно. Однако мы сумеем найти $\bar{y}$ приближенно. При этом ошибка будет тем меньше, чем меньше максимальное значение $f'''(x)$ на отрезке, содержащем все $x_1, x_2, \ldots, x_n$.

Поступим следующим образом: заменим $f(x)$ на ее второй многочлен Тейлора в точке a:

$$f(x) \approx T_2(x) = f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2.$$

Тогда

$$\bar{y} \approx \frac{1}{n} \sum_{i=1}^{n} \left(f(a) + f'(a)(x_i - a) + \frac{1}{2} f''(a)(x_i - a)^2 \right) =$$

$$= \frac{1}{n} nf(a) + f'(a) \left[\frac{1}{n} \sum_{i=1}^{n} x_i - \frac{1}{n} na \right] + \frac{1}{2n} f''(a) \sum_{i=12}^{n} (x_i - a)^2.$$

С учетом того, что $\frac{1}{2n} \sum_{i=1}^{n} (x_i - a)^2 = \frac{1}{2} \sigma^2$, получим

$$\bar{y} \approx f(a) + \frac{1}{2} f''(a) \sigma^2. \tag{2.91}$$

П р и м е р 2.59. Для чисел $x_1, \ldots, x_{100}$ известны среднее арифметическое a = 2 и среднее квадратичное отклонение σ = 0,1. Найти приближенно сумму кубов $x_1^3 + \ldots + x_{100}^3$.

Р е ш е н и е. Для функции $y = f(x) = x^3$ найдем среднее арифметическое по формуле (2.91):

$$\bar{y} = \frac{x_1^3 + \ldots + x_{100}^3}{100} \approx a^3 + \frac{1}{2} 6a\sigma^2 = 8{,}06.$$

Откуда следует, что $x_1^3 + \ldots + x_{100}^3 \approx 806$.

Пример 2.60. Для положительных чисел $x_1, ..., x_n$ известно среднее арифметическое a и среднее квадратичное отклонение σ. Найти (приближенно) среднее геометрическое $\sqrt[n]{x_1...x_n}$.

Решение. Среднее геометрическое можно представить как

$$\sqrt[n]{x_1...x_n} = e^{\frac{1}{n}(\ln x_1 + ... + \ln x_n)}.$$

Используя формулу (2.91) для функции $f(x) = \ln x$, получим

$$\frac{1}{n}(\ln x_1 + ... + \ln x_n) \approx \ln a - \frac{\sigma^2}{2a^2}.$$

Следовательно, среднее геометрическое

$$\sqrt[n]{x_1...x_n} \approx ae^{-\frac{\sigma^2}{2a^2}}. \qquad (2.92)$$

Пример 2.61. Пусть p_i – стоимость потребительской корзины на 31 декабря i-го года, $k_i = p_i / p_{i-1}$ – индекс потребительских цен за i-й год (i = 0, 1, ..., 10). Известно, что среднее арифметическое чисел k_1, ..., k_{10} равно 1, а среднее квадратичное отклонение σ = 0,1. Определить (приближенно) относительное изменение потребительских цен с 31 декабря нулевого года по 31 декабря десятого года.

Решение. Используя формулу (2.92), найдем приближенно среднее геометрическое

$$\sqrt[10]{k_1...k_{10}} \approx e^{-0,005}.$$

Далее, $\frac{p_{10}}{p_0} = k_1...k_{10} = \left(e^{-0,005}\right)^{10} = e^{-0,05} \approx 0,95.$ Следовательно, за десять лет цены уменьшились приблизительно на 5%.

ГЛАВА 3

ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 3.1. Частные производные

Для упрощения записи мы ограничимся случаем функций двух переменных. Все дальнейшее справедливо, однако, и в том случае, когда число переменных равно трем, четырем и т.д.

Итак, пусть в некоторой окрестности точки (x_0, y_0) определена функция $z = f(x, y)$. Определим приращения переменных x и y формулами:

$$\Delta x = x - x_0,\ \Delta y = y - y_0.$$

Тогда любая точка (x, y) из окрестности точки (x_0, y_0) может быть представлена как

$$(x, y) = (x_0 + \Delta x, y_0 + \Delta y).$$

При изменении x от x_0 до $x_0 + \Delta x$ (и постоянном $y = y_0$) функция z изменится на величину

$$\Delta_x z = f(x_0 + \Delta x, y_0) - f(x_0, y_0).$$

Эта разность называется *частным приращением функции z по x*. Частное приращение по y определяется аналогично:

$$\Delta_y z = f(x_0, y_0 + \Delta y) - f(x_0, y_0).$$

Определение. ***Частной производной*** *функции нескольких переменных по одной из этих переменных называется предел отношения частного приращения функции к приращению соответствующей независимой переменной, когда это приращение стремится к нулю.*

Частные производные функции $z = f(x, y)$ в точке (x_0, y_0) обозначаются так:

$$z'_x,\ \frac{\partial z}{\partial x},\ f'_x(x_0,y_0) \text{ – производная по } x;$$

$$z'_y,\ \frac{\partial z}{\partial y},\ f'_y(x_0,y_0) \text{ – производная по } y.$$

Имеем следующие равенства:

$$z'_x = \lim_{\Delta x \to 0} \frac{\Delta_x z}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x}, \tag{3.1}$$

$$z'_y = \lim_{\Delta y \to 0} \frac{\Delta_y z}{\Delta y} = \lim_{\Delta y \to 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y}. \tag{3.2}$$

Из определения частных производных следует, что для отыскания частной производной $f'_x(x, y)$ можно использовать правила дифференцирования функций одной переменной, считая константой y. Аналогично для нахождения $f'_y(x, y)$ константой следует считать x.

Пример 3.1. Найти частные производные функций: а) $z = \frac{x}{y}$; б) $z = x^y$.

Решение. а) Считая $y = \text{const}$, находим $z'_x = \frac{1}{y}$. Считая $x = \text{const}$, находим $z'_y = -\frac{x}{y^2}$. б) Производная $z'_x = yx^{y-1}$ вычисляется как производная степенной функции ($y = \text{const}$). Производная $z'_y = x^y \ln x$ вычисляется как производная показательной функции ($x = \text{const}$).

По аналогии с функциями одной переменной вводится понятие односторонней частной производной: *правой* (*левой*) *частной производной* функции называется предел отношения частного приращения функции к приращению соответствующей независимой переменной, когда это приращение стремится к нулю справа (слева).

Для функции $z = f(x, y)$ односторонние частные производные будем обозначать:

$$z'_{x+} = f'_{x+}(x_0, y_0) = \lim_{\Delta x \to 0+0} \frac{\Delta z}{\Delta x} \text{ и } z'_{x-} = f'_{x-}(x_0, y_0) = \lim_{\Delta x \to 0-0} \frac{\Delta z}{\Delta x}.$$

Соответственно, частные односторонние производные по y обозначаются $z'_{y\pm}$ или $f'_{y\pm}(x_0, y_0)$. Так как частная производная фактически является обычной производной (с той лишь оговоркой, что все независимые переменные, отличные от той, по которой выполняется дифференцирование, считаются константами), то для существования частной производной, скажем z'_x, в некоторой точке (x_0, y_0) необходимо и достаточно, чтобы в этой точке существовали и были равны обе односторонние производные:

$$z'_{x+} = z'_{x-}.$$

Пример 3.2. Для функции $z = |x - y|$ найти множество точек, в которых существуют обе частные производные z'_x и z'_y.

Решение. Используя правила дифференцирования функций одной переменной, находим при $x \neq y$:

$$z'_x = \frac{x - y}{|x - y|} \ (y = \text{const}) \quad \text{и} \quad z'_y = \frac{y - x}{|x - y|} \ (x = \text{const}).$$

Таким образом, обе частные производные существуют в каждой точке (x_0, y_0), для которой $x_0 \neq y_0$. Если же $x_0 = y_0$, то $z = 0$ и частные приращения соответственно равны $\Delta_x z = |\Delta x|$ и $\Delta_y z = |\Delta y|$. Поэтому частные односторонние производные будут:

$$z'_{x\pm} = \lim_{\Delta x \to 0\pm 0} \frac{|\Delta x|}{\Delta x} = \pm 1; \quad z'_{y\pm} = \lim_{\Delta x \to 0\pm 0} \frac{|\Delta y|}{\Delta y} = \pm 1.$$

Так как $z'_{x+} \neq z'_{x-}$ и $z'_{y+} \neq z'_{y-}$, то во всех точках (x_0, y_0), для которых $x_0 = y_0$, обе частные производные не существуют.

§ 3.2. Полный дифференциал и дифференцируемость функции

При одновременном изменении величин x и y функция $z = f(x, y)$ изменится на величину

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0). \qquad (3.3)$$

Величина Δz, заданная формулой (3.3), называется *полным приращением функции z в точке* (x_0, y_0). Так же как и в случае функции одной переменной, возникает задача о приближенной замене приращения Δz (которое, как правило, является нелинейной функцией от Δx и Δy) на линейную функцию от Δx и Δy. Роль линейного приближения выполняет *полный дифференциал функции*, который определяется как сумма произведений частных производных функции на приращения независимых переменных. Так, в случае функции от двух переменных, полный дифференциал определяется равенством

$$dz = z'_x \Delta x + z'_y \Delta y. \qquad (3.4)$$

В формуле (3.4) точка (x_0, y_0) явно не указана, однако следует помнить, что в различных точках (x_0, y_0) дифференциал будет различным.

Пример 3.3. Найти полный дифференциал функции $z = \dfrac{x}{y}$ в точках: а) (0; 2), б) (1; 1).

Решение. а) $dz\,|_{(0;2)} = \dfrac{1}{y}\,|_{(0;2)} \cdot \Delta x - \dfrac{x}{y^2}\,|_{(0;2)} \cdot \Delta y = \dfrac{1}{2}\Delta x.$

б) $dz\,|_{(1;1)} = \dfrac{1}{y}\,|_{(1;1)} \cdot \Delta x - \dfrac{x}{y^2}\,|_{(1;1)} \cdot \Delta y = \Delta x - \Delta y.$

В главе 2 было приведено несколько примеров, демонстрирующих полезность приближенного равенства $\Delta y \approx dy$ для функции $y = f(x)$ одной переменной. Успех применения приближенного равенства $\Delta y \approx dy$ объясняется тем, что его ошибка (для дифференцируемой функции $f(x)$) является бесконечно малой более высокого порядка, чем Δx, при $\Delta x \to 0$. Аналогичную роль для функции z от двух переменных играет приближенное равенство

$$\Delta z \approx dz. \tag{3.5}$$

При этом желательно, чтобы ошибка от замены Δz на dz была бесконечно малой более высокого порядка, чем расстояние ρ от точки (x, y) до точки (x_0, y_0), когда $\Delta x \to 0$, $\Delta y \to 0$. Для выполнения этого требования *недостаточно* (см. § 3.3) существования частных производных z'_x и z'_y в точке (x_0, y_0), поэтому вводится следующее понятие.

Определение. *Функция $z = f(x, y)$ называется* ***дифференцируемой*** *в точке (x_0, y_0), если ее полное приращение можно представить в виде*

$$\Delta z = f(x, y) - f(x_0, y_0) = f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y + \varepsilon\rho \tag{3.6}$$

или, короче,

$$\Delta z = dz + \varepsilon\rho,$$

где $\varepsilon = \varepsilon(\Delta x, \Delta y)$ – функция, бесконечно малая при $\Delta x \to 0$, $\Delta y \to 0$; $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ *– расстояние от точки (x, y) до точки (x_0, y_0).*

Если функция $z = f(x, y)$ дифференцируема в точке (x_0, y_0), то ее полное приращение можно представить в виде (3.6), откуда ошибка в приближенном равенстве (3.5) будет $\Delta z - dz = \varepsilon\rho$. Следовательно,

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{\Delta z - dz}{\rho} = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \varepsilon(\Delta x, \Delta y) = 0,$$

т.е. сформулированное выше требование малости ошибки выполняется для дифференцируемой функции.

Замечание. Дифференцируемость функции $z = f(x, y)$ в точке (x_0, y_0) предполагает наличие частных производных z'_x и z'_y в этой точке. Поэтому если хотя бы одна из указанных производных не существует, то функция не является дифференцируемой в точке (x_0, y_0).

Запишем линейный аналог уравнения (3.6), отбросив последнее слагаемое $\varepsilon\rho$:

$$z - f(x_0, y_0) = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0). \qquad (3.7)$$

Уравнение (3.7) в координатах x, y, z задает плоскость, которая называется *касательной плоскостью к графику функции $f(x, y)$ в точке $M(x_0, y_0, f(x_0, y_0))$*.

Можно доказать, что для любой сходящейся к M последовательности принадлежащих графику функции f и отличных от M точек $N_1, N_2, \ldots$, угол между прямой MN_i и касательной плоскостью (3.7) стремится к нулю.

Пример 3.4. Найти уравнение касательной плоскости к графику функции $z = xy$ в точке (3; 4).

Решение. Используя формулу (3.7), находим уравнение касательной плоскости:

$$z - 12 = 4(x - 3) + 3(y - 4),$$

или

$$4x + 3y - z = 12.$$

Теорема 3.1. *Если функция $z = f(x, y)$ дифференцируема в точке (x_0, y_0), то она непрерывна в этой точке.*

Доказательство. Необходимо проверить, что $\Delta z \to 0$, когда $\Delta x \to 0$ и $\Delta y \to 0$. Этот факт устанавливается следующим вычислением:

$$\lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \Delta z = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} (dz + \varepsilon\rho) = z'_x \cdot \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \Delta x + z'_y \cdot \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \Delta y + \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \varepsilon \cdot \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \rho =$$

$$= z'_x \cdot 0 + z'_y \cdot 0 + 0 \cdot 0 = 0.$$

§ 3.3. Достаточные условия дифференцируемости

Напомним, что дифференцируемость функции одной переменной $y = f(x)$ в точке x_0 равнозначна существованию конечной производной $f'(x_0)$. Поэтому можно предположить, что для функции двух переменных $z = f(x, y)$ дифференцируемость в точке (x_0, y_0) равнозначна существованию двух конечных частных производных $f'_x(x_0, y_0)$ и $f'_y(x_0, y_0)$. Однако, как показывает пример функции

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad (x^2 + y^2 \neq 0), \qquad f(0, 0) = 0,$$

это предположение неверно. Действительно, в точке (0;0) частные приращения $\Delta_x z$ и $\Delta_y z$ равны нулю при любых Δx и Δy. Поэтому частные производные z'_x и z'_y в точке (0;0) существуют и также равны нулю. Далее, при $x = y \neq 0$ функция $z = f(x, x) = \dfrac{1}{2}$, в то время как $f(0, 0) = 0$. Таким образом, $f(x, y)$ не стремится к $f(0, 0)$, когда $x \to 0$ и $y \to 0$, т.е. функция разрывная. Функция $z = f(x, y)$ недифференцируема в точке (0; 0), так как если бы она была дифференцируемой, то в силу теоремы 3.1 она была бы и непрерывной в этой точке.

Теорема 3.2 (***достаточное условие дифференцируемости***). *Если частные производные $f'_x(x, y)$ и $f'_y(x, y)$ определены в окрестности точки (x_0, y_0) и непрерывны в самой точке (x_0, y_0), то функция $z = f(x, y)$ дифференцируема в этой точке.*

Доказательство. Рассмотрим в координатной плоскости Oxy точки $P(x_0, y_0)$, $Q(x_0 + \Delta x, y_0)$ и $R(x_0 + \Delta x, y_0 + \Delta y)$ (рис. 3.1). Пусть частные производные определены в некоторой ε-окрестности точки P и точка R принадлежит данной окрестности.

Так как ε-окрестность точки P – круг радиуса ε с центром в P, то отрезки PQ и QR целиком содержатся в этой окрестности. Следовательно, функция $f(x, y)$ определена на отрезках PQ и QR.

Представим полное приращение функции $z = f(x, y)$ в точке (x_0, y_0) в виде

$$\Delta z = f(R) - f(P) = [f(R) - f(Q)] + [f(Q) - f(P)]. \qquad (3.8)$$

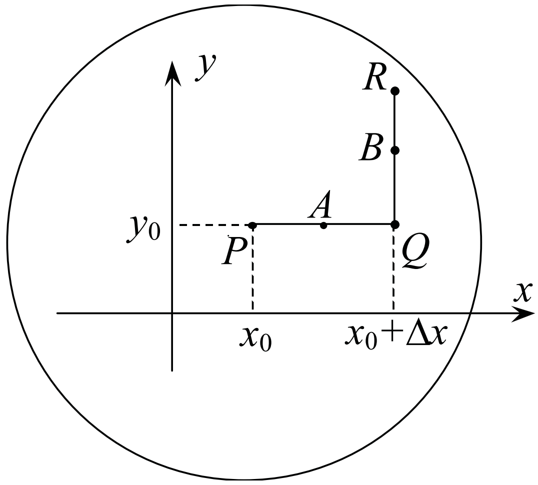

Рис. 3.1

Так как на отрезке PQ переменная y имеет постоянное значение $y = y_0$, то функция $f(x, y)$ на отрезке PQ является фактически функцией одной переменной x. Применяя формулу Лагранжа, получим

$$f(Q) - f(P) = f'_x(A)\Delta x \tag{3.9}$$

для некоторой точки A из отрезка PQ. Аналогично, на отрезке QR функция $f(x, y)$ зависит только от y. Поэтому на отрезке QR найдется точка B, для которой

$$f(R) - f(Q) = f'_y(B)\Delta y. \tag{3.10}$$

Учитывая равенства (3.9) и (3.10), запишем формулу (3.8) в следующем виде:

$$\Delta z = f'_x(A)\Delta x + f'_y(B)\Delta y.$$

Отсюда получаем

$$\begin{aligned}&\Delta z = [f'_x(P) + f'_x(A) - f'_x(P)]\Delta x + [f'_y(P) + f'_y(B) - f'_y(P)]\Delta y = \\ &= f'_x(P)\Delta x + f'_y(P)\Delta y + \alpha\Delta x + \beta\Delta y = \\ &= dz + \alpha\Delta x + \beta\Delta y,\end{aligned}$$

где $\alpha = f'_x(A) - f'_x(P)$, $\beta = f'_y(B) - f'_y(P)$.

Ясно, что при $\Delta x \to 0$, $\Delta y \to 0$ точки A и B стремятся к точке P. Так как частные производные непрерывны, то $\alpha \to 0$ и $\beta \to 0$, когда $\Delta x \to 0$, $\Delta y \to 0$. Вместе с α и β стремится к нулю и величина

$$\varepsilon = \frac{\alpha\Delta x + \beta\Delta y}{\sqrt{(\Delta x)^2 + (\Delta y)^2}}.$$

Поэтому из равенства

$$\Delta z = dz + \frac{\alpha\Delta x + \beta\Delta y}{\sqrt{(\Delta x)^2 + (\Delta y)^2}}\sqrt{(\Delta x)^2 + (\Delta y)^2} = dz + \varepsilon\rho$$

вытекает дифференцируемость функции $z = f(x, y)$ в точке (x_0, y_0).

Пример 3.5. Найти множество точек, в которых дифференцируемы функции: а) $z = |x - y|$; б) $z = (x^2 + y^2)^{3/4}$.

Решение. а) Частные производные z'_x и z'_y, найденные в примере 3.2, определены и непрерывны во всех точках (x_0, y_0), для которых $x_0 \neq y_0$. По теореме 3.2 функция z дифференцируема в таких точках. В других точках функция z недифференцируема, так как в них не существуют частные производные z'_x и z'_y. Искомое множество можно записать так: $\mathbb{R}^2 \setminus \{x = y\}$.

б) Находим частные производные:

$$z'_x = \frac{3}{2}x(x^2 + y^2)^{-\frac{1}{4}}, \qquad z'_y = \frac{3}{2}y(x^2 + y^2)^{-\frac{1}{4}}.$$

Эти функции являются элементарными и определены во всех точках, за исключением точки (0; 0). Вследствие теоремы 3.2 функция z дифференцируема во всех точках за исключением, быть может, точки (0; 0). Исследуем эту точку отдельно. Частные приращения будут $\Delta_x z = |x|^{3/2}$ и $\Delta_y z = |y|^{3/2}$. Следовательно,

$$z'_x = \lim_{\Delta x \to 0} \frac{|\Delta x|^{3/2}}{\Delta x} = 0, \ z'_y = \lim_{\Delta y \to 0} \frac{|\Delta y|^{3/2}}{\Delta y} = 0$$

и дифференциал dz также равен нулю: $dz = 0\cdot\Delta x + 0\cdot\Delta y = 0$. Так как в точке $(0;0)$ и значение $z = 0$ и дифференциал $dz = 0$, то полное приращение можно записать следующим образом:

$$\Delta z = z = dz + z = dz + \varepsilon\cdot\rho,$$

где $\rho = (x^2 + y^2)^{1/2}$, $\varepsilon = (x^2 + y^2)^{1/4} \to 0$, когда $\Delta x \to 0$, $\Delta y \to 0$. Поэтому функция z дифференцируема в точке (x_0, y_0). Итак, множество всех точек, в которых функция z дифференцируема, есть $\mathbb{R}^2$.

§ 3.4. Дифференцируемость сложной функции

Пусть $f(x, y)$ – функция от двух переменных x и y, а $\varphi(t)$ и $\psi(t)$ – функции от независимой переменной t. В этом случае говорят, что $F(t) = f(\varphi(t), \psi(t))$ – *сложная функция* от t.

Теорема 3.3. *Если функции $x = \varphi(t)$ и $y = \psi(t)$ дифференцируемы в точке t_0, а функция $z = f(x, y)$ дифференцируема в точке $(\varphi(t_0), \psi(t_0))$, то сложная функция $F(t) = f(\varphi(t), \psi(t))$ также дифференцируема в t_0. При этом производная сложной функции находится по формуле*

$$F'(t_0) = f'_x(\varphi(t_0),\psi(t_0))\varphi'(t_0) + f'_y(\varphi(t_0),\psi(t_0))\psi'(t_0),$$

или, короче,

$$z'_t = z'_x \cdot x'_t + z'_y \cdot y'_t\,. \tag{3.11}$$

Доказательство. Вследствие дифференцируемости $f(x, y)$ в точке (x_0, y_0), где $x_0 = \varphi(t_0)$, $y_0 = \psi(t_0)$, имеем

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) =$$

$$= z'_x\Delta x + z'_y\Delta y + \varepsilon(\Delta x, \Delta y)\sqrt{(\Delta x)^2 + (\Delta y)^2}. \tag{3.12}$$

Здесь Δx, Δy – произвольные (достаточно малые) приращения переменных x, y, а $\varepsilon(\Delta x, \Delta y)$ – бесконечно малая (при $\Delta x \to 0$, $\Delta y \to 0$)

функция. Поскольку равенство (3.12) справедливо при любых малых Δx, Δy, оно выполняется и для приращений вида

$$\Delta x = \varphi(t) - \varphi(t_0), \qquad \Delta y = \psi(t) - \psi(t_0). \tag{3.13}$$

Подставив равенства (3.13) в формулу (3.12), после деления на Δt получим

$$\frac{\Delta z}{\Delta t} = z'_x \frac{\Delta x}{\Delta t} + z'_y \frac{\Delta y}{\Delta t} + \varepsilon(\Delta x, \Delta y)\sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2}. \tag{3.14}$$

Так как $\varphi(t)$ и $\psi(t)$ дифференцируемы в t_0, то они непрерывны в t_0. Следовательно, $\Delta x \to 0$ и $\Delta y \to 0$, когда $\Delta t \to 0$. Далее,

$$\lim_{\Delta t \to 0} \varepsilon(\Delta x, \Delta y) \cdot \sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2} = 0 \cdot \sqrt{(x'_t)^2 + (y'_t)^2} = 0,$$

поэтому, переходя к пределу при $\Delta t \to 0$ в равенстве (3.14), получим формулу (3.11). Одновременно с этим доказано, что $F(t)$ дифференцируема в точке t_0.

Пример 3.6. Пусть $z = f(x, y)$, $x = t^3$, $y = t^2 + 1$. Тогда $z'_t = z'_x 3t^2 + z'_y 2t$.

Пример 3.7. Пусть $z = x^y$, $x = \cos t$, $y = \sin t$. Тогда по формуле (3.11) имеем

$$z'_t = z'_x \cdot x'_t + z'_y \cdot y'_t = yx^{y-1} \cdot (-\sin t) + x^y \ln x \cdot \cos t =$$
$$= -(\sin t)^2 (\cos t)^{\sin t - 1} + (\cos t)^{\sin t + 1} \ln \cos t =$$
$$= (\cos t)^{\sin t + 1} (\ln \cos t - \mathrm{tg}^2 t).$$

Рассмотрим теперь более сложный случай. Пусть $z = f(x, y)$ – функция от двух переменных x и y, которые, в свою очередь, зависят от двух других переменных $x = \varphi(u, v)$, $y = \psi(u, v)$. Тогда $z =$ $= f(\varphi(u, v), \psi(u, v))$ – сложная функция от двух независимых переменных u и v. Фиксируя сначала v, а затем u и используя формулу (3.11), получим следующие соотношения:

$$\begin{cases} z'_u = z'_x x'_u + z'_y y'_u, \\ z'_v = z'_x x'_v + z'_y y'_v. \end{cases} \qquad (3.15)$$

Справедливо утверждение: *если функции* $\varphi(u, v)$ *и* $\psi(u, v)$ *дифференцируемы в точке* (u_0, v_0), *а функция* $f(x, y)$ *дифференцируема в точке* $(\varphi(u_0, v_0), \psi(u_0, v_0))$, *то сложная функция*

$$z = f(\varphi(u, v), \psi(u, v))$$

также дифференцируема в точке (u_0, v_0).

§ 3.5. Производная по направлению. Градиент

Сначала необходимо уточнить понятие направления. Ясно, что всякий вектор $\vec{v}$ задает некоторое направление. При этом, однако, сам вектор $\vec{v}$ нельзя считать направлением, поскольку любой вектор $\vec{w} = \lambda\vec{v}$ $(\lambda > 0)$ задает то же самое направление. С другой стороны, в каждом семействе $\{\lambda\vec{v}\}_{\lambda>0}$ векторов, задающих некоторое направление, всегда имеется, и притом только один, вектор $\vec{e}$ единичной длины, $|\vec{e}\,| = 1$. Данное обстоятельство позволяет нам определить *направление* как вектор единичной длины.

Пусть $\vec{e} = (e_x, e_y)$ – двумерный единичный вектор,

$$|\vec{e}\,| = \sqrt{e_x^2 + e_y^2} = 1.$$

Определение. *Производной функции* $f(x, y)$ *в точке* (x_0, y_0) ***по направлению*** $\vec{e}$ *называется предел*

$$\frac{\partial f(x_0, y_0)}{\partial \vec{e}} = \lim_{t \to 0+0} \frac{f(x_0 + te_x, y_0 + te_y) - f(x_0, y_0)}{t}.$$

Говорят также, что $\partial f / \partial \vec{e}$ – это *скорость изменения* функции по направлению $\vec{e}$.

Производная по направлению является обобщением понятия правой частной производной. Действительно, для векторов $\vec{e}_1 = (1;\ 0)$ и $\vec{e}_2 = (0;\ 1)$ имеем

$$\frac{\partial f(x_0, y_0)}{\partial \vec{e}_1} = \lim_{t \to 0+0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t} = f'_{x+}(x_0, y_0),$$

$$\frac{\partial f(x_0, y_0)}{\partial \vec{e}_2} = \lim_{t \to 0+0} \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{t} = f'_{y+}(x_0, y_0).$$

Пример 3.8. Найти производную функции $z = |\,x - y\,|$ в точке $(0;\ 0)$ по направлению единичного вектора $\vec{e} = (e_x, e_y)$.

Решение. Исходя непосредственно из определения, имеем

$$\frac{\partial z}{\partial \vec{e}} = \lim_{t \to 0+0} \frac{|\,te_x - te_y\,| - |\,0\,|}{t} = \lim_{t \to 0+0} \frac{t\,|\,e_x - e_y\,|}{t} = |\,e_x - e_y\,|.$$

Пример 3.8 показывает, что функция может иметь в некоторой точке производные по любому направлению и одновременно не быть дифференцируемой в этой точке (пример 3.5а).

При постановке задачи об определении скорости изменения функции по направлению вектора вместо единичного вектора $\vec{e}$ часто задается какой-либо ненулевой вектор $\vec{v}$ или луч l. В этом случае в качестве вектора $\vec{e}$ следует взять вектор $\vec{e} = |\,\vec{v}\,|^{-1}\vec{v}$, где $\vec{v}$ – это либо заданный вектор, либо направляющий вектор луча l. Легко доказывается, что координаты вектора $\vec{e}$ совпадают с косинусами углов, которые вектор $\vec{v}$ (луч l) образует с соответствующими осями координат.

Действительно, пусть α – угол между вектором $\vec{v} = (v_x, v_y)$ и осью Ox, а β – угол между $\vec{v}$ и осью Oy. Тогда (рис. 3.2)

$$\cos\alpha = \frac{v_x}{|\,\vec{v}\,|}, \quad \cos\beta = \frac{v_y}{|\,\vec{v}\,|}.$$

Следовательно, $l = |\,\vec{v}\,|^{-1}\vec{v} = \left(\frac{v_x}{|\,\vec{v}\,|}, \frac{v_y}{|\,\vec{v}\,|}\right) = (\cos\alpha, \cos\beta).$

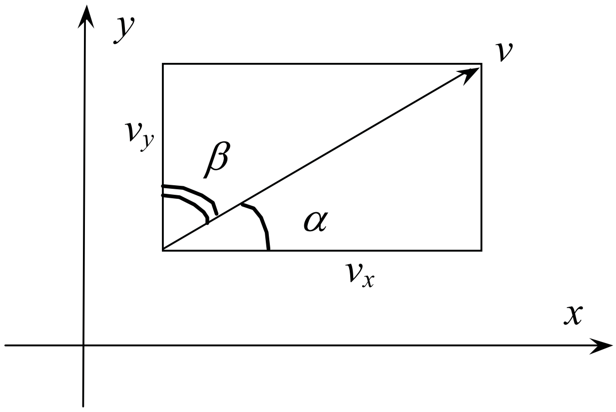

Рис. 3.2

Если функция $f(x, y)$ дифференцируема в точке $M(x_0, y_0)$, то при вычислении ее производной в направлении $\vec{e}$ в точке M вместо того, чтобы считать предел (как мы это проделали в примере 3.7), можно воспользоваться следующей теоремой.

Теорема 3.4. *Производная дифференцируемой функции* $z = f(x, y)$ *в точке* (x_0, y_0) *в направлении* $\vec{e}$ *находится по формуле*

$$\frac{\partial f(x_0, y_0)}{\partial \vec{e}} = f'_x(x_0, y_0) \cdot e_x + f'_y(x_0, y_0) \cdot e_y, \tag{3.16}$$

или

$$\frac{\partial z}{\partial \vec{e}} = z'_x \cdot e_x + z'_y \cdot e_y.$$

Доказательство. Определим дифференцируемые (по t) функции $\varphi(t)$ и $\psi(t)$, положив $\varphi(t) = x_0 + te_x$ и $\psi(t) = y_0 + te_y$. Определим также функцию

$$F(t) = f(\varphi(t),\ \psi(\text{t})) = f(x_0 + te_x, y_0 + te_y).$$

По теореме 3.3 находим производную $F(t)$ в нуле:

$$\begin{aligned} F'(0) &= f'_x(\varphi(0), \psi(0))\varphi'(0) + f'_y(\varphi(0), \psi(0))\psi'(0) = \\ &= f'_x(x_0, y_0)e_x + f'_y(x_0, y_0)e_y. \end{aligned} \tag{3.17}$$

С другой стороны,

$$F'(0) = \lim_{t\to 0}\frac{F(t)-F(0)}{t} = \lim_{t\to 0+0}\frac{F(t)-F(0)}{t} =$$
$$= \lim_{t\to 0+0}\frac{f(x_0+te_x, y_0+te_y)-f(x_0,y_0)}{t} = \frac{\partial f(x_0,y_0)}{\partial \vec{e}}. \qquad (3.18)$$

Сравнивая равенства (3.17) и (3.18), получаем формулу (3.16).

Определение. ***Градиентом функции в точке*** *M называется вектор, координаты которого равны соответствующим частным производным данной функции в точке M.*

Так, для функции двух переменных $f(x, y)$ имеем

$$\operatorname{grad} f(M) = (f'_x(M), f'_y(M)).$$

Для функции трех переменных $f(x, y, z)$ имеем

$$\operatorname{grad} f(M) = (f'_x(M), f'_y(M), f'_z(M))$$

и т.д. Используя градиент, формулу (3.16) можно записать короче:

$$\frac{\partial f(M)}{\partial \vec{e}} = (\operatorname{grad} f(M), \vec{e}), \qquad (3.19)$$

где $(\operatorname{grad} f(M), \vec{e})$ – скалярное произведение векторов.

Формула (3.19) отчетливо разграничивает роли двух факторов, от которых зависит производная по направлению: роль функции f и роль вектора $\vec{e}$.

Заметим, что ни число аргументов функции f, ни длина вектора $\vec{e}$ не играют существенной роли при выводе формулы (3.19). Это обстоятельство будет использоваться в дальнейшем, поэтому запишем более удобный вариант формулы (3.19) для функции f от n переменных и произвольного вектора $\vec{v}$.

Положим $\varphi(t) = f(\vec{p}+t\vec{v})$, где $\vec{p},\vec{v} \in \mathbb{R}^n$. Тогда выполняется следующее равенство:

$$\varphi'(0) = (\operatorname{grad} f(\vec{p}), \ \vec{v}). \qquad (3.20)$$

Пример 3.9. Найти производную функции $z = x^5y^6$ в точке (1; 1) по направлению вектора $\vec{v}$ = (3; – 2).

Решение. Вектор $\vec{v}$ задает направление – единичный вектор:

$$\vec{e} = |\vec{v}|^{-1}\vec{v} = \frac{1}{\sqrt{3^2+(-2)^2}}(3;-2) = \frac{1}{\sqrt{13}}(3;-2).$$

Градиент z в точке (x, y) равен $(z'_x; z'_y) = (5x^4y^6; 6x^5y^5)$. Поэтому

$$\text{grad}\, z(1;1) = (5; 6).$$

Используя формулу (3.19), находим

$$\frac{\partial z(1;1)}{\partial \vec{e}} = \left((5;6),\ \frac{1}{\sqrt{13}}(3;-2)\right) = \frac{3}{\sqrt{13}}.$$

Пусть $M(x_0, y_0)$ – точка, в которой вычисляется градиент функции $f(x, y)$, φ – угол между градиентом $\text{grad}f(M)$ и направлением $\vec{e}$. Так как

$$(\text{grad}f(M),\ \vec{e}) = |\text{grad}f(M)|\cdot|\vec{e}|\cdot\cos\varphi,$$

а $\vec{e}$ – единичный вектор, то из формулы (3.19) вытекает равенство

$$\frac{\partial f(M)}{\partial \vec{e}} = |\text{grad}\, f(M)|\cdot\cos\varphi.$$

Отсюда следует, что производная в направлении $\vec{e}$ принимает наибольшее значение при $\cos\varphi = 1$, т.е. когда направление $\text{grad}f(M)$ совпадает с $\vec{e}$. При этом $\dfrac{\partial f(M)}{\partial \vec{e}} = |\text{grad}\, f(M)|$.

Вывод: *градиент указывает направление наискорейшего роста функции, а максимальная скорость роста равна модулю градиента.*

§ 3.6. Касательные прямые и плоскости

Напомним, что *множеством уровня* функции n переменных $f(x_1, ..., x_n)$ называется множество всех точек $x \in \mathbb{R}^n$, координаты которых удовлетворяют уравнению $f(x_1, ..., x_n) = C$, где C – фиксированное число. В случае $n = 2$ множество уровня называется *линией уровня*, а в случае $n = 3$ – *поверхностью уровня* функции.

Отметим, что линия уровня может и не быть линией в привычном смысле этого слова. Например, множество уровня функции

$$f(x, y) = |x| + |1 - x| + |y| + |1 - y|,$$

заданное уравнением $f(x, y) = 2$, является квадратом.

В то же время если множество уровня функции двух переменных можно представить как график некоторой функции от одной переменной, то множество уровня действительно является *линией* уровня.

Примем без доказательства следующую теорему.

Теорема 3.5. *Пусть $f(x, y)$ – функция, имеющая непрерывные частные производные $f'_x(x, y)$, $f'_y(x, y)$ в некотором открытом множестве D; C – фиксированное число; $M(x_0, y_0) \in D$ – точка, такая, что $f(M) = C$ и $f'_y(M) \neq 0$. Тогда найдется такая дифференцируемая функция $\varphi(x)$, что в некоторой окрестности точки M множество решений уравнения $f(x, y) = C$ совпадает с графиком функции $y = \varphi(x)$.*

Замечание. Поскольку M – решение уравнения $f(x, y) = C$ и M содержится в любой своей окрестности, то для функции $\varphi(x)$ из теоремы 3.5 автоматически выполняется равенство $y_0 = \varphi(x_0)$.

Для функции $f(x, y)$, имеющей непрерывные частные производные, из теоремы 3.5 вытекает важное следствие: *всякая линия уровня в достаточно малой окрестности любой своей точки M может быть представлена как график дифференцируемой функции*, *если только* $\operatorname{grad} f(M) \neq 0$. Действительно, если $f'_y(M) \neq 0$, то линия уровня – это график функции $y = \varphi(x)$. Если $f'_x(M) \neq 0$, то, поменяв ролями x и y в теореме 3.5, получим, что линия уровня – это график функции $x = \psi(y)$. Кроме того, частные производные $f'_x(M)$ и

$f'_y(M)$ не могут одновременно обратиться в нуль, так как $\mathrm{grad} f(M) \neq 0$.

Поскольку локально линию уровня можно представить как график дифференцируемой функции, то теперь мы можем найти уравнение касательной к линии уровня. Пусть, как и выше, линия уровня задается уравнением

$$f(x, y) = C \tag{3.21}$$

и $M(x_0, y_0)$ – точка на этой линии уровня, т.е. $f(M) = C$. Предположим, для определенности, что $f'_y(M) \neq 0$. Тогда по теореме 3.5 уравнение (3.21) можно локально разрешить относительно y в виде $y = \varphi(x)$. Уравнение касательной к графику $y = \varphi(x)$ в точке M запишется следующим образом:

$$y = y_0 + \varphi'(x_0)(x - x_0). \tag{3.22}$$

Уравнение (3.22) – это и есть уравнение касательной к линии уровня (3.21) в точке M. Такая форма записи является достаточно эффективной, если $\varphi(x)$ – известная функция. В действительности знать явный вид функции $\varphi(x)$ необязательно, поскольку производную $\varphi'(x_0)$ можно выразить через частные производные функции $f(x, y)$.

Найдем для этого производную сложной функции $f(x, \varphi(x))$ двумя способами. С одной стороны,

$$\frac{d}{dx} f(x, \varphi(x)) = f'_x(x, \varphi(x)) \cdot 1 + f'_y(x, \varphi(x)) \cdot \varphi'(x),$$

с другой –

$$\frac{d}{dx} f(x, \varphi(x)) = \frac{d}{dx} C = 0.$$

Следовательно,

$$\varphi'(x_0) = -\frac{f'_x(M)}{f'_y(M)}. \tag{3.23}$$

Подставив выражение (3.23) в формулу (3.22), после алгебраических преобразований получим уравнение касательной к линии уровня в виде

$$f'_x(M)(x-x_0)+f'_y(M)(y-y_0)=0. \qquad (3.24)$$

Из аналитической геометрии известно, что вектор коэффициентов при x и y в уравнении прямой перпендикулярен к данной прямой и называется вектором нормали.

Вывод: *градиент функции* grad $f(M)$ *является вектором нормали касательной к линии уровня в точке* M (рис. 3.3).

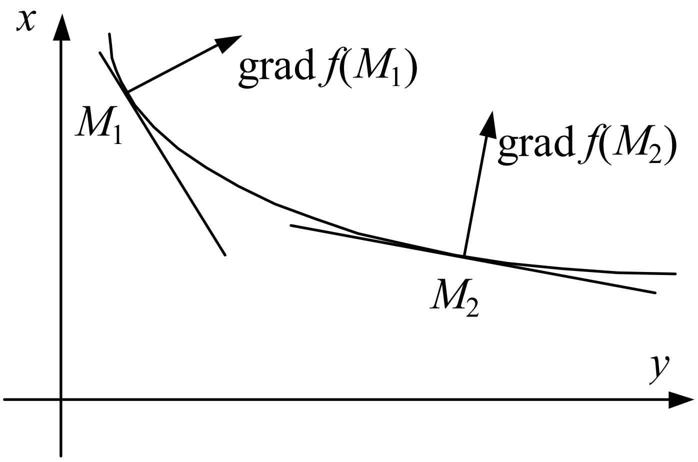

Рис. 3.3

Пример 3.10. Найти уравнение касательной к линии уровня функции $z=\sqrt[3]{xy^2}$ в точке (1; 1).

Решение. Используя формулу (3.24), получим уравнение

$$\frac{1}{3}(x-1)+\frac{2}{3}(y-1)=0, \text{ или } x+2y=3.$$

Рассмотрим теперь функцию трех переменных $f(x, y, z)$. Пусть уравнение

$$f(x,y,z)=C \quad (C=\text{const}) \qquad (3.25)$$

задает непустое множество уровня $S \neq \varnothing$. Пусть $M(x_0, y_0, z_0) \in S$ – какая-то точка из этого множества. Если частные производные функции $f(x, y, z)$ непрерывны и $\operatorname{grad} f(M) \neq 0$, то уравнение (3.25) в окрестности точки M можно разрешить либо в виде $z = \varphi(x, y)$, либо в виде $y = \psi(x, z)$, либо в виде $x = \chi(y, z)$, где φ, ψ, χ – дифференцируемые функции. Данное утверждение является аналогом теоремы 3.5 для функции трех переменных.

Следовательно, множество уровня S локально можно представить как график дифференцируемой функции двух переменных, что оправдывает его другое название – *поверхность уровня*.

С помощью рассуждений, аналогичных тем, что использовались при выводе уравнения касательной прямой к линии уровня, из уравнения (3.7) нетрудно получить уравнение *касательной плоскости к поверхности уровня* функции $f(x, y, z)$ в точке $M(x_0, y_0, z_0)$ следующего вида:

$$f'_x(M)(x - x_0) + f'_y(M)(y - y_0) + f'_z(M)(z - z_0) = 0. \qquad (3.26)$$

Вывод: *градиент* $\operatorname{grad} f(M)$ *функции* $f(x, y, z)$ *является вектором нормали касательной плоскости к поверхности уровня функции в точке* M.

Пусть, например, поверхность уровня S задается уравнением

$$x^2 + y^2 + z^2 = r^2,$$

т.е. S – сфера радиуса r с центром в начале координат. Для любой точки $M(x, y, z) \in S$ имеем $\operatorname{grad} f(M) = (2x, 2y, 2z) = 2\overrightarrow{OM}$. Следовательно, касательная плоскость перпендикулярна радиус-вектору точки касания.

Пример 3.11. Найти уравнение касательной плоскости к поверхности уровня функции $x^2 + 2y^2 + 3z^2$ в точке (1; 1; –1).

Решение. Используя формулу (3.26), получаем следующее уравнение:

$$2(1)(x - 1) + 4(1)(y - 1) + 6(-1)(z + 1) = 0, \quad \text{или} \quad x + 2y - 3z = 6.$$

§ 3.7. Предельная полезность и предельная норма замещения

Исходным понятием теории потребления является понятие *функции полезности* $U(x, y)$. Эта функция выражает меру полезности набора (x, y), где x – количество товара X, а y – количество товара Y. Чувствительность набора (x, y) к незначительному изменению x при фиксированном y называется *предельной полезностью* X и определяется как частная производная U'_x. Аналогично *предельная полезность* Y определяется как U'_y. Чаще всего линии уровня функции полезности (их еще называют кривыми безразличия) являются графиками убывающих функций. Это означает, что для точек

$$A(x_0, y_0) \text{ и } B(x_0 + \Delta x, y_0 + \Delta y),$$

расположенных на одной линии уровня, приращения Δx и Δy имеют разные знаки (рис. 3.4).

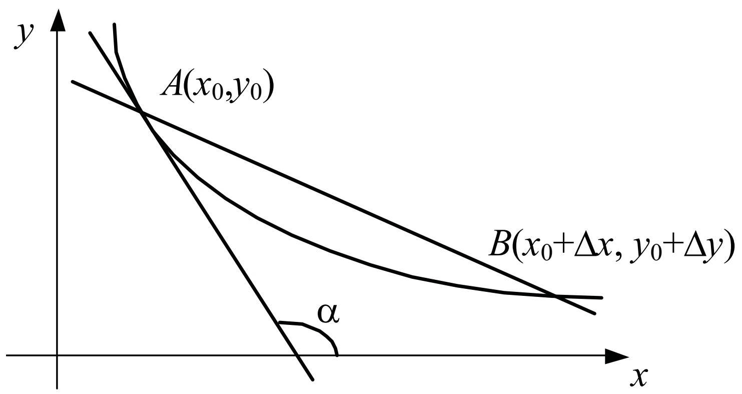

Рис. 3.4

Пусть, для определенности, $\Delta x > 0$, а $\Delta y < 0$. В этом случае говорят, что Δx единиц первого товара замещается на $(-\Delta y)$ единиц второго товара (имеется в виду переход из B в A). *Предельной нормой замещения* X *на* Y *в точке* $A(x_0, y_0)$ называется предел отношения $\frac{-\Delta y}{\Delta x}$, когда точка B стремится к A, оставаясь на одной с A линии уровня функции $U(x, y)$. Предельная норма замещения обозначается MRS_{xy}, или $MRS_{xy}(A)$, если необходимо явно указать ее зависимость от точки A.

Пусть l – касательная к линии уровня функции $U(x, y)$ в точке A. Из рис. 3.4 ясно, что секущая AB стремится к l, когда B стремится к A, поэтому

$$MRS_{xy}(A) = -\operatorname{tg}\alpha, \tag{3.27}$$

где α – угол наклона касательной l. Согласно формуле (3.24) мы можем записать уравнение l в виде

$$U'_x(A)(x - x_0) + U'_y(A)(y - y_0) = 0,$$

или

$$y = y_0 - \frac{U'_x(A)}{U'_y(A)}(x - x_0). \tag{3.28}$$

Из уравнения (3.28) следует, что $\operatorname{tg}\alpha$ (угловой коэффициент касательной) равен $-\dfrac{U'_x(A)}{U'_y(A)}$. Используя равенство (3.27), получаем формулу

$$MRS_{xy}(A) = \frac{U'_x(A)}{U'_y(A)}. \tag{3.29}$$

Вывод: *предельная норма замещения одного товара другим равна отношению их предельных полезностей*.

Пример 3.12. Найти предельную норму замещения x на y для функции полезности $U(x, y) = \ln x + \ln y$ в точках: а) (3; 12), б) (2; 1).

Решение. а) По формуле (3.29) получаем

$$MRS_{xy} = \frac{1/x}{1/y} = \frac{y}{x},$$

поэтому $MRS_{xy}(3; 12) = 4$.

б) Аналогично находим $MRS_{xy}(2; 1) = 0{,}5$.

Экономический смысл предельной нормы замещения можно показать на следующем примере. Предположим, что $MRS_{xy} = 2$, т.е. одна единица товара X по вкладу в общую полезность набора эквивалентна 2 единицам товара Y. Пусть p – цена X, q – цена Y. Тогда стоимость набора (x, y) будет $px + qy$. При замене в этом наборе одной единицы X на две единицы Y стоимость набора изменится на $2q - p$. Если же, наоборот, две единицы Y заменить на одну единицу X, то стоимость набора изменится на $p - 2q$. Предположим теперь, что (x^*, y^*) – самый дешевый набор при заданном уровне полезности $U(x, y) = \text{const}$. Тогда получаем, что указанные выше изменения стоимости должны быть неотрицательными: $2q - p \geq 0$, $p - 2q \geq 0$. Следовательно, $p = 2q$. Ясно, что двойка в приведенных выше рассуждениях может быть заменена любым другим числом, так что фактически мы обосновали равенство

$$MRS_{xy}(x^*, y^*) = \frac{p}{q}.$$

Это же самое равенство, хотя и при других предположениях, еще раз будет получено в § 3.18, где оно используется для вывода функций спроса.

§ 3.8. Эластичность функции нескольких переменных

В § 2.8 было введено понятие эластичности функции одной переменной. Аналогично вводится понятие *эластичности функции нескольких переменных*. Пусть, например, $z = f(x, y)$ – функция двух переменных;

$$\Delta_x z = f(x_0 + \Delta x, y_0) - f(x_0, y_0), \ \Delta_y z = f(x_0, y_0 + \Delta y) - f(x_0, y_0)$$

– ее частные приращения.

Определение. ***Эластичностью*** *функции* $z = f(x, y)$ *в точке* (x_0, y_0) *по* x *называется предел*

$$E_{zx}(x_0, y_0) = \lim_{\Delta x \to 0}\left(\frac{\Delta_x z}{z} : \frac{\Delta x}{x}\right).$$

Эластичностью *z по y в той же точке называется предел*

$$E_{zy}(x_0, y_0) = \lim_{\Delta y \to 0}\left(\frac{\Delta_y z}{z} : \frac{\Delta y}{y}\right).$$

Говорят, что E_{zx} – *коэффициент эластичности z по x*, а E_{zy} – *коэффициент эластичности z по y* (обозначение точки часто опускается).

Из определения вытекают следующие формулы:

$$E_{zx}(x, y) = \frac{x}{z} z'_x = x(\ln z)'_x,$$

$$E_{zy}(x, y) = \frac{y}{z} z'_y = y(\ln z)'_y. \tag{3.30}$$

Пример 3.13. Найти коэффициенты эластичности по x и по y функции $z = x^y$ в точке (2; 3).

Решение. Согласно формулам (3.30) имеем

$$E_{zx}(x, y) = x(\ln z)'_x = x(y \ln x)'_x = y,$$
$$E_{zy}(x, y) = y(\ln z)'_y = y(y \ln x)'_y = y \ln x.$$

Следовательно, $E_{zx}(2; 3) = 3$, $E_{zy}(2; 3) = 3\ln 2$.

Формулы (3.30) полностью аналогичны формулам, которые использовались при выводе свойств 1…3 эластичности в § 2.8. Поэтому первые три свойства эластичности справедливы и в случае функции нескольких переменных. Четвертое и пятое свойства также сохраняются, но форма их записи становится сложнее. Остановимся подробнее на этих свойствах.

Свойство 4. *Для функций* $z = f(x, y)$, $x = \varphi(t)$ и $y = \psi(t)$ *эластичность z по t в точке* t_0 *находится по формуле*

$$E_{zt} = E_{zx} E_{xt} + E_{zy} E_{yt}, \tag{3.31}$$

где E_{zx} *и* E_{zy} *– эластичности z по x и по y в точке* $(\varphi(t_0), \psi(t_0))$*, а* E_{xt} *и* E_{yt} *– эластичности x и y по t в точке* t_0.

Доказательство. Используя формулы (3.30) и (3.31), находим

$$E_{zt} = \frac{t}{z} z'_t = \frac{t}{z}(z'_x x'_t + z'_y y'_t) = \frac{x}{z} z'_x \frac{t}{x} x'_t + \frac{y}{z} z'_y \frac{t}{y} y'_t = E_{zx} E_{xt} + E_{zy} E_{yt}.$$

Для того чтобы сформулировать аналог пятого свойства эластичности, необходимо следующее определение.

Определение[1]**.** *Пусть* $X \subset \mathbb{R}^2$. *Отображение* $x_1 = g_1(y_1, y_2)$, $x_2 = g_2(y_1, y_2)$ *называется* ***обратным*** *для отображения* $y_1 = f_1(x_1, x_2)$, $y_2 = f_2(x_1, x_2)$, $(x_1, x_2) \in X$, *если для любой точки* (x_1, x_2) *из* X *выполняются равенства*:

$$x_1 = g_1(f_1(x_1, x_2), f_2(x_1, x_2));$$
$$x_2 = g_2(f_1(x_1, x_2), f_2(x_1, x_2)).$$

Пример 3.14. Найти обратное отображение для отображения $y_1 = x_1 + x_2^2$, $y_2 = x_1 - 2x_2$, где $(x_1, x_2) \in \mathbb{R}_+^2$.

Решение. Из равенства $y_2 = x_1 - 2x_2$ следует, что $x_1 = y_2 + 2x_2$. Поэтому $y_1 = y_2 + 2x_2 + x_2^2 = y_2 - 1 + (1 + x_2)^2$. Находим $x_2 = \pm\sqrt{y_1 - y_2 + 1} - 1$, однако, учитывая, что $(x_1, x_2) \in \mathbb{R}_+^2$, имеем $x_2 = \sqrt{y_1 - y_2 + 1} - 1$. Следовательно, $x_1 = y_2 + 2\sqrt{y_1 - y_2 + 1} - 2$.

Для любой пары функций $y_1 = f_1(x_1, x_2)$, $y_2 = f_2(x_1, x_2)$ имеем 4 коэффициента эластичности $E_{y_i x_j}$ $(i = 1, 2; j = 1, 2)$. Записав их в виде таблицы, получим матрицу размером 2×2:

$$E_{yx} = \begin{pmatrix} E_{y_1 x_1} & E_{y_1 x_2} \\ E_{y_2 x_1} & E_{y_2 x_2} \end{pmatrix}.$$

Элементы этой матрицы, расположенные вне главной диагонали, называются *перекрестными коэффициентами эластичности*.

[1] Данное определение является частным случаем определения обратного отображения $g = f^{-1} : Y \to X$ для отображения $f : X \to Y$, где $X, Y \subset \mathbb{R}^n$.

Свойство 5. *Пусть* $x_1 = g_1(y_1, y_2)$, $x_2 = g_2(y_1, y_2)$ – *обратное отображение для отображения* $y_1 = f_1(x_1, x_2)$, $y_2 = f_2(x_1, x_2)$. *Тогда матрица коэффициентов эластичности* E_{xy} *является обратной к матрице* E_{yx}.

Доказательство. Умножим матрицы E_{xy} и E_{yx}, используя полученную выше формулу (3.31):

$$E_{xy}E_{yx} = \begin{pmatrix} E_{x_1y_1}E_{y_1x_1} + E_{x_1y_2}E_{y_2x_1} & E_{x_1y_1}E_{y_1x_2} + E_{x_1y_2}E_{y_2x_2} \\ E_{x_2y_1}E_{y_1x_1} + E_{x_2y_2}E_{y_2x_1} & E_{x_2y_1}E_{y_1x_2} + E_{x_2y_2}E_{y_2x_2} \end{pmatrix} =$$

$$= \begin{pmatrix} E_{x_1x_1} & E_{x_1x_2} \\ E_{x_2x_1} & E_{x_2x_2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Так как в результате умножения получилась единичная матрица, то E_{xy} и E_{yx} – взаимно обратные матрицы.

Коэффициенты эластичности используются для анализа функций спроса при любом числе различных товаров. В качестве примера рассмотрим случай с двумя товарами. Пусть x_i – количество i-го товара, p_i – его цена ($i = 1, 2$). Для пары дополняющих товаров (например, чай и сахар) или замещающих товаров (например, масло и маргарин) естественно считать, что спрос на каждый товар зависит от обеих цен p_1 и p_2:

$$\begin{cases} x_1 = D_1(p_1, p_2), \\ x_2 = D_2(p_1, p_2). \end{cases} \tag{3.32}$$

Предположим, что не только цены определяют спрос, но и, напротив, спрос определяет цены. Иными словами, будем считать, что систему (3.32) можно разрешить относительно p_1 и p_2 в следующем виде:

$$\begin{cases} p_1 = p_1(x_1, x_2), \\ p_2 = p_2(x_1, x_2). \end{cases} \tag{3.33}$$

Системы (3.32) и (3.33) определяют взаимно обратные отображения $(p_1, p_2) \mapsto (x_1, x_2)$ и $(x_1, x_2) \mapsto (p_1, p_2)$. Согласно свойству 5 матрица коэффициентов эластичности цен по спросу может быть найдена как обратная матрица $E_{px} = (E_{xp})^{-1}$ к матрице E_{xp} коэффициентов эластичности спроса по ценам. Заметим, что когда перекрестные коэффициенты не равны нулю, $E_{p_i x_i} \neq \dfrac{1}{E_{x_i p_i}}$. Действительно, используя известную из линейной алгебры формулу для обратной матрицы, находим:

$$E_{p_1 x_1} = \frac{E_{x_2 p_2}}{E_{x_1 p_1} E_{x_2 p_2} - E_{x_1 p_2} E_{x_1 p_2}},$$

$$E_{p_2 x_2} = \frac{E_{x_1 p_1}}{E_{x_1 p_1} E_{x_2 p_2} - E_{x_1 p_2} E_{x_1 p_2}}.$$

§ 3.9. Однородные функции. Формула Эйлера

Пусть $D \subset \mathbb{R}^n$ – область в $\mathbb{R}^n$, содержащая вместе с каждой своей точкой $(x_1,...,x_n)$ и все точки вида $(tx_1,...,tx_n)$ при $t > 0$.

Функция $f(x_1,...,x_n)$ с такой областью определения D называется *однородной* степени α, если для любого $t > 0$ выполняется равенство

$$f(tx_1, ..., tx_n) = t^{\alpha} f(x_1, ..., x_n).$$

Однородный многочлен степени n является, очевидно, однородной функцией той же степени однородности. Так, например, многочлен $z = x^2 - 2xy + 3y^2$ является однородной функцией степени 2.

Степень однородности α может быть любым действительным числом. Например, функция

$$z = (x^2 + y^2)^{\pi} \operatorname{tg} \frac{x}{y}$$

является однородной функцией степени 2π от переменных x и y.

Примером однородной функции степени 1 может служить производственная функция Кобба–Дугласа[1] $f(K,L)=CK^{1/4}L^{3/4}$.

Предположим, что дифференцируемая функция $f(x, y)$ является одновременно и однородной функцией степени α. Фиксируя произвольную точку (x, y), для любого $t>0$ имеем

$$f(tx,ty)=t^{\alpha}f(x,y).$$

Продифференцируем левую и правую части этого равенства по t (левую часть – по правилу дифференцирования сложной функции, правую часть – как степенную функцию). В результате приходим к тождеству

$$f'_x(tx,ty)x+f'_y(tx,ty)y=\alpha t^{\alpha-1}f(x,y).$$

Положив здесь $t = 1$, получим *формулу Эйлера*:

$$f'_x(x,y)x+f'_y(x,y)y=\alpha f(x,y).$$

Аналогично записывается формула Эйлера для однородной функции от любого числа аргументов. Так, например, формула Эйлера для функции трех переменных $u = f(x, y, z)$ выглядит следующим образом:

$$f'_x(x,y,z)x+f'_y(x,y,z)y+f'_z(x,y,z)z=\alpha f(x,y,z),$$

или, короче,

$$u'_x x+u'_y y+u'_z z=\alpha u. \qquad (3.34)$$

Предположим, что функция $u = f(x, y, z)$ не обращается в нуль в некоторой точке (x_0, y_0, z_0). Разделив тогда левую и правую части

---

[1] Производственной функцией называется функция, выражающая зависимость объема выпуска от затрат ресурсов. В данном случае переменная K отражает затраты капитала, L – затраты труда.

равенства (3.34) на значение функции в этой точке, получим формулу

$$E_{ux} + E_{uy} + E_{uz} = \alpha, \tag{3.35}$$

где E_{ux}, E_{uy}, E_{uz} – коэффициенты эластичности u по x, по y и по z в точке (x_0, y_0, z_0).

Пример 3.15. Проверить соотношение (3.35) для функции

$$u = \frac{x}{y+z}.$$

Решение. Находим $E_{ux} = 1$, $E_{uy} = \frac{-y}{y+z}$, $E_{uz} = \frac{-z}{y+z}$. Следовательно, $E_{ux} + E_{uy} + E_{uz} = 0$. что согласуется с равенством (3.35), поскольку степень однородности функции u равна 0.

§ 3.10. Частные производные высших порядков

Пусть D – открытое множество в $\mathbb{R}^2$, $f(x, y)$ – определенная на множестве D функция. Предположим, что в каждой точке $M \in D$ существуют частные производные f'_x и f'_y. Тогда частные производные $f'_x(x, y)$ и $f'_y(x, y)$ естественно считать функциями с областью определения D. Они называются *частными производными первого порядка*. Частные производные от функций $f'_x(x, y)$ и $f'_y(x, y)$ называются *частными производными второго порядка* (или *вторыми частными производными*) от функции $f(x, y)$. Частные производные от частных производных второго порядка называются *частными производными третьего порядка* (или *третьими частными производными*) и т.д.

Если первая производная функции $z = f(x, y)$ была взята, скажем, по x, то ее частные производные в (x_0, y_0) обозначаются так:

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 f(x_0, y_0)}{\partial x^2} = (z'_x)'_x = z''_{xx} = f''_{xx}(x_0, y_0),$$

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 f(x_0, y_0)}{\partial y \partial x} = (z'_x)'_y = z''_{xy} = f''_{xy}(x_0, y_0).$$

Аналогичные обозначения используются и для других частных производных. Например, $\dfrac{\partial^2 z}{\partial x \partial y} = z''_{xy}$, $\dfrac{\partial^3 z}{\partial x \partial y \partial x} = z'''_{xyx}$ и т.д.

Определение. *Частные производные второго порядка* z''_{xy} *и* z''_{yx} *называются* ***смешанными частными производными***.

Пример 3.16. Найти все частные производные второго порядка от функций: а) $z = x^3 + y^2 + 5x^2 y$; б) $z = \operatorname{arctg}\dfrac{x}{y}$.

Решение. а) $z'_x = 3x^2 + 10xy$, $z'_y = 2y + 5x^2$. Следовательно,

$$z''_{xx} = 6x + 10y,\ z''_{yy} = 2,\ z''_{xy} = z''_{yx} = 10x.$$

б) Имеем $z'_x = \dfrac{y}{x^2 + y^2}$, $z'_y = \dfrac{-x}{x^2 + y^2}$. Следовательно,

$$z''_{xx} = -\frac{2xy}{(x^2 + y^2)^2}, \quad z''_{yy} = \frac{2xy}{(x^2 + y^2)^2}, \quad z''_{xy} = z''_{yx} = \frac{x^2 - y^2}{(x^2 + y^2)^2}.$$

В примерах 3.16 а) и б) смешанные частные производные от одной и той же функции z совпадают. Являются ли данные совпадения случайными или они – следствие какого-то общего правила? Следующая теорема дает ответ на этот вопрос.

Теорема 3.6. *Если производные* $f''_{xy}(x, y)$ *и* $f''_{yx}(x, y)$ *существуют в некоторой окрестности точки* $M(x_0,\ y_0)$ *и непрерывны в самой точке* M, *то имеет место равенство*

$$f''_{xy}(M) = f''_{yx}(M). \tag{3.36}$$

Доказательство. Определим выражение W формулой

$$W = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)] - [f(x_0, y_0 + \Delta y) - f(x_0, y_0)].$$

Это выражение совпадает с приращением $\Delta\varphi$ на отрезке $[x_0, x_0 + \Delta x]$ функции $\varphi(x) = f(x, y_0 + \Delta y) - f(x, y_0)$.

По теореме Лагранжа о конечных приращениях найдется точка $x_1 \in (x_0, x_0 + \Delta x)$, в которой выполняется равенство

$$W = \Delta\varphi = \varphi'(x_1)\Delta x$$

или

$$W = [f'_x(x_1, y_0 + \Delta y) - f'_x(x_1, y_0)]\Delta x.$$

Применяя теорему Лагранжа теперь к приращению $\Delta\psi$ функции $\psi(y) = f'_x(x_1, y)$ на отрезке $[y_0, y_0 + \Delta y]$, найдем точку $y_1 \in (y_0, y_0 + \Delta y)$, в которой $\Delta\psi = \psi'(y_1)\Delta y$. Следовательно,

$$W = [f'_x(x_1, y_0 + \Delta y) - f'_x(x_1, y_0)]\Delta x =$$

$$= \Delta\psi\Delta x = \psi'(y_1)\Delta x\Delta y = f''_{xy}(x_1, y_1)\Delta x\Delta y. \tag{3.37}$$

С другой стороны, для функции

$$\xi(y) = f(x_0 + \Delta x, y) - f(x_0, y)$$

выражение W можно представить как приращение $\xi(y)$ в точке y_0:

$$\Delta\xi(y_0) = [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)] -$$

$$-[f(x_0 + \Delta x, y_0) - f(x_0, y_0)] = W.$$

Аналогично находим, что

$$W = f''_{yx}(x_2, y_2)\Delta x\Delta y, \tag{3.38}$$

где $x_2 \in (x_0, x_0 + \Delta x)$, $y_2 \in (y_0, y_0 + \Delta y)$. Сравнивая формулы (3.37) и (3.38), получим

$$f''_{xy}(x_1, y_1) = f''_{yx}(x_2, y_2), \tag{3.39}$$

причем $x_1 \to x_0$, $x_2 \to x_0$, когда $\Delta x \to 0$, и $y_1 \to y_0$, $y_2 \to y_0$, когда $\Delta y \to 0$. Поскольку смешанные производные по предположению непрерывны, то, переходя к пределу в равенстве (3.39) при $\Delta x \to 0$, $\Delta y \to 0$, получим равенство (3.36).

§ 3.11. Формула Тейлора для функций нескольких переменных

Представим дифференциал функции $f(x_1,...,x_n)$ в точке $\vec{x}$ как однородный многочлен первой степени от переменных $v_1,...,v_n$:

$$df = df_{\vec{x}}(v_1,...,v_n) = \sum_{i=1}^{n} \frac{\partial f(\vec{x})}{\partial x_i} v_i . \tag{3.40}$$

Вместо переменных Δx_i в стандартной записи дифференциала мы используем переменные v_i, чтобы подчеркнуть тот факт, что переменные $v_1, ..., v_n$ могут принимать произвольные значения вне всякой зависимости от значений переменных $x_1, ..., x_n$. Рассматривая $v_1,...,v_n$ как параметры, т.е. полагая дифференциалы переменных v_i равными нулю ($dv_i = 0$), определим *высшие дифференциалы*:

$$d^2 f = d(df)),\ d^3 f = d(d^2 f)),\ d^4 f = d(d^3 f)) \text{ и т. д.}$$

Нетрудно убедиться в том, что

$$d^2 f_{\vec{x}}(v_1,...,v_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f(\vec{x})}{\partial x_i \partial x_j} v_i v_j ,$$

$$d^3 f_{\vec{x}}(v_1,...,v_n) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^3 f(\vec{x})}{\partial x_i \partial x_j \partial x_k} v_i v_j v_k ,$$

и, вообще,

$$d^r f_{\vec{x}}(v_1,...,v_n) = \sum_{i_1=1}^{n} ... \sum_{i_r=1}^{n} \frac{\partial^r f(\vec{x})}{\partial x_{i_1} ... \partial x_{i_r}} v_{i_1} ... v_{i_r} . \tag{3.41}$$

Дифференциал $d^2 f$ называется *вторым дифференциалом*, $d^3 f$ – *третьим*, …, $d^r f$ – *дифференциалом порядка r*.

Сравнивая (3.40) и (3.41), видим, что обычный дифференциал df является дифференциалом первого порядка (или первым дифференциалом). Из (3.41) следует также, что $d^r f$ – однородный многочлен степени r относительно переменных $v_1,...,v_n$. В частности $d^2 f$ – квадратичная форма.

Поскольку дифференциал переменной x_i равен

$$d(x_i)=\sum_{j=1}^{n}\frac{\partial x_i}{\partial x_j}v_j=v_i\ ,$$

вместо переменных v_i в записи дифференциалов $d^s f$ часто используются переменные dx_i, i = 1, …, n. Такие обозначения несут в себе определенную неоднозначность: с одной стороны, dx_i – символ переменной (эквивалент v_i); с другой, dx_i – дифференциал (т.е. линейная функция от $dx_1,...,dx_n$).

Впрочем, возможно, данная неоднозначность и объясняет популярность подобных обозначений.

Пример 3.17. Найти второй дифференциал функции $z=xy^2$ в точке $P(1;1)$. Будет ли квадратичная форма $d^2 z_P$ положительно определенной?

Решение. Сначала находим $d^2 z$ в произвольной точке:

$$d^2 z=d(dz))=d(y^2dx+2xydy)=4ydxdy+2x(dy)^2,$$

затем – в точке P: $d^2 z_P=4dxdy+2(dy)^2$. Данная квадратичная форма настолько проста, что ответ на вопрос о ее положительной определенности угадывается и без привлечения общих методов. Положив $dx=-1, dy=1$, находим $d^2 z_P(-1;1)=-2<0$ – форма не является положительно определенной.

Наиболее простой вид имеют высшие дифференциалы для функций одной переменной $f(x)=f(x_1)$ (случай n = 1):

$$d^r f_x(v)=f^{(r)}(x)v^r, \tag{3.42}$$

где $f^{(r)}(x)$ – r-я производная f в точке x.

Формула (3.42) позволяет записать k-й многочлен Тейлора функции f в точке a следующим образом:

$$T(x) = f(a) + \sum_{s=1}^{k} \frac{f^{(s)}(a)}{s!}(x-a)^s = f(a) + \sum_{s=1}^{k} \frac{d^s f_a(x-a)}{s!}. \quad (3.43)$$

Для функции f от n переменных ($n > 1$) определим ее k-й многочлен Тейлора в точке $\vec{a} = (a_1,...,a_n)$ по аналогии с (3.43):

$$T(x_1,...,x_n) = f(a_1,...,a_n) + \sum_{s=1}^{k} \frac{1}{s!} d^s f_{\vec{a}}(x_1 - a_1, \dots, x_n - a_n),$$

или в векторной записи:

$$T(\vec{x}) = f(\vec{a}) + \sum_{s=1}^{k} \frac{1}{s!} d^s f_{\vec{a}}(\vec{x} - \vec{a}).$$

Теорема 3.7. *Пусть функция* $f(x_1, ..., x_n)$ *имеет в* ε-*окрестности точки* $\vec{a}$ *непрерывные частные производные порядка* $k+1$. *Тогда для любой точки* $\vec{x}$ *из этой окрестности найдется точка* $\vec{c} = \vec{a} + \theta(\vec{x} - \vec{a})$, $0 < \theta < 1$, *такая, что*

$$f(\vec{x}) = T(\vec{x}) + \frac{d^{k+1} f_{\vec{c}}(\vec{x} - \vec{a})}{(k+1)!}, \quad (3.44)$$

где $T(x)$ – k-*й многочлен Тейлора функции* f *в точке* $\vec{a}$.

Доказательство. Положим $\varphi(t) = f(\vec{a} + t(\vec{x} - \vec{a}))$. Поскольку для любого $t \in [0;1]$ точка $\vec{a} + t(\vec{x} - \vec{a})$ содержится в ε-окрестности точки $\vec{a}$, функция $\varphi(t)$ определена и имеет $(k+1)$ производных на отрезке $[0;1]$. Для функции $\varphi(t)$ справедлива формула Маклорена

$$\varphi(1) = \varphi(0) + \frac{\varphi'(0)}{1!} + \frac{\varphi''(0)}{2!} + \dots + \frac{\varphi^{(k+1)}(\theta)}{(k+1)!}, \quad (3.45)$$

где $0 < \theta < 1$ – некоторое число.

Из определения $\varphi(t)$ следует, что

$$\varphi(1)=f(\vec{x}),\quad \varphi(0)=f(\vec{a}),\quad \varphi'(0)=df_{\vec{a}}(\vec{x}-\vec{a}),\quad \varphi''(0)=d^2f_{\vec{a}}(\vec{x}-\vec{a}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\varphi^{(k)}(0)=d^kf_{\vec{a}}(\vec{x}-\vec{a}),\quad \varphi^{(k+1)}(0)=d^{k+1}f_{\vec{c}}(\vec{x}-\vec{a}).$$

Подставив эти равенства в (3.45), получим утверждение теоремы.

Так же, как и в случае формулы Тейлора для функции одной переменной, последнее слагаемое в правой части (3.44) называется *остаточным членом в форме Лагранжа*.

Так же, как и в случае функций одной переменной, формула Тейлора служит обоснованием приближенных равенств

$$f(\vec{x})\approx T(\vec{x}) \tag{3.46}$$

для точек $\vec{x}$ из малой окрестности точки $\vec{a}$.

Явный вид остаточного члена позволяет получать оценки абсолютной погрешности в (3.46).

Пример 3.18. Для функции $z=\frac{1}{2}(e^{x+y-2}+e^{x-y})$ найти ее первый многочлен Тейлора $T(x,y)$ в точке $P(1;1)$ и оценить абсолютную ошибку в приближенном равенстве $z\approx T(x,y)$ при условии, что $|x-1|<\varepsilon, |y-1|<\varepsilon,\ \varepsilon=0{,}001.$

Решение. Имеем:

$$dz=\frac{1}{2}e^{x+y-2}(dx+dy)+\frac{1}{2}e^{x-y}(dx-dy),$$

$$d^2z=\frac{1}{2}e^{x+y-2}(dx+dy)^2+\frac{1}{2}e^{x-y}(dx-dy)^2,$$

$$dz_P=dx,\quad T(x,y)=1+(x-1)=x.$$

Согласно формуле Тейлора для любой точки $M(x,y)$ на отрезке PM найдется точка $Q(1+\theta(x-1), 1+\theta(y-1)), 0<\theta<1$, для которой

$$z(M)=T(x,y)+\frac{1}{2}d^2z_Q(x-1,y-1).$$

Отсюда следует, что

$$\frac{1}{2}(e^{x+y-2}+e^{x-y})\approx T(x,y)=x, \tag{3.47}$$

причем в заданной (квадратной) окрестности точки P абсолютная погрешность $|d^2z_Q|$ равенства (3.47) оценивается следующим образом:

$$|d^2z_Q|\le\frac{1}{2}\left[\frac{1}{2}e^{2\varepsilon}(2\varepsilon)^2+\frac{1}{2}e^{2\varepsilon}(2\varepsilon)^2\right]=2e^{2\varepsilon}\varepsilon^2<2{,}1\cdot 10^{-6}.$$

Фактически, в данном примере мы получили оценку для остаточного члена в форме Лагранжа в некотором частном случае. Получим теперь оценку остаточного члена в общем случае. Поскольку частные производные f порядка $k+1$ непрерывны в точке $\vec{a}$, найдется такая постоянная M, что для любой точки $\vec{x}$ из некоторой ε-окрестности $\vec{a}$ одновременно выполняется неравенство

$$\left|\frac{\partial^{k+1}f(\vec{x})}{\partial x_{i_1}\ldots\partial x_{i_{k+1}}}\right|<M$$

для всех частных производных f порядка $k+1$. Для оценки остаточного члена достаточно получить оценку дифференциала

$$d^{k+1}f_{\vec{c}}(\vec{x}-\vec{a})=\sum_{i_1=1}^{n}\ldots\sum_{i_{k+1}=1}^{n}\frac{\partial^{k+1}f(\vec{c})}{\partial x_{i_1}\ldots\partial x_{i_{k+1}}}(x_{i_1}-a_{i_1})\ldots(x_{i_{k+1}}-a_{i_{k+1}}). \tag{3.48}$$

Для каждого слагаемого в (3.48) имеем

$$\left|\frac{\partial^{k+1}f(\vec{c})}{\partial x_{i_1}\ldots\partial x_{i_{k+1}}}(x_{i_1}-a_{i_1})\ldots(x_{i_{k+1}}-a_{i_{k+1}})\right|\le M|\vec{x}-\vec{a}|^{k+1}.$$

Поскольку число слагаемых в (3.48) равно n^{k+1},

$$\left|d^{k+1}f_{\vec{c}}(\vec{x}-\vec{a})\right|\le Mn^{k+1}|\vec{x}-\vec{a}|^{k+1}.$$

В итоге получаем оценку остаточного члена:

$$\left|\frac{1}{(k+1)!}d^{k+1}f_{\vec{c}}(\vec{x}-\vec{a})\right|\le\frac{n^{k+1}M}{(k+1)!}|\vec{x}-\vec{a}|^{k+1},$$

или

$$\left|\frac{1}{(k+1)!}d^{k+1}f_{\vec{c}}(\vec{x}-\vec{a})\right|\le\alpha(x)|\vec{x}-\vec{a}|^{k},$$

где $\alpha(x)=\dfrac{n^{k+1}M}{(k+1)!}|\vec{x}-\vec{a}|$. Поскольку $\alpha(x)\to 0$ при $\vec{x}\to\vec{a}$, последнее неравенство записывается обычно в виде

$$\frac{1}{(k+1)!}d^{k+1}f_{\vec{c}}(\vec{x}-\vec{a})=o\left(|\vec{x}-\vec{a}|^{k}\right).$$

С учетом данной записи из формулы Тейлора (3.44) получим

$$f(\vec{x})=T(\vec{x})+o(|\,\vec{x}-\vec{a}\,|^{k}).$$

Это равенство называется формулой Тейлора с остаточным членом в *форме Пеано*. Данный вариант формулы Тейлора справедлив и без предположения о наличии частных производных порядка $k+1$ – достаточно лишь потребовать, чтобы функция $f(\vec{x})$ была k раз дифференцируемой в точке $\vec{a}$. Доказательство этого утверждения мы не приводим.

§ 3.12. Локальный экстремум функции нескольких переменных

Пусть функция $f(\vec{x})$ определена и непрерывна в некоторой окрестности точки $\vec{a} \in \mathbb{R}^n$.

Определение. *Точка* $\vec{a}$ *называется* ***точкой локального максимума*** (***минимума***) *функции* $f(\vec{x})$, *если существует такая* ε*-окрестность* $U_\varepsilon(\vec{a}) = \{\vec{x} \in \mathbb{R}^n : |\vec{x} - \vec{a}| < \varepsilon\}$ *точки* $\vec{a}$, *в которой для любой точки* $\vec{x} \in U_\varepsilon(\vec{a})$ *выполняется неравенство*

$$f(\vec{x}) \le f(\vec{a}) \quad (f(\vec{x}) \ge f(\vec{a})).$$

Точки локального максимума и локального минимума называются *точками локального экстремума*, или просто *точками экстремума*.

1. Необходимые условия первого порядка

Теорема 3.8. *Для того чтобы дифференцируемая функция* $f(\vec{x})$ *имела локальный экстремум в точке* $\vec{a}$, *необходимо, чтобы все ее частные производные первого порядка в этой точке были равны* 0.

Доказательство. Для каждого $i = 1, 2, \ldots, n$ рассмотрим функцию одной переменной

$$\varphi(x) = f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n).$$

Нетрудно видеть, что эта функция имеет локальный экстремум в точке a_i, а ее производная

$$\varphi'(x) = f'_{x_i}(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n).$$

По теореме 2.13 производная функции $\varphi(x)$ в точке a_i равна 0. Отсюда

$$\varphi'(a_i) = f'_{x_i}(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n) = f'_{x_i}(\vec{a}) = 0.$$

Теорема доказана.

Итак, “подозрительными” на экстремум являются те точки $\vec{a}$, в которых функция дифференцируема, а ее частные производные первого порядка обращаются в нуль: $\operatorname{grad} f(\vec{a}) = 0.$ Как и в случае функции одной переменной, такие точки называются *стационарными*. Обычно стационарные точки (в отличие от точек экстремума) достаточно просто находятся. Так, для функции двух переменных $f(x, y)$ их можно найти, решив систему уравнений

$$\begin{cases} f'_x(x, y) = 0, \\ f'_y(x, y) = 0. \end{cases} \tag{3.49}$$

Пример 3.19. Найти стационарные точки функции

$$z = x^3 - 3x + y^4 - 2y^2.$$

Решение. Система (3.49) имеет вид

$$\begin{cases} z'_x = 3x^2 - 3 = 0, \\ z'_y = 4y^3 - 4y = 0. \end{cases}$$

Из первого уравнения находим $x = \pm 1$, из второго получаем $y = 0, \pm 1$. Следовательно, имеется шесть стационарных точек: (1; 0), (1; 1), (1; –1), $(1; 0)$, $(-1; 1)$ и $(-1; -1)$.

Условия первого порядка, как будет доказано ниже, являются достаточными условиями экстремума в случае выпуклых и вогнутых функций. Однако в общем случае это не так. Например, для функции $f(x, y) = x^2 - y^2$ частные производные первого порядка равны нулю в точке (0; 0), однако эта точка не является точкой локального экстремума. Действительно, в любой окрестности точки (0; 0) существуют точки $(x, 0)$ и $(0, y)$, в которых

$$f(x, 0) > f(0, 0) > f(0, y).$$

Поэтому (0; 0) не является ни точкой локального максимума, ни точкой локального минимума.

Для проведения более тонкого анализа предполагаемых точек локального экстремума используются критерии, связанные со вторыми частными производными (условия второго порядка).

2. Необходимые условия второго порядка

Теорема 3.9. *Пусть функция $f(\vec{x})$ имеет в окрестности точки своего локального экстемума $\vec{a}$ непрерывные частные производные второго порядка. Тогда*:

1) *если $\vec{a}$ – точка локального минимума, то $d^2 f_{\vec{a}}$ – неотрицательно определенная квадратичная форма*;

2) *если же $\vec{a}$ – точка локального максимума, то $d^2 f_{\vec{a}}$ – неположительно определенная квадратичная форма.*

Доказательство. 1) Пусть O – такая окрестность $\vec{a}$, что для $\vec{x} \in O$ имеем $f(\vec{x}) \geq f(\vec{a})$. Предположим противное: существует вектор $\vec{v}$, такой что $d^2 f_{\vec{a}}(\vec{v}) < 0$. Из непрерывности вторых частных производных следует, что величина $d^2 f_{\vec{x}}(\vec{v})$ непрерывно зависит от $\vec{x}$, поэтому в некоторой ε-окрестности $U \subset O$ точки $\vec{a}$ выполняется неравенство

$$d^2 f_{\vec{x}}(\vec{v}) < 0, \quad \vec{x} \in U.$$

Выберем $t > 0$ так, чтобы точка $\vec{b} = \vec{a} + t\vec{v} \in U$, и запишем формулу Тейлора с учетом того, что $\vec{a}$ – точка локального минимума и $df_{\vec{a}} = 0$:

$$f(\vec{b}) = f(\vec{a}) + \frac{1}{2} d^2 f_{\vec{c}}(\vec{b} - \vec{a}).$$

Поскольку $\vec{b} - \vec{a} = t\vec{v}$ и $\vec{c} \in U$ ($\vec{c}$ – точка отрезка $\vec{a}\vec{b}$), имеем

$$f(\vec{b}) = f(\vec{a}) + \frac{t^2}{2} d^2 f_{\vec{c}}(\vec{v}) < f(\vec{a}),$$

что противоречит существованию локального минимума в точке $\vec{a}$. Полученное противоречие доказывает, что для любого $\vec{v} \in \mathbb{R}^n$ выполняется неравенство $d^2 f_{\vec{a}}(\vec{v}) \geq 0,$ т.е. квадратичная форма $d^2 f_{\vec{a}}$ неотрицательно определена.

Пункт 2) сводится к 1) заменой f на $-f$. Теорема доказана.

Пусть $f''(\vec{a})$ – матрица частных производных второго порядка функции $f(\vec{x})$ в точке $\vec{a}$. Поскольку $f''(\vec{a})$ – матрица квадратичной формы $d^2 f_{\vec{a}},$ вся работа по определению знакового типа второго дифференциала производится именно с этой матрицей. Вследствие симметричности матрицы $f''(\vec{a})$ все ее собственные значения $\alpha_1, \alpha_2, \ldots, \alpha_n$ – действительные числа[1]; если они известны, то тип квадратичной формы $d^2 f_{\vec{a}}$ находится следующим образом:

1) все $\alpha_i > 0 \quad \Leftrightarrow \quad d^2 f_{\vec{a}}$ – положительно определена;
2) все $\alpha_i \geq 0 \quad \Leftrightarrow \quad d^2 f_{\vec{a}}$ – неотрицательно определена;
3) все $\alpha_i < 0 \quad \Leftrightarrow \quad d^2 f_{\vec{a}}$ – отрицательно определена;
4) все $\alpha_i \leq 0 \quad \Leftrightarrow \quad d^2 f_{\vec{a}}$ – неположительно определена;
5) существуют $\alpha_i > 0$ и $\alpha_j < 0 \quad \Leftrightarrow \quad d^2 f_{\vec{a}}$ знаконеопределена, т.е. для некоторых $\vec{v}, \vec{w} \in \mathbb{R}^n$ имеем $d^2 f_{\vec{a}}(\vec{v}) > 0, d^2 f_{\vec{a}}(\vec{w}) < 0.$

Замечание 1. В случае невырожденной матрицы $f''(a),$ а такой случай встречается достаточно часто, $\alpha_i \neq 0,$ вследствие чего неотрицательная определенность $d^2 f_{\vec{a}}$ эквивалентна ее положительной определенности, а неположительная определенность $d^2 f_{\vec{a}}$ эквивалентна ее отрицательной определенности.

Замечание 2. Без вычисления собственных значений положительную определенность квадратичной формы можно выяснить по критерию Сильвестра:

квадратичная форма положительно определена в том и только в том случае, если все угловые миноры ее матрицы больше 0.

---

[1] Симметрическая матрица порядка n имеет n собственных значений с учетом их кратности. Для симметрической матрицы кратность собственного значения – это одновременно и кратность корня характеристического уравнения и размерность подпространства собственных векторов.

Установить неотрицательную определенность можно с помощью следующего критерия:

квадратичная форма неотрицательно определена в том и только в том случае, если все миноры ее матрицы, симметричные относительно главной диагонали, неотрицательны.

3. Достаточные условия существования локального экстремума

Теорема 3.10. *Пусть функция n переменных $f(\vec{x})$ имеет в окрестности своей стационарной точки $\vec{a}$ непрерывные частные производные второго порядка. Тогда:*

1) *если квадратичная форма $d^2 f_{\vec{a}}$ положительно определена, то $\vec{a}$ – точка локального минимума f;*

2) *если квадратичная форма $d^2 f_{\vec{a}}$ отрицательно определена, то $\vec{a}$ – точка локального максимума f.*

Доказательство. 1) Рассмотрим зависящие от точки $\vec{x}$ матрицы:

$$A_k(\vec{x}) = \begin{pmatrix} \dfrac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_1} & \dots & \dfrac{\partial^2 f(\vec{x})}{\partial x_1 \partial x_k} \\ \dots & \dots & \dots \\ \dfrac{\partial^2 f(\vec{x})}{\partial x_k \partial x_1} & \dots & \dfrac{\partial^2 f(\vec{x})}{\partial x_k \partial x_k} \end{pmatrix}, \quad k = 1, 2, \dots, n.$$

Согласно критерию Сильвестра положительная определенность $d^2 f_{\vec{a}}$ означает, что $\det A_k(\vec{a}) > 0, k = 1, 2, \dots, n$. Из непрерывной зависимости $\det A_k(\vec{x})$ от $\vec{x}$ следует, что найдется ε-окрестность U точки $\vec{a}$, в которой выполняются неравенства:

$$\det A_k(\vec{x}) > 0, \vec{x} \in U, k = 1, 2, \dots, n.$$

Применяя критерий Сильвестра еще раз, получаем, что второй дифференциал f в произвольной точке $\vec{c} \in U$ – положительно определенная квадратичная форма. С учетом того, что $\vec{a}$ – стационарная точка, формула Тейлора имеет вид

$$f(\vec{x}) = f(\vec{a}) + \frac{1}{2} d^2 f_{\vec{c}}(\vec{x} - \vec{a}).$$

Поскольку $\vec{c} \in U$ ($\vec{c}$ – точка отрезка $\vec{a}\vec{x}$), квадратичная форма $d^2 f_{\vec{c}}$ положительно определена. Следовательно, $f(\vec{x}) > f(\vec{a})$, т.е. $\vec{a}$ – точка локального минимума f.

Пункт 2) сводится к 1) заменой f на $-f$. Теорема доказана.

В качестве иллюстрации доказанных теорем рассмотрим подробнее условия локального экстремума функции двух переменных.

Теорема 3.11. *Пусть функция $f(x, y)$ имеет непрерывные частные производные второго порядка в некоторой окрестности своей стационарной точки P. Положим*

$$\Delta = \det f''(P) = f''_{xx}(P) f''_{yy}(P) - (f''_{xy}(P))^2.$$

Тогда:

1) *если $\Delta > 0$, то в точке P функция имеет локальный экстремум, причем при $f''_{xx}(P) < 0$ – локальный максимум, при $f''_{xx}(P) > 0$ – локальный минимум*;

2) *если $\Delta < 0$, то в точке P нет экстремума.*

Доказательство. 1. Пусть $f''_{xx}(P) > 0$. Тогда согласно критерию Сильвестра второй дифференциал $d^2 f_P$ – положительно определенная квадратичная форма. По теореме 3.10 P – точка локального минимума. Предположим теперь, что $f''_{xx}(P) < 0$. Рассмотрим функцию $g(x, y) = -f(x, y)$. Для этой функции $g''_{xx}(P) > 0$ и $\det g''(P) = (-1)^2 \det f''(P) > 0$. Следовательно, P – точка локального минимума для $g(x, y)$ и одновременно точка локального максимума для $f(x, y)$.

2. Пусть α_1, α_2 – все собственные значения матрицы $f''(P)$, при этом $\alpha_1 = \alpha_2$ в случае кратного собственного значения. Из линейной алгебры известно, что определитель симметрической мат-

рицы равен произведению всех ее собственных значений с учетом кратности. Отсюда $\alpha_1\alpha_2 = \Delta < 0$ и, следовательно, собственные значения имеют разные знаки. Таким образом, второй дифференциал $d^2 f_P$ является знаконеопределенным. По теореме 3.9 в точке P нет локального экстремума.

Пример 3.20. Исследовать на локальный экстремум функцию

$$z = x^3 - 3x - (x + 1 + \operatorname{arctg} y)^2.$$

Решение. Найдем стационарные точки, решив систему уравнений

$$\begin{cases} z'_x = 3x^2 - 3 - 2(x + 1 + \operatorname{arctg} y) = 0, \\ z'_y = -2(x + 1 + \operatorname{arctg} y)(1 + y^2)^{-1} = 0. \end{cases}$$

Из второго уравнения вытекает, что $x + 1 + \operatorname{arctg} y = 0$. Следовательно, $z'_x = 3x^2 - 3 = 0$, откуда $x = \pm 1$. Так как $|\operatorname{arctg} y| < \pi / 2 < 2$, то $x + 1 + \operatorname{arctg} y \neq 0$ в случае $x = 1$. Если же $x = -1$, то

$$x + 1 + \operatorname{arctg} y = 0 \Rightarrow y = 0.$$

Итак, имеется только одна стационарная точка $P(-1; 0)$. Находим частные производные второго порядка:

$$z''_{xx}(P) = -8; \; z''_{yy}(P) = -2; \; z''_{xy}(P) = -2 .$$

Поскольку $z''_{xx}(P) < 0$ и $\Delta = (-8) \cdot (-2) - (-2)^2 = 12 > 0$, по теореме 3.11 точка P является точкой локального максимума. Интересно отметить следующее: несмотря на то, что точка P является *единственной* стационарной точкой всюду дифференцируемой функции z, значение функции в этой точке *не является* наибольшим. Действительно, $z(P) = 2$ однако существуют точки, в которых значение z больше. Например, $z(10; 0) = 849$.

4. Экономические приложения

Пусть z – количество продукции, выпущенной некоторой фирмой; x, y – затраты ресурсов двух видов; $z = Q(x, y)$ – дифференцируемая функция, устанавливающая связь x, y и z. Предположим, что величины x, y, z заданы в натуральных единицах и p_x, p_y, p_z – соответствующие этим единицам *постоянные* цены. Тогда выручка (валовой доход) будет $R(x, y) = p_z Q(x, y),$ а функция прибыли

$$\pi(x, y) = R(x, y) - p_x x - p_y y. \tag{3.50}$$

Пусть z^* – оптимальный (с точки зрения прибыли) выпуск продукции; x^*, y^* – соответствующие затраты ресурсов. Если точка $M(x^*, y^*)$ является внутренней точкой области определения функции $\pi(x, y),$ то M – точка ее локального максимума. Согласно необходимому признаку локального экстремума в точке M обращаются в нуль частные производные первого порядка:

$$\begin{cases} \pi'_x(M) = R'_x(M) - p_x = 0, \\ \pi'_y(M) = R'_y(M) - p_y = 0, \end{cases}$$

или

$$\begin{cases} R'_x(M) = p_x, \\ R'_y(M) = p_y. \end{cases} \tag{3.51}$$

Вывод: *в точке локального максимума прибыли предельная выручка от каждого ресурса совпадает с его ценой.*

Этот вывод сохраняется и в общем случае, когда цена p_z зависит от объема выручки: $p_z = p_z(Q)$. Действительно, тогда $R(x, y) = p_z(Q(x, y))Q(x, y),$ однако частные производные π'_x, π'_y и система (3.51) имеют прежний вид.

Рассмотрим теперь фирму-монополию, которая продает свою продукцию на двух независимых рынках. Пусть p_i, q_i – соответственно цена и количество продукции, проданной монополией на i-м рынке (i = 1, 2). Независимость рынков означает, что цена p_1 не зависит от q_2, т.е. $p_1 = p_1(q_1)$. Аналогично $p_2 = p_2(q_2)$.

Пусть $C(q)$ – дифференцируемая функция издержек. Тогда функция прибыли имеет вид

$$\pi = p_1 q_1 + p_2 q_2 - C(q_1 + q_2). \tag{3.52}$$

В точке локального максимума прибыли имеем

$$\begin{cases} \pi'_{q_1} = p'_1 q_1 + p_1 - C' = p_1(E_{p_1 q_1} + 1) - C' = 0, \\ \pi'_{q_2} = p'_2 q_2 + p_2 - C' = p_2(E_{p_2 q_2} + 1) - C' = 0. \end{cases}$$

Здесь $E_{p_1 q_1}$ и $E_{p_2 q_2}$ – коэффициенты эластичности (см. § 3.8). Из данной системы уравнений следует, что $p_1(E_{p_1 q_1} + 1) = p_2(E_{p_2 q_2} + 1)$. Отсюда получаем отношение цен:

$$\frac{p_2}{p_1} = \frac{1 + E_{p_1 q_1}}{1 + E_{p_2 q_2}}. \tag{3.53}$$

Пусть e_{ij} – эластичность спроса на i-м рынке по цене на j-м рынке (i, j = 1, 2). Так как рынки по предположению независимы, то, используя свойства эластичности функции одной переменной, получим

$$E_{p_i q_i} = (E_{q_i p_i})^{-1} = e_{ii}^{-1}.$$

Следовательно, отношение цен имеет вид

$$\frac{p_2}{p_1} = \frac{1 + e_{11}^{-1}}{1 + e_{22}^{-1}}. \tag{3.54}$$

Пример 3.21. На сколько процентов цена на втором из двух независимых рынков выше, если эластичность спроса на первом рынке (–2), а на втором (–1,5)?

Решение. Используя формулу (3.54), находим

$$\frac{p_2}{p_1} = \frac{1-(-2)^{-1}}{1-(-1{,}5)^{-1}} = \frac{1-0{,}5}{1-2/3} = 1{,}5.$$

Следовательно, на втором рынке цена на 50% больше.

Замечание. В случае зависимых рынков для отношения цен можно получить формулу, аналогичную (3.54), используя свойство 5′ эластичности функции двух переменных:

$$\frac{p_2}{p_1} = \frac{\Delta + e_{22} + \dfrac{q_1 e_{12}}{q_2}}{\Delta + e_{11} + \dfrac{q_2 e_{21}}{q_1}},$$

где $\Delta = e_{11}e_{22} - e_{12}e_{21}$.

§ 3.13. Условный экстремум

Пусть функция $f(\vec{x})$ от n переменных определена на множестве $D \subset \mathbb{R}^n$, X – некоторое подмножество в D.

Определение. *Точка $\vec{x}^* \in X$ называется точкой* ***условного локального максимума*** (***минимума***) *функции f, если для всех достаточно близких к ней точек $\vec{x} \in X$ выполняется неравенство*

$$f(\vec{x}) \le f(\vec{x}^*) \quad (f(\vec{x}) \le f(\vec{x}^*)). \qquad (3.55)$$

Точки условного локального максимума и точки условного локального минимума называются точками ***условного локального экстремума***, *или просто точками* ***условного экстремума***.

Отличие условного экстремума от обычного состоит в том, что неравенство (3.55) должно выполняться не для всех вообще $\vec{x}$ из некоторой окрестности $U_\varepsilon(\vec{x}^*) = \{\vec{x} \in \mathbb{R}^n : |\vec{x} - \vec{x}^*| < \varepsilon\}$, а только для тех $\vec{x}$, которые принадлежат $U_\varepsilon(\vec{x}^*) \cap X$, – в этом и состоит *условность* экстремума.

Для приложений наиболее важен случай, когда множество X задается с помощью некоторой системы уравнений и неравенств. Ограничимся пока случаем уравнений. Итак, пусть X задано системой

$$\begin{cases} g_1(x_1, ..., x_n) = 0, \\ \quad \cdots\cdots\cdots\cdots \\ g_s(x_1, ..., x_n) = 0, \end{cases} \tag{3.56}$$

где $g_1, ..., g_s$ – некоторые функции от n переменных.

Уравнения (3.56) называются обычно *уравнениями связи*, так как они *связывают* значения переменных x_1, ..., x_n. Таким образом, возникает задача о нахождении экстремумов функции $f(x_1,...,x_n)$ при условии, что переменные x_1, ..., x_n связаны уравнениями (3.56). Если бы переменные не были связаны, то эта задача, по крайней мере в случае дифференцируемой функции f, решалась бы путем исследования ее стационарных точек. Оказывается, что и при наличии связей задача также сводится к поиску стационарных точек. Однако в данном случае точкам условного экстремума исходной функции f соответствуют стационарные точки другой функции.

Теорема 3.12 (*необходимое условие условного экстремума*). *Пусть функции f и g_1, ..., g_s определены и имеют непрерывные частные производные в окрестности точки $\vec{x}^*$, причем векторы*

$$\operatorname{grad} g_1(\vec{x}^*), ..., \operatorname{grad} g_s(\vec{x}^*) \tag{3.57}$$

линейно независимы. Тогда если $\vec{x}^*$ – точка условного экстремума функции f при условиях (3.56), *то найдутся числа $\lambda_1, ..., \lambda_s$, для которых $\vec{x}^*$ – стационарная точка функции*

$$L(\vec{x}) = f(\vec{x}) + \lambda_1 g_1(\vec{x}) + ... + \lambda_s g_s(\vec{x}).$$

Функция L называется *функцией Лагранжа*, а числа $\lambda_1, ..., \lambda_s$ – *множителями Лагранжа*.

Доказывать теорему в общем случае не будем, а ограничимся частным случаем, когда число переменных равно 3, а число связей – 2. Более того, примем, что $\vec{x}^* = (0; 0; 0)$, а уравнения связи имеют вид

$$\begin{cases} x_1 = 0, \\ x_2 = 0. \end{cases}$$

Следовательно, множество X всех решений этой системы есть ось Ox_3, а ограничение функции f на X – функция одной переменной $f(0; 0; x_3)$. Так как $(0; 0; 0)$ – точка условного экстремума функции трех переменных $f(x_1, x_2, x_3)$, то 0 – точка обычного экстремума функции одной переменной $f(0; 0; x_3)$. Поэтому $f'_{x_3}(\vec{x}^*) < 0$. Для функции Лагранжа

$$L(x_1, x_2, x_3) = f(x_1, x_2, x_3) + \lambda_1 x_1 + \lambda_2 x_2$$

частные производные имеют вид:

$$\frac{\partial L}{\partial x_1} = \frac{\partial f}{\partial x_1} + \lambda_1, \quad \frac{\partial L}{\partial x_2} = \frac{\partial f}{\partial x_2} + \lambda_2, \quad \frac{\partial L}{\partial x_3} = \frac{\partial f}{\partial x_3}.$$

Как было доказано выше, в точке $\vec{x}^*$ частная производная $\partial f / \partial x_3 = 0$, поэтому и частная производная $\partial L / \partial x_3$ также равна нулю в этой точке. Частные производные функции Лагранжа по x_1 и по x_2 обращаются в нуль в точке $\vec{x}^*$, если положить

$$\lambda_1 = -\frac{\partial f}{\partial x_1}(0;0;0), \quad \lambda_2 = -\frac{\partial f}{\partial x_2}(0;0;0). \tag{3.58}$$

Следовательно, $\vec{x}^*$ – стационарная точка функции Лагранжа, когда множители Лагранжа принимают значения (3.58). В частном случае теорема доказана. Для искушенного в математическом анализе читателя должно быть ясно, что данное доказательство, по сути, является общим, поскольку функции g_1, ..., g_s можно принять за новые координаты в окрестности точки $\vec{x}^*$.

Замечание 1. Если имеется только одно уравнение связи

$$g(x_1, ..., x_n) = 0,$$

то условие линейной независимости градиентов (3.57) сводится к условию $\operatorname{grad} g(\vec{x}^*) \neq 0$, которое означает, что по крайней мере одна частная производная $\partial g / \partial x_i$ отлична от нуля в точке $\vec{x}^*$.

Замечание 2. Пусть P^* – точка, в которой функция f принимает наибольшее (наименьшее) значение на множестве X. Тогда P^* – точка условного локального экстремума функции $f(P)$ при условии $P \in X$. Следовательно, P^* – стационарная точка функции Лагранжа (при определенных значениях множителей Лагранжа) либо для P^* нарушаются условия теоремы 3.12.

Пример 3.22. Найти наибольшее значение функции $f = x + y + z$ при условии

$$9x^2 + 4y^2 + z^2 = 36. \qquad (3.59)$$

Решение. Пусть $X \subset \mathbb{R}^3$ – множество всех точек $P(x, y, z)$, координаты которых удовлетворяют уравнению (3.59). Поскольку из уравнения (3.59) вытекают неравенства $|x| \leq 2$, $|y| \leq 3$, $|z| \leq 6$, X – ограниченное множество. Так как левая часть уравнения (3.59) – непрерывная функция, то X – еще и замкнутое множество. Из непрерывности функции $f = x + y + z$ следует, что она достигает наибольшего значения на X в некоторой точке P^*. Чтобы найти точку P^*, воспользуемся замечаниями 1 и 2. Запишем условие (3.59) в виде уравнения $g(x, y, z) = 0$, где

$$g(x, y, z) = 36 - 9x^2 - 4y^2 - z^2.$$

Таким образом, функция Лагранжа будет

$$L(x, y, z) = x + y + z + \lambda(36 - 9x^2 - 4y^2 - z^2).$$

Приравнивая нулю ее частные производные, получим систему уравнений, задающую стационарные точки:

$$\begin{cases} 1-18\lambda x=0, \\ 1-8\lambda y=0, \\ 1-2\lambda z=0. \end{cases}$$

Следовательно, координаты стационарных точек следующим образом выражаются через множитель Лагранжа:

$$x=\frac{1}{18\lambda}, \quad y=\frac{1}{8\lambda}, \quad z=\frac{1}{2\lambda}.$$

Подставив эти выражения в уравнение связи (3.59), найдем значения множителя Лагранжа $\lambda=\pm\dfrac{7}{72}$. Отсюда получим две стационарные точки:

$$P_1\left(\frac{4}{7};\frac{9}{7};\frac{36}{7}\right) \text{ и } P_2\left(-\frac{4}{7};-\frac{9}{7};-\frac{36}{7}\right).$$

Проверим теперь условия теоремы 3.12 для произвольной точки $P_0(x_0,y_0,z_0)\in X$. Во-первых, вычисляя частные производные первого порядка функций f и g, нетрудно убедиться в том, что все они непрерывны в любой точке $P_0\in\mathbb{R}^3$. Во-вторых,

$$\operatorname{grad} g(P_0)=(-18x_0;-8y_0;-2z_0)\neq 0,$$

так как если точка $P_0\in X$, то хотя бы одна из ее координат отлична от нуля. Итак, для любой точки множества X выполняются все условия теоремы 3.12, следовательно, точками условного экстремума f могут быть только точки P_1 и P_2. Поскольку $f(P_1)=7>$ $>f(P_2)=-7$, то $P^*=P_1$. Таким образом, наибольшее значение f на X будет $f(P^*)=7$.

Пример 3.23. Пусть U – полезность набора товаров, состоящего из x единиц первого товара, y – второго и z единиц третьего товара. Найти стоимость наиболее дешевого набора товаров с заданным значением полезности $U = 10000$, если цена первого товара – 4, второго – 25, третьего – 20, а функция полезности имеет вид $U = xyz^2$.

Решение. Пусть $S = 4x + 25y + 20z$ – стоимость набора, $X \subset \mathbb{R}^3$ – множество всех точек $P(x, y, z)$, координаты которых удовлетворяют системе ограничений

$$\begin{cases} xyz^2 = 10000, \\ x \ge 0, y \ge 0, z \ge 0. \end{cases}$$

Необходимо найти минимальное значение S на множестве X. В отличие от предыдущего примера, множество X неограничено. Поэтому теперь мы будем рассуждать иначе. Найдем какую-нибудь точку A с заданным значением полезности $U = 10000$. Например, можно взять точку $A(10; 10; 10)$. Стоимость набора A будет $S(A) = 490$. Затем удалим из X все наборы, стоимость которых заведомо больше стоимости самого дешевого набора, скажем, больше 500. Полученное множество обозначим X', другими словами, X' задается системой ограничений

$$\begin{cases} xyz^2 = 10000, \\ x \ge 0, y \ge 0, z \ge 0, \\ 4x + 25y + 20z \le 500. \end{cases}$$

Левые части ограничений данной системы – непрерывные функции, поэтому X' – замкнутое множество. Кроме того, для любой точки $P(x, y, z) \in X'$ выполняются неравенства

$$0 \le x \le 125, \quad 0 \le y \le 20, \quad 0 \le z \le 25.$$

Отсюда X' – замкнутое и ограниченное множество. Вследствие непрерывности функции S найдется точка $P^*(x^*, y^*, z^*) \in X'$, в которой S принимает наименьшее значение на X' (а значит, и на X). Поскольку $x^*y^*z^{*2} = 10000$, имеем $x^* > 0$, $y^* > 0$, $z^* > 0$. Кроме того,

$$S(P^*) \le S(A) = 490 < 500.$$

Следовательно, и для любой достаточно близкой к P^* точке $P(x, y, z)$ также будут выполняться неравенства

$$x > 0, y > 0, z > 0 \quad \text{и} \quad 4x + 25y + 20z < 500.$$

Таким образом, если достаточно близкая к P^* точка $P(x, y, z)$ удовлетворяет условию $U(P) = 10000$, то она принадлежит X'. Поэтому P^* – точка условного локального минимума функции S при условии $U = 10000$.

Найдем точку P^*, используя теорему 3.12. Функция Лагранжа имеет вид

$$L = 4x + 25y + 20z + \lambda(10000 - xyz^2).$$

Приравнивая нулю частные производные, получаем систему уравнений для ее стационарных точек

$$\begin{cases} 4 - \lambda yz^2 = 0, \\ 25 - \lambda xz^2 = 0, \\ 20 - 2\lambda xyz = 0. \end{cases}$$

С учетом уравнения связи из этой системы находим:

$$x = 2500\lambda, \quad y = 400\lambda, \quad z = 1000\lambda.$$

Подставив данные выражения в уравнение $xyz^2 = 10000$, найдем $\lambda = 0{,}01$. Отсюда $P^* = (25; 4; 10)$. Поэтому стоимость самого дешевого набора $S(P^*)$ будет $4 \cdot 25 + 25 \cdot 4 + 20 \cdot 10 = 400$.

§ 3.14. Выпуклые функции нескольких переменных

Напомним, что *прямой* в $\mathbb{R}^n$ называется множество точек $\vec{x}(t)$ следующего вида:

$$\vec{x}(t) = \vec{a} + t\vec{v}, \quad t \in \mathbb{R}, \tag{3.60}$$

где $\vec{a}, \vec{v} \in \mathbb{R}^n$ фиксированы, причем $\vec{v} \neq 0$. В координатной записи равенство (3.60) превращается в систему равенств:

$$x_1(t) = a_1 + tv_1\,, \dots, x_n(t) = a_n + tv_n.$$

Отрезком называется подмножество прямой, состоящее из точек $\vec{x}(t)$, для которых $t \in [0; 1]$. Точки $\vec{x}(0), \vec{x}(1)$ называются *концами* данного отрезка. Известно, что всякий отрезок $\vec{a}\vec{b}$ совпадает с множеством точек вида $\alpha\vec{a} + (1-\alpha)\vec{b}$, где $\alpha \in [0; 1]$. Множество $X \subset \mathbb{R}^n$ называется *выпуклым*, если оно вместе с любыми своими точками $\vec{a}$ и $\vec{b}$ содержит и весь отрезок $\vec{a}\vec{b}$.

Пусть функция $f(\vec{x})$ определена на выпуклом множестве $X \subset \mathbb{R}^n$. По аналогии с функциями одной переменной функция n переменных $f(\vec{x})$ называется *выпуклой* на X, если для любых $\vec{a}, \vec{b} \in X$ и любого $\alpha \in [0; 1]$ выполняется неравенство

$$f(\alpha\vec{a} + (1-\alpha)\vec{b}) \leq \alpha f(\vec{a}) + (1-\alpha) f(\vec{b}). \tag{3.61}$$

Функция $f(\vec{x})$ называется *строго выпуклой* на X, если для любых различных $\vec{a}, \vec{b} \in X$ и любого $\alpha \in (0; 1)$ неравенство (3.61) выполняется как строгое неравенство. Функция $f(\vec{x})$ называется *вогнутой* (*строго вогнутой*) на выпуклом множестве X, если функция $-f(\vec{x})$ является выпуклой (строго выпуклой) на X. Для того чтобы функция была вогнутой на выпуклом множестве X, необходимо и достаточно, чтобы для любых $\vec{a}, \vec{b} \in X$ и любого $\alpha \in [0; 1]$ выполнялось неравенство

$$f(\alpha\vec{a} + (1-\alpha)\vec{b}) \geq \alpha f(\vec{a}) + (1-\alpha) f(\vec{b}). \tag{3.62}$$

Из данных определений элементарно следует, что для выпуклой (строго выпуклой) функции $f(\vec{x})$ на выпуклом множестве X функция $\alpha f(\vec{x})$ также является выпуклой (строго выпуклой), если $\alpha > 0$. Если же $\alpha < 0$, то функция $\alpha f(\vec{x})$ будет вогнутой (строго вогнутой) на X.

Рассмотрим линейную функцию n переменных вида

$$l(\vec{x}) = c_1 x_1 + ... + c_n x_n + c_0.$$

Пусть $\vec{a}, \vec{b} \in \mathbb{R}^n$ и $\alpha, \beta \in \mathbb{R}, \alpha + \beta = 1.$ Тогда

$$\begin{aligned} l(\alpha\vec{a} + \beta\vec{b}) &= c_1[\alpha a_1 + \beta b_1] + ... + c_n[\alpha a_n + \beta b_n] + c_0 = \\ &= \alpha(c_1 a_1 + ... + c_k a_k + c_0) + \beta(c_1 b_1 + ... + c_k b_k + c_0) = \\ &= \alpha l(\vec{a}) + \beta l(\vec{b}). \end{aligned} \tag{3.63}$$

Из равенства (3.63) вытекают одновременно два неравенства (3.61) и (3.62) для функции $f(\vec{x}) = l(\vec{x})$. Отсюда – вывод: *линейная функция n переменных является одновременно выпуклой и вогнутой функцией в* $\mathbb{R}^n$.

Рассмотрим теперь ряд критериев выпуклости (строгой выпуклости) сложных функций и функций, являющихся суммами сложных функций.

Из свойств непрерывных функций следует, что множество значений $f(X)$ непрерывной функции $f(\vec{x})$ на выпуклом множестве X является промежутком (возможно, бесконечным). Для линейной функции $l(\vec{x})$ множество ее значений $l(X)$ на X также является промежутком. Это утверждение можно получить либо из непрерывности $l(\vec{x})$, либо непосредственно из равенства (3.63).

Теорема 3.13. *Пусть* $X \subset \mathbb{R}^n$ *– выпуклое множество*; $l_i(\vec{x})$ *– линейная функция от* $\vec{x} \in \mathbb{R}^n$, $f_i(t)$ *– функция на промежутке* $l_i(X)$, $i = 1, 2, ..., k$. *Определим функцию*

$$g(\vec{x}) = f_1(l_1(\vec{x})) + ... + f_k(l_k(\vec{x})), \quad \vec{x} \in X.$$

Тогда:

1) *если каждая функция* $f_i(t)$ ($i = 1, 2, ..., k$) *выпукла* (*вогнута*) *на* $l_i(X)$, *то* $g(\vec{x})$ *выпукла* (*вогнута*) *на* X;

2) *если каждая функция* $f_i(t)$ ($i = 1, 2, ..., k$) *строго выпукла* (*строго вогнута*) *на* $l_i(X)$ *и любая точки* $\vec{a} \in X$ *однозначно определяется набором чисел* $l_1(\vec{a}), ..., l_k(\vec{a})$, *то* $g(\vec{x})$ *строго выпукла* (*строго вогнута*) *на* X.

Доказательство. 1) Пусть $\vec{a}, \vec{b} \in X$ и $\alpha, \beta \in [0;1], \alpha + \beta = 1$. Для упрощения записей будем считать $k = 2$. Пусть $f_1(t)$, $f_2(t)$ – выпуклые функции. Тогда, используя формулу (3.63), получим

$$g(\alpha\vec{a} + \beta\vec{b}) = f_1(l_1(\alpha\vec{a} + \beta\vec{b})) + f_2(l_2(\alpha\vec{a} + \beta\vec{b})) =$$

$$= f_1(\alpha l_1(\vec{a}) + \beta l_1(\vec{b})) + f_2(\alpha l_2(\vec{a}) + \beta l_2(\vec{b})) \le$$

$$\le \alpha f_1(l_1(\vec{a})) + \beta f_1(l_1(\vec{b})) + \alpha f_2(l_2(\vec{a})) + \beta f_2(l_2(\vec{b})) = \alpha g(\vec{a}) + \beta g(\vec{b}),$$

т.е. $g(\vec{x})$ – выпуклая функция на X. Если $f_1(t)$, $f_2(t)$ – вогнутые функции, то доказательство аналогично.

2) Как и выше, считаем $k = 2$. Пусть $f_1(t)$ строго выпукла на $l_1(X)$, а $f_2(t)$ строго выпукла на $l_2(X)$. Тогда для различных точек $\vec{a}, \vec{b} \in X$ и $\alpha \in (0;1)$ имеем неравенства:

$$f_1(l_1(\alpha\vec{a} + \beta\vec{b})) \le \alpha f_1(l_1(\vec{a})) + \beta f_1(l_1(\vec{b})), \tag{3.64}$$

$$f_2(l_2(\alpha\vec{a} + \beta\vec{b})) \le \alpha f_2(l_2(\vec{a})) + \beta f_2(l_2(\vec{b})). \tag{3.65}$$

При этом неравенство (3.64) (соответственно (3.65)) является строгим, если $l_1(\vec{a}) \ne l_1(\vec{b})$ (соответственно, если $l_2(\vec{a}) \ne l_2(\vec{b})$). Так как $\vec{a} \in X$ однозначно определяется числами $l_1(\vec{a}), l_2(\vec{a})$, а $\vec{b} \ne \vec{a}$ и $\vec{b} \in X$, то либо $l_1(\vec{a}) \ne l_1(\vec{b})$, либо $l_2(\vec{a}) \ne l_2(\vec{b})$. Складывая неравенства (3.64) и (3.65) (по крайней мере, одно из них является строгим), получим

$$g(\alpha\vec{a} + \beta\vec{b}) < \alpha g(\vec{a}) + \beta g(\vec{b}),$$

т.е. $g(\vec{x})$ – строго выпуклая функция на X. В случае строго вогнутых функций $f_i(t)$ строгая вогнутость $g(\vec{x})$ доказывается аналогично.

Замечание. Если функции $f_i(t)$ выпуклы (строго выпуклы, вогнуты или строго вогнуты) на всей числовой прямой $\mathbb{R}$, то условие выпуклости (строгой выпуклости, вогнутости или строгой вогнутости) функции $f_i(t)$ на промежутке $l_i(X)$ для любого выпуклого множества $X \subset \mathbb{R}^n$ будет выполнено автоматически.

Пример 3.24. а) Пусть $f_1(t), ..., f_n(t)$ – строго выпуклые (строго вогнутые) функции на числовой прямой. Доказать, что функция n переменных $z = f_1(x_1) + ... + f_n(x_n)$ является строго выпуклой (строго вогнутой) в $\mathbb{R}^n$.

б) Доказать строгую выпуклость функции $z = x_1^2 + x_2^2$ в $\mathbb{R}^2$.

в) Доказать строгую вогнутость функции $z = x_1 - e^{x_1} - e^{x_2}$ в $\mathbb{R}^2$.

Решение. а) Пусть $l_i(\vec{x}) = x_i, i = 1, ..., n.$ Числа $l_1(\vec{a}), ..., l_n(\vec{a}),$ очевидно, полностью определяют $\vec{a} \in \mathbb{R}^n,$ поэтому z является строго выпуклой (строго вогнутой) функцией по теореме 3.13.

б) Частный случай а). Действительно, $f(t) = t^2$ – строго выпуклая функция, так как $(t^2)'' = 2>0$, а z можно представить в виде $z = f_1(x_1) + f_2(x_2)$, где $f_1(t) = f_2(t) = t^2$.

в) Положим $f_1(t) = t - e^t$, $f_2(t) = -e^t$. Поскольку $f_1''(t) = -e^t < 0$ и $f_2''(t) = -e^t < 0,$ $f_1(t)$ и $f_2(t)$ – строго вогнутые функции в $\mathbb{R}$. Функция $z = f_1(x_1) + f_2(x_2)$ является строго вогнутой в $\mathbb{R}^2$ как частный случай функции z из примера 3.24 а).

Пример 3.25. Пусть X – плоскость в $\mathbb{R}^3$, заданная уравнением $x_1 + x_2 + x_3 = 1.$ Доказать строгую выпуклость функции $z = x_1^2 + x_2^2$ на плоскости X.

Решение. Имеем $z = f(l_1(\vec{x})) + f(l_2(\vec{x}))$, где $f(t) = t^2$ – строго выпуклая функция в $\mathbb{R}$, а $l_1(\vec{x}) = x_1$, $l_2(\vec{x}) = x_2$ – координатные функции. Числа $l_1(\vec{a}), l_2(\vec{a})$ однозначно определяют $\vec{a} \in X$, поэтому z является строго выпуклой функцией по теореме 3.13.

Если теорема 3.13 применяется к функциям $f_i(\vec{x})$, которые обладают определенными свойствами выпуклости на промежутках, отличных от всей числовой прямой, то необходимо определить множества значений $l_i(X)$. Для произвольной линейной функции n переменных вида

$$l(\vec{x}) = c_1 x_1 + ... + c_n x_n + c_0$$

перечислим возможные типы промежутков $l(X)$ в следующих случаях:

1) $X = \mathbb{R}^n$ – все пространство;

2) $X = \mathbb{R}^n_+$ – множество неотрицательных векторов;

3) X – выпуклый многогранник.

В первом случае все просто: если хотя бы один из коэффициентов $c_i \neq 0$, то $l(X) = \mathbb{R}$.

Во втором случае множество значений зависит от знаков чисел $c_1, ..., c_n$. Так, если все $c_i \geq 0$ и хотя бы одно из этих чисел строго больше нуля, то $l(X) = [c_0, +\infty)$. Если все $c_i \leq 0$ и хотя бы одно $c_i < 0$, то $l(X) = (-\infty, c_0]$.

В третьем случае существуют: точка $\vec{a}$, в которой $l(\vec{x})$ принимает наименьшее значение на X, и точка $\vec{b}$, в которой $l(\vec{x})$ принимает наибольшее значение на X. Тогда $l(X)$ – отрезок $[l(\vec{a}), l(\vec{b})]$.

Пример 3.26. Проверить строгую вогнутость функций:

а) $z = \sqrt{x_1} + \sqrt{x_2}$ на множестве $X = \mathbb{R}^2_+$;

б) $z = \sin x_1 + \sin x_2$ на множестве $X = \{x_1 + x_2 \leq \pi,\ x_1 \geq 0,\ x_2 \geq 0\}$.

Решение. а) Для линейных функций x_1 и x_2 множества их значений на X таковы: $x_1(X) = [0, +\infty)$ и $x_2(X) = [0, +\infty)$. Далее, $\left(\sqrt{t}\right)'' = -\frac{1}{4} t^{-3/2} < 0$ при $t > 0$, поэтому функция $\sqrt{t}$ строго вогнута на промежутке $[0, +\infty)$. По теореме 3.13 функция $z = \sqrt{x_1} + \sqrt{x_2}$ строго вогнута на X.

б) Наименьшее значение x_1 на X равно 0, а наибольшее значение, достигаемое в точке $(\pi; 0)$, равно π. Аналогично наименьшее значение x_2 на X равно 0, а наибольшее значение, достигаемое в точке $(0; \pi)$, также равно π. Далее, $(\sin t)'' = -\sin t < 0$ при $t \in (0; \pi)$, поэтому функция $\sin t$ строго вогнута на отрезке $[0; \pi]$. По теореме 3.13 функция $z = \sin x_1 + \sin x_2$ строго вогнута на X.

Теорема 3.14. *Пусть* $X \subset \mathbb{R}^n$ *– выпуклое множество,* $f(\vec{x})$ *– выпуклая (вогнутая) функция на* X, $\psi(t)$ *– возрастающая выпуклая (вогнутая) функция на множестве значений* $f(X)$. *Тогда функция* $g(\vec{x}) = \psi(f(\vec{x}))$ *– строго выпукла (вогнута) на* X. *Если же* $f(\vec{x})$ *строго выпукла (строго вогнута) на* X, *то и* $g(\vec{x})$ *также строго выпукла (строго вогнута) на* X.

Доказательство. Рассмотрим случай строго выпуклой функции $f(\vec{x})$ и выпуклой функции $\psi(t)$. Для $\alpha, \beta \in (0;1)$, $\alpha + \beta = 1$ и любых различных $\vec{a}, \vec{b} \in X$ имеем

$$g(\alpha\vec{a} + \beta\vec{b}) = \psi(f(\alpha\vec{a} + \beta\vec{b})) < \psi(\alpha f(\vec{a}) + \beta f(\vec{b})) \le$$
$$\le \alpha\psi(f(\vec{\alpha})) + \beta\psi(f(\vec{b})) = \alpha g(\vec{\alpha}) + \beta g(\vec{b}),$$

т.е. $g(\vec{x})$ – строго выпуклая функция на X. В остальных случаях доказательство аналогично.

Пример 3.27. Проверить строгую выпуклость в $\mathbb{R}^2$ следующей функции: $z = (x_1^2 + x_2^2)^3$.

Решение. В $\mathbb{R}^2$ функция $f(x_1, x_2) = x_1^2 + x_2^2$ является строго выпуклой (пример 3.24, б). Множество ее значений в $\mathbb{R}^2$ – это множество всех неотрицательных чисел, т.е. $f(\mathbb{R}^2) = [0, +\infty)$. Далее, для функции $\psi(t) = t^3$ имеем $\psi'(t) = 3t^2 > 0$ и $\psi''(t) = 6t > 0$ при $t > 0$, поэтому $\psi(t)$ – возрастающая строго выпуклая функция на промежутке $[0, +\infty)$. По теореме 3.14 $z = \psi(f(x_1, x_2))$ – строго выпуклая функция в $\mathbb{R}^2$.

Пусть $\vec{a} \ne \vec{b}$ – две различные точки в $\mathbb{R}^n$. Прямую, соединяющую эти точки, можно представить в виде $l = \{\vec{x}(t) = \vec{a} + t(\vec{b} - \vec{a})\}$. Пусть $f(\vec{x})$ – функция n переменных. *Ограничением* функции $f(\vec{x})$ на прямую l называется функция $\varphi(t) = f(\vec{x}(t))$. Областью определения $\varphi(t)$ является множество всех t, для которых точка $\vec{x}(t)$ принадлежит области определения функции $f(\vec{x})$.

Теорема 3.15. *Для того чтобы функция* $f(\vec{x})$ *была выпуклой* (*строго выпуклой*, *вогнутой*, *строго вогнутой*) *на выпуклом множестве* $X \subset \mathbb{R}^n$, *необходимо и достаточно*, *чтобы для любых различных* $\vec{a}, \vec{b} \in X$ *функция* $\varphi(t) = f(\vec{a} + t(\vec{b} - \vec{a}))$ *была выпуклой* (*строго выпуклой*, *вогнутой*, *строго вогнутой*) *на отрезке* [0; 1].

Доказательство. *Необходимость*. Пусть $\vec{v} = \vec{b} - \vec{a}$. Тогда $\vec{x}(t) = \vec{a} + t\vec{v}$ и для любых чисел t_1, t_2 и $\alpha + \beta = 1$ выполняются равенства:

$$\vec{x}(\alpha t_1 + \beta t_2) = \vec{a} + (\alpha t_1 + \beta t_2)\vec{v} =$$
$$= \alpha\vec{a} + \alpha t_1\vec{v} + \beta\vec{a} + \beta t_2\vec{v} = \alpha\vec{x}(t_1) + \beta\vec{x}(t_2).$$

Положив $\vec{x}_1 = \vec{x}(t_1), \vec{x}_2 = \vec{x}(t_2)$, имеем

$$\varphi(\alpha t_1 + \beta t_2) = f(\vec{x}(\alpha t_1 + \beta t_2)) = f(\alpha\vec{x}_1 + \beta\vec{x}_2).$$

Пусть $f(\vec{x})$ – выпуклая функция на *X*. Тогда для любых t_1, t_2 и $\alpha + \beta = 1, \alpha \in [0;1]$ выполняется неравенство

$$\varphi(\alpha t_1 + \beta t_2) = f(\alpha\vec{x}_1 + \beta\vec{x}_2) \le \alpha f(\vec{x}_1) + \beta f(\vec{x}_2) = \alpha\varphi(t_1) + \beta\varphi(t_2).$$

Следовательно, $\varphi(t)$ – выпуклая функция на отрезке [0; 1]. Аналогично рассматриваются случаи, когда $f(\vec{x})$ – строго выпуклая, вогнутая или строго вогнутая функция.

Достаточность. Пусть функция $\varphi(t) = f(\vec{a} + t(\vec{b} - \vec{a}))$ является выпуклой для любых $\vec{a}, \vec{b} \in X$, $\vec{a} \ne \vec{b}$. Тогда для всяких $\alpha \in [0;1]$, $\alpha + \beta = 1$ имеем

$$f(\alpha\vec{a} + \beta\vec{b}) = f(\alpha\vec{x}(0) + \beta\vec{x}(1)) = f(\vec{x}(\alpha \cdot 0 + \beta \cdot 1)) =$$
$$= \varphi(\alpha \cdot 0 + \beta \cdot 1) \le \alpha \cdot \varphi(0) + \beta \cdot \varphi(1) = \alpha f(\vec{a}) + \beta f(\vec{b}),$$

т.е. $f(\vec{x})$ – выпуклая функция на *X*. Случаи, когда $\varphi(t)$ – строго выпуклая, вогнутая или строго вогнутая функция, рассматриваются аналогично.

Теорема 3.16. *Пусть* $X = \mathbb{R}_+^n$ *– множество неотрицательных векторов*; $\alpha_1, ..., \alpha_n$ *– неотрицательные числа*, *сумма которых* $\alpha \le 1$. *Тогда функция* $f(\vec{x}) = x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n}$ *вогнута на X*.

Доказательство. Пусть $\vec{p}, \vec{q} \in X$ и $\vec{p} \ne \vec{q}$. Положим $\vec{v} = \vec{q} - \vec{p}$ и рассмотрим функцию $y = \varphi(t) = f(\vec{x}(t))$, где $\vec{x}(t) = \vec{p} + t\vec{v}$. Если хотя бы одна из линейных функций $x_i(t) = p_i + tv_i$ $(i = 1, ..., n)$ обращается тождественно в нуль (т.е. $p_i = v_i = 0$), то и функция $\varphi(t)$ также тождественно равна нулю. Далее считаем, что все $x_i(t)$ являются ненулевыми функциями.

Положим $t_i = -\dfrac{p_i}{v_i}$, если $v_i \ne 0$ $(i = 1, ..., n)$. Тогда слева и справа от точки t_i на числовой прямой функция $x_i(t)$ принимает значения разных знаков. Так как $\vec{x}(0) = \vec{p} \ge 0$ и $\vec{x}(1) = \vec{q} \ge 0$, то ни одно из чисел t_i не может быть внутренней точкой отрезка [0; 1], т.е. $t_i \notin (0; 1)$. Следовательно, все $x_i(t) > 0$ на интервале (0; 1). Поэтому

$$\varphi(t) = x_1(t)^{\alpha_1} x_2(t)^{\alpha_2} ... x_n(t)^{\alpha_n} > 0$$

на интервале (0; 1). Так как $y = \varphi(t) > 0$, то, используя логарифмическую производную, запишем y' в виде $y' = y(\ln y)'$. Дифференцируя последнее равенство по *t*, получим

$$y'' = y([(\ln y)']^2 + [\ln y]'').$$

Находим квадрат логарифмической производной:

$$[(\ln y)']^2 = [(\alpha_1 \ln(p_1 + tv_1)' + ... + (\alpha_n \ln(p_n + tv_n)']^2 =$$
$$= \left(\frac{\alpha_1 v_1}{p_1 + tv_1} + ... + \frac{\alpha_n v_n}{p_n + tv_n} \right)^2.$$

Далее,

$$[\ln y]'' = \left[\frac{\alpha_1 v_1}{p_1 + tv_1} + \ldots + \frac{\alpha_n v_n}{p_n + tv_n}\right]' = -\left[\frac{\alpha_1 v_1^2}{(p_1 + tv_1)^2} + \ldots + \frac{\alpha_n v_n^2}{(p_n + tv_n)^2}\right].$$

Пусть $\alpha = \alpha_1 + \ldots. + \alpha_n = 1$. Из неравенства Йенсена для выпуклой функции $\varphi(s) = s^2$ (§ 2.18) получаем

$$[\alpha_1 s_1 + \ldots + \alpha_n s_n]^2 \le \alpha_1 s_1^2 + \ldots + \alpha_n s_n^2,$$

или

$$[\alpha_1 s_1 + \ldots + \alpha_n s_n]^2 - [\alpha_1 s_1^2 + \ldots + \alpha_n s_n^2] \le 0.$$

Положив $s_i = \dfrac{v_i}{p_i + tv_i}$, имеем $y([(\ln y)']^2 + [\ln y]'') \le 0$. Поэтому $y'' \le 0$. Итак, мы доказали, что: либо функция $y = \varphi(t)$ тождественно равна нулю, либо $\varphi''(t) \le 0$ на интервале (0; 1). В обоих случаях $\varphi(t)$ – вогнутая функция на отрезке [0; 1]. По теореме 3.15 функция $f(\vec{x})$ вогнута на $X = \mathbb{R}_+^n$.

Если же $\alpha < 1$, то функцию $f(\vec{x})$ можно представить как сложную функцию $f(\vec{x}) = [g(\vec{x})]^\alpha$, где

$$g(\vec{x}) = x_1^{\beta_1} \ldots x_n^{\beta_n},\ \beta_1 = \frac{\alpha_1}{\alpha},\ \ldots,\ \beta_n = \frac{\alpha_n}{\alpha},\ \beta_1 + \ldots + \beta_n = 1.$$

Как было доказано, $g(\vec{x})$ вогнута на $\mathbb{R}_+^n$. Кроме того, $\psi(t) = t^\alpha$ – возрастающая вогнутая функция на промежутке [0; +∞), поскольку $\psi'(t) = \alpha t^{\alpha-1} > 0$ и $\psi''(t) = \alpha(\alpha-1)t^{\alpha-2} < 0$ при $t > 0$. По теореме 3.15 функция $f(\vec{x}) = \psi(g(\vec{x}))$ является вогнутой на $X = \mathbb{R}_+^n$.

С помощью теоремы 3.16 доказывается вогнутость производственных функций Кобба–Дугласа. Например, функция $Y = K^{1/4}L^{3/4}$

является вогнутой, поскольку сумма степеней переменных равна $\frac{1}{4}+\frac{3}{4}=1.$

Теорема 3.17. *Сумма любого числа выпуклых (вогнутых) на выпуклом множестве $X \subset \mathbb{R}^{n\ n}$ функций также является выпуклой (вогнутой) функцией на X. Если при этом, хотя бы одна из суммируемых функций является строго выпуклой (строго вогнутой) на X, то и вся сумма будет строго выпуклой (строго вогнутой) на X.*

Доказательство теоремы 3.17, по существу, повторяет доказательство теоремы 2.19, поэтому мы его не приводим.

§ 3.15. Квадратичные формы и выпуклые функции

Теорема 3.18. *Пусть функция $f(\vec{x})$ имеет непрерывные частные производные второго порядка на открытом выпуклом множестве $X \subset \mathbb{R}^n$. Тогда:*

1) *для того чтобы функция f была строго выпукла на X, достаточно, чтобы квадратичная форма $d^2 f_{\vec{a}}$ была положительно определена в каждой точке $\vec{a} \in X$;*

2) *для того чтобы f была выпукла на X, необходимо и достаточно, чтобы квадратичная форма $d^2 f_{\vec{a}}$ была неотрицательно определена в каждой точке $\vec{a} \in X$.*

Доказательство. 1) Пусть $\vec{a}, \vec{b} \in X, \vec{a} \neq \vec{b}.$ Для каждого $t \in \mathbb{R}$ положим $\vec{c}(t) = \vec{a} + t(\vec{b} - \vec{a}).$ Из выпуклости X следует, что $\vec{c}(t) \in X$ для $t \in [0;1]$. Определим функцию $\varphi(t) = f(\vec{c}(t)).$ Поскольку

$$\varphi''(t) = d^2 f_{\vec{c}(t)}(\vec{b} - \vec{a}) > 0,$$

функция $\varphi(t)$ строго выпукла на $[1;0]$ вследствие теоремы 2.21. По теореме 3.15 функция $f(\vec{x})$ строго выпукла на X.

2) *Достаточность*. Следует из теорем 2.21 и 3.15 аналогично 1).

Необходимость. Предположим противное: найдутся $\vec{a} \in X$ и $\vec{v} \in \mathbb{R}^n$ такие, что $d^2 f_{\vec{a}}(\vec{v}) < 0$. Положим $\vec{c}(t) = \vec{a} + t\vec{v}$. Из непрерывности вторых частных производных f следует, что существует $\varepsilon > 0$, такое что $d^2 f_{\vec{c}(t)}(\vec{v}) < 0$ для $t \in [-\varepsilon; \varepsilon]$; поэтому функция $\varphi(t) = f(\vec{c}(t))$ строго вогнута на отрезке $[-\varepsilon; \varepsilon]$. В частности, $f(\vec{a}) > \frac{1}{2}(f(\vec{a} - \varepsilon\vec{v}) + f(\vec{a} + \varepsilon\vec{v}))$, что противоречит выпуклости f. Полученное противоречие доказывает необходимость неотрицательной определенности $d^2 f$.

В качестве иллюстрации рассмотрим достаточное условие строгой выпуклости (строгой вогнутости) для функции двух переменных.

Теорема 3.19. *Пусть D – выпуклое открытое подмножество координатной плоскости Oxy; $f(x, y)$ – функция, имеющая в D непрерывные частные производные второго порядка. Положим*

$$\Delta(x, y) = f''_{xx}(x, y) f''_{yy}(x, y) - [f''_{xy}(x, y)]^2.$$

Тогда функция $f(x, y)$ является строго выпуклой (строго вогнутой) на множестве D, если в каждой точке $P \in D$ выполняются неравенства:

1) $\Delta(P) > 0$;

2) $f''_{xx}(P) > 0 \quad (f''_{xx}(P) < 0)$.

Доказательство. Из условий 1) и 2) вытекает положительная (отрицательная) определенность второго дифференциала $d^2 f_P$, $P \in D$. По теореме 3.18 функция $f(x, y)$ является строго выпуклой (строго вогнутой) на множестве D.

Пример 3.28. Проверить строгую выпуклость в $\mathbb{R}^2$ функции

$$z = x^2 + xy + y^2 + \sin x.$$

Решение. Имеем $z'_x = 2x + y + \cos x$, $z'_y = x + 2y$. Далее,

$$z''_{xx} = 2 - \sin x,\ z''_{xy} = 1,\ z''_{yy} = 2,\ \Delta = z''_{xx} z''_{yy} - (z''_{xy})^2 = 3 - 2\sin x.$$

В любой точке $P(x, y) \in \mathbb{R}^2$ выполняются неравенства:

$$z''_{xx}(P) = 2 - \sin x > 0 \text{ и } \Delta(P) = 3 - 2\sin x > 0.$$

По теореме 3.19 функция z является строго выпуклой в $\mathbb{R}^2$.

Теорема 3.20. *Квадратичная форма* $q(\vec{x}), \vec{x} \in \mathbb{R}^n$, *является выпуклой* (*строго выпуклой*) *функцией в том и только в том случае*, *если* $q(\vec{x})$ *неотрицательно* (*положительно*) *определена*.

Доказательство. Нетрудно убедиться в том, что второй дифференциал квадратичной формы совпадает с ней самой: $d^2 q_{\vec{a}}(\vec{v}) = q(\vec{v})$. Отсюда и из утверждения 2) теоремы 3.18 вытекает эквивалентность понятий выпуклости и неотрицательной определенности для квадратичной формы. Аналогично, из утверждения 1) теоремы 3.18 вытекает строгая выпуклость $q(\vec{x})$ как следствие ее положительной определенности.

Докажем теперь обратное: строгая выпуклость $q(\vec{x})$ влечет ее положительную определенность. Для фиксированного ненулевого вектора $\vec{x}$ рассмотрим функцию $\varphi(t) = q(t\vec{x}) = t^2 q(\vec{x})$. Строгая выпуклость $q(\vec{x})$ означает и строгую выпуклость $\varphi(t)$. Последнее возможно лишь в том случае, если $q(\vec{x}) > 0$, что и означает положительную определенность $q(\vec{x})$.

Пример 3.29. При каких значениях параметра a выпукла функция

$$f(x, y, z) = x^2 + 5y^2 + az^2 + 4xy + 8yz \text{ ?}$$

Решение. По теореме 3.20 вопрос о выпуклости f эквивалентен вопросу о неотрицательной определенности f. Матрица f имеет вид

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 4 \\ 0 & 4 & a \end{pmatrix}.$$

Находим угловые миноры, необходимые для критерия Сильвестра:

$$\Delta_1 = 1, \quad \Delta_2 = \begin{vmatrix} 1 & 2 \\ 2 & 5 \end{vmatrix} = 1, \quad \Delta_3 = \begin{vmatrix} 1 & 2 & 0 \\ 2 & 5 & 4 \\ 0 & 4 & a \end{vmatrix} = a - 16.$$

Условия положительной определенности $\Delta_1 > 0, \Delta_2 > 0, \Delta_3 > 0$ приводят к неравенству $a > 16$. При $a < 16$ определитель $\Delta_3 < 0$ – квадратичная форма f знаконеопределена. Остается исследовать случай $a = 16$. Поскольку f непрерывно зависит от a и $f(x, y, z) \geq 0$ при любых x, y, z и $a > 16$, переходя к пределу справа при $a \to 16$, получим $f(x, y, z) \geq 0$ и в случае $a = 16$. Таким образом, f неотрицательно определена, если $a = 16$. Этот же факт можно было установить и без предельного перехода с помощью критерия неотрицательной определенности, сформулированного в замечании 2 из п.2 § 3.12.

О т в е т : f выпукла $\Leftrightarrow$ $a \geq 16$.

§ 3.16. Экстремумы и стационарные точки выпуклых функций

1. Глобальные и локальные экстремумы

О п р е д е л е н и е . *Точкой* ***глобального минимума*** (***глобального максимума***) *функции* $f(x_1, ..., x_n)$ *на множестве* $X \subset \mathbb{R}^n$ *называется точка* $P \in X$, *в которой функция* $f(x_1, ..., x_n)$ *принимает наименьшее* (*наибольшее*) *значение на* X. *Точки глобального минимума и точки глобального максимума называются точками* ***глобального экстремума***. *Точки локального условного экстремума* f *при условии* $(x_1, ..., x_n) \in X$ *называются также точками* ***локального экстремума*** f ***на*** X.

Ясно, что каждый глобальный экстремум f на X одновременно является и локальным экстремумом f на X. Обратное утверждение в общем случае неверно. Тем не менее, как показывает следующая теорема, для достаточно широкого класса функций f и множеств X всякий локальный экстремум f на X является глобальным.

Теорема 3.21. *Пусть P – локальный (условный) экстремум функции f на выпуклом множестве X. Тогда*:

1) *если P – минимум, а f выпукла, то P – глобальный минимум f на X*;
2) *если P – максимум, а f вогнута, то P – глобальный максимум f на X.*

Доказательство. 1) Предположим противное: существует точка $Q \in X$, такая что $f(Q) < f(P)$. Каждую точку A отрезка PQ, отличную от P, можно представить в виде $A = (1-\alpha)P + \alpha Q$, где $\alpha \in (0;1]$. Из выпуклости f следует, что

$$f(A) \le (1-\alpha) f(P) + \alpha f(Q) = f(P) + \alpha(f(Q) - f(P)) < f(P).$$

Таким образом, $f(A) < f(P)$, причем точка A может быть выбрана сколь угодно близко к точке P. Последнее неравенство противоречит тому, что P – точка локального минимума f на X. Данное противоречие доказывает, что для любой точки $Q \in X$ выполняется неравенство $f(Q) \ge f(P)$, т.е. P – точка глобального минимума f на X.

Утверждение 2) доказывается совершенно аналогично.

2. Стационарные точки и глобальные экстремумы

Пусть $f(x_1, ..., x_n)$ – функция n переменных, определенная в некоторой окрестности точки $P \in \mathbb{R}^n$. Напомним, что точка P называется *стационарной* точкой функции f, если f дифференцируема в P и $\operatorname{grad} f(P) = 0$, Для функции f, определенной на множестве $X \subset \mathbb{R}^n$ или на некотором расширении этого множества, будем рассматривать стационарные точки двух видов:

1) P – внутренняя точка X,
2) P – граничная точка X.

В случае 2) считается, что область определения функции f содержит не только множество X, но и некоторую окрестность точки P.

Теорема 3.22. *Пусть* $X \subset \mathbb{R}^n$ *– выпуклое множество,* f *– вогнутая (выпуклая) функция на* X, $P \in X$ *– стационарная точка* f. *Тогда* $f(P)$ *– наибольшее (наименьшее) значение* f *на* X.

Доказательство. Пусть $Q \in X$, $Q \neq P$. Рассмотрим функцию одной переменной $\varphi(t) = f(P + t\overrightarrow{PQ})$. Используя формулу (3.20), находим ее производную в нуле

$$\varphi'(0) = (\operatorname{grad} f(P), \overrightarrow{PQ}).$$

Градиент $\operatorname{grad} f(P) = 0$, поскольку P – стационарная точка. Следовательно,

$$\varphi'(0) = (0, \overrightarrow{PQ}) = 0,$$

т.е. 0 – стационарная точка функции $\varphi(t)$. Пусть функция f вогнута на X. Вследствие теоремы 3.15, функция $\varphi(t)$ вогнута на отрезке [0; 1]. Так как 0 – ее стационарная точка, то по теореме 2.24 $\varphi(0)$ – наибольшее значение функции $\varphi(t)$ на отрезке [0; 1]. Следовательно,

$$f(P) = \varphi(0) \geq \varphi(1) = f(Q),$$

т.е. $f(P)$ – наибольшее значение f на X.

Случай выпуклой функции f рассматривается аналогично.

3. Единственность экстремума выпуклой функции

Теорема 3.23. *Строго выпуклая (строго вогнутая) функция на выпуклом множестве* $X \subset \mathbb{R}^n$ *не может иметь более одной точки глобального минимума (максимума).*

Доказательство этой теоремы почти дословно повторяет доказательство теоремы 2.25, поэтому мы его не приводим.

Пример 3.30. Предположим, что некоторая монополия продает свою продукцию на двух независимых рынках. Пусть x, p_x – количество и цена продукции, проданной на 1-м рынке; y, p_y – количество и цена продукции, проданной на 2-м рынке; $D_1(p_x)$, $D_2(p_y)$ – функции спроса; $C(q)$ – функция издержек. Найти точку (x^*, y^*), в которой значение функции прибыли

$$\pi(x, y) = p_x x + p_y y - C(x + y)$$

максимально, если

$$D_1(p_x) = \left(\frac{160}{p_x}\right)^2, \; D_2(p_y) = \left(\frac{320}{p_y}\right)^2, \; C(q) = q^2.$$

Решение. Из уравнений $x = D(p_x)$, $y = D(p_y)$ находим:

$$p_x = 160x^{-0,5}, \quad p_y = 320y^{-0,5}.$$

Следовательно, выручка

$$R(x, y) = p_x x + p_y y = 160\sqrt{x} + 320\sqrt{y}.$$

С помощью теоремы 3.13 легко проверяется, что $R(x, y)$ – строго вогнутая, а $-C(x + y)$ – вогнутая функции в первой четверти координатной плоскости Oxy. По теореме 3.17 их сумма $\pi(x, y)$ будет строго вогнутой. Найдем стационарные точки $\pi(x, y)$, решив систему уравнений:

$$\begin{cases} \pi'_x = \dfrac{80}{\sqrt{x}} - 2(x + y) = 0, \\ \pi'_y = \dfrac{160}{\sqrt{y}} - 2(x + y) = 0. \end{cases}$$

Имеем $\dfrac{80}{\sqrt{x}} = \dfrac{160}{\sqrt{y}} = 2(x+y)$. Откуда $y = 4x$, $40 = 5x^{3/2}$. Следовательно,

$$x = 4, \quad y = 16, \quad \pi(4;\, 16) = 1200.$$

Из теорем 3.22, 3.23 получаем, что 1200 – максимальное значение прибыли, а точка $(x^*; y^*) = (4; 16)$ – единственная, в которой это значение достигается.

§ 3.17. Теорема Куна–Таккера

Доказанная в предыдущем параграфе теорема 3.22 означает, что стационарная точка функции f является точкой глобального экстремума f на любом выпуклом множестве, на котором f – выпуклая (или вогнутая) функция. Позволяя получать простые решения некоторых важных с практической точки зрения задач, данная теорема оказывается бесполезной при решении задач общего вида. Объясняется это тем, что глобальный экстремум функции f на X часто достигается в граничной точке P множества X, а в этом случае, как правило, градиент f в точке P не равен 0, т.е. P – не стационарная точка.

Ниже при исследовании экстремумов функции на выпуклом множестве мы используем терминологию, принятую в теории оптимизации.

Определение. *Задача об отыскании экстремумов функции f на множестве X называется **задачей математического программирования**. Функция f при этом называется **целевой функцией**, а X – **допустимым множеством**.*

Задача на минимум коротко записывается следующим образом:

$$f(x_1, ..., x_n) \to \min, \quad (x_1, ..., x_n) \in X. \tag{3.66}$$

Задача на максимум записывается аналогично:

$$f(x_1, ..., x_n) \to \max, \quad (x_1, ..., x_n) \in X. \tag{3.67}$$

Глобальным решением задачи (3.66) (соответственно (3.67)) называется точка глобального минимума (соответственно максимума) функции f на X. *Локальным решением* задачи (3.66) (соответственно (3.67)) называется точка локального условного минимума (соответственно максимума) функции $f(x_1, ..., x_n)$ при условии, что $(x_1, ..., x_n) \in X$.

Трудность задачи математического программирования, очевидно, определяется типом целевой функции и геометрическими свойствами допустимого множества. Методы решения разнообразных классов задач математического программирования изучаются в теории оптимизации. Здесь же мы ограничимся рассмотрением только одного важного для приложений класса.

1. Постановка задачи выпуклого программирования

Определение. *Задача* (3.66) (*соответственно* (3.67)) *называется задачей* ***выпуклого программирования****, если целевая функция* f *является выпуклой* (*соответственно вогнутой*), *а допустимое множество* X *выпукло и задано системой неравенств или равенств*.

Замечание. Из теоремы 3.21 следует, что всякое локальное решение задачи выпуклого программирования является ее глобальным решением.

Поскольку всякая система ограничений, содержащая как равенства, так и неравенства, эквивалентна системе, содержащей одни лишь неравенства, будем считать, что допустимое множество X задано системой вида

$$\begin{cases} g_1(x_1,...,x_n) \geq 0, \\ \quad \dots\dots\dots\dots\dots \\ g_s(x_1,...,x_n) \geq 0. \end{cases} \tag{3.68}$$

Фундаментальную роль в исследовании задачи выпуклого программирования играет уже знакомая читателю функция Лагранжа:

$$L(x_1, ..., x_n, \lambda_1, ..., \lambda_s) = f(x_1, ..., x_n) + \sum_{j=1}^{s} \lambda_j g_j(x_1, ..., x_n).$$

Используя векторные обозначения $\vec{x}=(x_1,...,x_n)$, $\vec{\lambda}=(\lambda_1,...,\lambda_s)$, функции f, g_j и L можно записать в виде

$$f(\vec{x})=f(x_1,...,x_n), \quad g_j(\vec{x})=g_j(x_1,...,x_n),$$
$$L(\vec{x},\vec{\lambda})=f(\vec{x})+\lambda_1 g_1(\vec{x})+...+\lambda_s g_s(\vec{x}).$$

Чтобы получить достаточно простые условия экстремума, предположим, что допустимое множество X удовлетворяет следующему условию регулярности.

Определение. *Система ограничений* (3.68), *задающая выпуклое множество* X, *называется* ***регулярной***, *если все функции* $g_1,...,g_s$ *вогнуты на X и существует точка* $\vec{x}^+\in X$, *в которой все нелинейные среди* $g_1,...,g_s$ *функции принимают положительные значения.*

Замечание. Система (3.68), состоящая из одних только линейных неравенств, считается регулярной. Необходимость выбора точки $\vec{x}^+$ в этом случае отпадает.

Пример 3.31. Проверить регулярность системы ограничений

$$\begin{cases} x\ge 0, y\ge 0, z\ge 0, \\ x+y+z-36\ge 0, \\ -x-y-z+36\ge 0, \\ \sqrt[3]{x}+\sqrt[3]{y}+\sqrt[3]{z}-5\ge 0. \end{cases} \tag{3.69}$$

Решение. Пусть M – множество всех решений системы (3.69), а K – множество всех решений подсистемы в (3.69), состоящей из первых 5 неравенств. Нетрудно видеть, что K – треугольник с вершинами $(36;0;0)$, $(0;36;0)$ и $(0;0;36)$; M – подмножество в K, заданное неравенством $\sqrt[3]{x}+\sqrt[3]{y}+\sqrt[3]{z}\ge 5$. Так как функция $\sqrt[3]{x}+\sqrt[3]{y}+\sqrt[3]{z}$ вогнута в $\mathbb{R}^3_+$, то она вогнута и на K. Отсюда следует выпуклость M. Левые части первых пяти неравенств – линейные функции, поэтому они вогнуты в $\mathbb{R}^3$. Левая часть последнего неравенства вогнута в $\mathbb{R}^3_+$. Таким образом, все левые части вогнуты на *M*. Осталось указать точку из *M*, в которой все нелинейные ограничения обращаются в строгие неравенства. В качестве такой точки можно взять $(1; 8; 27)$.

2. Необходимое условие экстремума

Нетрудно видеть, что замена целевой функции f на $-f$ меняет ролями задачи (3.66) и (3.67), поэтому при рассмотрении задач выпуклого программирования достаточно ограничиться задачами какого-то одного типа, например задачами на максимум.

Теорема 3.24. *Пусть* $\vec{x}^* \in X$ *– точка локального максимума функции* $f(\vec{x})$ *на выпуклом множестве* X, *заданном регулярной системой ограничений* (3.68). *Если все функции* $f, g_1, ..., g_s$ *дифференцируемы в точке* $\vec{x}^*$, *то существует такой вектор* $\vec{\lambda}^* = (\lambda_1^*, ..., \lambda_s^*) \geq 0$, *что*

$$L'_{x_i}(\vec{x}^*, \vec{\lambda}^*) = 0 \quad (i = 1, 2, ..., n), \tag{A}$$

$$\lambda_j^* g_j(\vec{x}^*) = 0 \quad (j = 1, 2, ..., s). \tag{B}$$

Доказательство. В точке $\vec{x}^*$ все функции $g_j \geq 0$, т.е. для каждого j либо $g_j(\vec{x}^*) = 0$, либо $g_j(\vec{x}^*) > 0$. Введем обозначения для соответствующих множеств номеров:

$$J_0 = \{j : g_j(\vec{x}^*) = 0\}, \quad J_+ = \{j : g_j(\vec{x}^*) > 0\}.$$

Для каждого $j \in J_0$ функция g_j может быть или не быть линейной, поэтому J_0 разбивается на два подмножества $J_0 = \bar{J}_0 \cup \tilde{J}_0$, где $\bar{J}_0$ – подмножество номеров *линейных* функций среди g_j $(j \in J_0)$, а $\tilde{J}_0$ – подмножество номеров *нелинейных* функций среди тех же функций g_j.

Доказательство теоремы построено на анализе неравенств, связывающих следующие линейные однородные функции от переменных $z_1, z_2, ..., z_n$:

$$\frac{\partial f}{\partial x_1}(\vec{x}^*)z_1 + \frac{\partial f}{\partial x_2}(\vec{x}^*)z_2 + ... + \frac{\partial f}{\partial x_n}(\vec{x}^*)z_n,$$

$$\frac{\partial g_j}{\partial x_1}(\vec{x}^*)z_1 + \frac{\partial g_j}{\partial x_2}(\vec{x}^*)z_2 + ... + \frac{\partial g_j}{\partial x_n}(\vec{x}^*)z_n.$$

Чтобы немного сократить обозначения, представим данные функции в виде скалярных произведений $(f'(\vec{x}^*), \vec{z})$ и $(g'_j(\vec{x}^*), \vec{z})$, где $f'(\vec{x}^*)$ – градиент функции f в точке $\vec{x}^*$, а $g'_j(\vec{x}^*)$ – градиент функции g_j в той же точке $\vec{x}^*$, $\vec{z} = (z_1, z_2, ..., z_n)$.

Дальнейшую часть доказательства разобьем на три этапа.

Этап 1. Рассмотрим систему линейных неравенств

$$\begin{cases} (f'(\vec{x}^*), \vec{z}) > 0, \\ (g'_j(\vec{x}^*), \vec{z}) \geq 0, \quad j \in \bar{J}_0, \\ (g'_j(\vec{x}^*), \vec{z}) > 0, \quad j \in \tilde{J}_0, \end{cases} \tag{3.70}$$

где неизвестным является вектор $\vec{z}$. Убедимся, что система (3.70) несовместна.

Действительно, если бы существовало решение $\vec{z} = \vec{a}$, то в силу первого неравенства производная функции $\varphi(t) = f(\vec{x}^* + t\vec{a})$ в точке $t = 0$ была бы больше 0, и при достаточно малом $t > 0$ мы имели бы

$$f(\vec{x}^* + t\vec{a}) > f(\vec{x}^*),$$

и, аналогично,

$$g_j(\vec{x}^* + t\vec{a}) > g_j(\vec{x}^*) = 0 \quad (j \in \tilde{J}_0).$$

Для линейных g_j при любом $t > 0$ имеем

$$g_j(\vec{x}^* + t\vec{a}) = g_j(\vec{x}^*) + (g'_j(\vec{x}^*),\ t\vec{a}) \geq g_j(\vec{x}^*) \geq 0 \quad (j \in \bar{J}_0).$$

Что же касается номеров $j \in J_+$, то для таких j выполняется неравенство $g_j(\vec{x}^*) > 0$, а в силу непрерывности функций g_j в точке $\vec{x}^*$ при достаточно малом $t > 0$ выполняется также и неравенство $g_j(\vec{x}^* + t\vec{a}) > 0$. Итак, при достаточно малом $t > 0$ имеем $f(\vec{x}^* + t\vec{a}) > f(\vec{x}^*)$ и $g_j(\vec{x}^* + t\vec{a}) \geq 0$ для всех $j = 1, 2, ..., s$. Это означает, что $\vec{x}^*$ не является точкой локального максимума f на X вопреки условию. Полученное противоречие завершает этап 1.

Этап 2. Покажем, что несовместна и более слабая система

$$\begin{cases} (f'(\vec{x}^*), \vec{z}) > 0, \\ (g_j'(\vec{x}^*), \vec{z}) \geq 0, \ \ j \in J_0. \end{cases} \tag{3.71}$$

Ослабление по сравнению с (3.70) состоит в том, что часть строгих неравенств заменяется нестрогими.

Рассмотрим приращение функции g_j, связанное с переходом из точки $\vec{x}^*$ в точку $\vec{x} = \vec{x}^* + \Delta\vec{x}$:

$$\Delta g_j = g_j(\vec{x}^* + \Delta\vec{x}) - g_j(\vec{x}^*).$$

Как и для любой дифференцируемой вогнутой функции, приращение функции g_j не превосходит ее дифференциала

$$dg_j = \sum_{i=1}^{n} \frac{\partial g_j}{\partial x_i}(\vec{x}^*)\Delta x_i = (g_j'(\vec{x}^*), \Delta\vec{x}) \geq \Delta g_j. \tag{3.72}$$

Согласно условию регулярности можем выбрать точку $\vec{x} \in X$ так, чтобы все нелинейные функции g_j в этой точке были больше нуля. Фиксируем точку $\vec{x}$ и соответствующий вектор $\Delta\vec{x}$ до конца доказательства.

Для любого $j \in \widetilde{J}_0$ с учетом выбора $\Delta\vec{x}$ неравенство (3.72) приобретает вид

$$(g_j'(\vec{x}^*), \Delta\vec{x}) \geq g_j(\vec{x}) - g_j(\vec{x}^*) = g_j(\vec{x}) > 0 \quad (j \in \widetilde{J}_0). \tag{3.73}$$

Аналогичным образом для $j \in \bar{J}_0$ имеем $g_j(\vec{x}) \geq 0,$ поэтому

$$(g'_j(\vec{x}^*), \Delta\vec{x}) \geq 0 \quad (j \in \bar{J}_0). \tag{3.74}$$

Рассуждая от противного, предположим, что система (3.71) имеет решение $\vec{b}$. Покажем, что вектор $\vec{z} = \vec{b} + t\Delta\vec{x}$ при достаточно малом $t > 0$ удовлетворяет системе (3.70).

Так как вектор $\vec{b}$ – решение (3.71), то $(f'(\vec{x}^*), \vec{z}) > 0.$ Поэтому при достаточно малом $t > 0$ имеем

$$(f'(\vec{x}^*), \vec{b} + t\Delta\vec{x}) = (f'(\vec{x}^*), \vec{b}) + t(f'(\vec{x}^*), \Delta\vec{x}) > 0.$$

Кроме того, поскольку $\vec{b}$ – решение (3.71), для каждого $j \in J_0$ еще выполняется неравенство $(g'_j(\vec{x}^*), \vec{b}) \geq 0.$ Вследствие (3.73), для $t > 0$ и $j \in \tilde{J}_0$ имеем

$$(g'_j(\vec{x}^*), \vec{b} + t\Delta\vec{x}) = (g'_j(\vec{x}^*), \vec{b}) + t(g'_j(\vec{x}^*), \Delta\vec{x}) > 0.$$

Аналогично, вследствие (3.74), для $t > 0$ и $j \in \bar{J}_0$ получаем

$$(g'_j(\vec{x}^*), \vec{b} + t\Delta\vec{x}) \geq 0.$$

Итак, при достаточно малом $t > 0$ вектор $\vec{z} = \vec{b} + t\Delta\vec{x}$ удовлетворяет несовместной системе (3.70). Полученное противоречие доказывает, что система (3.71) также не имеет решений.

Этап 3. Завершение доказательства теоремы.

Несовместность системы (3.71) означает, что ни одно решение системы

$$(g'_j(\vec{x}^*), \vec{z}) \geq 0, \ \ j \in J_0, \tag{3.75}$$

не удовлетворяет неравенству $(f'(\vec{x}^*), \vec{z}) > 0.$ Следовательно, всякое решение системы (3.75) удовлетворяет неравенству $(f'(\vec{x}^*), \vec{z}) \le 0$ или, что то же,

$$(-f'(\vec{x}^*), \vec{z}) \ge 0. \qquad (3.76)$$

Таким образом, неравенство (3.76) – следствие системы (3.75). По теореме Фаркаша (см. ч.1 учебника) существуют такие числа λ_j^* $(j \in J_0)$, что

$$-f'(\vec{x}^*) = \sum_{j \in J_0} \lambda_j^* g_j'(\vec{x}^*).$$

Условия (B) теоремы будут выполнены, если положить $\lambda_j^* = 0$ для $j \in J_+$. Чтобы проверить условия (A), достаточно заметить, что

$$-f'(\vec{x}^*) = \sum_{j=1}^{s} \lambda_j^* g_j'(\vec{x}^*),$$

или

$$\frac{\partial f}{\partial x_i}(\vec{x}^*) + \sum_{j=1}^{s} \lambda_j^* \frac{\partial g_i}{\partial x_i}(\vec{x}^*) = 0 \quad (i = 1, 2, ..., n),$$

поэтому

$$L'_{x_i}(\vec{x}^*, \vec{\lambda}^*) = \frac{\partial f}{\partial x_i}(\vec{x}^*) + \sum_{j=1}^{s} \lambda_j^* \frac{\partial g_i}{\partial x_i}(\vec{x}^*) = 0.$$

3. Достаточное условие экстремума

Теорема 3.25. *Пусть* $X \subset \mathbb{R}^n$ *– выпуклое множество;* $f, g_1, ..., g_s$ *– функции, определенные на* X, *такие, что* f *вогнута, а* $g_1, ..., g_s$ *неотрицательны на* X. *Предположим, что все функции* $f, g_1, ..., g_s$ *дифференцируемы в некоторой точке* $\vec{x}^* \in X$ *и существует неотрицательный вектор* $\vec{\lambda}^* \in \mathbb{R}^s$, *для которого выполня-*

ются условия (A) *и* (B) *из теоремы* 3.24. *Тогда* $\vec{x}^*$ *– точка глобального максимума* f *на* X.

Доказательство. Как и выше, положим

$$J_0 = \{j : g_j(\vec{x}^*) = 0\}, \quad J_+ = \{j : g_j(\vec{x}^*) > 0\}.$$

Для произвольной точки $\vec{x} \in X$ рассмотрим функции одной переменной

$$\varphi_j(t) = g_j(\vec{x}^* + t\Delta\vec{x}), \quad j = 1,2,...,s,$$

где $\Delta\vec{x} = \vec{x} - \vec{x}^*$. В силу выпуклости X для любого $t \in [0;1]$ точка $\vec{x}^* + t\Delta\vec{x} \in X$, поэтому $\varphi_j(t) = g_j(\vec{x}^* + t\Delta\vec{x}) \geq 0$. По условию теоремы все функции g_j дифференцируемы в точке $\vec{x}^*$, следовательно, все функции $\varphi_j(t)$ дифференцируемы в нуле. Для $j \in J_0$ имеем $\varphi_j(0) = g_j(\vec{x}^*) = 0$. Отсюда

$$\varphi_j'(0) = \lim_{t\to 0}\frac{\varphi(t) - \varphi(0)}{t} = \lim_{t\to 0}\frac{\varphi(t)}{t} = \lim_{t\to 0+0}\frac{\varphi(t)}{t} \geq 0 \quad (j \in J_0).$$

С другой стороны, по правилу дифференцирования сложной функции

$$\varphi_j'(0) = \sum_{i=1}^{n}\frac{\partial g_j}{\partial x_i}(\vec{x}^*)\Delta x_i = (g_j'(\vec{x}^*), \Delta\vec{x}),$$

где $g_j'(\vec{x}^*)$ – градиент функции g_j в точке $\vec{x}^*$, поэтому

$$(g_j'(\vec{x}^*), \Delta\vec{x}) \geq 0 \quad (j \in J_0). \qquad (3.77)$$

Далее, из (B) следует $\lambda_j^* = 0$ для $j \in J_+$. Отсюда для градиента функции f в точке $\vec{x}^*$ из (A) находим

$$f'(\vec{x}^*) = -\sum_{j\in J_0}\lambda_j^* g_j'(\vec{x}^*).$$

Ввиду вогнутости f, с учетом (3.77), имеем

$$f(\vec{x}) - f(\vec{x}^*) \le (f'(\vec{x}^*), \Delta\vec{x}) = -\sum_{j \in J_0} \lambda_j^* (g'_j(\vec{x}^*), \Delta\vec{x}) \le 0.$$

Таким образом, $f(\vec{x}) \le f(\vec{x}^*)$ для $\vec{x} \in X$, т.е. $\vec{x}^*$ – точка глобального максимума f на X. Теорема 3.25 доказана.

Теоремы 3.24, 3.25 демонстрируют связь выпуклого программирования с двойственностью в линейном программировании (см. ч.1 учебника). Так, доказательства теоремы 3.24 и основной теоремы двойственности построены на общем основании – теореме Фаркаша. При этом множители Лагранжа $\lambda_1, ..., \lambda_s$ играют в доказательстве ту же роль, что и неизвестные двойственной задачи линейного программирования. Данная аналогия прослеживается и в другой важной теореме теории двойственности – теореме равновесия: ее формулировка напоминает условия (В) из теоремы 3.24.

Пример 3.32. Найти наибольшее значение функции

$$f = \ln x_1 + \ln x_2 + \ln x_3$$

при условии $x_1 + 2x_2 + 3x_3 \le 90.$

Решение. Найдем сначала точку $\vec{x}^*$ глобального максимума функции f на выпуклом множестве X (полуоткрытой пирамиде), заданном системой

$$\begin{cases} x_1 > 0, \\ x_2 > 0, \\ x_3 > 0, \\ x_1 + x_2 + x_3 \le 90. \end{cases}$$

Вообще говоря, для поиска глобального экстремума удобнее использовать необходимые условия (т.е. теорему 3.24), однако в данном примере множество X задано ограничениями, часть из которых являются строгими неравенствами, и мы не можем использовать теорему 3.24. Попробуем найти глобальный максимум, используя достаточные условия его существования (теорему 3.25).

Составим для этого функцию Лагранжа:

$$L = \ln x_1 + \ln x_2 + \ln x_3 + \lambda(90 - x_1 - 2x_2 - 3x_3).$$

Затем запишем условия (A) и (B) из теоремы 3.24:

$$\frac{1}{x_1} - \lambda = 0,$$
$$\frac{1}{x_2} - 2\lambda = 0,$$
$$\frac{1}{x_3} - 3\lambda = 0,$$
$$\lambda(90 - x_1 - 2x_2 - 3x_3) = 0.$$

Из первых трех уравнений следует, что $x_1 = \dfrac{1}{\lambda},\ x_2 = \dfrac{1}{2\lambda},\ x_3 = \dfrac{1}{3\lambda}$, причем $\lambda \neq 0$. Поэтому из четвертого уравнения получаем

$$90 - x_1 - 2x_2 - 3x_3 = 90 - \frac{3}{\lambda} = 0.$$

Следовательно, $\lambda = \dfrac{1}{30},\ x_1 = 30,\ x_2 = 15,\ x_3 = 10.$ Поскольку функция f вогнута, по теореме 3.25 получаем, что $\vec{x}^*$ = (30; 15; 10) – точка глобального максимума f на X. Отсюда $\max f = f(\vec{x}^*) = \ln 4500$.

Пример 3.33. Решить задачу выпуклого программирования:

$$f = 20x + 20y - 10z \to \max \text{ при условии } x^2 + y^2 + z^2 \le 9. \quad (3.78)$$

Решение. Допустимое множество данной задачи (шар радиуса 3) – выпуклое, а целевая функция вогнута, как и всякая линейная функция, поэтому координаты точки максимума f на допустимом множестве удовлетворяют условиям (A) и (B) из теоремы 3.24 для функции Лагранжа $L = 20x + 20y - 10z + \lambda(9 - x^2 - y^2 - z^2)$:

$$L'_x = 20 - 2\lambda x = 0,$$
$$L'_y = 20 - 2\lambda y = 0,$$
$$L'_z = -10 - 2\lambda z = 0,$$
$$\lambda(9 - x^2 - y^2 - z^2) = 0.$$

Из первых трех уравнений следует, что

$$\lambda \neq 0, x = \frac{10}{\lambda}, y = \frac{10}{\lambda}, z = -\frac{5}{\lambda}.$$

Из четвертого уравнения находим

$$9 - \left(\frac{10}{\lambda}\right)^2 - \left(\frac{10}{\lambda}\right)^2 - \left(-\frac{5}{\lambda}\right)^2 = 0.$$

Отсюда $\lambda^2 = 25.$ С учетом неотрицательности λ имеем:

$$\lambda = 5, x = 2, y = 2, x = -1.$$

Согласно теореме 3.25 точка $(2;2;-1)$ является глобальным решением задачи выпуклого программирования (3.78).

Рассмотрим теперь пример с экономическим содержанием.

Пример 3.34. Фирма, производящая продукцию на трех заводах, решила выпускать в месяц не менее 210 ед. продукции при наименьших суммарных затратах. Сколько продукции ежемесячно следует выпускать на каждом заводе, если затраты завода $i = 1, 2, 3$ по выпуску x_i единиц продукции в месяц имеют вид:

$$C_1(x_1) = x_1 + \frac{1}{20}x_1^2 \text{ – для первого завода;}$$

$$C_2(x_2) = x_2 + \frac{1}{40}x_2^2 \text{ – для второго завода;}$$

$$C_3(x_3) = 2x_3 + \frac{1}{60}x_3^2 \text{ – для третьего завода?}$$

Р е ш е н и е. Необходимо решить задачу выпуклого программирования:

$$\begin{cases} x_1 + x_2 + x_3 - 210 \geq 0, \\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0, \end{cases} \tag{3.79}$$
$$f(\vec{x}) = C_1(x_1) + C_2(x_2) + C_3(x_3) \to \min.$$

Данная задача эквивалентна задаче на максимум для функции $-f(\vec{x})$. Функция Лагранжа для $-f(\vec{x})$ выглядит следующим образом:

$$L = -x_1 - \frac{1}{20}x_1^2 - x_2 - \frac{1}{40}x_2^2 - 2x_3 - \frac{1}{60}x_3^2 + \\ + \lambda(x_1 + x_2 + x_3 - 210) + \mu_1 x_1 + \mu_2 x_2 + \mu_3 x_3.$$

Запишем условия (A) и (B) из теоремы 3.24:

$$\left.\begin{aligned} -1 - \frac{1}{10}x_1 + \lambda + \mu_1 = 0, \\ -1 - \frac{1}{20}x_2 + \lambda + \mu_2 = 0, \\ -2 - \frac{1}{30}x_3 + \lambda + \mu_3 = 0, \\ \lambda(x_1 + x_2 + x_3 - 210) = 0, \\ \mu_1 x_1 = 0, \mu_2 x_2 = 0, \mu_3 x_3 = 0. \end{aligned}\right\} \tag{3.80}$$

Вследствие теорем 3.24 и 3.25 поиск решения задачи (3.79) сводится к поиску неотрицательного решения системы (3.80). Для этого, вообще говоря, необходимо перебрать все комбинации случаев $x_i > 0$ и $x_i \geq 0.$. Данный перебор лучше начинать с наиболее правдоподобных случаев.

Случай $x_1 > 0, x_2 > 0, x_3 > 0.$ Из (3.80) последовательно находим:

$$\mu_1 = \mu_2 = \mu_3 = 0, \quad \lambda > 0, \quad x_1 + x_2 + x_3 - 210 = 0,$$

$$x_1 = 10\lambda - 10,\ x_2 = 20\lambda - 20,\ x_3 = 30\lambda - 60,$$

$$x_1 + x_2 + x_3 - 210 = 60\lambda - 90 - 210 = 0,$$

$$\lambda = 5, \quad x_1 = 40,\ x_2 = 80,\ x_3 = 90.$$

Следовательно, точка $\vec{x}^* = (40; 80; 90)$ – глобальное решение задачи (3.79). Так как f – строго выпуклая функция, то других точек глобального минимума f на допустимом множестве в силу теоремы 3.23 не существует.

4. Включение уравнений в систему ограничений

Мы уже отмечали аналогию между множителями Лагранжа и двойственными неизвестными в задаче линейного программирования. Эта аналогия имеет дальнейшее развитие: подобно тому, как двойственные неизвестные, соответствующие уравнениям, не имеют условий неотрицательности, так и множители Лагранжа, соответствующие ограничениям типа «равенство», также не имеют подобных ограничений.

Рассмотрим множество $X \subset \mathbb{R}^n$, заданное системой ограничений, содержащей, помимо r неравенств, еще и $s - r$ уравнений:

$$\begin{cases} g_j(x_1,\dots,x_n) \ge 0,\ j = 1,\dots,r; \\ g_j(x_1,\dots,x_n) = 0,\ j = r+1,\dots,s. \end{cases} \tag{3.81}$$

Условие регулярности для системы (3.81) дословно повторяет прежнее условие регулярности, сформулированное для системы, состоящей из одних неравенств:

необходимо, чтобы все функции $g_1,\dots,g_s$ были вогнутыми и существовала точка $\vec{x}^+ \in X$, в которой все нелинейные среди $g_1,\dots,g_s$ функции были больше нуля.

Регулярность системы (3.81), очевидно, означает, что при $j > r$ функции g_j линейные.

Следующая теорема утверждает, что экстремумы функции f на X тесным образом связаны со стационарными (по $\vec{x}$) точками функции Лагранжа

$$L(\vec{x},\vec{\lambda}) = f(\vec{x}) + \lambda_1 g_1(\vec{x}) + \dots + \lambda_s g_s(\vec{x}).$$

Теорема 3.26 (***теорема Куна–Таккера***). *Пусть* $X \subset \mathbb{R}^n$ *– выпуклое множество, заданное регулярной системой ограничений* (3.81); $f, g_1, ..., g_s$ *– дифференцируемые и вогнутые на X функции. Для того чтобы точка* $\vec{x}^* \in X$ *была точкой глобального максимума f на X, необходимо и достаточно, чтобы существовал вектор* $\vec{\lambda}^* \in \mathbb{R}^s$, *удовлетворяющий условиям*:

$$L'_{x_i}(\vec{x}^*, \vec{\lambda}^*) = 0 \quad (i = 1, 2, ..., n), \tag{A}$$

$$\lambda_j^* g_j(\vec{x}^*) = 0 \quad (j = 1, 2, ..., r), \tag{B}$$

$$\lambda_j^* \geq 0 \quad (j = 1, 2, ..., r). \tag{C}$$

Отметим, что условия неотрицательности (C) распространяются на множители Лагранжа, соответствующие неравенствам в системе ограничений.

Доказательство. Как уже отмечалось выше, из регулярности системы (3.81) вытекает линейность каждого из условий $g_j(\vec{x}) = 0$, поэтому множество X может быть задано с помощью регулярной системы

$$\begin{cases} g_j(\vec{x}) \geq 0, \ j = 1, ..., r; \\ g_j(\vec{x}) \geq 0, \ j = r+1, ..., s; \\ -g_j(\vec{x}) \geq 0, \ j = r+1, ..., s. \end{cases} \tag{3.82}$$

Рассмотрим задачу выпуклого программирования

$$f(\vec{x}) \to \max \text{ при условиях (3.82).} \tag{3.83}$$

Функция Лагранжа, соответствующая задаче (3.83), имеет вид

$$F(\vec{x}, \vec{\alpha}, \vec{\beta}, \vec{\gamma}) = f(\vec{x}) + \sum_{j=1}^{r} \alpha_j g_j(\vec{x}) + \sum_{j=r+1}^{s} \beta_j g_j(\vec{x}) - \sum_{j=r+1}^{s} \gamma_j g_j(\vec{x}).$$

Необходимость. Пусть $\vec{x}^* \in X$ – точка максимума *f* на *X*. Тогда $\vec{x}^*$ – решение задачи (3.83) и по теореме 3.24 найдутся неотрицательные векторы $\vec{\alpha}^* \in \mathbb{R}^r$, $\vec{\beta}^* \in \mathbb{R}^{s-r}$, $\vec{\gamma}^* \in \mathbb{R}^{s-r}$ такие, что

$$F'_{x_i}(\vec{x}^*, \vec{\alpha}^*, \vec{\beta}^*, \vec{\gamma}^*) = 0 \quad (i = 1, ..., n), \tag{3.84}$$

$$\alpha_j^* g_j(\vec{x}^*) = 0 \quad (j = 1, ..., r), \tag{3.85}$$

$$\beta_{j-r}^* g_j(\vec{x}^*) = 0 \quad (j = r+1, ..., s), \tag{3.86}$$

$$\gamma_{j-r}^* g_j(\vec{x}^*) = 0 \quad (j = r+1, ..., s). \tag{3.87}$$

Отсюда следует, что вектор $\vec{\lambda}^* \in \mathbb{R}^s$ с координатами

$$\lambda_j^* = \alpha_j^* \ (j = 1, ..., r); \quad \lambda_j^* = \beta_{j-r}^* - \gamma_{j-r}^* \ (j = r+1, ..., s)$$

удовлетворяет условиям (A), (B) и (C) теоремы Куна–Таккера.

Достаточность. Пусть вектор $\vec{\lambda}^* \in \mathbb{R}^s$ удовлетворяет условиям (A), (B) и (C). Пусть $\vec{\alpha}^* \in \mathbb{R}^r, \vec{\beta}^* \in \mathbb{R}^{s-r}, \vec{\gamma}^* \in \mathbb{R}^{s-r}$ – неотрицательные векторы с координатами:

$$\alpha_j^* = \lambda_j^* \ \ (j = 1, ..., r),$$

$$\beta_{j-r}^* = \frac{|\lambda_j^*| + \lambda_j^*}{2}, \quad \gamma_{j-r}^* = \frac{|\lambda_j^*| - \lambda_j^*}{2} \ \ (j = r+1, ..., s).$$

Нетрудно видеть, что для $\vec{\alpha}^*, \vec{\beta}^*, \vec{\gamma}^*$ выполняются уравнения (3.84), (3.85), (3.86) и (3.87). По теореме 3.25 $\vec{x}^*$ – точка глобального максимума *f* на *X*. Теорема Куна–Таккера доказана.

Пример 3.35. Найти точку, в которой функция

$$f = \frac{1}{2}(x^2 + y^2 + z^2 + u^2)$$

принимает наименьшее значение при условиях:

$$x + y + z + u = 60,$$
$$x + 2y + 3z + 4u = 90,$$
$$x \geq 0, y \geq 0, z \geq 0, u \geq 0.$$

Р е ш е н и е. Для задачи на максимум с целевой функцией $-f$ функция Лагранжа имеет вид

$$L = -\frac{1}{2}(x^2 + y^2 + z^2 + u^2) + \lambda(x + y + z + u - 60) +$$
$$+ \mu(x + 2y + 3z + 4u - 90) + \alpha x + \beta y + \gamma z + \delta u.$$

Записываем условия (A), (B) и (C):

$$L'_x = -x + \lambda + \mu + \alpha = 0,$$
$$L'_y = -y + \lambda + 2\mu + \beta = 0,$$
$$L'_z = -z + \lambda + 3\mu + \gamma = 0,$$
$$L'_u = -u + \lambda + 4\mu + \delta = 0,$$
$$\alpha x = 0, \quad \beta y = 0, \quad \gamma z = 0, \quad \delta u = 0,$$
$$\lambda \geq 0, \quad \mu \geq 0.$$

Случай $x > 0,\ y > 0,\ z > 0, u > 0.$ Последовательно получаем:

$$\alpha = \beta = \gamma = \delta = 0,$$
$$x = \lambda + \mu, \quad y = \lambda + 2\mu, \quad z = \lambda + 3\mu, \quad u = \lambda + 4\mu,$$
$$x + y + z + u = 4\lambda + 10\mu,$$
$$x + 2z + 3y + 4u = 10\lambda + 30\mu,$$
$$\begin{cases} 4\lambda + 10\mu = 60, \\ 10\lambda + 30\mu = 90, \end{cases} \quad \begin{cases} \lambda = 45, \\ \mu = -12, \end{cases}$$
$$x = 33, \quad y = 21, \quad z = 9, \quad u = -3.$$

Пришли к противоречию.

Случай $x > 0, y > 0, z > 0, u = 0.$ Аналогично находим:

$$\alpha = \beta = \gamma = 0,$$

$$x = \lambda + \mu, \quad y = \lambda + 2\mu, \quad z = \lambda + 3\mu, \quad u = 0,$$

$$x + y + z + u = 3\lambda + 6\mu,$$

$$x + 2z + 3y + 4u = 6\lambda + 14\mu,$$

$$\begin{cases} 3\lambda + 6\mu = 60, \\ 6\lambda + 14\mu = 90, \end{cases} \quad \begin{cases} \lambda = 50, \\ \mu = -15, \end{cases}$$

$$x = 35, \quad y = 20, \quad z = 5, \quad u = 0.$$

Поскольку найденные значения неизвестных удовлетворяют всем необходимым условиям, $(35; 20; 5; 0)$ – точка минимума *f*, причем единственная, так как *f* – строго выпуклая функция (теорема 3.23).

5. Связь с седловыми точками

Пусть функция $L(x, y)$ определена на декартовом произведении множеств $X \times Y$.

Определение. *Точка* $(x^*, y^*) \in X \times Y$ *называется* ***седловой точкой*** *функции L, если для любых* $x \in X, y \in Y$ *выполняются неравенства*

$$L(x, y^*) \le L(x^*, y^*) \le L(x^*, y).$$

Предположим, что для функций $f, g_1, ..., g_s$ выполняются условия теоремы 3.26 и $L(\vec{x}, \vec{\lambda})$ – соответствующая функция Лагранжа. Функцию L можно считать определенной на декартовом произведении множеств $X \times \Lambda$, где X – выпуклое множество, заданное системой ограничений (3.81), а Λ – множество всех векторов $(\lambda_1, ..., \lambda_s)$, таких, что $\lambda_i \ge 0,\ i = 1, ..., r.$ Из теоремы Куна–Таккера вытекает следующее следствие.

Следствие. *Для того чтобы в условиях теоремы* 3.26 *точка* $\vec{x}^* \in X$ *была точкой глобального максимума f на X, необходимо и достаточно, чтобы для некоторого* $\vec{\lambda}^* \in \Lambda$ *точка* $(\vec{x}^*, \vec{\lambda}^*)$ *была седловой точкой функции Лагранжа* $L(\vec{x}, \vec{\lambda})$ *на* $X \times \Lambda$.

Доказательство. *Достаточность*. Пусть $(\vec{x}^*, \vec{\lambda}^*)$ – седловая точка функции Лагранжа

$$L(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \lambda_1 g_1(\vec{x}) + ... + \lambda_s g_s(\vec{x}),$$

т.е.

$$L(\vec{x}, \vec{\lambda}^*) \le L(\vec{x}^*, \vec{\lambda}^*), \quad \vec{x} \in X, \tag{3.88}$$

$$L(\vec{x}^*, \vec{\lambda}^*) \le L(\vec{x}^*, \vec{\lambda}), \quad \vec{x} \in \Lambda. \tag{3.89}$$

Так как $L(\vec{x}^*, \vec{\lambda})$ – линейная функция по λ_j, то из (3.89) следуют неравенства (В) теоремы Куна–Таккера. Отсюда с учетом (3.88) имеем

$$f(\vec{x}^*) - f(\vec{x}) = L(\vec{x}^*, \lambda^*) - f(\vec{x}) \ge L(\vec{x}^*, \lambda^*) - L(\vec{x}, \lambda^*) \ge 0,$$

т.е. $\vec{x}^*$ – точка максимума *f* на *X*.

Необходимость. Пусть $\vec{x}^*$ – точка максимума *f* на *X*. По теореме Куна–Таккера найдется $\vec{\lambda}^* \in \Lambda$, для которого выполняются условия (А) и (В). Поскольку $L(\vec{x}, \vec{\lambda}^*)$ – вогнутая по $\vec{x}$ функция, неравенство (3.88) следует из (А) с учетом теоремы 3.22. Неравенство (3.89) следует из (В), поскольку

$$L(\vec{x}^*, \vec{\lambda}) - L(\vec{x}^*, \vec{\lambda}^*) = \sum_{j=1}^{r} (\lambda_j - \lambda_j^*) g_j(\vec{x}^*) = \sum_{j=1}^{r} \lambda_j g_j(\vec{x}^*) \ge 0.$$

Следствие доказано.

§ 3.18. Функции спроса

Пусть p – цена товара X, q – цена товара Y, I – доход потребителя. Напомним, что функцией полезности $U(x, y)$ называется функция, задающая степень полезности (для потребителя) набора товаров, состоящего из x единиц товара X и y единиц товара Y. Будем считать, что потребитель может покупать только такие наборы (x, y), стоимость которых не превосходит его дохода, т.е. $px + qy \le I$.

Определение. *Пусть функция полезности $U(x, y)$ при любых положительных p, q и I имеет на множестве*

$$\{px + qy \le I,\ \ x \ge 0,\ \ y \ge 0\} \tag{3.90}$$

единственную точку глобального максимума (x^*, y^*). Тогда x^*, y^* – функции от p, q и I:

$$x^* = x^D(p, q, I),$$
$$y^* = y^D(p, q, I).$$

Эти функции называются ***функциями спроса***.

Смысл данного определения в том, что потребитель стремится к наибольшему удовлетворению от купленных им товаров при ограниченных средствах.

С геометрической точки зрения множество (3.90) – треугольник с вершинами $O(0;\,0)$, $A(I/p;\,0)$ и $B(0;\,I/q)$ (рис. 3.5). Как правило, функция $U(x, y)$ возрастает при увеличении x и y, поэтому наибольшее значение $U(x, y)$ достигается на отрезке AB, т.е. потребитель тратит на покупки весь свой доход.

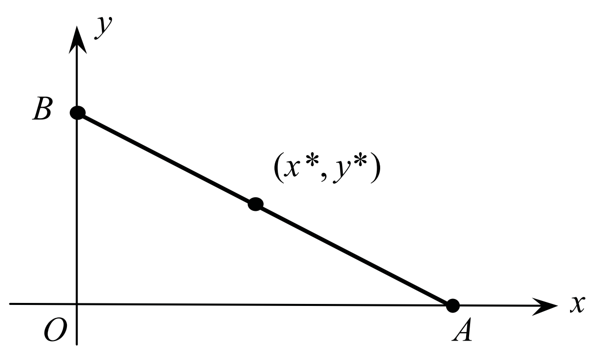

Рис. 3.5

Для любого $t > 0$ множество $\{tpx + tqy \le tI,\ x \ge 0,\ y \ge 0\}$ совпадает с множеством (3.90), поэтому функции спроса удовлетворяют следующим тождествам:

$$x^D(tp, tq, tI) = x^D(p, q, I),$$
$$y^D(tp, tq, tI) = y^D(p, q, I).$$

Таким образом, функции $x = x^D(p, q, I)$, $y = y^D(p, q, I)$ являются однородными функциями степени однородности 0. Следовательно, для дифференцируемых функций спроса выполняются тождества Эйлера:

$$\begin{aligned} px'_p + qx'_q + Ix'_I &= 0, \\ py'_p + qy'_q + Iy'_I &= 0, \end{aligned} \qquad (3.91)$$

а также следующие уравнения для эластичности:

$$E_{xp} + E_{xq} + E_{xI} = 0,$$
$$E_{yp} + E_{yq} + E_{yI} = 0.$$

Как правило, функция полезности считается дифференцируемой и строго вогнутой. В этом случае теорема Куна–Таккера помогает найти функции спроса. Пусть λ, μ, ν – множители Лагранжа, причем λ соответствует ограничению $px + qy \le I$, μ – неравенству $x \ge 0$, ν – неравенству $y \ge 0$. Тогда функция Лагранжа запишется так:

$$L = U(x, y) + \lambda(I - px - qy) + \mu x + \nu y.$$

Условия (A), (B) и (C) теоремы Куна–Таккера запишем в виде системы:

$$\begin{cases} U'_x(x, y) - \lambda p + \mu = 0, \\ U'_y(x, y) - \lambda q + \nu = 0, \\ \lambda(I - px - qy) = 0, \\ \mu x = 0, \quad \nu y = 0, \\ \lambda \ge 0, \ \mu \ge 0, \ \nu \ge 0. \end{cases} \qquad (3.92)$$

Если заранее известно, что функции спроса не обращаются в нуль, то из четвертого и пятого уравнений системы (3.92) следует, что $\mu = \nu = 0$. В этом случае система (3.92) выглядит проще:

$$\begin{cases} U'_x(x,y) - \lambda p = 0, \\ U'_y(x,y) - \lambda q = 0, \\ \lambda(I - px - qy) = 0, \\ \lambda \geq 0. \end{cases} \tag{3.93}$$

Если $U'_x > 0$ или $U'_y > 0$ (чаще всего выполняются оба условия), то из первых двух уравнений системы (3.93) следует, что $\lambda > 0$. Но тогда λ можно исключить из системы. В итоге получаем систему уравнений

$$\begin{cases} MRS_{xy} = \dfrac{U'_x(x,y)}{U'_y(x,y)} = \dfrac{p}{q}, \\ px + qy = I. \end{cases} \tag{3.94}$$

Пример 3.36. Найти функции спроса x^D, y^D в случае функции полезности $U(x,y) = \ln x + \ln y - \ln(x+y)$.

Решение. Для заданной функции полезности частные производные первого порядка таковы:

$$U'_x = \frac{y}{x(x+y)}, \; U'_y = \frac{x}{y(x+y)}.$$

Система уравнений (3.94) имеет вид

$$\begin{cases} \dfrac{U'_x}{U'_y} = \dfrac{y^2}{x^2} = \dfrac{p}{q}, \\ px + qy = I. \end{cases}$$

Функция $U(x,y)$ является строго вогнутой в области $\{x > 0, y > 0\}$, поскольку при положительных x и y выполняются неравенства:

$$U''_{xx} = -\frac{1}{x^2} + \frac{1}{(x+y)^2} < 0 \;\text{ и }\; U''_{xx} \cdot U''_{yy} - (U''_{xy})^2 = \frac{2}{xy(x+y)^2} > 0.$$

Кроме того, $U'_x = \dfrac{y}{x(x+y)} > 0.$ Поэтому функции спроса таковы:

$$x^D = \frac{I}{p+\sqrt{pq}},\ y^D = \frac{I}{q+\sqrt{pq}}.$$

Если в примере 3.36 значение I = 100 фиксировано, то имеем функцию спроса $x^D(p,q) = \dfrac{100}{p+\sqrt{pq}}.$ Если зафиксировать I = 100 и, скажем, q = 9, то получим функцию спроса вида $x^D(p) = \dfrac{100}{p+3\sqrt{p}}$ и т.д.

§ 3.19. Уравнения Слуцкого

Рассмотрим модель, в которой потребитель может покупать n видов товаров X_1, …, X_n. Пусть $\vec{p} = (p_1, ..., p_n)$ – вектор цен, $\vec{x} = (x_1, ..., x_n)$ – приобретаемый потребителем набор товаров, где p_i – цена, а x_i – количество товара X_i (i=1, …, n). Как и в случае двух товаров, степень удовлетворения потребителя от набора $\vec{x}$ задается функцией полезности $u(\vec{x})$. Располагая доходом I, потребитель покупает товары, стремясь к максимальному удовлетворению своих потребностей. Пусть $\vec{x}^* = (x_1^*, ..., x_n^*)$ – единственная точка глобального максимума функции $u(\vec{x})$ на множестве

$$\begin{cases} p_1x_1 + ... + p_nx_n \le I, \\ x_1 \ge 0, ..., x_n \ge 0. \end{cases}$$

Функции спроса определяются соотношениями:

$$x_1^D(\vec{p}, I) = x_1^*,\ \ ...,\ \ x_n^D(\vec{p}, I) = x_n^*.$$

Все функции x_i^D имеют общую область определения – множество значений переменных p_1, …, p_n, I, для которых существует ровно одна точка максимума $\vec{x}^*$.

Рассмотрим теперь несколько иной подход к определению спроса. Пусть $\vec{x}^\circ = (x_1^\circ, ..., x_n^\circ)$ – точка минимума функции $p_1 x_1 + ... + p_n x_n$ на множестве

$$\begin{cases} u(x_1, ..., x_n) \geq U, \\ x_1 \geq 0, ..., x_n \geq 0. \end{cases}$$

Таким образом, $\vec{x}^\circ$ – самый дешевый товар, обеспечивающий уровень полезности U. Функции *компенсированного спроса* определяются формулами:

$$x_1^{CD}(\vec{p}, U) = x_1^\circ, \ ..., \ x_n^{CD}(\vec{p}, U) = x_n^\circ,$$

их общая область определения – это множество значений переменных $p_1, \ldots, p_n$ и U, для которых существует ровно одна точка минимума $\vec{x}^\circ$. Название функций x_i^{CD} связано с тем, что в случае роста цен сохранить постоянный уровень $u(\vec{x}) = U = \mathrm{const}$ возможно лишь при увеличении дохода I, компенсирующем данный рост цен.

Несмотря на формальное различие, одновременное применение обоих подходов к определению спроса не вызывает противоречий. Чтобы доказать это, предположим, что функция полезности непрерывна, не убывает по каждой переменной и возрастает при одновременном увеличении всех аргументов. Прежде всего, заметим, что

$$u(\vec{x}^{CD}(\vec{p}, U)) = U. \qquad (3.95)$$

Действительно, если бы выполнялось неравенство $u(\vec{x}^{CD}(\vec{p}, U)) > U$, то в силу непрерывности $u(\vec{x})$ это неравенство выполнялось бы и в некоторой окрестности точки $\vec{x}^{CD}(\vec{p}, U)$, но тогда в этой же окрестности нашелся бы набор товаров, более дешевый, чем набор $\vec{x}^{CD}(\vec{p}, U)$, что противоречит определению $\vec{x}^{CD}$. Данное противоречие доказывает (3.95).

Определим *компенсированный доход* как стоимость компенсированного спроса: $I^C(\vec{p},U) = p_1 x_1^{CD}(\vec{p},U) + ... + p_n x_n^{CD}(\vec{p},U)$.

Теорема 3.27. *Справедливы следующие соотношения*:

а) $\vec{x}^{CD}(\vec{p},U) = \vec{x}^D(\vec{p}, I^C(\vec{p},U))$, (3.96)

б) $\sum_{i=1}^{n} p_i x_i^{CD}(\vec{p}, u(\vec{x}^D(\vec{p}, I))) = I$, (3.97)

в) $\vec{x}^{CD}(\vec{p}, u(\vec{x}^D(\vec{p}, I))) = \vec{x}^D(p, I)$. (3.98)

Доказательство. а) Фиксируем $\vec{p}$, U, $\vec{x}^{CD} = \vec{x}^{CD}(\vec{p},U)$,

$$I^C = p_1 x_1^{CD} + ... + p_n x_n^{CD} \quad \text{и} \quad \vec{x}^D = \vec{x}^D(\vec{p}, I^C).$$

Необходимо доказать, что $\vec{x}^{CD} = \vec{x}^D$. Предположим, что $\vec{x}^{CD} \neq \vec{x}^D$. Рассмотрим точку $\vec{a}$ с координатами $a_i = \min\{x_i^{CD}, x_i^D\}$, $i = 1, ..., n$. Так как $u(\vec{x})$ не убывает, то $u(\vec{a}) \le u(\vec{x}^{CD}) = U$. Поскольку $\vec{x}^{CD} \neq \vec{x}^D$, а $\vec{x}^D$ – единственная точка максимума функции $u(\vec{x})$ на множестве

$$\begin{cases} p_1 x_1 + ... + p_n x_n \le I^C, \\ x_1 \ge 0, \ ..., \ x_n \ge 0, \end{cases}$$

имеем неравенство $u(\vec{x}^D) > u(\vec{x}^{CD}) = U$. В силу непрерывности функции $u(\vec{x})$ на отрезке $\vec{a}\vec{x}^D$ найдется (рис. 3.6) точка $\vec{b}$, для которой $u(\vec{b}) = U$, при этом $\vec{b} \neq \vec{x}^D$. С учетом положительности всех цен имеем

$$p_1 b_1 + ... + p_n b_n < p_1 x_1^D + ... + p_n x_n^D \le I^C.$$

Итак, $\vec{b}$ – более дешевый набор, чем $\vec{x}^{CD}$, несмотря на то, что $u(\vec{b}) = U$. Полученное противоречие доказывает, что $\vec{x}^{CD} = \vec{x}^D$.

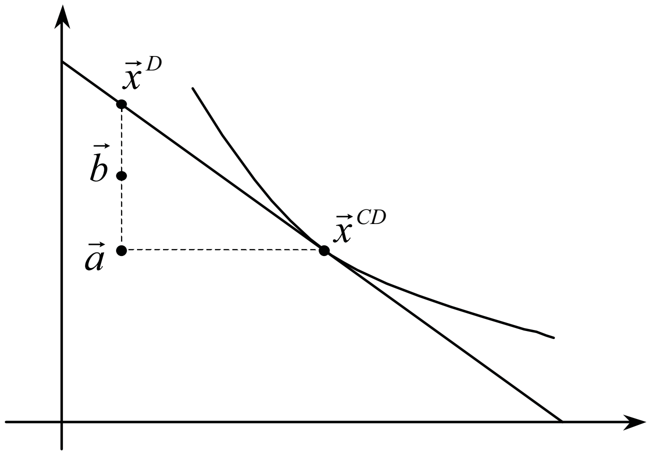

Рис. 3.6

б) Фиксируем $\vec{p}, I, \vec{x}^D = \vec{x}^D(\vec{p}, I)$, $U = u(\vec{x}^D)$, $x_i^{CD} = x_i^{CD}(p, I)$. Неравенство $\sum_{i=1}^{n} p_i x_i^{CD} < I$ противоречит тому, $\vec{x}^D$ – единственная точка глобального максимума функции $u(\vec{x})$ на множестве $p_1 x_1 + ... + p_n x_n \le I$, $x_1 \ge 0, ..., x_n \ge 0$. Противоположное неравенство $\sum_{i=1}^{n} p_i x_i^{CD} > I$ противоречит тому, что $\vec{x}^{CD}$ – точка глобального минимума функции $p_1 x_1 + ... + p_n x_n$ при условиях $u(\vec{x}) \ge U$, $x_1 \ge 0, ..., x_n \ge 0$. Таким образом, остается единственная возможность: $\sum_{i=1}^{n} p_i x_i^{CD} = I$, что означает справедливость (3.97).

в) Подставив $u(\vec{x}(\vec{p}, I))$ вместо U в (3.96), имеем

$$\vec{x}^{CD}(\vec{p}, u(\vec{x}(\vec{p}, I))) = \vec{x}^D(\vec{p}, \sum_{i=1}^{n} p_i \vec{x}_i^{CD}(\vec{p}, u(\vec{x}(\vec{p}, I)))).$$

Отсюда, используя (3.97), получаем (3.98).

Начиная с этого момента, мы предполагаем, что *функция полезности в дополнение к сформулированным выше условиям возрастания является* ***дифференцируемой и вогнутой функцией*** *в* $\mathbb{R}_+^n$.

Заметим, что отсюда немедленно следует выпуклость множества $\{\vec{x} \in \mathbb{R}_+^n : u(\vec{x}) \ge U\}$.

Следующая теорема утверждает, что производная компенсированного дохода по цене имеет неожиданно простой вид.

Теорема 3.28. *Если все функции компенсированного спроса* x_1^{CD}, …, x_n^{CD} *дифференцируемы в точке* $(\vec{p},U)$, *то для любого* $j = 1, \ldots, n$ *имеем*

$$\frac{\partial I^C(\vec{p},U)}{\partial p_j} = x_j^{CD}(\vec{p},U). \tag{3.99}$$

Доказательство. Фиксируем $\vec{p},U$ и $\vec{x}^{CD} = \vec{x}^{CD}(\vec{p},U)$. Точка $\vec{x}^{CD}$ является точкой максимума функции $-(p_1x_1 + \ldots + p_nx_n)$ при условиях:

$$u(\vec{x}) \geq U,\ x_1 \geq 0,\ \ldots,\ x_n \geq 0.$$

По теореме 3.24 найдутся такие неотрицательные числа $\lambda, \mu_1, \ldots, \mu_n$, что для функции Лагранжа

$$L = -(p_1x_1 + \ldots + p_nx_n) + \lambda(u(\vec{x}) - U) + \mu_1x_1 + \ldots + \mu_nx_n$$

выполняются равенства

$$\frac{\partial L}{\partial x_i}(\vec{x}^{CD}, \lambda, \vec{\mu}) = -p_i + \lambda\frac{\partial u}{\partial x_i}(\vec{x}^{CD}) + \mu_i = 0, \tag{3.100}$$

$$\mu_i x_i^{CD} = 0,\ (i = 1, \ldots, n). \tag{3.101}$$

Из (3.100) следует, что

$$\frac{\partial u}{\partial x_i}(\vec{x}^{CD}) = \lambda^{-1}(p_i - \lambda_i).$$

Снимая фиксацию с $\vec{p}$ и $\vec{x}^{CD}$, находим

$$\frac{d}{dp_j}u(\vec{x}^{CD}) = \sum_{i=1}^{n}\frac{\partial u}{\partial x_i}\frac{\partial x_i^{CD}}{\partial p_j} = \lambda^{-1}\sum_{i=1}^{n}(p_i - \mu_i)\frac{\partial x_i^{CD}}{\partial p_j}.$$

С другой стороны, с учетом (3.95)

$$\frac{d}{dp_j}u(\vec{x}^{CD}) = \frac{d}{dp_j}U = 0 \Rightarrow \sum_{i=1}^{n} p_i \frac{\partial x_i^{CD}}{p_j} = \sum_{i=1}^{n} \mu_i \frac{\partial x_i^{CD}}{p_j}.$$

Рассмотрим каждое слагаемое последней суммы. Если $x_i^{CD} > 0,$ то вследствие (3.101) имеем $\mu_i = 0.$ Если же $x_i^{CD} = 0,$ то поскольку x_i^{CD} – неотрицательная дифференцируемая функция, $\partial x_i^{CD}/\partial p_j = 0.$ Отсюда следует, что

$$\sum_{i=1}^{n} p_i \frac{\partial x_i^{CD}}{p_j} = 0. \tag{3.102}$$

Окончательно имеем

$$\frac{\partial I^C}{\partial p_j} = \frac{d}{dp_j}\sum_{i=1}^{n} p_i x_i^{CD} = x_j^{CD} + \sum_{i=1}^{n} p_i \frac{\partial x_i^{CD}}{\partial p_j}.$$

Теорема 3.29. *Если функции спроса дифференцируемы в точке* $(\vec{p}, I)$, *а функция компенсированного спроса дифференцируема в точке* $(\vec{p}, U)$, *где* $U = u(\vec{x}^D(\vec{p}, I))$, *то выполняются следующие соотношения*:

$$\frac{\partial x_i^D}{\partial p_j}(\vec{p}, I) = \frac{\partial x_i^{CD}}{\partial p_j}(\vec{p}, U) - \frac{\partial x_j^D}{\partial I}(\vec{p}, I)\cdot x_i^D(\vec{p}, I),\ i, j = 1, 2, ..., n. \tag{3.103}$$

Доказательство. Дифференцируя (3.96) по p_j, с учетом (3.99) находим

$$\frac{\partial x_i^{CD}}{\partial p_j}(\vec{p}, U) = \frac{d}{dp_j} x_i^D(\vec{p}, I^C(\vec{p}, U)) =$$

$$= \frac{\partial x_i^D}{\partial p_j}(\vec{p}, I^C(\vec{p}, U)) + \frac{\partial x_i^D}{\partial I}\vec{x}^D(\vec{p}, I^C(\vec{p}, U))\cdot x_j^D(\vec{p}, U).$$

Заменяя U на $u(\vec{x}^D(\vec{p},I))$ и используя (3.97) и (3.98), получим уравнение

$$\frac{\partial x_i^{CD}}{\partial p_j}(\vec{p},U)=\frac{\partial x_i^D}{\partial p_j}(\vec{p},I)+\frac{\partial x_i^D}{\partial I}(\vec{p},I)\cdot x_j^D(\vec{p},I),$$

которое, очевидно, эквивалентно (3.103).

Уравнения (3.103) называются *уравнениями Слуцкого*. Производная $\partial x_i^{CD}/\partial p_j$ из правой части уравнения Слуцкого называется *компенсационной производной* и обозначается $(\partial x_i^D/\partial p_j)_{comp}$. Уравнения Слуцкого в сокращенной записи выглядят так:

$$\frac{\partial x_i}{\partial p_j}=\left(\frac{\partial x_i}{\partial p_j}\right)_{comp}-x_i\cdot\frac{\partial x_j}{\partial I}. \qquad (3.104)$$

Считается, что компоненты правой части уравнения Слуцкого (3.104) соответствуют определенным эффектам изменения потребительского спроса:

$$\left(\frac{\partial x_i}{\partial p_j}\right)_{comp} \quad \text{– «эффекту замены»,}$$

$$x_i\frac{\partial x_j}{\partial I} \quad \text{– «эффекту дохода».}$$

Докажем симметричность «эффекта замены» для любых двух товаров X_i, X_j, предполагая непрерывность и дифференцируемость всех функций, которые появятся в процессе доказательства.

Сначала перепишем (3.100) в виде

$$p_l=\lambda\frac{\partial u}{\partial x_l}+\mu_l,\quad l=1,\dots,n,$$

после чего продифференцируем данное уравнение по p_j:

$$\frac{\partial p_l}{\partial p_j}=\frac{\partial \lambda}{\partial p_j}\frac{\partial u}{\partial x_l}+\lambda\sum_k\frac{\partial^2 u}{\partial x_l\partial x_k}\frac{\partial x_k^{CD}}{\partial p_j}+\frac{\partial \mu_l}{\partial p_j}. \qquad (3.105)$$

Далее умножим (3.105) на $\partial x_l^{CD}/\partial p_i$ и просуммируем по l:

$$\sum_l \frac{\partial x_l^{CD}}{\partial p_i}\frac{\partial p_l}{\partial p_j} = \frac{\partial \lambda}{\partial p_j}\sum_l \frac{\partial u}{\partial x_l}\frac{\partial x_l^{CD}}{\partial p_i} +$$

$$+\lambda \sum_{k,l}\frac{\partial^2 u}{\partial x_l \partial x_k}\frac{\partial x_k^{CD}}{\partial p_j}\frac{\partial x_l^{CD}}{\partial p_i} + \sum_l \frac{\partial \mu_l}{\partial p_j}\frac{\partial x_l^{CD}}{\partial p_i}. \qquad (3.106)$$

Поскольку $\partial p_l/\partial p_j = 0$, если $l \neq j$, и $\partial p_l/\partial p_j = 1$, если $l = j$, левая часть уравнения (3.106) равна $\partial x_j^{CD}/\partial p_i$. Первое слагаемое в правой части (3.106) равно 0, так как

$$\sum_l \frac{\partial u}{\partial x_l}\frac{\partial x_l^{CD}}{\partial p_i} = \frac{dU}{dp_i} = 0 \quad (U = \text{const}).$$

Далее, если $x_l^{CD} > 0$ в точке $\vec{p}$, то это соотношение остается в силе и в некоторой окрестности $\vec{p}$. Вследствие (3.101), $\mu_l = 0$ в окрестности $\vec{p}$ и $\partial \mu_l/\partial p_j = 0$ в самой точке $\vec{p}$. Если же $x_l^{CD} = 0$ в точке $\vec{p}$, то из неотрицательности x_l^{CD} следует, что $\partial x_l^{CD}/\partial p_j = 0$. Поэтому последнее слагаемое в правой части (3.106) также равно 0. В итоге имеем

$$\frac{\partial x_j^{CD}}{\partial p_i} = \lambda \sum_{k,l}\frac{\partial^2 u}{\partial x_l \partial x_k}\frac{\partial x_k^{CD}}{\partial p_j}\frac{\partial x_l^{CD}}{\partial p_i}. \qquad (3.107)$$

Поскольку правая часть (3.107) не меняется при замене $i \leftrightarrow j$,

$$\left(\frac{\partial x_j}{\partial p_i}\right)_{comp} = \left(\frac{\partial x_i}{\partial p_j}\right)_{comp}. \qquad (3.108)$$

Симметричность «эффекта замены» доказана.

Отсюда вытекает корректность следующего определения.

О п р е д е л е н и е. *Если на некотором множестве цен и доходов выполняется неравенство* $(\partial x_i/\partial p_j)_{comp} > 0,$ *то товары* X_i, X_j *называются* ***взаимозаменяемыми*** (*на данном множестве*). *Если же* $(\partial x_i/\partial p_j)_{comp} < 0,$ *то товары* X_i, X_j *называются* ***взаимодополняющими***.

Разумеется, можно было бы дать определение взаимозаменяемых и взаимодополняющих товаров и без ссылки на симметричность «эффекта замены», назвав товары X_i, X_j взаимозаменяемыми, если $(\partial x_i/\partial p_j)_{comp} > 0,$ $(\partial x_j/\partial p_i)_{comp} > 0,$ и взаимодополняющими, если выполняются противоположные неравенства.

Из (3.104), (3.108) следует, что функции спроса $x_1^D, x_2^D, \ldots, x_n^D$ удовлетворяют следующим уравнениям:

$$\frac{\partial x_i}{\partial p_j} + x_i \frac{\partial x_j}{\partial I} = \frac{\partial x_j}{\partial p_i} + x_j \frac{\partial x_i}{\partial I} \quad (i, j = 1, 2, \ldots, n),$$

которые, так же, как и (3.104), называются уравнениями Слуцкого.

Уравнения (3.107) в случае $i = j$ допускают интересную экономическую интерпретацию. Перепишем (3.107), используя второй дифференциал:

$$\frac{\partial x_i^{CD}}{\partial p_i} = \lambda d^2 u\left(\frac{\partial x_1^{CD}}{\partial p_i}, \frac{\partial x_2^{CD}}{\partial p_i}, \ldots, \frac{\partial x_n^{CD}}{\partial p_i} \right).$$

Так как $u(\vec{x})$ – вогнутая функция, то $d^2 u(\vec{v}) \le 0$ для любого $\vec{v}$ вследствие необходимого условия вогнутости (теорема 3.18). Отсюда $(\lambda > 0)$ имеем

$$\left(\frac{\partial x_i}{\partial p_i} \right)_{comp} = \frac{\partial x_i^{CD}}{\partial p_i} \le 0.$$

Таким образом, мы пришли к следующему утверждению: *компенсированный спрос на товар не возрастает при увеличении его цены*. Если предположить, что товар X_i является *ценным*, т.е. $\partial x_i^D/\partial I > 0,$ то последнее утверждение остается справедливым и в отношении некомпенсированного спроса. Действительно, из уравнения Слуцкого (3.104) имеем

$$\frac{\partial x_i^D}{\partial p_i} = \left(\frac{\partial x_i^D}{\partial p_i}\right)_{comp} - x_i^D \frac{\partial x_i^D}{\partial I} \le 0. \qquad (3.109)$$

Вместе с тем неравенство (3.109) может нарушаться для *малоценных* товаров (случай $\partial x_i^D / \partial I < 0$), т.е. при определенных условиях рост цены вызывает увеличение спроса (см. задачу ниже).

Пример 3.37. Полезность набора товаров X и Y определяется функцией $u(x, y) = \ln x + \ln y$. Для дохода I и цен на товары p_x, p_y требуется:

а) найти спрос на X и Y;

б) найти компенсированный спрос на X и Y;

в) выяснить, являются ли товары X и Y взаимозаменяемыми.

Решение. а) Чтобы найти точку максимума $u(x, y)$ при бюджетном ограничении $xp_x + yp_y \le I$, составим функцию Лагранжа

$$L = \ln x + \ln y + \lambda(I - xp_x - yp_y)$$

и приравняем ее частные производные нулю:

$$L'_x = x^{-1} - \lambda p_x = 0, \quad L'_y = y^{-1} - \lambda p_y = 0. \qquad (3.110)$$

Потребуем также, чтобы выполнялись условия:

$$\lambda(I - xp_x - yp_y) = 0, \quad \lambda \ge 0. \qquad (3.111)$$

Из (3.110), (3.111) находим: $\lambda = 2I^{-1}, \quad x = I(2p_x)^{-1}, \quad y = I(2p_y)^{-1}.$ Точка

$$(x^*, y^*) = (I(2p_x)^{-1},\ I(2p_x)^{-1})$$

удовлетворяет условиям (3.110), (3.111) и по теореме 3.25 является точкой глобального максимума функции полезности при бюджетном ограничении.

Таким образом, имеем функции спроса:

$$x^D = \frac{I}{2p_x}, \quad y^D = \frac{I}{2p_y}.$$

б) Чтобы найти минимум функции $xp_x + yp_y$ при условии $\ln x + \ln y \geq U$, составим функцию Лагранжа

$$L = -xp_x - yp_y + \lambda(\ln x + \ln y - U).$$

Потребуем, чтобы выполнялись условия:

$$L'_x = -p_x + \lambda / x = 0, \quad L'_y = -p_y + \lambda / y = 0, \qquad (3.112)$$

$$\lambda(\ln x + \ln y - U) = 0, \quad \lambda \geq 0. \qquad (3.113)$$

Из (3.112), (3.113) находим:

$$\lambda = e^{U/2}\sqrt{p_x p_y}, \quad x = e^{U/2}\sqrt{p_y / p_x}, \quad y = e^{U/2}\sqrt{p_x / p_y}.$$

Точка $(x_0, y_0) = (e^{U/2}\sqrt{p_y / p_x}, e^{U/2}\sqrt{p_x / p_y})$ удовлетворяет условиям (3.112), (3.113) и по теореме 3.25 является точкой глобального минимума функции $xp_x + yp_y$ при условии $\ln x + \ln y \geq U$. Функции компенсированного спроса имеют вид:

$$x^{CD} = e^{U/2}\sqrt{p_y / p_x}, \quad y^{CD} = e^{U/2}\sqrt{p_x / p_y}.$$

в) Имеем

$$\frac{\partial y^{CD}}{\partial p_x} = \frac{\partial x^{CD}}{\partial p_y} = \frac{e^{U/2}}{2\sqrt{p_x p_y}} > 0,$$

т.е. товары X и Y взаимозаменяемые.

Пример 3.38. Полезность товаров X, Y, Z задается функцией

$$u(x, y, z) = \ln x + \ln(\min\{y, z\}).$$

Будут ли товары Y, Z взаимозаменяемыми?

Р е ш е н и е. Пусть p_x, p_y, p_z – цены товаров X, Y, Z; (x_0, y_0, z_0) – точка минимума функции $xp_x + yp_y + zp_z$ при условии $u(x, y, z) \geq U$. Из вида функции полезности ясно, что $y_0 = z_0$, поэтому поиск точек (x_0, y_0, z_0) сводится к поиску точек минимума функции $xp_x + y(p_y + p_z)$ при условии $\ln x + \ln y \geq U$. Данная задача была решена в предыдущем примере, откуда имеем

$$(x_0, y_0, z_0) = \left(e^{U/2}\sqrt{\frac{p_y + p_z}{p_x}}, e^{U/2}\sqrt{\frac{p_x}{p_y + p_z}}, e^{U/2}\sqrt{\frac{p_x}{p_y + p_z}} \right).$$

Следовательно,

$$y^{CD} = e^{U/2}\sqrt{\frac{p_x}{p_y + p_z}}, \quad z^{CD} = e^{U/2}\sqrt{\frac{p_x}{p_y + p_z}},$$

$$\frac{\partial y^{CD}}{\partial p_z} = \frac{\partial z^{CD}}{\partial p_y} = -\frac{1}{2}e^{U/2}\sqrt{\frac{p_x}{(p_y + p_z)^3}} < 0,$$

т.е. товары Y, Z взаимодополняющие.

Следующая задача демонстрирует удивительную возможность: при некоторых условиях *увеличение цены* приводит к *увеличению спроса*! Функция полезности при этом недифференцируема, но ее можно «сгладить» с сохранением свойств функций спроса.

З а д а ч а. Покажите, что в модели с двумя товарами X, Y и функцией полезности $u(x, y) = \min\{2x + 2y - 8, y\}$ функция спроса $x^D(p_x, p_y, I)$ удовлетворяет следующим соотношениям:

$$x^D(1/3, 1, 3) = 3, \quad x^D(4/7, 1, 3) = 3{,}5.$$

У к а з а н и е. Используйте графический метод в соответствии с рис. 3.7, где:

AB – прямая, отделяющая область, в которой $u = 2x + 2y - 8$, от области, в которой $u = y$;

CE_1, CE_2 – отрезки, представляющие наборы товаров со стоимостью 3 в двух вариантах цен ($p_x = 1/3$, $p_y = 1$ и $p_x = 4/7$, $p_y = 1$);

D – точка спроса при ценах $p_x = 1/3$, $p_y = 1$;

FDG – линия уровня u, содержащая D.

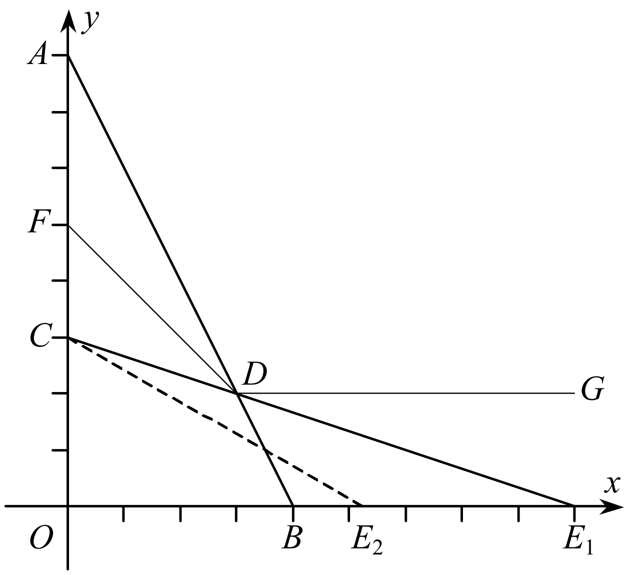

Рис. 3.7

Увеличение спроса на товар, вызванное увеличением его цены, называется *эффектом Гиффина*. Данный эффект может наблюдаться в экономической реальности. В качестве примера обычно приводят ситуацию в Ирландии конца XIX века, когда основной пищей для населения служил картофель. В случае, когда цены на картофель повышались, население отказывалось от потребления мяса, что приводило к увеличению спроса на картофель.

ГЛАВА 4

ОПРЕДЕЛЕННЫЙ ИНТЕГРАЛ И ЕГО ПРИЛОЖЕНИЯ

§ 4.1. Неопределенный интеграл и его свойства

1. Первообразная и неопределенный интеграл

В математике, как и в ее приложениях, часто возникает задача, обратная той, которая решалась в дифференциальном исчислении, а именно: дана функция $y = f(x)$, найти функцию $y = F(x)$, такую, что $F'(x) = f(x)$.

Пример 4.1. Динамика формирования оборотных средств – это процесс изменения во времени, который можно рассматривать как непрерывный. Пусть $K(t)$ – это зависимость объема оборотных средств от времени, тогда производная по времени dK/dt – это скорость формирования оборотных средств, которую можно рассматривать так же, как скорость потока денежных средств $I(t)$. Если дана функция скорости потока, то возникает задача отыскания функции оборотных средств, которая сводится к операции нахождения функции по ее производной.

Такая операция называется *интегрированием*, а раздел математики, изучающий методы поиска функции по ее производной, – *интегральным исчислением*.

Одним из главных понятий интегрального исчисления является понятие *первообразной*.

Определение. *Функция $y = F(x)$ называется* ***первообразной*** *для функции $y = f(x)$ на промежутке X, если для любого $x \in X$ выполняется равенство*

$$F'(x) = f(x).$$

Пример 4.2. Пусть $f(x) = \cos x$, тогда за первообразную можно взять $F(x) = \sin x$, поскольку $(\sin x)' = \cos x$.

Пример 4.3. Пусть $f(x) = x^2$, тогда можно положить $F(x) = \dfrac{x^3}{3}$, поскольку $\left(\dfrac{x^3}{3}\right)' = \dfrac{\left(x^3\right)'}{3} = \dfrac{3x^2}{3} = x^2$.

Заметим, что в примере 4.2 мы могли вместо $\sin x$ в качестве первообразной взять, например, $F_1(x) = \sin x + \pi$ или $F_2(x) = \sin x - 200$, поскольку $(\sin x + \pi)' = \cos x$, и точно так же $(\sin x - 200)' = \cos x$.

Теорема 4.1 (*об общем виде первообразной*). *Если $F(x)$ – первообразная для функции $y = f(x)$ на промежутке X, то все первообразные для функции $y = f(x)$ имеют вид $F(x) + C$, где C – произвольная постоянная.*

Доказательство. Пусть $F(x)$ – первообразная для $f(x)$. Тогда выполняется равенство $F'(x) = f(x)$. Для любой постоянной C

$$(F(x) + C)' = F'(x) + 0 = f(x),$$

а это означает, что $F(x) + C$ – также первообразная для $f(x)$.

Обратно, пусть наряду с данной первообразной $F(x)$ функция $F_1(x)$ – также первообразная для $f(x)$. Тогда выполняются равенства

$$F_1'(x) = F'(x) = f(x),$$

откуда $(F_1(x) - F(x))' = 0$. Тогда по теореме 2.10 разность этих двух первообразных будет тождественно равна константе

$$F_1(x) - F(x) = C,$$

или

$$F_1(x) = F(x) + C,$$

что завершает доказательство теоремы.

Эта теорема позволяет ввести основное понятие интегрального исчисления.

Определение. *Если* $F(x)$ – *первообразная для* $f(x)$, *то выражение* $F(x) + C$, *где* C – *произвольная постоянная*, *называется* ***неопределенным интегралом*** *от функции* $f(x)$.

Неопределенный интеграл обозначается $\int f(x)dx$ (читается “интеграл эф от икс дэ икс”). Таким образом,

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

Знак $\int$ называют *знаком интеграла*, функцию $f(x)$ – *подынтегральной функцией*, $f(x)dx$ – *подынтегральным выражением*, C – *постоянной интегрирования*. Символ $\int f(x)dx$, таким образом, обозначает множество всех первообразных данной функции.

Отыскание функции по ее производной или, что то же, отыскание неопределенного интеграла по данной подынтегральной функции, называется *интегрированием* данной функции. Интегрирование является операцией, обратной к дифференцированию. Чтобы проверить, правильно ли выполнено интегрирование, надо взять производную от полученного результата и убедиться, что получена подынтегральная функция.

Как всякая обратная операция, интегрирование – более сложное действие, чем дифференцирование. Мы приступаем к рассмотрению свойств неопределенного интеграла и методов интегрирования.

2. Свойства неопределенного интеграла

1. *Производная неопределенного интеграла равна подынтегральной функции*; *дифференциал от неопределенного интеграла равен подынтегральному выражению*, *т.е.*

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

Действительно,

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

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

2. *Неопределенный интеграл от дифференциала некоторой функции равен этой функции плюс произвольная постоянная*, *т.е*.

$$\int dF(x) = F(x)+C.$$

Поскольку $dF(x) = F'(x)dx$, то $\int dF(x) = \int F'(x)dx = F(x)+C.$

3. *Постоянный множитель можно вынести из-под знака неопределенного интеграла*, *точнее*, *если* $k \neq 0$, *то*

$$\int kf(x)dx = k\int f(x)dx.$$

4. *Неопределенный интеграл от суммы функций равен сумме интегралов от слагаемых*, *т.е*.

$$\int (f_1(x)+f_2(x))dx = \int f_1(x)dx + \int f_2(x)dx.$$

Действительно, дифференцируя левую часть равенства, получим по свойству 1 $\left(\int (f_1(x)+f_2(x))dx\right)' = f_1(x)+f_2(x)$, а производная правой части

$$\left(\int f_1(x)dx + \int f_2(x)dx\right)' = \left(\int f_1(x)dx\right)' + \left(\int f_2(x)dx\right)' = f_1(x)+f_2(x),$$

так что производные равны, что и требовалось проверить.

3. Таблица основных интегралов

Приведем таблицу *основных интегралов*. Часть формул этой таблицы непосредственно следует из таблицы производных.

I. $\int dx = x + C.$

II. $\int x^a dx = \dfrac{x^{a+1}}{a+1} + C,\ a \neq -1.$

III. $\int \dfrac{dx}{x} = \ln|x| + C.$

IV. $\int a^x dx = \dfrac{a^x}{\ln a} + C, \quad \int e^x dx = e^x + C.$

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

VI. $\int \sin x = -\cos x + C.$

VII. $\int \dfrac{dx}{\cos^2 x} = \operatorname{tg} x + C.$

VIII. $\int \dfrac{dx}{\sin^2 x} = -\operatorname{ctg} x + C.$

IX. $\int \dfrac{dx}{\sqrt{a^2 - x^2}} = \arcsin \dfrac{x}{a} + C.$

X. $\int \dfrac{dx}{a^2 + x^2} = \dfrac{1}{a} \operatorname{arctg} \dfrac{x}{a} + C.$

XI. $\int \dfrac{dx}{x^2 - a^2} = \dfrac{1}{2a} \ln \left| \dfrac{x-a}{x+a} \right| + C,\ a \neq 0.$

XII. $\int \dfrac{dx}{\sqrt{x^2 + k}} = \ln \left| x + \sqrt{x^2 + k} \right| + C.$

Проверим, например, справедливость формулы III. Действительно, при $x > 0$ $\ln|x| = \ln x$ и $(\ln x)' = 1/x$, а при $x < 0$ $\ln|x| = \ln(-x)$, так что $(\ln|x|)' = (-x)'/(-x) = 1/x$, что и требовалось получить.

Проверим еще справедливость формулы XI. Правую часть представим в виде

$$\frac{1}{2a}\left(\ln|x-a| - \ln|x+a|\right) + C$$

и после дифференцирования имеем

$$\frac{1}{2a}\left(\frac{1}{x-a} - \frac{1}{x+a}\right) = \frac{1}{2a}\frac{x+a-x+a}{x^2-a^2} = \frac{1}{x^2-a^2},$$

что совпадает с подынтегральной функцией.

Интегралы, содержащиеся в этой таблице, принято называть *табличными*.

Пример 4.4. Найти $\int \frac{2+3\sqrt[3]{x^2}+5\sqrt{x}}{\sqrt{x^3}}dx.$

Решение. Разделим почленно числитель на знаменатель и применим сначала свойства 3 и 4, а затем табличные интегралы II и III:

$$\int \frac{2+3\sqrt[3]{x^2}+5\sqrt{x}}{\sqrt{x^3}}dx = 2\int x^{-\frac{3}{2}}dx + 3\int x^{-\frac{5}{6}}dx + 5\int \frac{dx}{x} =$$

$$= 2\frac{x^{-\frac{1}{2}}}{-\frac{1}{2}} + 3\frac{x^{\frac{1}{6}}}{\frac{1}{6}} + 5\ln|x| + C = -4\sqrt{x} + 18\sqrt[6]{x} + 5\ln|x| + C.$$

Пример 4.5. Найти $\int \frac{dx}{x^4+x^2}$.

Решение. Сначала преобразуем подынтегральную функцию

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

Теперь запишем исходный интеграл как разность табличных интегралов II и X (при $a = 1$):

$$\int \frac{dx}{x^4+x^2} = \int \frac{dx}{x^2} - \int \frac{dx}{x^2+1} = \int x^{-2}dx - \int \frac{dx}{x^2+1} = -\frac{1}{x} - \operatorname{arctg} x + C.$$

Пример 4.6. Найти $\int \frac{\cos 2x dx}{\sin^2 x \cos^2 x}$.

Решение. Используя формулу $\cos 2x = \cos^2 x - \sin^2 x$, представим данный интеграл в виде разности табличных интегралов VIII и VII:

$$\int \frac{\cos 2x dx}{\sin^2 x \cos^2 x} = \int \frac{\cos^2 x - \sin^2 x}{\sin^2 x \cos^2 x}dx = \int \frac{dx}{\sin^2 x} - \int \frac{dx}{\cos^2 x} = -\operatorname{ctg} x - \operatorname{tg} x + C.$$

§ 4.2. Методы интегрирования

1. Интегрирование методом замены переменной

Во многих случаях введение новой переменной позволяет упростить подынтегральное выражение и свести интеграл к линейной комбинации табличных. Такой метод называется методом *замены переменной*. Он основан на следующей теореме.

Теорема 4.2. *Пусть функция $x = \varphi(t)$ определена и дифференцируема на промежутке T и X – множество ее значений, на котором определена функция $f(x)$. Тогда если $F(x)$ – первообразная для $f(x)$ на X, то $F(\varphi(t))$ – первообразная для $f(\varphi(t))\varphi'(t)$ на T, т.е. на множестве T выполняется равенство*

$$\int f(x)dx\Big|_{x=\varphi(t)} = \int f(\varphi(t))\varphi'(t)dt. \tag{4.1}$$

Доказательство. По правилу дифференцирования сложной функции производная левой части равенства равна

$$F_t'(\varphi(t)) = F_x'(\varphi(t))\cdot\varphi'(t) = f(\varphi(t))\varphi'(t),$$

что совпадает с подынтегральной функцией в правой части равенства, это и доказывает равенство (4.1).

Формула (4.1) называется *формулой замены переменной в неопределенном интеграле*.

Пример 4.7. Найти $\int\frac{\sqrt{x}}{1+x}dx.$

Решение. Для того чтобы избавиться от иррациональности, выполним замену переменной $x = t^2$, $t=\sqrt{x}$. Тогда $dx = 2tdt$ и наш интеграл примет вид

$$\int\frac{\sqrt{x}}{1+x}dx = \int\frac{t\cdot 2tdt}{1+t^2} = 2\int\frac{1+t^2-1}{1+t^2}dt = 2\left(\int dt - \int\frac{dt}{1+t^2}\right) = 2(t-\operatorname{arctg}t)+C.$$

Возвращаясь к переменной x, окончательно получим

$$\int\frac{\sqrt{x}}{1+x}dx = 2\left(\sqrt{x}-\operatorname{arctg}\sqrt{x}\right)+C.$$

Пример 4.8. Найти $\int\sqrt{4-x^2}\,dx.$

Решение. Здесь удобно выполнить *тригонометрическую замену* $x = 2\sin t,\ 0 \le x \le \frac{\pi}{2}$. Действительно, подкоренное выражение принимает вид

$$4 - x^2 = 4 - 4\sin^2 t = 4(1 - \sin^2 t) = 4\cos^2 t, \text{ а } dx = 2\cos t\,dt,$$

так что имеем

$$\int\sqrt{4-x^2}\,dx = \int 2\cos t \cdot 2\cos t dt = 2\int 2\cos^2 t dt.$$

Применяя тригонометрическую формулу "понижения степени": $2\cos^2 t = 1 + \cos 2t$, получим

$$\int\sqrt{4-x^2}\,dx = 2\int(1+\cos 2t)dt = 2t + \sin 2t + C.$$

Чтобы найти окончательный ответ, выражающий неопределенный интеграл через переменную x, выразим t из формулы замены через x: $t = \arcsin\frac{x}{2}$ и заметим, что

$$\sin 2t = 2\sin t\cos t = x\cdot\sqrt{1-\frac{x^2}{4}} = \frac{1}{2}x\cdot\sqrt{4-x^2}.$$

Таким образом,

$$\int\sqrt{4-x^2}\,dx = 2\arcsin\frac{x}{2} + \frac{1}{2}x\cdot\sqrt{4-x^2} + C.$$

Частным случаем теоремы 4.2 является следующая теорема, часто используемая на практике (*линейная замена переменной*).

Теорема 4.3. *Если* $\int f(x)dx = F(x) + C,$ *то*

$$\int f(ax+b)dx = \frac{1}{a}F(ax+b) + C.$$

Д о к а з а т е л ь с т в о . Положим $t = ax + b$, тогда $dt = adx$, откуда $dx = a^{-1}dt$, так что $\int (ax+b)dx = a^{-1}\int f(t)dt = a^{-1}F(t) + C$, и, выражая x через t, получим то, что и требовалось.

Пример 4.9. Найти $\int (2x+1)^4 dx$.

Р е ш е н и е . Можно, конечно, записать подынтегральную функцию как полином от x, возведя $2x + 1$ в четвертую степень, но это будет нерационально, поскольку проще выполнить линейную замену:

$$\int (2x+1)^4 dx = \frac{1}{2}\int (2x+1)^4 d(2x+1) = \frac{1}{2}\frac{(2x+1)^5}{5} + C = \frac{(2x+1)^5}{10} + C.$$

Пример 4.10. Найти $\int \frac{dx}{2x+5}$.

Р е ш е н и е . Воспользуемся табличным интегралом III:

$$\int \frac{dx}{x} = \ln|x| + C.$$

Тогда

$$\int \frac{dx}{2x+5} = \frac{1}{2}\ln|2x+5| + C.$$

Формулой замены переменной (4.1) пользуются и справа налево, и тогда этот метод называют иногда методом “подведения под дифференциал”.

Пример 4.11. Найти $\int \operatorname{tg} x dx$.

Р е ш е н и е . Имеем

$$\int \operatorname{tg} x dx = \int \frac{\sin x}{\cos x} dx = -\int \frac{d(\cos x)}{\cos x} = -\ln|\cos x| + C.$$

Аналогично находится интеграл от $\operatorname{ctg} x$.

2. Метод интегрирования по частям

Метод интегрирования по частям основан на формуле дифференцирования произведения двух функций.

Т е о р е м а 4.4. *Пусть* $u(x)$ *и* $v(x)$ *– две дифференцируемые функции на промежутке X. Тогда на X выполняется формула интегрирования по частям*

$$\int u dv = uv - \int v du. \tag{4.2}$$

Д о к а з а т е л ь с т в о. Имеем формулу для дифференциала произведения функций uv: $d(uv) = udv + vdu$. Интегрируя обе части равенства, получим слева uv по свойству 2 неопределенного интеграла, а справа сумму интегралов, так что

$$uv = \int u dv + \int v du,$$

откуда легко получается формула (4.2).

Эта формула позволяет свести нахождение неопределенного интеграла $\int u dv$ к неопределенному интегралу $\int v du$, который может оказаться более простым.

П р и м е р 4.12. Найти $\int x \cos x dx$.

Р е ш е н и е. Положим $u = x$, $dv = \cos x dx$. Тогда $du = dx$, а функция v находится интегрированием $v = \int \cos x dx$, так что можно положить $v = \sin x$. Подставляя в формулу (4.2), получим

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

П р и м е р 4.13. Найти $\int \ln x dx$.

Р е ш е н и е. Удобно все необходимые действия выполнять в одну строку, отделяя вспомогательные записи вертикальными чертами. Итак,

$$\int \ln x dx = \left| \begin{array}{l} u = \ln x,\ dv = dx \\ du = \dfrac{dx}{x},\ v = \int dx = x \end{array} \right| = \ln x \cdot x - \int x \frac{dx}{x} = x \ln x - x + C.$$

Иногда метод интегрирования по частям приходится комбинировать с методом замены переменной, как в случае интегрирования основных элементарных функций $y = \arcsin x$ и $y = \operatorname{arctg} x$.

Пример 4.14. Найти $\int \arcsin x dx$.

Решение. Имеем

$$\int \arcsin x dx = \left| \begin{array}{l} u = \arcsin x,\ dv = dx \\ du = \dfrac{dx}{\sqrt{1-x^2}},\ v = x \end{array} \right| = \arcsin x \cdot x - \int \frac{xdx}{\sqrt{1-x^2}}.$$

Для получившегося интеграла применим замену переменной $t = 1 - x^2$, $dt = -2xdx$, откуда $xdx = -\frac{1}{2}dt$. Поэтому

$$\int \frac{xdx}{\sqrt{1-x^2}} = -\frac{1}{2}\int \frac{dt}{\sqrt{t}} = -\sqrt{t} + C = -\sqrt{1-x^2} + C.$$

Окончательно имеем

$$\int \arcsin x dx = x \arcsin x + \sqrt{1-x^2} + C.$$

Заметим, что поскольку C – произвольная постоянная, то получающиеся функции от C можно обозначать по-прежнему C.

Иногда для вычисления интеграла метод интегрирования по частям приходится применять неоднократно.

Пример 4.15. Найти $\int x^2 e^x dx$.

Решение.

$$\int x^2 \cdot e^x dx = \left| \begin{array}{l} u = x^2,\ dv = e^x dx \\ du = 2xdx,\ v = \int e^x dx = e^x \end{array} \right| = x^2 \cdot e^x - \int e^x \cdot 2xdx =$$

$$= x^2 e^x - 2\int x \cdot e^x dx = \left| \begin{array}{l} u = x,\ dv = e^x dx \\ du = dx,\ v = e^x \end{array} \right| = x^2 e^x - 2\left(x \cdot e^x - \int e^x dx\right) =$$

$$= x^2 e^x - 2xe^x + 2e^x + C.$$

В заключение вычислим интеграл

$$I_n = \int \frac{dx}{\left(x^2+1\right)^n},$$

где n – натуральное число. При $n = 1$ мы имеем табличный интеграл X ($a = 1$):

$$I_1 = \operatorname{arctg} x + C.$$

Пусть $n > 1$. Представим 1 в числителе в виде $1 = x^2 + 1 - x^2$, так что интеграл I_n представляется в виде разности двух интегралов:

$$I_n = I_{n-1} - \int \frac{x^2 dx}{\left(x^2+1\right)^n}.$$

Во втором интеграле применим метод интегрирования по частям

$$\int \frac{x^2 dx}{\left(x^2+1\right)^n} = \left| \begin{array}{l} u = x, \;\; dv = \dfrac{xdx}{\left(x^2+1\right)^n} \\ du = dx, \;\; v = \displaystyle\int \frac{xdx}{\left(x^2+1\right)^n} = -\frac{1}{2(n-1)\left(x^2+1\right)^{n-1}} \end{array} \right| =$$

= (для нахождения v выполните замену $t = x^2 + 1$) =

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

Подставляя найденный интеграл в формулу для I_n, получим

$$I_n = I_{n-1} + \frac{x}{2(n-1)\left(x^2+1\right)^{n-1}} - \frac{1}{2(n-1)} I_{n-1},$$

т.е. интеграл I_n выражается через интеграл такого же вида с индексом на 1 меньше. Ясно, что за конечное число шагов мы дойдем до интеграла I_1 и, следовательно, выразим I_n в виде некоторой элементарной функции.

Формулы такого вида называются *рекуррентными*.

Пример 4.16. Найти $\int \frac{dx}{\left(x^2+1\right)^2}$.

Решение. Имеем

$$I_2 = I_1 + \frac{x}{2\left(x^2+1\right)} - \frac{1}{2} I_1 = \frac{1}{2}\operatorname{arctg} x + \frac{x}{2\left(x^2+1\right)} + C.$$

§ 4.3. Интегрирование некоторых классов функций

В предыдущих параграфах речь шла об общих приемах интегрирования. Далее мы рассмотрим интегрирование конкретных классов функций.

1. Интегрирование рациональных функций

Рассмотрим интеграл вида $\int R(x)dx$, где $R(x)$ – это рациональная функция, т.е. функция, которую можно записать в виде отношения двух многочленов:

$$R(x) = \frac{P(x)}{Q(x)}.$$

Если эта дробь *неправильная*, т.е. степень многочлена $P(x)$ не меньше степени многочлена $Q(x)$, то можно выполнить деление с остатком и представить $R(x)$ в виде суммы некоторого многочлена и *правильной* дроби.

Пример 4.17. Найти $\int \frac{x^3 dx}{x^2-4}$.

Решение. Рациональная функция $R(x)=\dfrac{x^3}{x^2-4}$ – неправильная дробь, поэтому сначала мы должны выполнить деление с остатком:

$$\begin{array}{r|l} -\;x^3 \qquad & x^2-4 \\ \cline{2-2} x^3-4x & x \\ \hline 4x & \end{array}$$

т.е. $x^3 = (x^2 - 4)\cdot x + 4x$, поэтому подынтегральную функцию можно представить в виде $\dfrac{x^3}{x^2-4}=x+\dfrac{4x}{x^2-4}$. Таким образом, получим

$$\int\frac{x^3dx}{x^2-4}=\int xdx+\int\frac{4xdx}{x^2-4}=\frac{x^2}{2}+2\ln\left|x^2-4\right|+C.$$

(Для нахождения интеграла от второго слагаемого выполните замену переменной $t = x^2 - 4$.)

Таким образом, можно считать, что рациональная функция $R(x)$ представлена в виде правильной дроби. В курсе алгебры доказывается следующая теорема.

Теорема 4.5. *Всякая правильная дробь может быть представлена в виде суммы* ***простейших*** *дробей вида*

$$\frac{A}{x-a},\ \frac{A}{(x-a)^n},\ \frac{Mx+N}{x^2+px+q},\ \frac{Mx+N}{\left(x^2+px+q\right)^n},$$

где A, M, N, a, p, q – действительные числа.

Пример 4.18. Найти $\displaystyle\int\frac{x^2+4}{x^3-4x}dx.$

Решение. Разложим знаменатель подынтегральной рациональной функции на множители: $x^3 - 4x = x(x^2 - 4x) = x(x - 2)(x + 2)$.

Согласно теореме 4.5 правильная дробь должна разлагаться в сумму простейших дробей

$$\frac{x^2+4}{x^3-4x}=\frac{A}{x}+\frac{B}{x-2}+\frac{C}{x+2}\,.$$

Найдем коэффициенты A, B, C. Воспользуемся для этого приемом, который называют *методом неопределенных коэффициентов*. А именно, приведем правую часть равенства к общему знаменателю и приравняем числители:

$$x^2 + 4 = A(x - 2)(x + 2) + Bx(x + 2) + Cx(x - 2).$$

Заметим, что у каждого слагаемого в правой части отсутствует в точности один сомножитель, так что при подстановке корней знаменателя все слагаемые правой части, кроме одного, обратятся в нуль:

$$\begin{array}{lll} x = 0: & 4 = A(-2)\cdot 2, & A = -1; \\ x = 2: & 8 = B\cdot 2\cdot 4, & B = 1; \\ x = -2: & 8 = C(-2)\cdot(-4), & C = 1. \end{array}$$

Таким образом, искомое разложение на простейшие дроби имеет вид:

$$\frac{x^2+4}{x^3-4x}=\frac{-1}{x}+\frac{1}{x-2}+\frac{1}{x+2}\,,$$

так что интеграл представляется в виде суммы интегралов, которые легко находятся

$$\int\frac{x^2+4}{x^3-4x}dx=-\int\frac{dx}{x}+\int\frac{dx}{x-2}+\int\frac{dx}{x+2}=-\ln|x|+\ln|x-2|+\ln|x+2|+C.$$

Пример 4.19. Найти $\displaystyle\int\frac{x^2+4x+4}{x^3-2x^2+x}dx.$

Решение. Разложим знаменатель на множители

$$x^3 - 2x^2 + x = x(x^2 - 2x + 1) = x(x - 1)^2.$$

Согласно теореме 4.5 подынтегральная функция разложится на сумму простейших дробей:

$$\frac{x^2+4x+4}{x(x-1)^2}=\frac{A}{x}+\frac{B}{x-1}+\frac{C}{(x-1)^2}.$$

Так же как в примере 4.18, приведем правую часть равенства к общему знаменателю и приравняем числители

$$x^2 + 4x + 4 = A(x - 1)^2 + Bx(x - 1) + Cx.$$

Подставим по очереди корни знаменателя $x = 0$ и $x = 1$ и найдем:

$$x = 0: \qquad 4 = A, \qquad A = 4;$$

$$x = 1: \qquad 1 + 4 + 4 = C, \qquad C = 9.$$

Для нахождения B корней не хватает! Однако поскольку многочлены в левой и правой частях равны тождественно, то у них равны коэффициенты при одинаковых степенях x. Приравнивая коэффициенты при x^2, получим

$$1 = A + B,$$

откуда $B = –3$. Таким образом, искомое разложение на простейшие дроби имеет вид

$$\frac{x^2+4x+4}{x(x-1)^2}=\frac{4}{x}+\frac{-3}{x-1}+\frac{9}{(x-1)^2},$$

так что интеграл равен сумме интегралов

$$\int\frac{x^2+4x+4}{x(x-1)^2}dx=4\int\frac{dx}{x}-3\int\frac{dx}{x-1}+9\int\frac{dx}{(x-1)^2}=$$

$$=4\ln|x|-3\ln|x-1|-\frac{9}{x-1}+C.$$

Пример 4.20. Найти $\int \frac{dx}{x\left(x^2+1\right)^2}$.

Решение. Рациональная дробь $\frac{1}{x\left(x^2+1\right)^2}$ – правильная, и ее разложение на простейшие дроби имеет вид

$$\frac{1}{x\left(x^2+1\right)^2}=\frac{A}{x}+\frac{Bx+C}{x^2+1}+\frac{Dx+E}{\left(x^2+1\right)^2}.$$

Сравнивая числители дробей в обеих частях равенства, получим

$$1 = A(x^2 + 1)^2 + (Bx + C)x(x^2 + 1) + (Dx + E)x.$$

В этом случае у нас имеется только один действительный корень $x = 0$, этого достаточно для нахождения только одного коэффициента A:

$$1 = A, \quad A = 1.$$

Для нахождения остальных коэффициентов раскроем скобки в правой части равенства и запишем ее в виде многочлена четвертой степени:

$$1 = (A + B)x^4 + Cx^3 + (2A + B + D)x^2 + (C + E)x + A.$$

Сравнивая коэффициенты при одинаковых степенях x в левой и правой частях, получим систему уравнений для нахождения неопределенных коэффициентов:

$$x^4:\ A+B=0; \quad x^3:\ C=0; \quad x^2:\ \ 2A+B+D=0;$$

$$x^1:\ C+E=0; \quad x^0:\ A=1.$$

Отсюда сразу находим $A = 1$; $B = -1$; $C = 0$; $D = -1$; $E = 0$. Искомое разложение имеет вид

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

Следовательно,

$$\int\frac{dx}{x\left(x^2+1\right)^2}=\int\frac{dx}{x}-\int\frac{xdx}{x^2+1}-\int\frac{xdx}{\left(x^2+1\right)^2}.$$

Второй и третий интеграл справа находим одинаковой заменой $t =$ $= x^2 + 1$, $dt = 2xdx$, и окончательно получим

$$\int\frac{dx}{x\left(x^2+1\right)^2}=\ln|x|-\frac{1}{2}\ln\left(x^2+1\right)+\frac{1}{2\left(x^2+1\right)}+C.$$

2. Интегрирование тригонометрических функций

Универсальная тригонометрическая подстановка. С помощью такой замены переменной интегралы вида

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

где R – некоторая рациональная функция, приводят к интегралу от рациональной функции, и, следовательно, к ним можно применить методы, рассмотренные в предыдущем пункте.

А именно, воспользуемся формулами, выражающими синус и косинус через тангенс половинного аргумента

$$\sin x=\frac{2\operatorname{tg}\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}},\ \cos x=\frac{1-\operatorname{tg}^2\frac{x}{2}}{1+\operatorname{tg}^2\frac{x}{2}},\ -\pi<x<\pi.$$

Если мы положим $t=\operatorname{tg}\frac{x}{2}$, то выражая x через t, получим

$$x=2\operatorname{arctg}t,\quad dx=\frac{2dt}{1+t^2}.$$

Поэтому

$$\int R(\sin x,\cos x)dx=\int R\left(\frac{2t}{1+t^2},\frac{1-t^2}{1+t^2}\right)\frac{2dt}{1+t^2}.$$

Таким образом, задача свелась к интегрированию рациональной функции. Разумеется, после нахождения интеграла справа нужно вернуться к переменной x, т.е. положить $t = \mathrm{tg}\dfrac{x}{2}$.

Пример 4.21. Найти $\displaystyle\int\frac{dx}{8-4\sin x+7\cos x}$.

Решение. Применим универсальную тригонометрическую подстановку:

$$\int\frac{dx}{8-4\sin x+7\cos x}=\int\frac{1}{8-4\dfrac{2t}{1+t^2}+7\dfrac{1-t^2}{1+t^2}}\frac{2dt}{1+t^2}=$$

$$=\int\frac{2dt}{t^2-8t+15}=\int\frac{2dt}{(t-4)^2-1}=2\cdot\frac{1}{2}\ln\left|\frac{t-4-1}{t-4+1}\right|+C=\ln\left|\frac{t-5}{t-3}\right|+C.$$

Здесь мы воспользовались табличным интегралом XI. Подставляя $t=\mathrm{tg}\dfrac{x}{2}$, окончательно получим

$$\int\frac{dx}{8-4\sin x+7\cos x}=\ln\left|\frac{\mathrm{tg}\dfrac{x}{2}-5}{\mathrm{tg}\dfrac{x}{2}-3}\right|+C.$$

Иногда, исходя из вида подынтегральной функции, проще воспользоваться другой заменой.

Пример 4.22. Найти $\displaystyle\int\frac{\cos^3 x}{\sin^4 x}dx.$

Решение. Выполним замену переменной $t = \sin x$. Числитель подынтегрального выражения можно представить следующим образом:

$$\cos^3 x dx = \cos^2 x\cdot\cos x dx = (1 - \sin^2 x)\, d(\sin x) = (1 - t^2)\, dt.$$

Поэтому имеем

$$\int\frac{\cos^3 x}{\sin^4 x}dx=\int\frac{1-t^2}{t^4}dt=\int\frac{dt}{t^4}-\int\frac{dt}{t^2}=\int t^{-4}dt-\int t^{-2}dt=\frac{t^{-3}}{-3}-\frac{t^{-1}}{-1}+C.$$

Возвращаясь к переменной x, получим

$$\int \frac{\cos^3 x}{\sin^4 x} dx = \frac{1}{\sin x} - \frac{1}{3\sin^3 x} + C.$$

Пример 4.23. Найти $\int \frac{dx}{\sin^2 x - 4\sin x \cos x + 5\cos^2 x}$.

Решение. Имеем

$$\int \frac{dx}{\sin^2 x - 4\sin x \cos x + 5\cos^2 x} =$$

$$= \int \frac{dx}{\cos^2 x\left(\operatorname{tg}^2 x - 4\operatorname{tg} x + 5\right)} = \int \frac{d(\operatorname{tg} x)}{\operatorname{tg}^2 x - 4\operatorname{tg} x + 5}.$$

Теперь ясно, что удобно выполнить замену переменной $t = \operatorname{tg} x$. Интеграл примет вид

$$\int \frac{dt}{t^2 - 4t + 5} = \int \frac{dt}{(t-2)^2 + 1} = \operatorname{arctg}(t-2) + C.$$

Таким образом,

$$\int \frac{dx}{\sin^2 x - 4\sin x \cos x + 5\cos^2 x} = \operatorname{arctg}(\operatorname{tg} x - 2) + C.$$

3. Использование справочников и математических процессоров. Неберущиеся интегралы

Из предыдущего ясно, что нахождение неопределенных интегралов – задача, существенно более сложная по сравнению с дифференцированием. Ее решение можно облегчить, применяя математические справочники и компьютерные программы, такие, например, как Scientific WorkPlace. В эту программу встроен символьный процессор, который позволяет, в частности, находить производные и первообразные функций. Ниже приведены примеры нахождения первообразных с помощью этой программы. Здесь слева указаны данные функции, а справа – их первообразные, найденные с помощью Scientific WorkPlace:

$$\int \frac{1}{(1+x)\cdot\sqrt{1+x-x^2}} = \operatorname{arctg}\frac{1+3x}{2\sqrt{1+x-x^2}},$$

$$\int \frac{dx}{1+\sin x} = -\frac{2}{1+\operatorname{tg}\frac{x}{2}}.$$

Задача нахождения первообразной элементарной функции – это вторая знаменитая математическая проблема, которая оказалась неразрешимой принципиально (первой была задача решения алгебраических уравнений в радикалах). Оказалось, что очень многие элементарные функции неинтегрируемы, т.е. первообразные таких функций не являются элементарными функциями. Таковы, например, функции e^{-x^2}, $\frac{1}{\ln x}$, $\frac{\sin x}{x}$ и т.д. Однако, как мы увидим далее в этой главе, такие первообразные существуют и играют значительную роль в математике и ее приложениях. Свойства этих функций хорошо изучены, существуют подробные таблицы их значений, можно рассмотреть их графики и т.п.

Если первообразная не является элементарной функцией, то говорят, что интеграл “не берется” в элементарных функциях.

§ 4.4. Определенный интеграл

1. Задача о вычислении площади криволинейной трапеции

Пусть дана непрерывная неотрицательная функция $y=f(x)$ на отрезке $[a, b]$. *Криволинейной трапецией* называется фигура $aABb$ на плоскости Oxy, ограниченная вертикальными прямыми $x = a$ и $x = b$, осью Ox и графиком функции $y=f(x)$ (рис. 4.1).

Как найти площадь этой трапеции? Напомним, что при вычислении площади круга в школьном курсе математики поступают следующим образом. Рассматривают вписанные и описанные правильные многоугольники с увеличивающимся числом сторон, вычисляют их площадь и затем принимают за площадь круга предел площадей этих многоугольников. По сути, этот метод, используемый еще со времен Архимеда и известный как “метод исчерпывания”, применим и в данной ситуации.

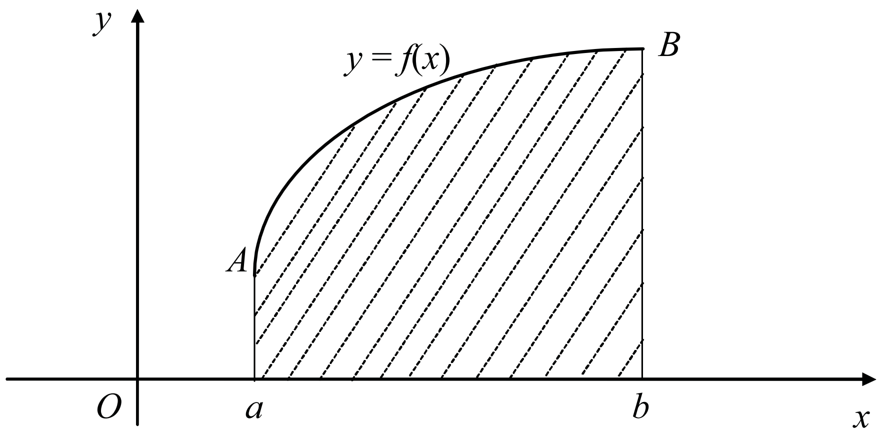

Рис. 4.1

Разобьем отрезок $[a, b]$ на мелкие части точками деления $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$. Затем найдем на каждом отрезке разбиения $[x_k, x_{k+1}]$ наименьшее значение функции $y = f(x)$, обозначим его m_k и построим прямоугольник с основанием $[x_k, x_{k+1}]$ и высотой m_k. Его площадь $S_k = m_k \cdot (x_{k+1} - x_k) = m_k \cdot \Delta x_k$.

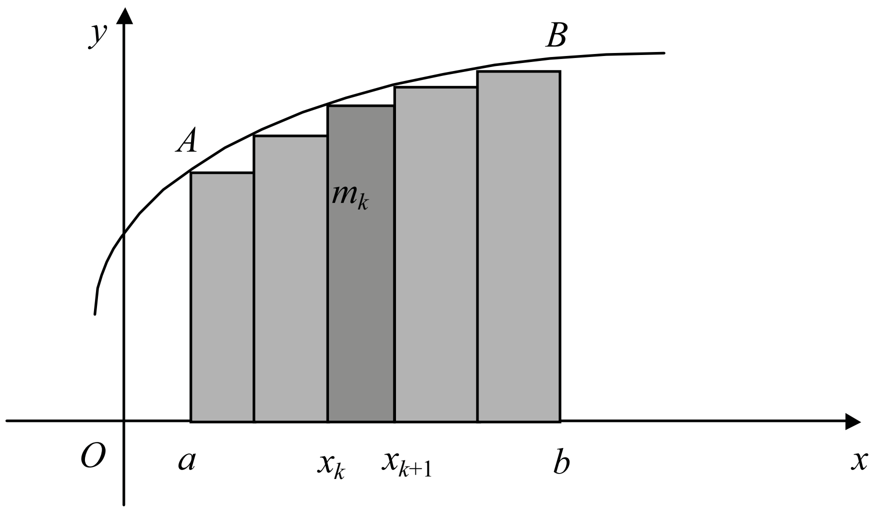

Рис. 4.2

Объединение всех таких прямоугольников даст ступенчатую фигуру, “вписанную” в криволинейную трапецию *aABb*, обозначим ее площадь s_T, где T обозначает выбранное *разбиение отрезка* $[a, b]$ (рис. 4.2).

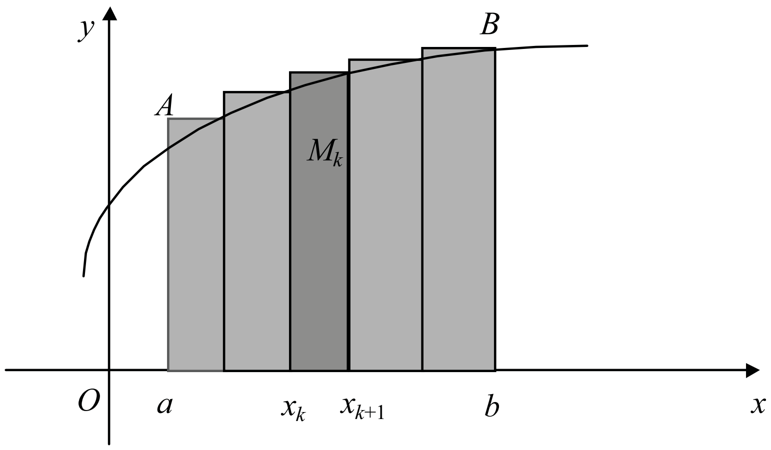

Рис. 4.3

Ясно, что S – площадь криволинейной трапеции *aABb* – будет не меньше площади вписанной ступенчатой фигуры s_T для любого разбиения T: $s_T \le S$. Если на каждом отрезке разбиения T выбрать наибольшее значение M_k функции $y = f(x)$, то, поступая аналогичным образом, получим ступенчатую фигуру, описанную вокруг криволинейной трапеции *aABb*; площадь этой фигуры обозначим S_T (рис. 4.3).

Очевидно, что для любого разбиения T отрезка $[a, b]$ выполняется двойное неравенство

$$s_T \le S \le S_T.$$

Интуитивно ясно, что если разбиения T сделать достаточно мелкими, то площади вписанной и описанной ступенчатых фигур будут мало различаться, поэтому естественно за площадь криволинейной трапеции S принять число, которое не меньше площади любой вписанной ступенчатой фигуры и не больше площади любой описанной ступенчатой фигуры. В рассматриваемом случае такое число должно быть единственным.

Понятно также, что искомая площадь S приближенно равна площади вписанной или описанной ступенчатой фигуры: $S \approx s_T$ или

$S \approx S_T$, причем точность равенства увеличивается с измельчением разбиения T. На практике отрезок $[a, b]$ разбивают на n равных частей, вместо обозначения s_T используют s_n, соответственно вместо $S_T - S_n$. Разность $S_n - s_n$ определяет точность. Чем больше n, тем выше точность; проще всего число точек удваивать и проверять на каждом шагу: достигнута ли искомая точность. Точное равенство получится, если в приближенном равенстве перейти к пределу.

$$S = \lim_{n\to\infty} s_n \text{, или } S = \lim_{n\to\infty} S_n .$$

Пример 4.24. С точностью до 0,01 вычислить площадь криволинейной трапеции, ограниченной кубической параболой $y = x^3$, осью Ox и прямыми $x = 0$ и $x = 1$ (рис. 4.4).

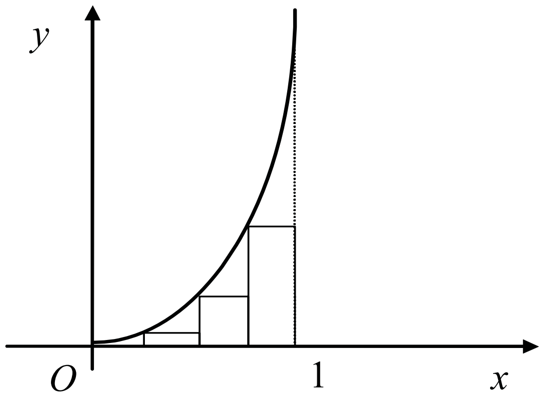

Рис. 4.4

Решение. Будем вычислять нижние и верхние суммы s_n и S_n, удваивая на каждом шаге число точек деления, т.е. полагая $n = 2$; 4; 8; ... и т.д. Результаты вычислений приведены в следующей таблице.

n	2	4	8	16	32	64	128	256	512	1024
s_n	0,063	0,141	0,191	0,22	0,235	0,242	0,246	0,248	0,249	0,25
S_n	0,563	0,391	0,316	0,282	0,266	0,258	0,254	0,252	0,251	0,25

Последнее табличное значение является точным. Ясно, что такой способ нахождения площади криволинейной трапеции сопряжен с большим объемом вычислений. Ниже эту задачу мы решим с минимальными вычислениями, однако только после того, как будут изложены необходимые теоретические сведения.

Пример 4.25. Пусть зависимость объема продаж данного товара от времени задана функцией $q = \dfrac{10t}{1+t^2}$ $(0 \le t \le 5)$ (рис. 4.5).

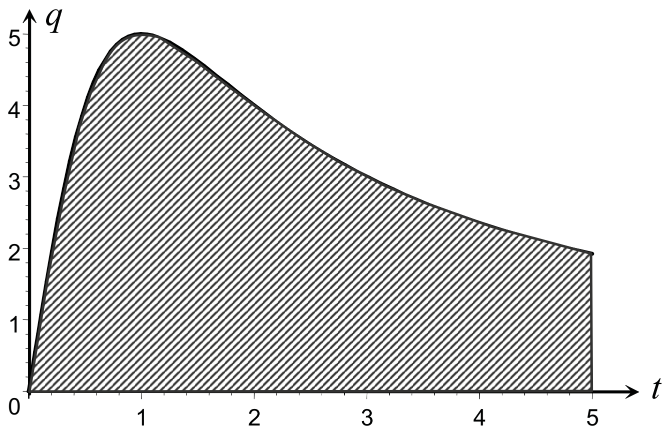

Рис. 4.5

Тогда общий объем продаж Q равен площади заштрихованной фигуры, т.е. опять площади криволинейной трапеции, и мы можем применить тот же метод исчерпывания для приближенного нахождения Q.

2. Понятие определенного интеграла

Два примера, рассмотренные в предыдущем пункте, показывают, что решение таких и многих аналогичных им задач приводит к построению нижних и верхних сумм, соответственно s_T и S_T , для подходящего разбиения T данного отрезка и нахождению числа, разделяющего множества $\{s_T\}$ и $\{S_T\}$. Рассмотрим данную задачу в общем виде.

Итак, пусть на отрезке $[a, b]$ дана ограниченная функция $y = f(x)$. Рассмотрим разбиение T отрезка $[a, b]$ точками деления

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

на каждом отрезке разбиения $[x_k, x_{k+1}]$ найдем нижнюю и верхнюю грани значений функции $y = f(x)$, соответственно m_k и M_k, и составим две суммы: *нижнюю сумму Дарбу* $s_T = \sum_{k=0}^{n-1} m_k \cdot \Delta x_k$ и *верхнюю сумму Дарбу* $S_T = \sum_{k=0}^{n-1} M_k \cdot \Delta x_k$. На рис. 4.2 и 4.3 этим суммам соответствуют площади вписанной и описанной ступенчатых фигур.

Эти суммы обладают следующими свойствами.

1. Для любого разбиения T выполняется неравенство $s_T \le S_T$, т.е. *для данного разбиения нижняя сумма Дарбу не превосходит верхней суммы*.

Это следует из того, что для любого k справедливо неравенство $m_k \le M_k$.

2. Если разбиение T_2 получается из разбиения T_1 добавлением нескольких новых точек, то $s_{T_1} \le s_{T_2}$ и $S_{T_1} \ge S_{T_2}$, т.е. *при измельчении разбиения нижние суммы Дарбу могут только увеличиться, а верхние суммы только уменьшиться*.

Идея доказательства совершенно очевидна из нижеприведенных графических иллюстраций. Действительно, на рис. 4.6 площадь фигуры x_kABx_{k+1} меньше суммарной площади фигуры x_kAC_1c и фигуры $cCDx_{k+1}$.

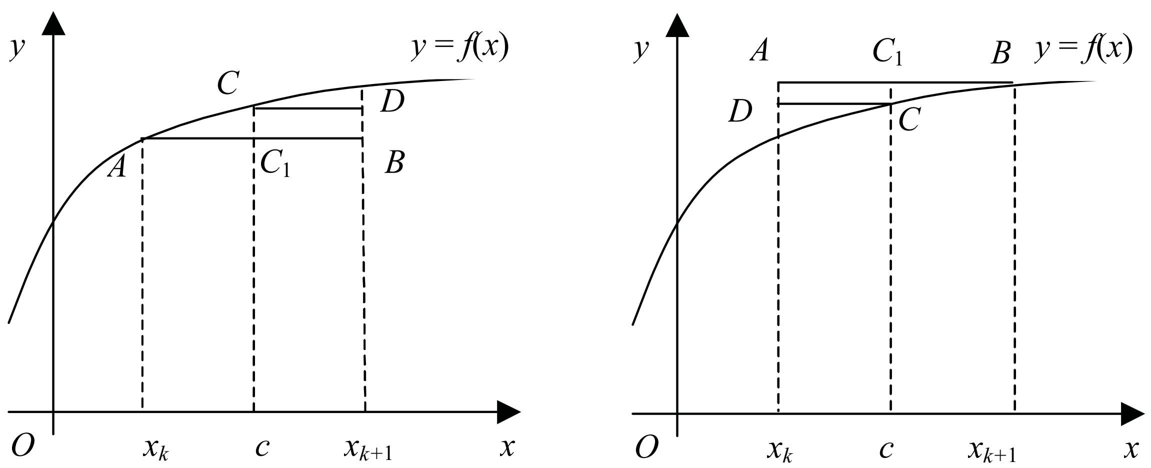

Рис. 4.6

3. Для любых разбиений T_1 и T_2 выполняется неравенство $s_{T_1} \le S_{T_2}$, или *любая нижняя сумма Дарбу не превосходит любой верхней суммы*.

Для доказательства составим разбиение отрезка T, которое включает все точки деления как разбиения T_1, так и T_2. Тогда по свойству 2 имеем $s_{T_1} \le s_T$, по свойству 1 $s_T \le S_T$ и, применяя еще раз свойство 2, получим $S_T \le S_{T_2}$. Таким образом, имеем

$$s_{T_1} \le s_T \le S_T \le S_{T_2},$$

что завершает доказательство.

Из свойства 3 следует, что множество X нижних сумм Дарбу находится на числовой оси левее множества Y верхних сумм Дарбу, поэтому существует хотя бы одно число I, разделяющее множества X и Y, т.е. для любого разбиения T отрезка $[a, b]$ выполняется двойное неравенство

$$s_T = \sum_{k=0}^{n-1} m_k \cdot \Delta x_k \le I \le \sum_{k=0}^{n-1} M_k \cdot \Delta x_k = S_T.$$

Определение. *Функция $y = f(x)$, ограниченная на отрезке $[a, b]$, называется* ***интегрируемой*** *на этом отрезке, если существует* ***единственное*** *число I, разделяющее множества нижних и верхних сумм Дарбу для всевозможных разбиений отрезка $[a, b]$. Если функция $y = f(x)$ интегрируема на отрезке $[a, b]$, то единственное число, разделяющее эти два множества, называют* ***определенным интегралом*** *функции $y = f(x)$ по отрезку $[a, b]$ и обозначают следующим образом*:

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

Знак определенного интеграла читается “интеграл от a до b”, числа a и b называют *нижним и верхним пределами интегрирования*. Обозначения были введены немецким ученым Г. Лейбницем, который вместе с И. Ньютоном был создателем дифференциального и интегрального исчислений. Лейбниц ввел знак интеграла $\int$ в виде вытянутой буквы S, которая обозначает знак суммирования. Самым замечательным результатом интегрального исчисления является формула Ньютона – Лейбница, которая устанавливает связь между определенным интегралом $\int_a^b f(x)dx$ и неопределенным интегралом $\int f(x)dx$, что делает оправданным употребление знака интеграла в обоих случаях.

Мы определили интеграл $\int_a^b f(x)dx$ для случая, когда $a < b$. Если $a > b$, положим $\int_a^b f(x)dx = -\int_b^a f(x)dx$, что можно считать естественным, поскольку при изменении направления каждая разность $\Delta x_k = x_{k+1} - x_k$ меняет знак, а тогда поменяют знак и суммы Дарбу и, следовательно, разделяющее их число, т.е. интеграл

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

Так как при $a = b$ все Δx_k обращаются в нуль, то положим

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

Геометрический смысл определенного интеграла. *Определенный интеграл равен площади криволинейной трапеции, ограниченной вертикальными прямыми $x = a$, $x = b$ при $a < b$, осью Ox и графиком неотрицательной и непрерывной функции $y = f(x)$.*

Экономические приложения определенного интеграла рассмотрим позже, однако один типичный случай уже был приведен в примере 4.25.

Рассмотрим теперь пример, показывающий, что существуют неинтегрируемые ограниченные функции. В качестве такой функции рассмотрим функцию Дирихле $y = D(x)$, которая определена на отрезке [0; 1] следующим образом:

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

Каково бы ни было разбиение T, в любом отрезке разбиения $[x_k, x_{k+1}]$ обязательно содержатся как рациональные, так и иррациональные точки, поэтому для любого отрезка Δx_k: $m_k = 0$ и $M_k = 1$. Тогда все нижние суммы Дарбу $\sum_{k=0}^{n-1} m_k \cdot \Delta x_k = 0$, поскольку все $m_k = 0$, и все верхние суммы Дарбу $\sum_{k=0}^{n-1} M_k \cdot \Delta x_k = 1$, так как $\sum_{k=0}^{n-1} M_k \cdot \Delta x_k = \sum_{k=0}^{n-1} 1 \cdot \Delta x_k = 1$ – длина отрезка [0, 1].

Таким образом, множество нижних сумм состоит из одного числа $X = \{0\}$, а множество верхних сумм – из одного числа $Y =$ $=\{1\}$, так что любое число из отрезка [0, 1] разделяет множества X и Y. Значит, функция Дирихле не является интегрируемой на отрезке [0, 1].

Сформулируем критерий интегрируемости ограниченной функции.

Теорема 4.6 (*критерий интегрируемости*). *Для того чтобы функция* $y = f(x)$, *определенная и ограниченная на отрезке* $[a, b]$, *была интегрируемой на этом отрезке*, *необходимо и достаточно*, *чтобы для любого* $\varepsilon > 0$ *существовало разбиение* T, *такое*, *что* $S_T - s_T < \varepsilon$.

Доказательство. Достаточность очевидна: положив $\varepsilon_n = 2^{-n}$, $n = 1, 2, ...$, получим систему стягивающихся отрезков $[s_{2^n}, S_{2^n}]$, сходящихся к единственной точке I, которая и будет единственным разделяющим числом.

Пусть, наоборот, известно, что разделяющее число единственно. Тогда для любого $\varepsilon > 0$ в интервал $(I - \varepsilon/2, I + \varepsilon/2)$ длины ε попадут точки как из $\{s_T\}$, так и из $\{S_T\}$. Поэтому найдутся разбиения T_1 и T_2, такие, что $S_{T_1} - s_{T_2} < \varepsilon$.

Возьмем в качестве T разбиение, которое включает точки из T_1 и из T_2. Тогда по свойству 2 сумм Дарбу имеем $S_T \le S_{T_1}$ и $s_T \ge s_{T_2}$. Отсюда $S_T - s_T < \varepsilon$, что завершает доказательство теоремы.

Поскольку

$$S_T - s_T = \sum_{k=0}^{n-1} M_k \Delta x_k - \sum_{k=0}^{n-1} m_k \Delta x_k = \sum_{k=0}^{n-1} (M_k - m_k)\Delta x_k,$$

условие $S_T - s_T < \varepsilon$ можно записать следующим образом:

$$\sum_{k=0}^{n-1} (M_k - m_k)\Delta x_k < \varepsilon. \tag{4.3}$$

Разность $M_k - m_k$ будем обозначать ω_k и называть *колебанием* функции на отрезке $[x_k, x_{k+1}]$. Неравенство (4.3) можно переписать следующим образом:

$$\sum_{k=0}^{n-1} \omega_k \Delta x_k < \varepsilon. \tag{4.4}$$

3. Интегрируемость непрерывной функции

В предыдущем пункте мы ввели понятие интегрируемой функции и установили критерий интегрируемости. Теперь покажем, что всякая функция, непрерывная на отрезке $[a, b]$, интегрируема на этом отрезке, т.е. что существует определенный интеграл $\int_a^b f(x)dx.$

Теорема 4.7. *Функция, непрерывная на отрезке* $[a, b]$, *интегрируема на этом отрезке.*

Доказательство. Возьмем произвольное $\varepsilon > 0$. По свойству *равномерной непрерывности* (см. приложение 3 к гл. 1) найдется такое разбиение T отрезка $[a, b]$, что для всех отрезков разбиения будет выполняться неравенство $\omega_k < \dfrac{\varepsilon}{b-a}$. Тогда

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

Согласно неравенству (4.4) это и означает интегрируемость функции на отрезке $[a, b]$.

4. Аддитивность определенного интеграла

Теорема 4.8. *Пусть функция $y = f(x)$ интегрируема на отрезках $[a, c]$ и $[b, c]$, $a < c < b$. Тогда она интегрируема на отрезке $[a, b]$ и выполняется равенство*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx. \tag{4.5}$$

Доказательство. Возьмем произвольное $\varepsilon > 0$. По условию функция $y = f(x)$ интегрируема на отрезке $[a, c]$. В силу критерия интегрируемости существует разбиение T_1 отрезка $[a, c]$, такое, что $S_{T_1} - s_{T_1} < \dfrac{\varepsilon}{2}$. Аналогично для отрезка $[c, b]$ найдется разбиение T_2, такое, что $S_{T_2} - s_{T_2} < \dfrac{\varepsilon}{2}$. Разбиения T_1 и T_2 в совокупности образуют разбиение T отрезка $[a, b]$, причем

$$S_T - s_T = \left(S_{T_1} + S_{T_2}\right) - \left(s_{T_1} + s_{T_2}\right) = \left(S_{T_1} - s_{T_1}\right) + \left(S_{T_2} - s_{T_2}\right) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Таким образом, для разбиения T отрезка $[a, b]$ выполнено условие критерия интегрируемости, а это и означает, что функция $f(x)$ интегрируема на этом отрезке.

Из неравенств $s_{T_1} \le \int\limits_a^c f(x)dx \le S_{T_1}$ и $s_{T_2} \le \int\limits_c^b f(x)dx \le S_{T_2}$ следует, что

$$s_T = s_{T_1} + s_{T_2} \le \int\limits_a^c f(x)dx + \int\limits_c^b f(x)dx \le S_{T_1} + S_{T_2} = S_T.$$

Поэтому число $\int\limits_a^c f(x)dx + \int\limits_c^b f(x)dx$ разделяет множество нижних и верхних сумм Дарбу для разбиений отрезка $[a, b]$, включающих точку c. Из свойства 2 сумм Дарбу следует, что тем более оно разделяет множество нижних и верхних сумм Дарбу для *всех* разбиений отрезка $[a, b]$. Поскольку такое число единственно, то оно равно $\int\limits_a^b f(x)dx$, что завершает проверку равенства (4.5) и доказательство теоремы.

В случае, когда $b < c$, имеем

$$\int\limits_a^c f(x)dx = \int\limits_a^b f(x)dx + \int\limits_b^c f(x)dx = \int\limits_a^b f(x)dx - \int\limits_c^b f(x)dx,$$

так что и в этом случае выполняется равенство (4.5). Аналогично рассматривается и случай, когда $c < a$.

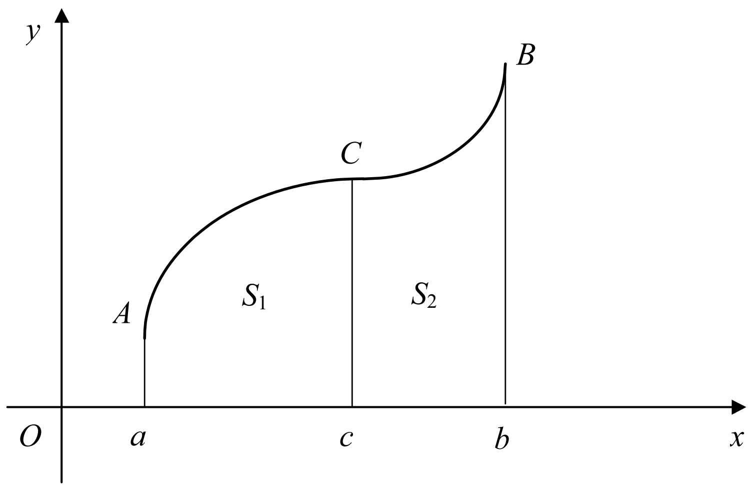

Рис. 4.7

Равенство (4.5) имеет наглядный геометрический смысл, оно выражает свойство *аддитивности площади плоской фигуры*. Так, площадь S криволинейной трапеции $aABb$ на рис. 4.7 равна сумме площадей S_1 трапеции $aACc$ и S_2 трапеции $cCBb$.

Но $S_1 = \int_a^c f(x)dx$, $S_2 = \int_c^b f(x)dx$ и $S = \int_a^b f(x)dx$, что еще раз подтверждает равенство (4.5).

5. Теорема о среднем для определенного интеграла

Теорема 4.9. *Пусть дана функция* $y = f(x)$, *непрерывная на отрезке* $[a, b]$. *Тогда на этом отрезке найдется точка* c, *такая, что выполняется равенство*

$$\int_a^b f(x)dx = f(c)(b-a). \tag{4.6}$$

Число $f(c) = \dfrac{1}{b-a}\int_a^b f(x)dx$ называют *средним значением* функции $f(x)$ на отрезке $[a, b]$.

Доказательство. Так как функция $f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на этом отрезке. Пусть m и M соответственно наименьшее и наибольшее значения функции $f(x)$ на отрезке $[a, b]$. Выражения $m(b - a)$ и $M(b - a)$ являются нижней и верхней суммой Дарбу для разбиения, состоящего только из одного отрезка, а именно из самого отрезка $[a, b]$. Поскольку $\int_a^b f(x)dx$ разделяет эти суммы, то выполняется двойное неравенство

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a).$$

Поделив все части неравенства на $b - a > 0$, получим

$$m \le \frac{1}{b-a}\int_a^b f(x)dx \le M.$$

Число $\frac{1}{b-a}\int\limits_a^b f(x)dx$ находится между наименьшим и наибольшим значениями непрерывной функции. По теореме о промежуточном значении (свойство 2 § 1.10) это значение достигается в некоторой точке c на отрезке $[a, b]$:

$$\frac{1}{b-a}\int\limits_a^b f(x)dx = f(c),$$

откуда получим равенство (4.6), что требовалось доказать.

Геометрический смысл теоремы виден из рис. 4.8. Площадь криволинейной трапеции $aABb$ равна площади прямоугольника с тем же основанием и высотой $f(c)$.

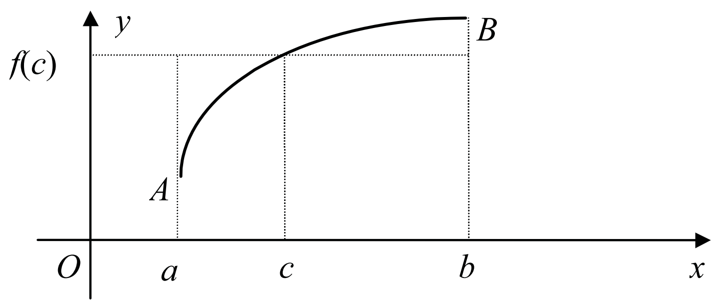

Рис. 4.8

§ 4.5. Формула Ньютона – Лейбница

До сих пор мы еще не дали ответа на важный вопрос: для каких функций существует первообразная. Частичный ответ на этот вопрос будет дан ниже, в п. 1 данного параграфа. Затем мы докажем основную формулу интегрального исчисления, устанавливающую связь между определенным интегралом и первообразной.

1. Интеграл с переменным верхним пределом

Если функция $y = f(x)$ интегрируема на отрезке $[a, b]$, то она интегрируема и на любом меньшем отрезке, и, следовательно, для любого $x \in [a, b]$ существует интеграл $\int_a^x f(x)dx$. Чтобы не смешивать обозначения верхнего предела и переменной интегрирования, будем записывать его в виде $\int_a^x f(t)dt$.

Определение. *Для функции* $y = f(x)$, *интегрируемой на отрезке* $[a, b]$, *интеграл вида*

$$\int_a^x f(t)dt,$$

где $x \in [a, b]$, *называется* ***интегралом с переменным верхним пределом***.

Для каждого $x \in [a, b]$ рассмотрим функцию $\Phi(x) = \int_a^x f(t)dt$. Оказывается, эта функция является первообразной для $y = f(x)$, что мы и докажем в следующей теореме.

Теорема 4.10. *Если функция* $y = f(x)$ *непрерывна на отрезке* $[a, b]$, *то функция* $\Phi(x) = \int_a^x f(t)dt$ *дифференцируема в любой внутренней точке этого отрезка, причем* $\Phi'(x) = f(x)$.

Иными словами, *интеграл с переменным верхним пределом является первообразной для непрерывной подынтегральной функции*.

Доказательство. Будем искать производную функции $\Phi(x) = \int_a^x f(t)dt$, пользуясь непосредственным определением производной, а именно

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

Для $x \in (a, b)$ выберем Δx столь малым, чтобы точка $x + \Delta x$ лежала внутри отрезка $[a, b]$; тогда $\Phi(x + \Delta x) = \int_a^{x+\Delta x} f(t)dt$. Имеем

$$\Phi(x+\Delta x) - \Phi(x) = \int_a^{x+\Delta x} f(t)dt - \int_a^{x} f(t)dt =$$

$$= \int_x^{x+\Delta x} f(t)dt + \int_a^{x} f(t)dt - \int_a^{x} f(t)dt = \int_x^{x+\Delta x} f(t)dt.$$

К последнему интегралу применим теорему о среднем предыдущего параграфа:

$$\Phi(x+\Delta x) - \Phi(x) = \int_x^{x+\Delta x} f(t)dt = f(c)\Delta x,$$

где промежуточная точка c находится между x и $x + \Delta x$, поэтому

$$\frac{\Phi(x+\Delta x) - \Phi(x)}{\Delta x} = \frac{f(c)\Delta x}{\Delta x} = f(c).$$

Так как функция $f(x)$ непрерывна и $c \to x$ при $\Delta x \to 0$, то $\lim\limits_{\Delta x \to 0} f(c) = f(x).$ Поэтому

$$\Phi'(x) = \lim_{\Delta x \to 0} \frac{\Phi(x+\Delta x) - \Phi(x)}{\Delta x} = \lim_{\Delta x \to 0} f(c) = f(x),$$

что и требовалось доказать.

Из этой теоремы вытекает, что функция $f(x)$, непрерывная на отрезке $[a, b]$, имеет на этом отрезке первообразную, а именно функцию

$$\Phi(x) = \int_a^{x} f(t)dt.$$

Поэтому доказанная теорема называется *теоремой о существовании первообразной для непрерывной функции*.

2. Формула Ньютона – Лейбница

Теорема 4.11. *Пусть функция $y = f(x)$ непрерывна на отрезке $[a, b]$ и $F(x)$ – первообразная для $f(x)$. Тогда*

$$\int_a^b f(x)dx = F(b) - F(a). \qquad (4.7)$$

Доказательство. Поскольку функция $f(x)$ непрерывна на отрезке $[a, b]$, то она интегрируема на этом отрезке и имеет на нем первообразную, а именно, функцию $\Phi(x)$ из предыдущего пункта. Сначала проверим справедливость формулы (4.7) для этой первообразной. Действительно, $\Phi(x) = \int_a^x f(t)dt,$ подставляя $x = b$, получим $\Phi(b) = \int_a^b f(t)dt,$ а подставляя $x = a$, получим $\Phi(a) = \int_a^a f(t)dt = 0.$ Поэтому

$$\int_a^b f(x)dx = \Phi(b) - \Phi(a).$$

Если $F(x)$ – другая первообразная для функции $f(x)$, то выполняется равенство

$$F(x) = \Phi(x) + C.$$

Имеем

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

что завершает доказательство теоремы.

Разность $F(b) - F(a)$ часто записывают в виде $F(x)\big|_a^b$, и *формула Ньютона – Лейбница* в этом случае принимает следующий вид:

$$\int_a^b f(x)dx = F(x)\big|_a^b = F(b) - F(a). \tag{4.8}$$

Отметим еще раз, что мы доказали формулу (4.8) для случая, когда $f(x)$ – *непрерывная* на $[a, b]$ функция. В действительности же эта формула справедлива для *любой* функции $f(x)$, имеющей первообразную функцию $F(x)$.

Формулу (4.8), доказанную в теореме 4.11, обычно называют *основной формулой* интегрального исчисления. Она позволяет сводить отыскание определенного интеграла к нахождению первообразной.

Отметим еще два варианта формул (4.7), (4.8):

$$\int_a^b F'(x)dx = F(b) - F(a), \tag{4.9}$$

$$\int_a^b dF(x)dx = F(b) - F(a). \tag{4.10}$$

Теперь нам не составит труда найти определенный интеграл из примера 4.24 предыдущего параграфа. Одной из первообразных функции $y = x^3$ будет функция $F(x) = \dfrac{x^4}{4}$, поэтому по формуле Ньютона – Лейбница получим

$$\int_0^1 x^3 dx = \left.\frac{x^4}{4}\right|_0^1 = \frac{1}{4} - 0 = 0{,}25.$$

3. Свойства определенного интеграла

Из формулы Ньютона – Лейбница легко получаются остальные свойства определенного интеграла. При формулировке этих свойств предполагается, что функции непрерывны на рассматриваемых промежутках.

1. *Интеграл от суммы двух функций* $f_1(x)$ *и* $f_2(x)$ *по отрезку* $[a, b]$ *равен сумме интегралов от этих функций по тому же отрезку*:

$$\int_a^b (f_1(x) + f_2(x))dx = \int_a^b f_1(x)dx + \int_a^b f_2(x)dx.$$

Доказательство. Из свойств неопределенного интеграла следует, что если $F_1(x)$ – первообразная для функции $f_1(x)$, а $F_2(x)$ – первообразная для функции $f_2(x)$, то первообразной для суммы функций $f_1(x) + f_2(x)$ будет служить сумма первообразных $F_1(x) + F_2(x)$. Следовательно,

$$\int_a^b (f_1(x) + f_2(x))dx = (F_1(b) + F_2(b)) - (F_1(a) + F_2(a)) =$$

$$= F_1(b) - F_1(a) + F_2(b) - F_2(a) = \int_a^b f_1(x)dx + \int_a^b f_2(x)dx.$$

Аналогично доказывается следующее свойство.

2. *Постоянный множитель можно выносить из-под знака интеграла*:

$$\int_a^b kf(x)dx = k\int_a^b f(x)dx, \quad k - \text{постоянная.}$$

С помощью этих двух свойств легко найти определенный интеграл от многочлена.

Пример 4.26. Вычислить $\int_{-1}^{2} (6x^2 - 4x + 5)dx.$

Р е ш е н и е . Имеем

$$\int_{-1}^{2}(6x^2-4x+5)dx=\int_{-1}^{2}6x^2dx+\int_{-1}^{2}-4xdx+\int_{-1}^{2}5dx=$$

$$=6\int_{-1}^{2}x^2dx-4\int_{-1}^{2}xdx+5\int_{-1}^{2}dx=6\frac{x^3}{3}\Big|_{-1}^{2}-4\frac{x^2}{2}\Big|_{-1}^{2}+5x\Big|_{-1}^{2}=$$

$$=2(2^3-(-1)^3)-2(2^2-(-1)^2)+5(2-(-1))=27.$$

3. *Интеграл от неотрицательной функции на отрезке* $[a, b]$ – *неотрицательное число*, *т.е. если* $f(x) \geq 0$ *на* $[a, b]$, *то*

$$\int_{a}^{b} f(x)dx \geq 0. \tag{4.11}$$

Действительно, если $f(x) \geq 0$ на $[a, b]$, то любая нижняя сумма Дарбу будет неотрицательной, тем более – значение интеграла.

Неравенство (4.11) имеет очевидный геометрический смысл: *площадь криволинейной трапеции*, *ограниченной сверху графиком неотрицательной функции*, *есть неотрицательное число*.

4. ***Интегрирование неравенств***. *Если на отрезке* $[a, b]$ *выполняется неравенство* $f(x) \leq g(x)$, *то такое же неравенство выполняется и для интегралов*:

$$\int_{a}^{b} f(x)dx \leq \int_{a}^{b} g(x)dx. \tag{4.12}$$

Действительно, если $f(x) \leq g(x)$, то $g(x) - f(x) \geq 0$, тогда согласно свойству 3 имеем $\int_{a}^{b}(g(x)-f(x))dx \geq 0$, т.е. $\int_{a}^{b} g(x)dx - \int_{a}^{b} f(x)dx \geq 0$, откуда следует искомое неравенство.

Геометрический смысл данного утверждения предоставляется выяснить читателю.

5. ***Оценка определенного интеграла.*** *Пусть m – наименьшее, а M – наибольшее значения непрерывной функции $y = f(x)$ на отрезке $[a, b]$. Тогда выполняется двойное неравенство*

$$m(b-a) \le \int_a^b f(x)dx \le M(b-a). \tag{4.13}$$

Действительно, левая и правая части неравенства (4.13) – это соответственно нижняя и верхняя суммы Дарбу для разбиения, состоящего из единственного отрезка $[a, b]$. Поскольку интеграл – это число, разбивающее *любые* нижние и верхние суммы, то неравенство (4.13) становится очевидным.

Пример 4.27. Оценить определенный интеграл $\int_2^4 \frac{dx}{x}$.

Решение. Функция $y = \frac{1}{x}$ убывает на промежутке [2, 4], поэтому $M = 0{,}5$, $m = 0{,}25$. Используя неравенства (4.13), получим

$$0{,}5 = 0{,}25(4-2) \le \int_2^4 \frac{dx}{x} \le 0{,}5 \cdot (4-2) = 1.$$

Взяв полусумму крайних значений, получим значение 0,75, тогда как точное значение интеграла, найденное по теореме Ньютона-Лейбница, равно

$$\ln 2 \approx 0{,}69.$$

4. Интегрирование по частям в определенном интеграле

Для определенного интеграла формула интегрирования по частям принимает следующий вид:

$$\int_a^b u dv = uv\Big|_a^b - \int_a^b v du. \tag{4.14}$$

Действительно, по формуле (4.10) имеем

$$uv\Big|_a^b = \int_a^b d(uv) = \int_a^b vdu + \int_a^b udv.$$

Перенося первое слагаемое налево с противоположным знаком, получим искомую формулу (4.14).

Пример 4.28. Вычислить $\int_1^e x^2 \ln x dx$.

Решение.

$$\int_1^e x^2 \ln x dx = \left| \begin{array}{l} u = \ln x,\ dv = x^2 dx \\ du = \dfrac{dx}{x},\ v = \int x^2 dx = \dfrac{x^3}{3} \end{array} \right| = \ln x \cdot \frac{x^3}{3}\Big|_1^e - \int_1^e \frac{x^3}{3}\frac{dx}{x} =$$

$$= \frac{e^3}{3} - \frac{x^3}{9}\Big|_1^e = \frac{e^3}{3} - \frac{e^3}{9} + \frac{1}{9} = \frac{2e^3}{9} + \frac{1}{9}.$$

Как и в случае неопределенного интеграла, успех в применении формулы (4.14) зависит от правильного выбора множителей u и dv.

5. Замена переменной в определенном интеграле

Теорема 4.12. *Пусть функция $y = f(x)$ непрерывна на отрезке $[a, b]$, а функция $x = \varphi(t)$ определена на отрезке $[\alpha, \beta]$ и имеет непрерывную производную внутри этого отрезка, причем $\varphi(\alpha) = a$, $\varphi(\beta) = b$ и $\varphi([\alpha, \beta]) = [a, b]$. Тогда*

$$\int_a^b f(x) = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt. \tag{4.15}$$

Доказательство. Пусть $F(x)$ – первообразная для функции $f(x)$, тогда $F'(x) = f(x)$ и $F(\varphi(t))' = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t)$. Заметим, что из условий теоремы следует, что последняя функция интегрируема на отрезке $[\alpha, \beta]$. Применяя формулу (4.9), получим

$$\int_{\alpha}^{\beta} f(\varphi(t))\varphi'(t)dt = F(\varphi(\beta)) - F(\varphi(\alpha)) = F(b) - F(a) = \int_{a}^{b} f(x)dx\,,$$

что завершает доказательство теоремы.

Пример 4.29. Вычислить $\int_0^2 \sqrt{4-x^2}\,dx.$

Решение. Применим тригонометрическую замену переменной $x = 2\sin t,\ 0 \le t \le \frac{\pi}{2}$. Легко проверить, что все условия теоремы 4.12 выполнены, так что можно применить формулу (4.15). Преобразуем подынтегральное выражение:

$$\sqrt{4-x^2} = \sqrt{4-4\sin^2 t} = 2\cos t,\ \ dx = 2\cos t dt.$$

Получим

$$\int_0^2 \sqrt{4-x^2}\,dx = \int_0^{\pi/2} 2\cos t \cdot 2\cos t dt = 2\int_0^{\pi/2} (1+\cos 2t)dt =$$

$$= 2\left(t + \frac{\sin 2t}{2}\right)\Bigg|_0^{\pi/2} = 2\frac{\pi}{2} = \pi.$$

Если воспользоваться геометрической интерпретацией определенного интеграла, то этот результат легко получить устно (рис. 4.9).

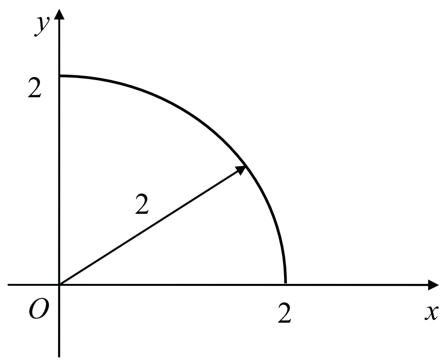

Рис. 4.9

Пример 4.30. Вычислить $\int\limits_4^9 \frac{\sqrt{x}dx}{\sqrt{x}-1}$.

Решение. Чтобы избавиться от иррациональности, выполним замену переменной $x = t^2$, $dx = 2tdt$, $2 \le t \le 3$. При этом все условия теоремы 4.12 выполнены, так что получим

$$\int\limits_4^9 \frac{\sqrt{x}dx}{\sqrt{x}-1} = \int\limits_2^3 \frac{t \cdot 2tdt}{t-1} = 2\int\limits_2^3 \frac{t^2-1+1}{t-1}dt = 2\int\limits_2^3 \left(t+1+\frac{1}{t-1}\right)dt =$$

$$= 2\left(\frac{t^2}{2}+t+\ln|t-1|\right)\Bigg|_2^3 = (9+6+2\ln 2)-(4+4+2\ln 1) = 7+2\ln 2.$$

Отметим некоторое "преимущество" в применении формулы замены переменной по сравнению со случаем неопределенного интеграла, а именно: не надо возвращаться к первоначальной переменной.

§ 4.6. Приложения определенного интеграла

1. Вычисление площадей плоских фигур

В предыдущем параграфе мы рассмотрели понятие определенного интеграла и выяснили, что для непрерывной неотрицательной функции $y = f(x)$, заданной на отрезке $[a, b]$, *определенный интеграл от данной функции по данному отрезку равен площади криволинейной трапеции aABb* (рис. 4.10).

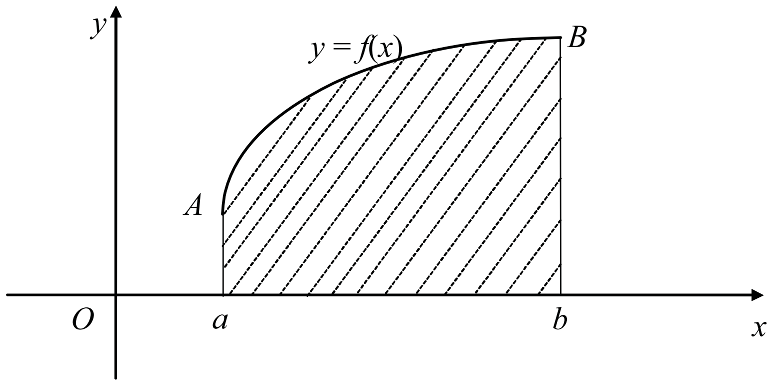

Рис. 4.10

Пример 4.31. Найти площадь криволинейной трапеции, ограниченной графиком функции $y = \sin^2 x$ на отрезке $\left[0, \frac{\pi}{2}\right]$.

Р е ш е н и е. Искомую площадь S находим по формуле

$$S = \int_0^{\frac{\pi}{2}} \sin^2 x dx.$$

Чтобы вычислить этот интеграл, воспользуемся тригонометрической формулой понижения степени $\sin^2 x = \frac{1-\cos 2x}{2}$. Тогда получим

$$S = \int_0^{\frac{\pi}{2}} \frac{1-\cos 2x}{2} dx = \frac{1}{2}\left(\int_0^{\frac{\pi}{2}} dx - \int_0^{\frac{\pi}{2}} \cos 2x dx\right) = \frac{1}{2}\left(x - \frac{\sin 2x}{2}\right)\Bigg|_0^{\frac{\pi}{2}} = \frac{1}{2}\cdot\frac{\pi}{2} = \frac{\pi}{4}.$$

Таким образом, искомая площадь $S = \frac{\pi}{4}$.

Интересно, что мы могли прийти к тому же результату без использования интеграла. Действительно, легко видеть, что S равна площади, ограниченной на том же отрезке функцией $y = \cos^2 x$, поскольку по формуле приведения графики этих функций получаются друг из друга сдвигом вдоль оси Ox. Однако $\sin^2 x + \cos^2 x = 1$, поэтому $S + S = \pi/2$, откуда еще раз получим $S = \pi/4$.

С помощью интеграла можно находить площади плоских фигур более сложного вида, чем криволинейная трапеция.

Рассмотрим фигуру, ограниченную слева прямой $x = a$, справа прямой $x = b$ $(a < b)$, снизу графиком неотрицательной функции $y = f_1(x)$ и сверху графиком функции $y = f_2(x)$ $(f_1(x) \le f_2(x))$ (рис. 4.11, на котором фигура $A_1A_2B_2B_1$ выделена штриховкой).

Площадь фигуры $A_1A_2B_2B_1$ можно вычислить следующим образом:

$$S = S_{aA_2B_2b} - S_{aA_1B_1b} = \int_a^b f_2(x)dx - \int_a^b f_1(x)dx = \int_a^b (f_2(x) - f_1(x))dx. \quad (4.16)$$

Условие неотрицательности функций $f_1(x)$ и $f_2(x)$ можно снять. Действительно, пусть m – наименьшее значение функции $f_1(x)$ на отрезке $[a, b]$.

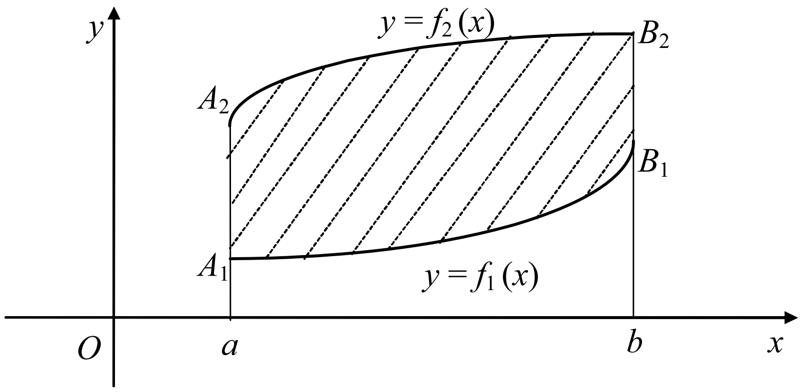

Рис. 4.11

Тогда функции $h_1(x) = f_1(x) + |m|$ и $h_2(x) = f_2(x) + |m|$ непрерывны и неотрицательны на отрезке $[a, b]$, и мы можем применить формулу (4.16):

$$S = \int_a^b (h_2(x) - h_1(x))dx = \int_a^b (f_1(x) + |m| - (f_2(x) + |m|))dx = \int_a^b (f_2(x) - f_1(x))dx.$$

Таким образом, формула (4.16) выполняется и в этом случае. Графическая иллюстрация приведена на рис. 4.12.

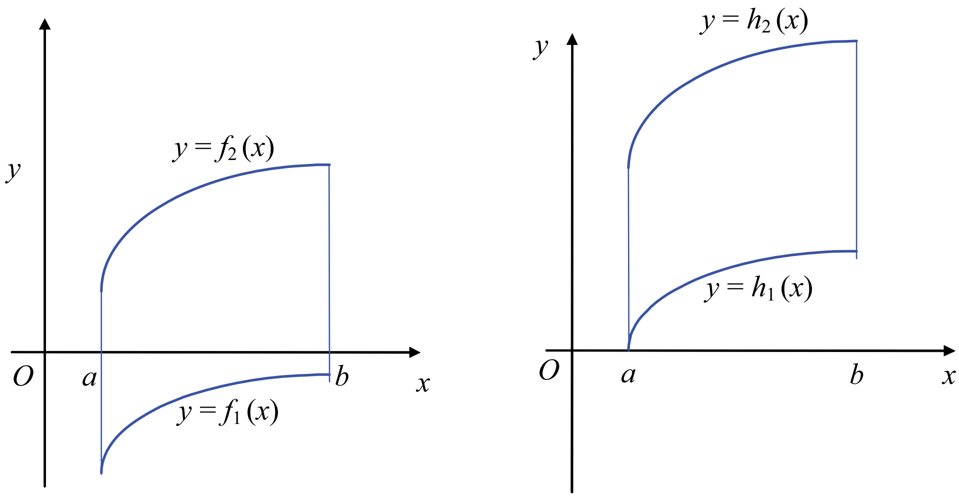

Рис. 4.12

Пример 4.32. Найти площадь фигуры, ограниченной прямой $y = x - 2$ и параболой $y = 4 - x^2$.

Решение. Прежде всего построим фигуру, площадь которой мы ищем (рис. 4.13).

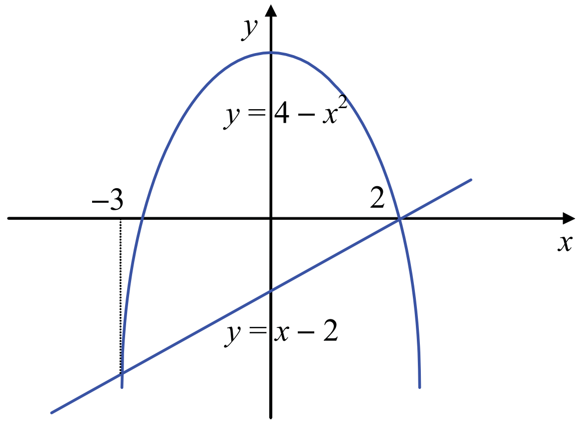

Рис. 4.13

Для этого, в частности, необходимо найти точки пересечения двух данных линий. Имеем $x - 2 = 4 - x^2$, откуда $x^2 + x - 6 = 0$, так что $x_1 = -3$, $x_2 = 2$. Поэтому пределы интегрирования будут $a = -3$, $b = 2$. Таким образом,

$$S = \int_{-3}^{2}\left(4 - x^2 - (x-2)\right)dx = \int_{-3}^{2}\left(6 - x - x^2\right)dx = 6x - \frac{x^2}{2} - \frac{x^3}{3}\Bigg|_{-3}^{2} =$$

$$= 12 - 2 - \frac{8}{3} - \left(-18 - \frac{9}{2} + 9\right) = 7\frac{1}{3} + 13\frac{1}{2} = 20\frac{5}{6}.$$

2. Вычисление объема тела вращения

Оказывается, что с помощью определенного интеграла можно вычислять не только площади, но и объемы фигур. Пусть функция $y = f(x)$ непрерывна и неотрицательна на отрезке $[a, b]$. Тогда тело, которое образуется при вращении вокруг оси Ox криволинейной трапеции $aABb$ (рис. 4.14), имеет объем

$$V = \pi\int_{a}^{b} f^2(x)dx. \qquad (4.17)$$

Чтобы доказать формулу (4.17), выполним разбиение T отрезка $[a, b]$ точками $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$. На каждом отрезке разбиения $[x_{i-1}, x_i]$ построим прямоугольник $MNQP$ (рис. 4.14), высота которого $MP = m_i$ равна наименьшему значению функции $y = f(x)$ на указанном отрезке. При вращении вокруг оси Ox прямоугольник $MNQP$ опишет цилиндр, объем которого будет равен $v_i = \pi m_i^2 \Delta x_i$, где $\Delta x_i = x_i - x_{i-1}$.

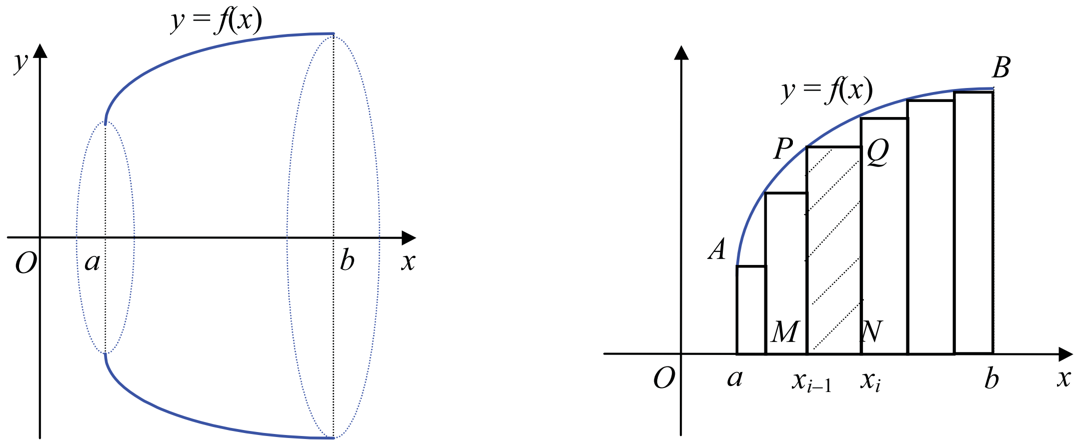

Рис. 4.14

Сумма объемов всех таких цилиндров для данного разбиения запишется в виде

$$v_T = \pi \sum_{i=1}^{n} m_i^2 \Delta x_i.$$

Если бы мы вместо наименьшего значения функции $f(x)$ брали наибольшее значение, то получили бы сумму

$$V_T = \pi \sum_{i=1}^{n} M_i^2 \Delta x_i.$$

С одной стороны, ясно, что при измельчении разбиения T обе эти суммы дают приближенное значение искомого объема V, а с другой стороны, эти суммы являются нижней и верхней суммой Дарбу для функции $y = \pi f^2(x)$. Последняя функция непрерывна, поскольку непрерывна функция $y = f(x)$ и, следовательно, интегрируема, откуда и следует формула (4.17).

П р и м е р 4.33. Окружность единичного радиуса с центром в точке $O(0, 2)$ вращается вокруг оси Ox. Найти объем тела вращения. (Это тело называется *тором.*)

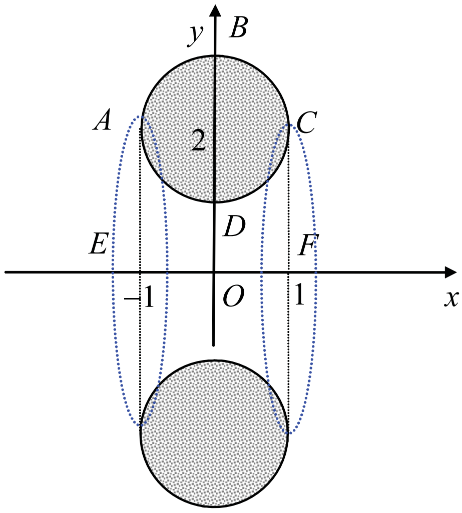

Рис. 4.15

Р е ш е н и е. Уравнение окружности имеет вид $x^2 + (y - 2)^2 = 1$. Выражая y через x, получим уравнение дуги ABC (рис. 4.15)

$$y = f_1(x) = 2 + \sqrt{1 - x^2}$$

и уравнение дуги ADC

$$y = f_2(x) = 2 - \sqrt{1 - x^2}.$$

Объем тора можно представить в виде разности объемов тел, полученных от вращения криволинейных трапеций $EABCF$ и $EADCF$. Используя формулу (4.17), получим

$$V = \pi \int_{-1}^{1} f_1^2(x)dx - \pi \int_{-1}^{1} f_2^2(x)dx = \pi \int_{-1}^{1} \left(f_1^2(x) - f_2^2(x)\right)dx =$$

$$= \pi \int_{-1}^{1} 4 \cdot 2\sqrt{1 - x^2}\, dx = 8\pi \int_{-1}^{1} \sqrt{1 - x^2}\, dx.$$

Похожий интеграл мы уже рассматривали в качестве примера 4.29 предыдущего параграфа и выяснили, что он дает площадь полукруга единичного радиуса, т.е. $\pi/2$. Подставляя это выражение в последний интеграл, получим искомый объем $V = 4\pi^2$.

3. Экономические приложения определенного интеграла

В курсе микроэкономики часто рассматривают так называемые *предельные величины*, т.е. для данной величины, представляемой некоторой функцией $y = f(x)$, рассматривают ее производную $f'(x)$. Например, если дана функция издержек C в зависимости от объема q выпускаемого товара $C = C(q)$, то *предельные издержки* будут задаваться производной этой функции $MC = C'(q)$. Ее экономический смысл – это издержки на производство дополнительной единицы выпускаемого товара. Поэтому часто приходится находить функцию издержек по данной функции предельных издержек.

Пример 4.34. Дана функция предельных издержек $MC = 3q^2 - 48q + 202$, $1 \le q \le 20$. Найти функцию издержек $C = C(q)$ и вычислить издержки в случае производства 10 единиц товара, если известно, что издержки для производства первой единицы товара составили 50 руб.

Решение. Функцию издержек находим интегрированием:

$$C(q) = \int_1^q MC dq + C_0,$$

где константа C_0 находится из данного условия $C(1) = 50$, так что $C_0 = 50$, поскольку интеграл обращается в нуль. Интегрируя, получим функцию издержек

$$C(q) = q^3 - 24q^2 + 202q + 50.$$

Подставляя $q = 10$ в полученную формулу, находим искомое значение

$$C(10) = 670.$$

Еще одним примером приложения определенного интеграла является нахождение *дисконтированной стоимости денежного потока*.

Допустим вначале, что для каждого дискретного момента времени $t = 1, 2, 3, \ldots$ задана величина денежного потока $R(t)$. Если ставку процента обозначить через p, то дисконтированную стоимость каждой из величин $R(1)$, $R(2)$, $R(3)$, ... найдем по известным формулам:

$$R(1)(1+p)^{-1},\ R(2)(1+p)^{-2},\ R(3)(1+p)^{-3}, \ldots .$$

Тогда дисконтированную стоимость денежного потока найдем, суммируя эти величины:

$$\Pi = \sum_{t=1}^{n} R(t)(1+p)^{-t}, \tag{4.18}$$

где n – общее число периодов времени.

В непрерывной модели время изменяется непрерывно, т.е. для каждого момента времени $0 \le t \le T$, где $[0, T]$ – рассматриваемый период времени, задана величина $I(t)$ – скорость изменения денежного потока (т.е. величина денежного потока за промежуток времени от t до $t + dt$ приближенно равна $I(t)dt$). Для получения величины П изменим формулу (4.18), а именно, знак суммирования заменим на знак определенного интеграла, формулы вычисления дисконтированной стоимости в дискретном случае заменим на их непрерывный аналог, и тогда формула (4.18) примет следующий вид:

$$\Pi = \int_0^T I(t)e^{-pt}dt. \tag{4.19}$$

Пример 4.35. Под строительство гидроэлектростанции задан непрерывный денежный поток со скоростью $I(t) = -t^2 + 20t + 5$ (млрд руб./год) в течение 20 лет с годовой процентной ставкой $p = 5\%$. Найти дисконтированную стоимость этого потока.

Решение. По формуле (4.19) имеем

$$\Pi = \int_0^{20} \left(-t^2 + 20t + 5\right)e^{-0{,}05t}dt.$$

Чтобы вычислить этот интеграл, выполним сначала замену переменной:

$$s = -0{,}05t, \quad t = -20s, \quad dt = -20ds.$$

При этом новые пределы интегрирования получаются подстановкой старых пределов в формулу замены: $s_0 = 0$, $s_1 = -1$. Имеем

$$\Pi = -20\int_0^{-1}\left(-400s^2 - 400s + 5\right)e^s ds = 20\int_{-1}^{0}\left(-400s^2 - 400s + 5\right)e^s ds.$$

К последнему интегралу применим формулу интегрирования по частям, полагая $u = -400s^2 - 400s + 5$, $du = (-800s - 400)ds$, $dv = e^s ds$, $v = e^s$. Поэтому

$$\Pi = 20\left(\left(-400s^2 - 400s + 5\right)e^s\Big|_{-1}^{0} + \int_{-1}^{0} e^s(800s + 400)ds\right).$$

В первом слагаемом подставим пределы интегрирования, а ко второму слагаемому еще раз применим формулу интегрирования по частям, полагая $u = 800s + 400$, $du = 800ds$. Имеем

$$\Pi = 20\left(5 - 5e^{-1} + (800s + 400)e^s\Big|_{-1}^{0} - \int_{-1}^{0} 800e^s ds\right) =$$

$$= 20(5 - 5e - 1 + 400 + (800 - 400)e \quad - 1 - 800 + 800e - 1) =$$

$$= 20(1195e - 1 - 395).$$

Окончательно получим $\Pi = 892$ (млрд руб.).

Далее рассмотрим некоторую модель экономического роста, предложенную Е.Д. Домаром. Основные допущения этой модели сформулированы ниже.

1. Всякое изменение величины скорости денежного потока $I(t)$ влияет как на совокупный спрос, так и на изменение объема производства.

2. Скорость изменения величины спроса $Y(t)$ пропорциональна производной скорости денежного потока с коэффициентом про-

порциональности $k = 1/s$, где s – предельная величина накопления. Это предположение можно записать в виде уравнения

$$\frac{dY}{dt} = \frac{1}{s}\frac{dI}{dt}. \tag{4.20}$$

3. Экономический потенциал κ (т.е. величина стоимости товаров, которые можно произвести) пропорционален объему оборотных средств K с коэффициентом пропорциональности ρ, т.е. $\kappa = \rho K$. Дифференцируя по t, получим

$$\frac{d\kappa}{dt} = \rho\frac{dK}{dt} = \rho I. \tag{4.21}$$

В модели Домара предполагается, что весь экономический потенциал полностью используется, иными словами, $Y = \kappa$. Дифференцируя по t, получим

$$\frac{dY}{dt} = \frac{d\kappa}{dt}. \tag{4.22}$$

Подставляя (4.20) и (4.21) в (4.22), имеем

$$\frac{1}{s}\frac{dI}{dt} = \rho I, \quad \frac{dI}{I} = \rho s dt. \tag{4.23}$$

Чтобы найти функцию $I(t)$ из уравнения (4.23), проинтегрируем обе части последнего равенства по t от 0 до t. Получим

$$\int_0^t \frac{dI}{I} = \int_0^t \rho s dt, \quad \text{или} \quad \ln|I(t)|\Big|_0^t = \rho st\Big|_0^t,$$

откуда $\ln|I| = \ln|I(0)| + \rho st$. Потенцируя последнее равенство, получим окончательное выражение для $I(t)$:

$$I(t) = I(0)e^{\rho st}, \tag{4.24}$$

где $I(0)$ – это скорость денежного потока в начальный момент времени.

Таким образом, для того чтобы поддерживать равновесие между объемом производимых благ и совокупным спросом на них, скорость денежного потока должна расти с экспоненциальной скоростью согласно формуле (4.24).

Модель Домара – это типичный пример модели роста, записываемой в виде одного или нескольких уравнений, в которые входят производные неизвестных величин. Такие уравнения называются *дифференциальными*, они будут рассмотрены ниже в этой книге. Существенным компонентом решения таких уравнений является применение интегрирования.

§ 4.7. Несобственные интегралы

Вводя понятие определенного интеграла, мы предполагали, что отрезок интегрирования конечен, а подынтегральная функция ограничена на этом отрезке. В противном случае множество сумм Дарбу не будет ограниченным. Однако возможны случаи, когда одно или оба этих условия не выполняются. В таком случае соответствующие интегралы называются *несобственными*. Будем различать два случая: промежуток интегрирования бесконечен или подынтегральная функция не ограничена.

1. Несобственные интегралы с бесконечными пределами интегрирования

Начнем со случая, когда промежутком интегрирования является луч $[a, +\infty)$.

Определение. *Пусть функция $y = f(x)$ интегрируема на каждом конечном отрезке $[a, b]$ $(b > a)$. Тогда за **несобственный интеграл** $\int\limits_a^{+\infty} f(x)dx$ принимают предел функции $I(b) = \int\limits_a^b f(x)dx$, когда b стремится к бесконечности*:

$$\int\limits_a^{+\infty} f(x)dx = \lim_{b\to+\infty}\int\limits_a^b f(x)dx. \qquad (4.25)$$

Если предел $\lim\limits_{b\to+\infty}\int\limits_a^b f(x)dx$ существует и конечен, то говорят, что несобственный интеграл *сходится*.

Если предел в правой части (4.25) не существует, то несобственный интеграл *расходится*.

По аналогии с определением (4.25) можно рассмотреть несобственный интеграл с бесконечным нижним пределом, а именно,

$$\int_{-\infty}^{b} f(x)dx = \lim_{a\to-\infty}\int_{a}^{b} f(x)dx. \tag{4.26}$$

Наконец, можно рассмотреть несобственный интеграл с бесконечными нижним и верхним пределами $\int_{-\infty}^{\infty} f(x)dx.$ Для этого возьмем произвольную точку c. Она разобьет числовую ось на два луча $(-\infty, c]$ и $[c, +\infty)$. Если существуют несобственные интегралы $\int_{-\infty}^{c} f(x)dx$ и $\int_{c}^{+\infty} f(x)dx,$ то говорят, что существует и несобственный интеграл $\int_{-\infty}^{\infty} f(x)dx.$ В этом случае полагают

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{c} f(x)dx + \int_{c}^{+\infty} f(x)dx. \tag{4.27}$$

Нетрудно показать, что правая часть формулы (4.27) не зависит от выбора промежуточной точки c.

Геометрический смысл несобственного интеграла. Пусть дана неотрицательная функция $y = f(x)$, непрерывная на луче $[a, +\infty)$. Для каждого $b > a$ определенный интеграл $\int_{a}^{b} f(x)dx$ дает площадь криволинейной трапеции $aABb$. Мысленно перемещая отрезок Bb вправо, мы получим в качестве значения несобственного интеграла $\int_{a}^{+\infty} f(x)dx$ площадь “треугольника” $aA\infty$ (рис. 4.16).

П р и м е р 4.36. Найти несобственный интеграл $\int\limits_1^{\infty}\frac{dx}{x^2}$.

Р е ш е н и е . По определению имеем

$$\int\limits_1^{\infty}\frac{dx}{x^2}=\lim_{b\to\infty}\int\limits_1^{b}x^{-2}dx=\lim_{b\to\infty}-\frac{1}{x}\Big|_1^b=\lim_{b\to\infty}\left(-\frac{1}{b}+1\right)=1.$$

Таким образом, несобственный интеграл сходится и равен 1.

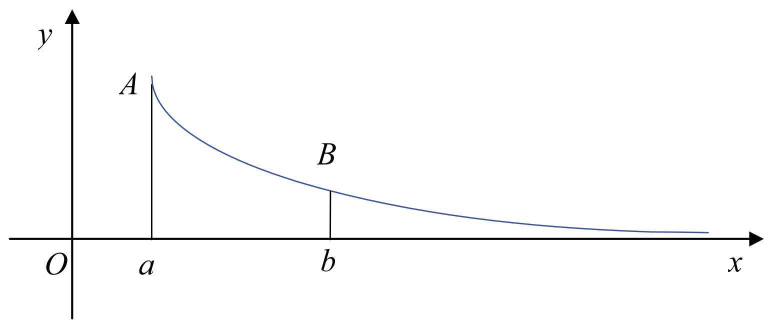

Рис. 4.16

П р и м е р 4.37. Найти несобственный интеграл $\int\limits_1^{\infty}\frac{dx}{\sqrt{x}}$.

Р е ш е н и е . Имеем

$$\int\limits_1^{\infty}\frac{dx}{\sqrt{x}}=\lim_{b\to\infty}\int\limits_1^{b}x^{-\frac{1}{2}}dx=\lim_{b\to\infty}2\sqrt{x}\Big|_1^b=\lim_{b\to\infty}\left(2\sqrt{b}-2\right)=\infty.$$

В данном случае несобственный интеграл расходится (предел равен бесконечности).

П р и м е р 4.38. Найти несобственный интеграл $\int\limits_0^{\infty}\cos 2x dx$.

Р е ш е н и е . В данном случае

$$\int\limits_0^{\infty}\cos 2x dx=\lim_{b\to\infty}\int\limits_0^{b}\cos 2x dx=\lim_{b\to\infty}\frac{\sin 2x}{2}\Bigg|_0^b=\lim_{b\to\infty}\frac{\sin 2b}{2}x$$

и интеграл расходится, поскольку данный предел не существует.

Пример 4.39. Рассмотрим ситуацию, когда денежный поток не прекращается никогда, например в случае эксплуатации земельного участка. Если r – непрерывная процентная ставка, а $R(t)$ – соответствующая рента, то нахождение дисконтированной стоимости земельного участка приводит к формуле, включающей несобственный интеграл

$$\Pi = \int_0^{\infty} R(t) e^{-rt} dt. \qquad (4.28)$$

Пусть $R(t) = 5e^{-0,7t}$ (млн руб./год) – рента, получаемая от земельного участка, $r = 10\%$ – процентная ставка. Определим дисконтированную стоимость земельного участка по формуле (4.28):

$$\Pi = \int_0^{\infty} 5e^{-0,7t} \cdot e^{-0,1t} dt = 5\int_0^{\infty} e^{-0,8t} dt = 5\left(\left. \frac{e^{-0,8t}}{-0,8} \right|_0^{\infty} \right) = \frac{5}{0,8} = 6,25 \text{ (млн руб.)}.$$

Интересно сравнить эту величину со стоимостью участка в данный момент времени, которая равна $R(0) = 5$ (млн руб.).

2. Несобственные интегралы от неограниченных функций

Пусть функция $y = f(x)$ определена на промежутке $[a, b)$. Точку $x = b$ будем называть *особой*, если функция $f(x)$ неограничена в любой окрестности этой точки, но ограничена на любом отрезке, заключенном в промежутке $[a, b)$ (рис. 4.17).

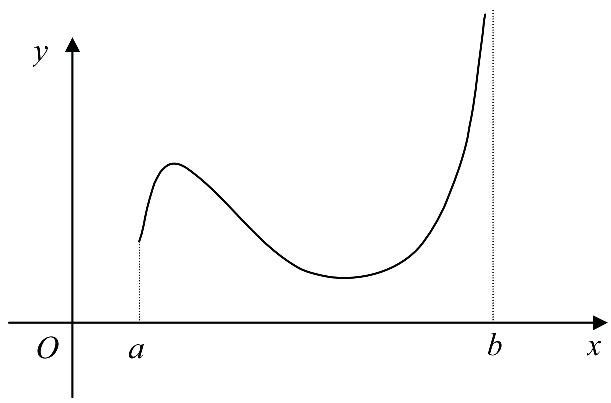

Рис. 4.17

Определение. *Пусть функция* $y = f(x)$ *не ограничена на отрезке* $[a, b]$, *однако интегрируема на любом меньшем отрезке* $[a, b - \varepsilon]$, *где* $\varepsilon > 0$. *Тогда если существует предел* $\lim\limits_{\varepsilon \to 0+0} \int\limits_a^{b-\varepsilon} f(x)dx$, *его принимают за* ***несобственный интеграл от неограниченной функции*** $f(x)$:

$$\int\limits_a^b f(x)dx = \lim_{\varepsilon \to 0+0} \int\limits_a^{b-\varepsilon} f(x)dx. \quad (4.29)$$

Если предел в правой части (4.29) конечен, то говорят, что несобственный интеграл *сходится*. Если же предел не существует или бесконечен, то несобственный интеграл *расходится*.

Аналогично если $x = a$ – особая точка, то несобственный интеграл определяют так:

$$\int\limits_a^b f(x)dx = \lim_{\varepsilon \to 0+0} \int\limits_{a+\varepsilon}^{b} f(x)dx. \quad (4.30)$$

Если же $x = c$ – *единственная* внутренняя особая точка на отрезке $[a, b]$, то полагают

$$\int\limits_a^b f(x)dx = \int\limits_a^c f(x)dx + \int\limits_c^b f(x)dx \quad (4.31)$$

при условии, что оба несобственных интеграла справа сходятся.

Если особых точек на отрезке $[a, b]$ несколько, то отрезок разбивают таким образом, чтобы в каждом отрезке разбиения было не более одной особой точки, и используют определение (4.31).

Пусть $F(x)$ – первообразная для функции $f(x)$. Положим

$$F(a + 0) = \lim_{\varepsilon \to 0+0} F(a + \varepsilon), \quad F(b - 0) = \lim_{\varepsilon \to 0+0} F(b - \varepsilon),$$

если эти пределы существуют.

Аналогом формулы Ньютона – Лейбница для сходящихся несобственных интегралов, у которых особыми точками являются только $x = a$ и $x = b$, будет следующая формула:

$$\int_a^b f(x)dx = F(b-0) - F(a+0). \tag{4.32}$$

В случае непрерывной функции $F(x)$ получим формулу, по виду полностью совпадающую с формулой Ньютона – Лейбница:

$$\int_a^b f(x)dx = F(b) - F(a). \tag{4.33}$$

Пример 4.40. Найти интеграл $\int_0^1 \frac{dx}{\sqrt{x}}$.

Решение. Подынтегральная функция $f(x) = \frac{1}{\sqrt{x}}$ имеет единственную особую точку $x = 0$ на отрезке интегрирования [0; 1]. Первообразной для этой функции будет $F(x) = 2\sqrt{x}$, которая непрерывна на этом отрезке. По формуле (4.33) имеем

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_0^1 = 2 - 0 = 2.$$

Таким образом, данный несобственный интеграл сходится и равен 2.

Пример 4.41. Найти интеграл $\int_0^1 \frac{dx}{x}$.

Решение. Подынтегральная функция $f(x) = \frac{1}{x}$ имеет на промежутке интегрирования [0; 1] единственную особую точку $x = 0$. Первообразной для нее будет $F(x) = \ln x$. Поэтому имеем

$$\int_0^1 \frac{dx}{x} = \lim_{\varepsilon \to 0+0} \int_\varepsilon^1 \frac{dx}{x} = \lim_{\varepsilon \to 0+0} \ln x\Big|_\varepsilon^1 = \lim_{\varepsilon \to 0+0} (\ln 1 - \ln \varepsilon) = -\lim_{\varepsilon \to 0+0} \ln \varepsilon = +\infty.$$

Таким образом, данный несобственный интеграл расходится.

Имеется тесная связь между несобственными интегралами с бесконечными пределами и интегралами от неограниченных функций. Покажем это на предыдущем примере. Выполним замену переменной

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

Верхний предел интегрирования не изменится, а нижний предел интегрирования будет равен $+\infty$, поскольку $\lim\limits_{x \to 0+0} \frac{1}{x} = +\infty$. Таким образом,

$$\int_0^1 \frac{dx}{x} = \int_{+\infty}^1 -\frac{tdt}{t^2} = \int_1^{+\infty} \frac{dt}{t} = \int_1^{+\infty} \frac{dx}{x}.$$

Поэтому последний несобственный интеграл с бесконечным верхним пределом также расходится.

3. Признаки сходимости несобственных интегралов

До сих пор мы исследовали сходимость несобственных интегралов, исходя из определения. В некоторых случаях удобно действовать по-другому. Рассмотрим вначале случай несобственных интегралов с бесконечным верхним пределом от неотрицательной функции. Итак, пусть рассматривается

$$I = \int_a^{\infty} f(x)dx, \tag{4.34}$$

где $f(x) \geq 0$ для $x \geq a$.

Лемма. *Несобственный интеграл* (4.34) *либо сходится*, *либо равен* $+\infty$.

Доказательство. Напомним, что согласно определению (4.27) сходимость интеграла (4.34) зависит от существования предела функции $I(b)=\int\limits_a^b f(x)dx$ при $b\to\infty$. Функция $I(b)$ не убывает при $b > a$, поскольку если $b_2 > b_1$, то

$$\int\limits_a^{b_2} f(x)dx=\int\limits_a^{b_1} f(x)dx+\int\limits_{b_1}^{b_2} f(x)dx\ge\int\limits_a^{b_1} f(x)dx.$$

Всякая такая функция либо неограничена при $b\to\infty$, либо ограничена и тогда в силу монотонности имеет предел. Это заканчивает доказательство леммы.

Теорема 4.13 (***признак сравнения***). *Пусть даны функции $f(x)$ и $g(x)$, интегрируемые на любом отрезке $[a, b]$, причем для любого $x > a$ выполняются неравенства*

$$0\le f(x)\le g(x). \tag{4.35}$$

Тогда если сходится несобственный интеграл от «большей» функции $\int\limits_a^\infty g(x)dx$, то сходится интеграл от «меньшей» функции $\int\limits_a^\infty f(x)dx$. И наоборот, если интеграл от меньшей функции расходится, то расходится и интеграл от большей функции.

Доказательство. Для любой функции $y = f(x)$, интегрируемой на каждом отрезке $[a, b]$, обозначим

$$I_f(b)=\int\limits_a^b f(x)dx.$$

Пусть несобственный интеграл $\int\limits_a^\infty g(x)dx$ сходится. Применим лемму к функции $I_{g-f}(b)$. Если существует конечный предел $\lim\limits_{b\to\infty} I_{g-f}(b)$, а по условию существует конечный предел $\lim\limits_{b\to\infty} I_g(b)$, то из свойств пределов существует конечный предел $\lim\limits_{b\to\infty} I_f(b)$, т.е. интеграл от меньшей функции сходится.

Напротив, предположим, что $\lim\limits_{b\to\infty} I_{g-f}(b)=\infty$. Тогда существует последовательность $b_1<b_2<\ldots<b_n<\ldots$, для которой $I_{g-f}(b_n)>n$. Однако $I_{g-f}(b_n)\le I_g(b_n)$, так что для этой последовательности подавно $I_g(b_n)>n$. Это противоречит предположению о сходимости несобственного интеграла от функции g. Доказательство теоремы закончено.

Пример 4.42. Исследовать на сходимость несобственный интеграл

$$\int\limits_1^\infty \frac{dx}{x^2\sqrt{1+x}}.$$

Решение. Для $x\ge 1$ выполняется неравенство

$$\frac{1}{x^2\sqrt{1+x}}\le\frac{1}{x^2}.$$

Нетрудно проверить, что интеграл от большей функции сходится. Поскольку выполнены все условия теоремы 4.13, то данный интеграл сходится.

Данный признак может быть полезен для исследования несобственных интегралов на сходимость для функций, принимающих значения разных знаков.

Определение. *Несобственный интеграл* (4.34) *сходится* ***абсолютно***, *если сходится интеграл от модуля функции*

$$\int\limits_a^\infty |f(x)|dx. \tag{4.36}$$

Заметим, что в определении абсолютной сходимости мы ничего не говорили о сходимости интеграла (4.34). Это предположение является излишним, как показывает следующая теорема.

Теорема 4.14. *Если* $f(x)$ *интегрируема на любом отрезке* $[a, b]$, $a < b$, *и интеграл* (4.34) *сходится абсолютно*, *то он сходится*.

Доказательство. Вместе с данной функцией $f(x)$ рассмотрим две функции:

$$f_+(x)=\begin{cases} f(x), & f(x)\geq 0, \\ 0, & f(x)<0, \end{cases} \qquad f_-(x)=\begin{cases} 0, & f(x)\geq 0, \\ -f(x), & f(x)<0. \end{cases}$$

Ясно, что обе эти функции принимают только неотрицательные значения. Выполняются тождества

$$|f(x)|=f_+(x)+f_-(x),\ f(x)=f_+(x)-f_-(x).$$

Поэтому если несобственный интеграл для $|f(x)|$ сходится, то он сходится по признаку сравнения и для каждого слагаемого $f_+(x)$, $f_-(x)$. Поскольку $f(x)$ есть разность двух интегрируемых[1] функций, то несобственный интеграл сходится и для нее.

Пример 4.43. Исследовать на сходимость несобственный интеграл

$$\int_1^{\infty}\cos(x^2)dx.$$

Решение. Выполним замену переменной:

$$t=x^2,\ x=\sqrt{t},\ dx=\frac{dt}{2\sqrt{t}},$$

при этом пределы интегрирования останутся прежними. Получим

$$\frac{1}{2}\int_1^{\infty}\frac{\cos t dt}{\sqrt{t}}.$$

---

[1] Интегрируемость $f_+(x)$ и $f_-(x)$ требует доказательства, см. теорему 6.9.

Выполним интегрирование по частям, полагая

$$u = \frac{1}{\sqrt{t}},\ du = -\frac{dt}{2t^{3/2}},\ dv = \cos t\,dt,\ v = \sin t.$$

Получим выражение

$$\frac{1}{2}\left(\frac{\sin t}{\sqrt{t}}\Bigg|_1^\infty + \frac{1}{2}\int_1^\infty \frac{\sin t}{t^{3/2}}\,dt \right).$$

Нетрудно видеть, что предел первого слагаемого в скобках равен $-\sin 1$. Поэтому сходимость данного интеграла зависит от сходимости интеграла

$$\int_1^\infty \frac{\sin t}{t^{3/2}}\,dt.$$

Последний интеграл сходится абсолютно, поскольку $|\sin t| \le 1$, а интеграл $\int_1^\infty \frac{dt}{t^{3/2}}$ сходится. Таким образом, исходный интеграл сходится. Можно найти его приближенное значение, равное –0,28.

§ 4.8. Приближенное вычисление определенных интегралов

В том случае, когда первообразная подынтегральной функции известна, например, является элементарной функцией, определенный интеграл вычисляют по формуле Ньютона – Лейбница. Однако часто возникает ситуация, когда первообразная либо не выражается в виде элементарной функции, либо выражение для первообразной выглядит слишком сложно. В этом случае применяют приближенные формулы вычисления определенных интегралов. Мы познакомимся с двумя из них – *формулой прямоугольников и формулой Симпсона*.

Основная идея получения этих формул состоит в замене подынтегральной функции на функцию более простого вида, например, многочлен, интеграл от которого находят непосредственно по формуле Ньютона – Лейбница.

1. Формула прямоугольников

Пусть требуется вычислить определенный интеграл $\int_a^b f(x)dx$, где $f(x)$ – интегрируемая функция на отрезке $[a, b]$. Разобьем отрезок интегрирования на $2n$ равных частей:

$$a = x_0 < x_1 < x_2 < \ldots < x_{2k-2} < x_{2k-1} < x_{2k} < \ldots < x_{2n-2} < x_{2n-1} < x_{2n} = b.$$

На каждом отрезке $[x_{2k-2}, x_{2k}]$ как на основании cd построим прямоугольник высотой $f(x_{2k-1})$ – $cCDd$ (рис. 4.18). Площадь криволинейной трапеции $aABb$ заменим на сумму площадей всех таких прямоугольников. Таким образом, определенный интеграл приближенно равен следующей сумме:

$$\int_a^b f(x)dx \approx f(x_1)\cdot 2h + f(x_1)\cdot 2h + \ldots + f(x_{2n-1})\cdot 2h =$$

$$= 2h(f(x_1) + f(x_3) + \ldots + f(x_{2n-1})). \qquad (4.37)$$

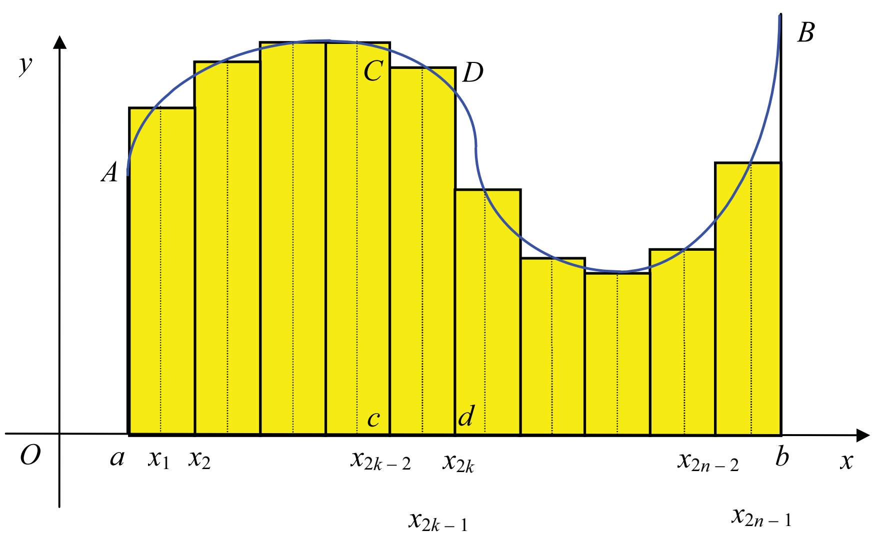

Рис. 4.18

Здесь h – длина отрезка разбиения, $h = \frac{b-a}{2n}$, абсциссы x_{2k-1}, $1 \le k \le n$, вычисляют последовательно:

$$x_1 = a + h, \quad x_{2k+1} = x_{2k-1} + 2h, \quad k = 1, 2, ..., n-1.$$

Формулу (4.37) называют *формулой прямоугольников*.

Пример 4.44. Вычислить с точностью до 10^{-3} определенный интеграл $\int\limits_{-1}^{1} e^{-x^2}\,dx$, используя формулу прямоугольников.

Решение. Полагаем $f(x) = e^{-x^2}$, $a = -1$, $b = 1$, тогда $h = \frac{1}{n}$. Расчетные формулы выглядят следующим образом:

$$x_1 = a + h;\ s_0 = 0;\ 1 \le i \le n, x_{i+1} = x_i + 2h, s_i = s_{i-1} + f(x_i);\ S_n = 2h \cdot s_n.$$

Результаты вычислений приведены в следующей таблице.

n	2	4	8	16	32	64	128
S_n	2	1,558	1,509	1,497	1,495	1,494	1,494

Как видно из таблицы, значение интеграла с заданной точностью получается уже на шестом шаге ($n = 64$).

Доказано, что погрешность применения формулы (4.37) не больше чем

$$\Delta = \frac{(b-a)^2}{24n^2} M,$$

где M – максимум значений второй производной на данном отрезке. Обычно вычисления заканчивают, когда модуль разности $|S_{2n} - S_n|$ становится меньше заданной точности.

2. Формула Симпсона

Для получения приближенной формулы вычисления определенного интеграла более высокой точности, чем формула прямоугольников, будем вписывать вместо прямолинейных отрезков куски парабол.

Лемма 1. *Через любые три точки* $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$ *с различными абсциссами можно провести единственную параболу*

$$y = ax^2 + bx + c. \tag{4.38}$$

Доказательство. Для нахождения неизвестных коэффициентов a, b, c подставим в уравнение (4.38) координаты данных точек M_1, M_2, M_3. Получим систему уравнений:

$$\begin{cases} ax_1^2 + bx_1 + c = y_1, \\ ax_2^2 + bx_2 + c = y_2, \\ ax_3^2 + bx_3 + c = y_3. \end{cases}$$

Определитель этой системы

$$\Delta = \begin{vmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ x_3^2 & x_3 & 1 \end{vmatrix} = (x_2 - x_1)(x_3 - x_2)(x_3 - x_1)$$

не равен нулю, поскольку все числа x_1, x_2, x_3 по условию различны. Поэтому эта система имеет единственное решение, что и требовалось доказать.

Лемма 2. *Площадь* s *криволинейной трапеции, ограниченной параболой* $y = ax^2 + bx + c$, *проходящей через точки* $M_1(-h, y_1)$, $M_2(0, y_2)$, $M_3(h, y_3)$ (рис. 4.19), *выражается формулой*

$$s = \frac{h}{3}(y_1 + 4y_2 + y_3). \tag{4.39}$$

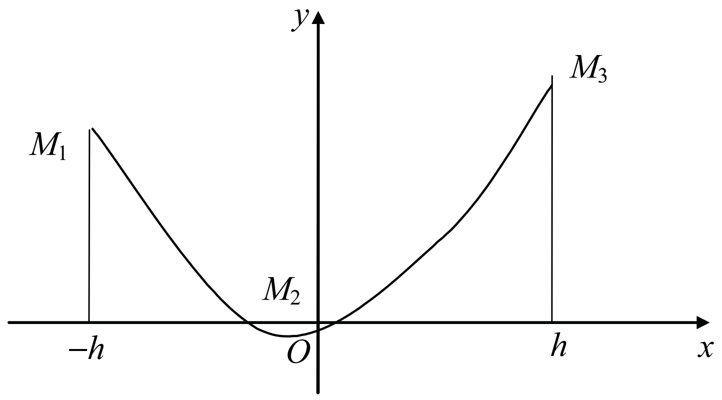

Рис. 4.19

Д о к а з а т е л ь с т в о. В этом случае из (4.38) следует

$$\begin{cases} ah^2 - bh + c = y_1, \\ c = y_2, \\ ah^2 + bh + c = y_3, \end{cases}$$

откуда $2ah^2 + 2c = y_1 + y_3$. Поэтому

$$s = \int_{-h}^{h} \left(ax^2 + bx + c\right)dx = \int_{-h}^{h} \left(ax^2 + c\right)dx + b\int_{-h}^{h} xdx = 2\int_{0}^{h} \left(ax^2 + c\right)dx =$$

$$= 2\left(a\frac{x^3}{3} + cx\right)\Bigg|_0^h = \frac{h}{3}\left(2ah^2 + 6c\right) = \frac{h}{3}\left(y_1 + 4y_2 + y_3\right).$$

Рассмотрим криволинейную трапецию, ограниченную графиком функции $y = f(x)$, и разобьем отрезок интегрирования $[a, b]$ на $2n$ равных частей точками

$$a = x_0 < x_1 < x_2 < \ldots < x_{2k} < x_{2k+1} < x_{2k+2} < \ldots < x_{2n-2} < x_{2n-1} < x_{2n} = b.$$

Через каждую из троек

$$M_0M_1M_2,\ \ldots,\ M_{2k}M_{2k+1}M_{2k+2},\ \ldots,\ M_{2n-2}M_{2n-1}M_{2n}$$

проведем параболу (лемма 1). В результате получим n криволинейных трапеций, ограниченных сверху параболами. По лемме 2 площадь каждой такой частичной криволинейной трапеции, опирающейся на отрезок $[x_{2k}, x_{2k+2}]$, равна

$$\int_{x_{2k}}^{x_{2k+2}} f(x)dx \approx \frac{b-a}{6n}(y_{2k} + 4y_{2k+1} + y_{2k+2}),$$

где $y_k = f(x_k)$, $0 \le k \le 2n$. Складывая почленно эти приближенные равенства, получим приближенную формулу

$$\int_a^b f(x)dx \approx \frac{b-a}{6n}\sum_{k=0}^{n-1}(y_{2k} + 4y_{2k+1} + y_{2k+2}),$$

или в развернутом виде

$$\int_a^b f(x)dx \approx \frac{b-a}{6n}(y_0 + y_{2n} + 2(y_2 + ... + y_{2n-2}) + 4(y_1 + ... + y_{2n-1})). \quad (4.40)$$

Соотношение (4.40) называется *формулой Симпсона*. В этой формуле значения функции $f(x)$ в нечетных точках разбиения x_1, ..., x_{2n-1} имеют коэффициент 4, в четных точках x_2, ..., x_{2n-2} – коэффициент 2, в граничных точках – коэффициент 1.

Можно показать, что погрешность формулы (4.40) не больше чем

$$M\frac{(b-a)^5}{2880n^4},$$

где M – наибольшее значение модуля четвертой производной данной функции на отрезке $[a, b]$.

Пример 4.45. Вычислить приближенно определенный интеграл из примера 4.44, используя формулу Симпсона.

Решение. Чтобы сравнить результаты вычислений по формулам прямоугольников и Симпсона, будем брать одинаковые разбиения отрезка [–1; 1], причем число n подбирать таким образом, чтобы наиболее трудоемкая часть вычислений – нахождение значений функции – была одинаковой. Результаты вычислений приведены в следующей таблице.

$2n$	4	8	16	32
S_{2n}	1,49436	1,49371	1,49365	1,49365

Сравним эффективность вычислений для этих двух методов. Мы видим, что формула Симпсона дает три верных знака после запятой при разбиении отрезка на 4 интервала, при использовании формулы прямоугольников понадобилось 64. Мы получим 5 верных знаков после запятой при $2n = 16$!

Таким образом, формула Симпсона дает значительный выигрыш в точности при незначительном усложнении вычислительной схемы и примерно том же объеме вычислений.

В заключение отметим, что в некоторых случаях (очень дорогие или трудно доступные значения функции) применяют другие вычислительные схемы Гаусса и Чебышева, которые позволяют за счет тщательного выбора значений функции находить приближенно значения определенных интегралов еще более экономным способом.

ГЛАВА 5

ЧИСЛОВЫЕ И СТЕПЕННЫЕ РЯДЫ

§ 5.1. Понятие числового ряда

1. Основные определения

Определение. *Пусть дана числовая последовательность*

$$a_1, a_2, a_3, ..., a_n, ...$$

Выражение вида

$$a_1 + a_2 + a_3 + ... + a_n + ... = \sum_{n=1}^{\infty} a_n \tag{5.1}$$

называют ***числовым рядом****, или просто* ***рядом****.*

Числа a_1, a_2, ..., a_n называют *членами ряда*, число a_n с общим номером n называют *общим членом ряда*.

Суммы конечного числа первых членов ряда

$$S_1 = a_1,$$
$$S_2 = a_1 + a_2,$$
$$S_3 = a_1 + a_2 + a_3,$$
$$...$$
$$S_n = a_1 + a_2 + a_3 + ... + a_n,$$

называют *частичными суммами ряда* (5.1). Так как число членов ряда бесконечно, то частичные суммы образуют числовую последовательность

$$S_1, S_2, S_3, ..., S_n, ... \tag{5.2}$$

Определение. *Ряд* (5.1) *называют* ***сходящимся****, если последовательность* (5.2) *его частичных сумм сходится к некоторому числу S. В этом случае число S называют* ***суммой ряда*** (5.1). *В противном случае ряд* (5.1) *называют* ***расходящимся.***

В случае сходимости ряда (5.1) его сумму записывают в виде символического равенства

$$S = a_1 + a_2 + a_3 + \ldots + a_n + \ldots, \text{ или } S = \sum_{n=1}^{\infty} a_n.$$

Пример 5.1. Мы знакомы с рядами уже давно, по крайней мере с того времени, когда впервые познакомились с действительными числами. Числа e и π нам знакомы. Записи $\pi = 3{,}14\ldots$ и $e = 2{,}1718\ldots$ означают, что

$$\pi = 3 + \frac{1}{10} + \frac{4}{10^2} + \ldots,$$

$$e = 2 + \frac{1}{10} + \frac{7}{10^2} + \frac{1}{10^3} + \frac{8}{10^4} + \ldots.$$

По аналогии любое десятичное разложение действительного числа представляет собой сходящийся числовой ряд, аналогичный указанным для чисел π и e, а частичные суммы S_n – это приближенные значения числа с заданной точностью.

Пример 5.2. Если в качестве последовательности $\{a_n, n = 1, 2, 3, \ldots\}$ мы рассмотрим *бесконечную геометрическую прогрессию* с первым членом, равным $b \neq 0$, и знаменателем q, то получим ряд

$$b + bq + bq^2 + \ldots + bq^{n-1} + \ldots. \tag{5.3}$$

Частичная сумма S_n этого ряда при $q \neq 1$ имеет вид

$$S_n = b + bq + bq^2 + \ldots + bq^{n-1} = \frac{b(1-q^n)}{1-q} = \frac{b}{1-q} - \frac{bq^n}{1-q}.$$

Отсюда:

1) если $|q| < 1$, то $\lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty}\dfrac{b}{1-q} - \lim\limits_{n\to\infty}\dfrac{bq^n}{1-q} = \dfrac{b}{1-q}$, т.е. ряд (5.3) сходится и его сумма $S = \dfrac{b}{1-q}$;

2) если $|q| > 1$, то $\lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty}\dfrac{b\left(1-q^n\right)}{1-q} = \infty$, т.е. ряд (5.3) расходится;

3) если $q = 1$, то $\lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} nb = \infty$ при $b \neq 0$, так что ряд расходится. Если же $q = -1$, то ряд (5.3) принимает вид

$$b - b + b - b + \dots .$$

Частичные суммы этого ряда выглядят следующим образом:

$$S_1 = b,\ S_2 = 0,\ S_3 = b,\ S_4 = 0,\ \dots .$$

Поэтому предел частичных сумм не существует, и ряд (5.3) расходится и в этом случае.

Мы видим, что ряд (5.3) сходится, когда $|q| < 1$, и расходится при $|q| \geq 1$. В случае $|q| < 1$ прогрессию называют *бесконечно убывающей*.

Пример 5.3. Приведем пример сходящегося ряда, отличного от геометрической прогрессии. Рассмотрим ряд

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \dots + \frac{1}{n(n+1)} + \dots .$$

Так как $\dfrac{1}{n(n+1)} = \dfrac{1}{n} - \dfrac{1}{n+1}$, то для n-й частичной суммы ряда получаем выражение

$$S_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \dots + \left(\frac{1}{n} - \frac{1}{n+1}\right).$$

После раскрытия скобок все слагаемые, кроме первого и последнего, взаимно уничтожаются, и в результате получаем $S_n = 1 - \dfrac{1}{n+1}$. Отсюда следует, что $\lim\limits_{n\to\infty} S_n = 1$. Итак, ряд сходится и его сумма равна 1.

2. Свойства сходящихся рядов

1. *Если сходится ряд* $a_1 + a_2 + a_3 + ... + a_n + ...$, *то сходится и любой ряд, полученный из него отбрасыванием конечного числа членов.*

Ограничимся случаем, когда отбрасываются первые k членов ряда (на самом деле можно рассмотреть общий случай, важно только, чтобы отбрасываемых членов было конечное число). Будем называть оставшийся ряд $a_{k+1} + a_{k+2} + ...$ *остатком* (или более точно k-м остатком) исходного ряда. Обозначим сумму первых k членов через σ_k, частичную сумму исходного ряда – через S_n, частичную сумму остатка – через $\overline{S}_n$, тогда при $n > k$ имеем

$$S_n = (a_1 + a_2 + ... + a_k) + (a_{k+1} + a_{k+2} + ... + a_n) = \sigma_k + \overline{S}_{n-k}.$$

Из этого равенства следует, что если существует предел частичных сумм исходного ряда $\lim\limits_{n\to\infty} S_n$, то существует и предел $\lim\limits_{n\to\infty} \overline{S}_n$, и наоборот. В частности, выполняется равенство, связывающее суммы исходного ряда и остатка:

$$S = \overline{S} + \sigma_k.$$

Над сходящимися рядами можно выполнять обычные арифметические действия.

2. *Если ряд* $a_1 + a_2 + ... + a_n + ...$ *сходится и его сумма равна* S, *а* c *– некоторое число, то сходится ряд* $ca_1 + ca_2 + ... + ca_n + ...$ *и его сумма равна* cS.

Пусть S_n – частичная сумма исходного ряда, а σ_n – частичная сумма ряда $ca_1 + ca_2 + ca_3 + ... + ca_n + ...$, тогда очевидно, что $\sigma_n = cS_n$. Переходя к пределу при $n \to \infty$, получаем

$$\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} cS_n = c\lim_{n\to\infty} S_n = cS,$$

т.е. последовательность частичных сумм $\{\sigma_n\}$ сходится к числу cS.

Пусть даны два ряда

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots, \qquad (5.4)$$

$$b_1 + b_2 + b_3 + \ldots + b_n + \ldots. \qquad (5.5)$$

Составим из них новый ряд

$$(a_1 + b_1) + (a_2 + b_2) + (a_3 + b_3) + \ldots + (a_n + b_n) + \ldots. \qquad (5.6)$$

Говорят, что ряд (5.6) получен *почленным сложением* (5.4) и (5.5). При этом выполняется следующее свойство.

3. *Если оба ряда* (5.4) *и* (5.5) *сходятся, а их суммы равны соответственно S и T, то и ряд* (5.6) *сходится и его сумма Q равна $S + T$.*

Очевидно, что для любого n имеем $Q_n = S_n + T_n$, где S_n, T_n, Q_n – соответственно частичные суммы для (5.4), (5.5) и (5.6). Переходя к пределу при $n \to \infty$, получаем, что существует предел $\lim\limits_{n\to\infty} Q_n$ и $Q = S + T$.

4. *Если сходится ряд $a_1 + a_2 + a_3 + \ldots + a_n + \ldots$, то сходится и любой ряд, полученный из него группировкой слагаемых, причем суммы обоих рядов одинаковы.*

Например, сходится ряд

$$(a_1 + a_2) + (a_3 + a_4) + \ldots + (a_{2n-1} + a_{2n}) + \ldots,$$

полученный из исходного ряда группировкой членов по два. На самом деле группы могут состоять из произвольного числа членов, зависящего от ее номера.

Назовем ряд, полученный из данного группировкой членов, “группированным”. Ясно, что k-я частичная сумма $\overline{S}_k$ группированного ряда совпадает с некоторой n-й частичной суммой данного ряда S_n для некоторого $n \ge k$:

$$\overline{S}_k = S_n. \qquad (5.7)$$

Поэтому если $k \to \infty$, то и $n \to \infty$. Тогда из равенства (5.7) следует, что сумма группированного ряда $\overline{S}$ равна сумме S данного ряда.

3. Необходимый признак сходимости ряда

При рассмотрении рядов возникают две задачи: 1) исследовать ряд на сходимость и 2) зная, что ряд сходится, найти его сумму. Установим сначала необходимое условие сходимости.

Теорема 5.1 (*необходимый признак сходимости*). *Если ряд сходится, то предел его общего члена равен нулю.*

Эквивалентная формулировка: *Если предел общего члена ряда не равен нулю или не существует, то данный ряд расходится.*

Доказательство. Пусть данный ряд сходится и его сумма равна S. Для любого натурального n имеем $S_n = S_{n-1} + a_n$, или

$$a_n = S_n - S_{n-1}. \tag{5.8}$$

При $n \to \infty$ обе частичные суммы S_n и S_{n-1} стремятся к пределу S, поэтому из равенства (5.8) следует, что

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0.$$

Подчеркнем еще раз, что мы установили только необходимое условие сходимости ряда, т.е. условие, при нарушении которого ряд не может сходиться. С помощью этого признака можно доказывать только расходимость ряда.

Пример 5.4. Исследовать на сходимость ряд

$$\frac{1}{2} + \frac{2}{3} + \ldots + \frac{n}{n+1} + \ldots\ .$$

Решение. Имеем $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \dfrac{n}{n+1} = 1 \neq 0,$ поэтому данный ряд расходится.

Пример 5.5. Исследовать на сходимость ряд

$$1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \ldots\ .$$

Решение. В этом случае предел общего члена ряда, очевидно, равен нулю, однако ряд расходится. Действительно, если бы данный ряд сходился, то сходился бы и ряд

$$1+\left(\frac{1}{2}+\frac{1}{2}\right)+\left(\frac{1}{3}+\frac{1}{3}+\frac{1}{3}\right)+\ldots,$$

полученный из данного ряда группировкой членов. Но общий член последнего ряда равен 1, и для него не выполнен необходимый признак сходимости.

Имеется замечательный ряд, расходимость которого не так очевидна. Его называют *гармоническим*.

Пример 5.6. Исследовать на сходимость гармонический ряд

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots+\frac{1}{n}+\ldots\ .$$

Решение. Очевидно, что для гармонического ряда выполнено необходимое условие сходимости, так как $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty}\frac{1}{n} = 0.$ Если бы данный ряд сходился, то, обозначая его сумму через S, мы бы имели

$$\lim_{n\to\infty}(S_{2n}-S_n)=\lim_{n\to\infty}S_{2n}-\lim_{n\to\infty}S_n=S-S=0. \qquad (5.9)$$

Но

$$S_{2n}-S_n=\frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{2n}>\frac{1}{2n}+\frac{1}{2n}+\ldots+\frac{1}{2n}=n\frac{1}{2n}=\frac{1}{2},$$

т.е. $S_{2n}-S_n > ½$, что противоречит равенству (5.9).

Тем не менее гармонический ряд расходится очень медленно, что можно увидеть из следующих значений его частичных сумм:

$$S_{10} = 2{,}929;\ S_{100} = 5{,}187;\ S_{1000} = 7{,}485;\ S_{10000}= 9{,}788.$$

§ 5.2. Ряды с положительными членами. Признаки сходимости

В данном параграфе рассматриваются ряды $a_1 + a_2 + ...$, все члены которых положительны. Вопрос о сходимости таких рядов решается особенно просто.

1. Критерий сходимости

Теорема 5.2. *Для того чтобы ряд с положительными членами сходился, необходимо и достаточно, чтобы последовательность его частичных сумм была ограничена.*

Доказательство. *Необходимость*. Пусть ряд $a_1 + a_2 + ... + a_n + ...$ сходится. Тогда существует предел его частичных сумм. Из свойств пределов следует, что последовательность частичных сумм ограничена.

Достаточность. Поскольку все члены данного ряда положительны и для любого n $S_n = S_{n-1} + a_n$, то последовательность его частичных сумм монотонно возрастает. Однако известно, что ограниченная сверху монотонная последовательность имеет предел.

Указанный критерий имеет скорее теоретическое, чем практическое значение. Однако он является основой, на которой базируются другие признаки сходимости. Все они являются только достаточными.

2. Достаточные признаки сходимости

Теорема 5.3 (*первый признак сравнения*). *Пусть даны два ряда с положительными членами*:

$$a_1 + a_2 + ... + a_n + ... , \tag{5.10}$$

$$b_1 + b_2 + ... + b_n + ... , \tag{5.11}$$

причем члены первого ряда не превосходят соответствующих членов второго: $a_n \le b_n$ (n =1, 2, ...). *Тогда из сходимости второго ряда («большего») следует сходимость первого ряда («меньшего»). Эквивалентно из расходимости меньшего ряда следует расходимость большего ряда.*

Д о к а з а т е л ь с т в о. Обозначим соответственно S_n и T_n частичные суммы меньшего и большего рядов (5.10) и (5.11). Тогда для всех n выполняются неравенства

$$S_n \le T_n. \tag{5.12}$$

Поскольку больший ряд сходится, по критерию сходимости последовательность его частичных сумм $\{T_n\}$ ограничена, т.е. существует такое число C, что для всех n выполняются неравенства $T_n \le C$. Тогда из (5.12) следует, что такие же неравенства выполняются и для частичных сумм S_n, т.е. последовательность $\{S_n\}$ также является ограниченной. По теореме 5.2 меньший ряд сходится, что и требовалось доказать.

П р и м е р 5.7. Рассмотрим ряд

$$\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{(n+1)^2}+\ldots, \tag{5.13}$$

который мы будем сравнивать со сходящимся рядом из примера 5.3:

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\ldots+\frac{1}{n(n+1)}+\ldots .$$

Поскольку $\frac{1}{(n+1)^2}<\frac{1}{n(n+1)}$, то из первого признака сравнения следует, что ряд (5.13) сходится. Заметим, что сумма ряда (5.13) выражается через число π, а именно, $S=\frac{\pi^2}{6}-1$, однако вывод этого равенства выходит за рамки данного курса.

П р и м е р 5.8. Доказать расходимость ряда

$$\frac{1}{\sqrt{1}}+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\ldots+\frac{1}{\sqrt{n}}+\ldots . \tag{5.14}$$

Р е ш е н и е. Будем сравнивать данный ряд с гармоническим рядом

$$\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}+\ldots,$$

расходимость которого была установлена в § 5.1. Поскольку $\frac{1}{\sqrt{n}}>\frac{1}{n}$, то ряд (5.14) тоже расходится.

Т е о р е м а 5.4 (*второй признак сравнения*). *Если для рядов* (5.10) *и* (5.11) *с положительными членами существует отличный от нуля предел отношения*

$$\lim_{n\to\infty}\frac{a_n}{b_n}=u, \tag{5.15}$$

то ряды (5.10) *и* (5.11) *сходятся или расходятся одновременно*.

Д о к а з а т е л ь с т в о**.** Пусть ряд (5.11) сходится; докажем, что сходится и ряд (5.10). Выберем некоторое число v, большее, чем u. Из условия (5.15) следует существование такого номера n_0, что для всех $n > n_0$ выполняется неравенство $\frac{a_n}{b_n}<v,$ так что

$$a_n < vb_n. \tag{5.16}$$

Отбросив первые n_0 членов рядов (5.10) и (5.11), что не влияет на сходимость, можно считать, что неравенства (5.16) выполняются для всех натуральных n. По условию ряд (5.11) сходится, тогда сходится и ряд с общим членом vb_n. По первому признаку сходимости сходится и ряд (5.10).

Пусть теперь сходится ряд (5.10); докажем сходимость ряда (5.11). Для этого надо переписать условие (5.15) в виде

$$\lim_{n\to\infty}\frac{b_n}{a_n}=\frac{1}{u}$$

и поменять ролями ряды (5.10) и (5.11). Теорема доказана полностью.

П р и м е р 5.9. Исследовать на сходимость ряд с общим членом

$$a_n = \frac{2n+3}{n^3+3n^2+5n+1}.$$

Р е ш е н и е. Хорошо известно, что предел a_n зависит от отношения старших степеней числителя и знаменателя. Рассмотрим ряд с общим членом $b_n = \frac{2n}{n^3} = \frac{2}{n^2}$, полученным именно таким способом. Последний ряд только множителем отличается от сходящегося ряда примера 5.7 данного параграфа, поэтому сходится. Так как

$$\lim_{n\to\infty}\frac{a_n}{b_n} = \lim_{n\to\infty}\frac{n^2(2n+3)}{2n^3+3n^2+5n+1} = 1,$$

то по теореме 5.4 сходится и исходный ряд.

Для решения вопроса о сходимости того или иного ряда обычно пытаются его сравнить с одним из “стандартных” рядов. Таким “стандартным” рядом в случае проверки сходимости служит геометрическая прогрессия, в случае же проверки расходимости – гармонический ряд. Сравнение с геометрической прогрессией лежит в основе следующего часто применяемого признака сходимости.

Т е о р е м а 5.5 (*признак Даламбера*). *Если для ряда с положительными членами* $a_1 + a_2 + ... + a_n + ...$ *существует такое число* $q < 1$, *что при всех n (или, начиная с некоторого n) выполняется неравенство*

$$\frac{a_{n+1}}{a_n} \le q, \tag{5.17}$$

то ряд сходится. Если же $\frac{a_{n+1}}{a_n} \ge 1$ *для всех или, начиная с некоторого n, то ряд расходится.*

Д о к а з а т е л ь с т в о. Отбросив, если необходимо, несколько первых членов ряда, можно считать, что неравенство (5.17) выполняется для всех $n = 1, 2, ...$ Перепишем это неравенство в виде $a_{n+1} \le q a_n$.

Отсюда имеем $a_2 \le qa_1$, $a_3 \le qa_2 \le q^2 a_1$, $a_4 \le qa_3 \le q^3 a_1$ и т.д.; вообще, при любом n справедливо неравенство

$$a_n \le q^{n-1} a_1. \tag{5.18}$$

Это показывает, что члены ряда $a_1 + a_2 + ... + a_n + ...$ не превосходят соответствующих членов бесконечной геометрической прогрессии

$$a_1 + a_1 q + a_1 q^2 + ... + a_1 q^{n-1} + ...$$

Поскольку по условию $0 \le q < 1$, эта прогрессия сходится. В силу первого признака сравнения сходится и данный ряд.

В случае, когда $\frac{a_{n+1}}{a_n} \ge 1$, имеем неравенство $a_{n+1} \ge a_n$, т.е. члены ряда образуют неубывающую последовательность, и поэтому не выполняется необходимый признак сходимости ряда, что доказывает теорему полностью.

Признак Даламбера часто применяется в следующей предельной форме.

Теорема 5.6. *Если существует предел*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = d, \tag{5.19}$$

то ряд сходится в случае $d < 1$ *и расходится в случае* $d > 1$.

Доказательство. Пусть $d < 1$. Возьмем некоторое число q между d и 1: $d < q < 1$. Из условия (5.19) следует, что, начиная с некоторого номера n, будет выполняться неравенство $\frac{a_{n+1}}{a_n} \le q$. На основании теоремы 5.5 отсюда следует сходимость ряда. Случай $d > 1$ разбирается аналогично.

В формулировке теоремы 5.6 ничего не говорится о случае $d = 1$. При $d = 1$ возможна как сходимость, так и расходимость ряда. Например, гармонический ряд, для которого $d=1$, расходится, а ряд $\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{(n+1)^2}+\ldots$, для которого также $d = 1$, сходится.

Полезно заметить, что признак Даламбера дает также и оценку остатка ряда; из неравенства (5.18) следует, что

$$R_n = a_{n+1} + a_{n+2} + a_{n+3} + \ldots \leq a_{n+1} + a_{n+1}q + a_{n+1}q^2 + \ldots ,$$

т.е.

$$R_n \leq \frac{a_{n+1}}{1-q}. \qquad (5.20)$$

Пример 5.10. Исследовать на сходимость ряд

$$1+\frac{3}{1!}+\frac{3^2}{2!}+\ldots+\frac{3^n}{n!}+\ldots .$$

Решение. Имеем $a_n = \frac{3^n}{n!}$, $a_{n+1} = \frac{3^{n+1}}{(n+1)!}$, поэтому $\frac{a_{n+1}}{a_n}=\frac{3^{n+1}}{(n+1)!}:\frac{3^n}{n!}=\frac{3}{n+1}$. Очевидно, что последнее отношение будет оставаться меньше единицы для всех $n > 3$. По теореме 5.5 данный ряд сходится.

Мы увидим в дальнейшем, что сумма этого ряда выражается через число e, а именно, $S = e^3$.

Пример 5.11. Исследовать на сходимость ряд

$$1+\frac{2^2}{2!}+\frac{3^3}{3!}+\ldots+\frac{n^n}{n!}+\ldots .$$

Решение. Имеем

$$a_n = \frac{n^n}{n!}, \quad a_{n+1} = \frac{(n+1)^{n+1}}{(n+1)!}=\frac{(n+1)^n}{n!}.$$

Составим отношение последующего члена ряда к предыдущему:

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)^n}{n!} : \frac{n^n}{n!} = \frac{(n+1)^n}{n^n} = \left(1+\frac{1}{n}\right)^n.$$

Поскольку предел отношения d, очевидно, равен числу $e > 1$, то по теореме 5.6 данный ряд расходится.

Пример 5.12. Доказать сходимость ряда

$$\frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \ldots + \frac{n}{2^n} + \ldots$$

и оценить остаток R_{10}.

Имеем

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

Предел этого отношения равен ½, т.е. ряд сходится. Для оценки остатка ряда $R_{10} = a_{11} + a_{12} + \ldots$ заметим, что если $n > 10$, то

$$\frac{a_{n+1}}{a_n} = \frac{1}{2}\left(1+\frac{1}{n}\right) < \frac{1}{2} \cdot 1{,}1 = 0{,}55.$$

Поэтому

$$R_{10} < \frac{a_{11}}{1-q} = \frac{11/2^{11}}{1-0{,}55} = \frac{11}{2^{11} \cdot 0{,}45} < \frac{1}{2^{10} \cdot 0{,}08} < \frac{1}{80}.$$

Любую числовую последовательность a_1, a_2, ..., a_n, ... можно рассматривать как функцию, определенную на множестве натуральных чисел: $a_n = f(n)$, $n = 1, 2, \ldots$. Следовательно, всякий числовой ряд можно представить в виде

$$f(1) + f(2) + \ldots + f(n) + \ldots \qquad (5.21)$$

В том случае, когда функция $f(x)$ монотонно убывает, может оказаться полезным следующий признак сходимости.

Теорема 5.7 (*интегральный признак сходимости*). *Пусть неотрицательная функция* $y = f(x)$ *определена и монотонно убывает для* $x \geq 1$. *Тогда для сходимости ряда* (5.21) *необходимо и достаточно, чтобы сходился несобственный интеграл*

$$\int_{1}^{\infty} f(x)dx. \tag{5.22}$$

Доказательство. Сходимость интеграла (5.22) означает существование предела при $n \to \infty$ интеграла

$$\int_{1}^{n} f(x)dx. \tag{5.23}$$

Следовательно, сходится ряд

$$\int_{1}^{2} f(x)dx + \int_{2}^{3} f(x)dx + \ldots + \int_{n-1}^{n} f(x)dx + \ldots\ , \tag{5.24}$$

для которого интеграл (5.23) является частичной суммой. Поэтому задача состоит в том, чтобы доказать одновременную сходимость или расходимость рядов (5.21) и (5.24).

Поскольку функция $f(x)$ монотонно убывает на каждом отрезке $[n, n + 1]$, то имеем $f(n + 1) \leq f(x) \leq f(n)$. Интегрируя по отрезку $[n, n + 1]$, получаем $f(n+1) \int_{n}^{n+1} dx \leq \int_{n}^{n+1} f(x)dx \leq f(n) \int_{n}^{n+1} dx$, т.е.

$$f(n+1) \leq \int_{n}^{n+1} f(x)dx \leq f(n). \tag{5.25}$$

Если сходится ряд (5.21), то из второго неравенства в (5.25) согласно первому признаку сравнения вытекает сходимость ряда

(5.24). Обратно, если сходится ряд (5.24), то из первого неравенства в (5.25) по первому признаку сравнения вытекает сходимость ряда $f(2) + f(3) + ... + f(n + 1) + ...$, а, значит, и ряда (5.21). Теорема доказана полностью.

Пример 5.13. Выяснить, при каких $\alpha > 0$ сходится ряд

$$1+\frac{1}{2^\alpha}+\frac{1}{3^\alpha}+\ldots+\frac{1}{n^\alpha}+\ldots \ . \tag{5.26}$$

Решение. Положим $f(x) = \dfrac{1}{x^\alpha}$ $(x \geq 1)$. Данная функция монотонно убывает, поскольку ее производная $f'(x) = -\dfrac{\alpha}{x^{\alpha+1}}$ отрицательна. Поэтому сходимость ряда (5.26) эквивалентна сходимости несобственного интеграла $\int\limits_1^\infty \dfrac{dx}{x^\alpha}$. Легко проверить, что этот интеграл сходится при $\alpha > 1$ и расходится при $0 < \alpha \leq 1$. Значит, и ряд (5.26) сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$. Например, при $\alpha = 1{,}5$ имеем сходящийся ряд

$$1+\frac{1}{2\sqrt{2}}+\frac{1}{3\sqrt{3}}+\ldots+\frac{1}{n\sqrt{n}}+\ldots,$$

а при $\alpha = 1/3$ имеем расходящийся ряд

$$1+\frac{1}{\sqrt[3]{2}}+\frac{1}{\sqrt[3]{3}}+\ldots+\frac{1}{\sqrt[3]{n}}+\ldots \ .$$

3. Оценка остатка ряда

Неравенства (5.25) позволяют получить оценку k-го остатка ряда (5.21). Первое из них

$$f(n+1) \leq \int\limits_n^{n+1} f(x)dx \ \ (n = 1, 2, ...) \tag{5.27}$$

позволяет оценить сверху k-й остаток ряда (5.21), т.е. число

$$R_k = f(k+1) + f(k+2) + \ldots .$$

Чтобы получить такую оценку, достаточно сложить почленно неравенства (5.27) для всех $n = k$, $k+1$, $k+2$, ... Тогда получим

$$R_k \le \int_k^{\infty} f(x)dx.$$

Аналогично получается оценка остатка снизу. Для этого воспользуемся вторым неравенством в (5.25), а именно,

$$\int_n^{n+1} f(x)dx \le f(n) \ \ (n = 1, 2, \ldots). \qquad (5.28)$$

Если сложить неравенства (5.28), записанные для $n = k+1$, $k+$ $+2$, $k+3$, ..., то получим $\int_{k+1}^{\infty} f(x)dx \le R_k$. В результате находим двусторонние оценки остатка:

$$\int_{k+1}^{\infty} f(x)dx \le R_k \le \int_k^{\infty} f(x)dx. \qquad (5.29)$$

Пример 5.14. Сколько членов ряда $1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{n^2} + \ldots$ надо сложить, чтобы найти его сумму с точностью до 0,01?

Решение. В данном случае $f(x) = \dfrac{1}{x^2}$; оценки (5.29) дают

$$\int_{k+1}^{\infty} \frac{dx}{x^2} \le R_k \le \int_k^{\infty} \frac{dx}{x^2}, \text{ или } \frac{1}{k+1} \le R_k \le \frac{1}{k}.$$

Второе из этих неравенств показывает, что при $k \ge 100$ остаток R_k не превосходит 0,01. Таким образом, для вычисления суммы ряда с требуемой точностью следует сложить 100 членов ряда. Как видим, сходимость этого ряда очень медленная.

§ 5.3. Знакопеременные ряды

В предыдущем параграфе изучались ряды с положительными членами. Рассмотрим теперь ряды с членами произвольного знака, или, как принято говорить, *знакопеременные ряды*.

1. Знакочередующиеся ряды. Теорема Лейбница

Особенно часто среди знакопеременных рядов встречаются ряды, члены которых имеют чередующиеся знаки, т.е. *знакочередующиеся ряды*.

Условимся считать первый член ряда положительным; тогда знакочередующийся ряд в общем виде запишется так:

$$a_1 - a_2 + a_3 - a_4 + \ldots + (-1)^{n-1} a_n + \ldots, \qquad (5.30)$$

где все a_n положительны ($n = 1, 2, \ldots$).

Теорема 5.8 (*теорема Лейбница*). *Если члены знакочередующегося ряда убывают по абсолютной величине и стремятся к нулю, когда $n \to \infty$, то*: 1) *ряд сходится*; 2) *любой остаток ряда не превосходит по абсолютной величине первого из своих членов и имеет одинаковый с ним знак*.

Доказательство. Пусть дан ряд (5.30) и известно, что $a_n > a_{n+1}$ для всех n и $a_n \to 0$ при $n \to \infty$.

Рассмотрим частичную сумму ряда (5.30) с четным числом членов

$$S_{2n} = a_1 - a_2 + a_3 - a_4 + \ldots + a_{2n-1} - a_{2n} =$$
$$= (a_1 - a_2) + (a_3 - a_4) + \ldots + (a_{2n-1} - a_{2n}).$$

В силу первого условия все разности в скобках положительны, поэтому последовательность частичных сумм $\{S_{2n}\}$ возрастающая. Докажем, что она является ограниченной. Для этого представим S_{2n} в виде

$$S_{2n} = a_1 - [(a_2 - a_3) + (a_4 - a_5) + \ldots + (a_{2n-2} - a_{2n-1}) + a_{2n}].$$

Выражение в квадратных скобках положительно, поэтому $S_{2n} < a_1$ для любого n, т.е. последовательность $\{S_{2n}\}$ ограничена.

Итак, последовательность $\{S_{2n}\}$ возрастающая и ограниченная, следовательно, она имеет предел $\lim\limits_{n\to\infty} S_{2n} = S$.

Так как $S_{2n+1} = S_{2n} + a_{2n+1}$ и по условию

$$\lim_{n\to\infty} a_{2n+1} = 0 \text{, то } \lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} S_{2n} = S.$$

Мы доказали, что ряд (5.30) сходится и его сумма удовлетворяет неравенствам

$$0 < S < a_1. \tag{5.31}$$

Докажем теперь второе утверждение. Рассмотрим остаток ряда (5.30) с четным номером $2k$: $R_{2k} = a_{2k+1} - a_{2k+2} + \ldots$. Этот ряд является знакочередующимся и удовлетворяет всем условиям теоремы, поэтому выполняются оценки

$$0 < R_{2k} < a_{2k+1}. \tag{5.32}$$

Что касается остатков ряда с нечетными номерами, то любой из них можно записать в виде $R_{2k+1} = -a_{2k+2} + a_{2k+3} - \ldots = -(a_{2k+2} - a_{2k+3} + \ldots)$. Ряд в скобках снова удовлетворяет условиям теоремы, поэтому $0 < -R_{2k+1} < a_{2k+2}$, или

$$-a_{2k+2} < R_{2k+1} < 0. \tag{5.33}$$

Сходимость ряда (5.30) вместе с неравенствами (5.31), (5.32) и (5.33) полностью доказывает теорему.

Пример 5.15. Доказать сходимость ряда

$$1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \ldots + \frac{(-1)^{n-1}}{n^2} + \ldots \; . \tag{5.34}$$

Сколько членов ряда надо взять, чтобы вычислить сумму ряда с точностью до 0,01?

Р е ш е н и е. Сходимость ряда следует непосредственно из теоремы Лейбница. Для вычисления суммы ряда с заданной точностью воспользуемся оценкой остатка $|R_k| < a_{k+1}$. В данном случае получаем неравенство $|R_k| < \dfrac{1}{(k+1)^2}$. Таким образом, достаточно, чтобы выполнялось неравенство $\dfrac{1}{(k+1)^2} \le \dfrac{1}{100}$, т.е. $k \ge 9$. Для нахождения суммы ряда с нужной точностью достаточно сложить только девять членов. Это на порядок величины отличается от случая, рассмотренного в предыдущем параграфе. Знакочередующиеся ряды сходятся быстрее, чем соответствующие ряды с положительными членами. Поэтому они удобнее для вычислений.

2. Ряды с членами произвольного знака. Плюс- и минус-ряды для данного ряда

Рассмотрим теперь ряды, члены которых имеют произвольные знаки. С самого начала условимся считать все члены ряда отличными от нуля (нулевые члены всегда можно отбросить – это не отразится ни на сходимости, ни на его сумме). С каждым знакопеременным рядом

$$a_1 + a_2 + \ldots + a_n + \ldots \qquad (5.35)$$

можно связать два ряда с положительными членами; один из них будем называть *плюс-рядом*, а другой *минус-рядом* для данного ряда (5.35). Чтобы получить плюс-ряд, следует оставить в (5.35) только положительные члены (а отрицательные отбросить); для получения минус-ряда нужно оставить только отрицательные члены ряда (5.35), сменив в них знаки – на +. Например, для знакочередующегося ряда

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{(-1)^{n-1}}{n} + \ldots$$

плюс- и минус-ряды соответственно имеют вид

$$1 + \frac{1}{3} + \frac{1}{5} + \ldots \text{ и } \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \ldots .$$

Любая частичная сумма S_n ряда (5.35) включает некоторую часть плюс-ряда и некоторую часть минус-ряда; поэтому

$$S_n = S_{n_1}^{+} - S_{n_2}^{-}, \tag{5.36}$$

где $S_{n_1}^{+}$ – некоторая частичная сумма плюс-ряда и $S_{n_2}^{-}$ – частичная сумма минус-ряда.

3. Абсолютно сходящиеся ряды и их свойства

Пусть дан знакопеременный ряд (5.35). Рассмотрим ряд, составленный из абсолютных величин его членов:

$$|a_1| + |a_2| + \ldots + |a_n| + \ldots . \tag{5.37}$$

Очевидно, что (5.37) есть ряд с положительными членами.

Определение. *Ряд* (5.35) *называют* ***абсолютно сходящимся***, *если сходится ряд* (5.37), *составленный из модулей его членов*.

Справедливо следующее утверждение.

Теорема 5.9. *Всякий абсолютно сходящийся ряд сходится. Сумма такого ряда равна разности между суммой его плюс-ряда и суммой минус-ряда*.

Доказательство. Пусть ряд (5.35) сходится абсолютно, т.е. сходится ряд (5.37). Частичные суммы ряда (5.37) обозначим T_n. Аналогично равенству (5.36) имеем

$$T_n = S_{n_1}^{+} + S_{n_2}^{-}. \tag{5.38}$$

Ввиду сходимости ряда (5.37) его частичные суммы ограничены некоторым числом C. Тогда из равенства (5.38) следует, что $S_{n_1}^{+} \le C$ и $S_{n_2}^{-} \le C$, т.е. частичные суммы плюс- и минус-рядов также ограничены сверху числом C. Согласно критерию сходимости рядов с положительными членами отсюда вытекает сходимость плюс- и минус-рядов, т.е. существуют пределы $S^{+} = \lim\limits_{k\to\infty} S_k^{+}$ и

$S^- = \lim_{l\to\infty} S_l^-$. Если теперь в равенстве (5.36) перейти к пределу при $n \to \infty$, то получим $\lim_{n\to\infty} S_n = S^+ - S^-$, что требовалось доказать.

Пример 5.16. Доказать абсолютную сходимость ряда

$$\sum_{n=1}^{\infty} (-1)^n \frac{\sin\frac{\pi}{n}}{n}. \tag{5.39}$$

Решение. Составим ряд из абсолютных величин членов данного ряда:

$$\sum_{n=1}^{\infty} \frac{\sin\frac{\pi}{n}}{n}. \tag{5.40}$$

Так как выполняется неравенство $\sin\frac{\pi}{n} < \frac{\pi}{n}$, то члены ряда (5.40) не превосходят соответствующих членов сходящегося ряда $\sum_{n=1}^{\infty} \frac{\pi}{n^2}$. Следовательно, ряд (5.40) сходится, а ряд (5.39) сходится абсолютно.

Поскольку абсолютная сходимость ряда (5.35) означает сходимость ряда (5.37), члены которого положительны, то для установления абсолютной сходимости можно пользоваться любым из признаков сходимости из предыдущего параграфа.

4. Условно сходящиеся ряды

Определение. *Ряд* $a_1 + a_2 + \ldots + a_n + \ldots$ *называют* ***условно сходящимся***, *если он сходится, а ряд, составленный из модулей его членов, расходится.*

Таким образом, все сходящиеся ряды можно разделить на два класса: *абсолютно сходящиеся*, т.е. такие, для которых ряд из модулей их членов сходится, и *условно сходящиеся*, для которых ряд из абсолютных величин их членов расходится.

Примером условно сходящегося ряда может служить ряд

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots+\frac{(-1)^{n-1}}{n}+\ldots\ . \tag{5.41}$$

Действительно, по теореме Лейбница этот ряд сходится; в то же время ряд из модулей его членов

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots+\frac{1}{n}+\ldots$$

(гармонический ряд), как известно, расходится.

Свойства условно сходящихся рядов резко отличаются от свойств абсолютно сходящихся рядов. В частности, для таких рядов не выполняется теорема 5.9, поскольку и плюс-, и минус-ряды расходятся. Еще одно необычное свойство условно сходящихся рядов демонстрируется в следующем примере.

Пример 5.17. В условно сходящемся ряде (5.41) переставим и сгруппируем члены следующим образом:

$$\left(1-\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{3}-\frac{1}{6}-\frac{1}{8}\right)+\ldots+\left(\frac{1}{2n-1}-\frac{1}{4n-2}-\frac{1}{4n}\right)+\ldots\ .$$

Если затем внутри каждой скобки сложить первые два слагаемых, то получим ряд

$$\left(\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{6}-\frac{1}{8}\right)+\ldots\left(\frac{1}{4n-2}-\frac{1}{4n}\right)+\ldots,$$

сумма которого равна

$$\frac{1}{2}-\frac{1}{4}+\frac{1}{6}-\frac{1}{8}+\frac{1}{10}-\frac{1}{12}+\ldots=\frac{1}{2}\left(1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots\right).$$

Таким образом, при перестановке членов исходного ряда (5.41) его сумма уменьшилась вдвое!

Справедлива следующая теорема, которую мы приведем без доказательства.

Теорема 5.10 (*теорема Римана*). *Если ряд сходится условно, то в результате перестановки его членов можно получить ряд, имеющий любую сумму, а также расходящийся ряд.*

§ 5.4. Степенные ряды

1. Степенной ряд. Теорема Абеля

Определение. *Ряд вида*

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots , \tag{5.42}$$

где a_0, a_1, a_2, ..., a_n, ... *– некоторая числовая последовательность, называют* ***степенным рядом***.

Числа a_0, a_1, a_2, ..., a_n, ... называют коэффициентами степенного ряда.

Придавая переменной x различные числовые значения, будем получать различные числовые ряды, которые могут сходиться или расходиться. Множество тех значений x, при которых ряд (5.42) сходится, называют *областью сходимости*. Область сходимости любого степенного ряда не пуста, поскольку очевидно, что ряд (5.42) сходится для $x = 0$.

Важную роль при изучении области сходимости степенных рядов играет следующее утверждение.

Теорема 5.11 (*теорема Абеля*). 1) *Если степенной ряд* (5.42) *сходится при некотором* $x = x_0$, *не равном нулю*, *то он сходится*, *и притом абсолютно*, *при всех* x, *удовлетворяющих условию* $|x| < |x_0|$; 2) *если ряд* (5.42) *расходится при некотором* $x = x_1$, *то он расходится при всех* x, *удовлетворяющих условию* $|x| > |x_1|$.

Доказательство. 1. По условию числовой ряд

$$a_0 + a_1x_0 + a_2x_0^2 + \ldots + a_nx_0^n + \ldots$$

сходится, поэтому его общий член $a_n x_0^n$ стремится к нулю при $n \to \infty$. Следовательно, сходящаяся последовательность $\{a_n x_0^n\}$ ограничена, т.е. существует константа M, такая, что $|a_n x_0^n| < M$ для всех $n = 0, 1, 2, \ldots$.

Рассмотрим теперь ряд, составленный из абсолютных величин членов ряда (5.42):

$$|a_0| + |a_1 x| + |a_2 x^2| + \ldots + |a_n x^n| + \ldots . \qquad (5.43)$$

Пусть $|x| < |x_0|$, тогда имеем $|a_n x^n| = |a_n x_0^n| \left|\dfrac{x}{x_0}\right|^n < M \left|\dfrac{x}{x_0}\right|^n$, причем $\left|\dfrac{x}{x_0}\right| < 1$. Поэтому члены ряда (5.43) не превосходят соответствующих членов сходящегося ряда

$$M + M\left|\frac{x}{x_0}\right| + M\left|\frac{x}{x_0}\right|^2 + \ldots + M\left|\frac{x}{x_0}\right|^n + \ldots$$

– суммы бесконечно убывающей геометрической прогрессии. Следовательно, ряд (5.43) сходится, а ряд (5.42) сходится абсолютно.

2. Докажем вторую часть теоремы. Предположим противное, т.е. ряд (5.42) расходится при $x = x_1$, в то же время для некоторого x, удовлетворяющего условию $|x| > x_1$, ряд (5.42) сходится. По первой части теоремы ряд (5.42) сходится абсолютно при $x = x_1$, что противоречит условию. Теорема доказана полностью.

Геометрическая интерпретация теоремы Абеля очень проста. Если x_0 – точка сходимости ряда (5.42), то во всех точках интервала $(-|x_0|, |x_0|)$ ряд сходится абсолютно, а если при $x = x_1$ ряд (5.42) расходится, то он расходится и во всех точках, расположенных вне отрезка $[-|x_1|, |x_1|]$ (рис. 5.1).

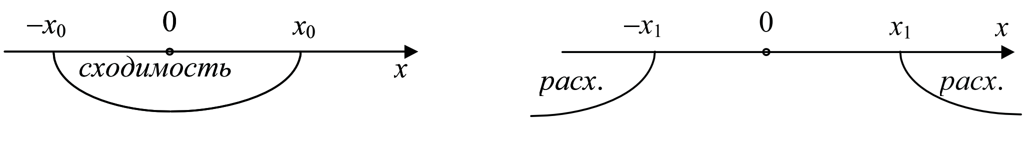

Рис. 5.1

2. Область сходимости степенного ряда

Теорема Абеля позволяет в принципе описать область сходимости степенного ряда.

Теорема 5.12. *Для степенного ряда* (5.42) *возможны только три случая*:

1) *ряд сходится в единственной точке* $x = 0$;

2) *ряд сходится для всех значений* x;

3) *существует такое* $R > 0$, *что ряд сходится для всех значений* x *из интервала* $(-R, R)$ *и расходится для всех значений* x *вне отрезка* $[-R, R]$.

Доказательство. Обозначим через X множество всех положительных значений x, для которых ряд (5.42) сходится.

Если X пусто, то ряд сходится только при $x = 0$. Это дает случай 1.

Пусть X не пусто. Если X не ограничено сверху, то по теореме Абеля ряд (5.42) сходится для всех x. Это дает случай 2.

Остается рассмотреть случай, когда X непусто и ограничено сверху. Существует точная верхняя грань множества X, обозначим ее через R. Число R обладает тем свойством, что в любой его окрестности имеются точки, как принадлежащие, так и не принадлежащие X. Иными словами, в любой окрестности R имеются точки, в которых ряд (5.42) сходится, и точки, в которых он расходится. Тогда по теореме Абеля для всех x, таких, что $|x| < R$, ряд (5.42) сходится, а при всех x, для которых $|x| > R$, – расходится. Это дает случай 3, что доказывает теорему полностью.

Итак, *областью сходимости степенного ряда* (5.42) *является или единственная точка* $x = 0$, *или вся числовая прямая*, *или интервал* $(-R, R)$, *к которому могут присоединяться один или оба конца*.

Определение. *Интервал* $(-R, R)$ *называют* ***интервалом сходимости*** *ряда* (5.42), *а число* R – ***радиусом сходимости*** *этого ряда*.

Понятие радиуса сходимости будем распространять на все три случая в теореме 5.12: в случае 1 положим $R = 0$, в случае 2 $R = \infty$.

3. Отыскание радиуса сходимости степенного ряда

Радиус сходимости степенного ряда чаще всего находят с помощью признака Даламбера.

Теорема 5.13. *Если существует предел* $D = \lim\limits_{n\to\infty}\left|\dfrac{a_{n+1}}{a_n}\right|$, *отличный от нуля, то радиус сходимости степенного ряда* $a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots$ *равен*

$$R = \frac{1}{D} = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|. \qquad (5.44)$$

Доказательство. Рассмотрим ряд, составленный из модулей членов степенного ряда (5.42). По условию существует предел $D = \lim\limits_{n\to\infty}\left|\dfrac{a_{n+1}}{a_n}\right|$, не равный нулю. Обозначим его через $\dfrac{1}{R}$. По признаку Даламбера ряд (5.42) сходится абсолютно для данного значения x, если

$$\lim_{n\to\infty}\left|\frac{a_{n+1}x^{n+1}}{a_nx^n}\right| = |x|\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \frac{|x|}{R} < 1,$$

т.е. при $|x| < R$.

По теореме о сходимости знакопеременных рядов ряд (5.42) сходится при этих значениях x. При $|x| > R$ имеем

$$\lim_{n\to\infty}\left|\frac{a_{n+1}x^{n+1}}{a_nx^n}\right| = \frac{|x|}{R} > 1,$$

поэтому ряд (5.42) расходится, так как общий член ряда a_nx^n не стремится к нулю.

Таким образом, данный ряд сходится внутри интервала $(-R, R)$ и расходится вне отрезка $[-R, R]$, т.е. R – радиус сходимости ряда (5.42), и формула (5.44) доказана.

Пример 5.18. Найти область сходимости ряда

$$x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots + \frac{x^n}{n} + \ldots\ .$$

Р е ш е н и е . В данном случае $a_n = \dfrac{1}{n}$, $a_{n+1} = \dfrac{1}{n+1}$, поэтому

$$R = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty}\frac{n+1}{n} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right) = 1.$$

Следовательно, по теореме 5.13 данный ряд сходится на интервале (−1; 1). Исследуем поведение ряда на концах интервала, т.е. при $x = \pm 1$. При $x = 1$ получаем гармонический ряд, следовательно, ряд расходится. При $x = -1$ получаем знакочередующийся ряд $-1+\dfrac{1}{2}-\dfrac{1}{3}+\ldots+\dfrac{(-1)^n}{n}+\ldots$, который сходится по признаку Лейбница. Таким образом, областью сходимости данного степенного ряда является полуинтервал [−1, 1). Позже мы увидим, что этот ряд в области сходимости равен функции $-\ln(1-x)$.

П р и м е р 5.19. Найти область сходимости степенного ряда

$$1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots+\frac{x^n}{n!}+\ldots\ .$$

Р е ш е н и е . Напомним, что $n!$ (“эн-факториал”) обозначает произведение всех натуральных чисел до n включительно. В данном случае

$$a_n = \frac{1}{n!},\ a_{n+1} = \frac{1}{(n+1)!} = \frac{1}{n!(n+1)},$$

поэтому $R = \lim\limits_{n\to\infty}\dfrac{1}{n!}:\dfrac{1}{n!(n+1)} = \lim\limits_{n\to\infty}(n+1) = \infty$. Следовательно, данный степенной ряд сходится абсолютно на всей числовой прямой. Опять-таки, для каждого x сумма данного ряда дает значение функции e^x.

П р и м е р 5.20. Найти область сходимости степенного ряда

$$x + (2x)^2 + (3x)^3 + \ldots + (nx)^n + \ldots\ .$$

Р е ш е н и е. Общий член ряда равен $(nx)^n$, поэтому каково бы ни было значение $x \neq 0$, при $n > \dfrac{1}{|x|}$ имеем $(n|x|)^n > 1$, т.е. ряд расходится. Область сходимости данного степенного ряда состоит из одной точки $x = 0$, радиус сходимости $R = 0$.

4. Свойства степенных рядов

Пусть функция $f(x)$ является суммой степенного ряда

$$f(x) = a_0 + a_1x + a_2x^2 + ... + a_nx^n + ... , \tag{5.45}$$

интервал сходимости которого $(-R, R)$.

В этом случае говорят, что *функция* $f(x)$ *на интервале* $(-R, R)$ *разлагается в степенной ряд.*

Выполняются следующие теоремы о свойствах степенных рядов, которые мы приведем без доказательства.

Т е о р е м а 5.14 (*о почленном дифференцировании степенного ряда*). *Пусть функция* $f(x)$ *разлагается на интервале* $(-R, R)$ *в степенной ряд* (5.45). *Рассмотрим степенной ряд*

$$a_1 + 2a_2x + ... + na_nx^{n-1} + ... , \tag{5.46}$$

полученный почленным дифференцированием ряда (5.45). *Тогда*:

1) *ряд* (5.46) *имеет тот же радиус сходимости* R, *что и ряд* (5.45);

2) *на всем интервале* $(-R, R)$ *функция* $f(x)$ *имеет производную* $f'(x)$, *которая разлагается в степенной ряд* (5.46).

Применяя теорему 5.14 повторно, заключаем, что вторая производная $f''(x)$ также существует и равна сумме ряда, полученного двукратным почленным дифференцированием ряда (5.45). Аналогичное заключение справедливо и для третьей производной, и т.д. В итоге получаем следствие.

С л е д с т в и е. *Функция* $f(x)$, *которая разлагается в степенной ряд* (5.45) *на интервале* $(-R, R)$, *бесконечно дифференцируема на этом интервале. Разложение в степенной ряд для любой производной получается почленным дифференцированием ряда* (5.45).

При этом радиусы сходимости соответствующих рядов равны радиусу сходимости ряда (5.45).

Теорема 5.15 (*о почленном интегрировании степенного ряда*). *Если функция* $f(x)$ *разлагается в степенной ряд на интервале* $(-R, R)$, *то она интегрируема в этом интервале. Интеграл от суммы ряда равен сумме интегралов от членов ряда.*

Иными словами, если $[x_1, x_2] \subset (-R, R)$, то

$$\int_{x_1}^{x_2} f(x)dx = \int_{x_1}^{x_2}\left(a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots\right)dx =$$

$$= \int_{x_1}^{x_2} a_0dx + \int_{x_1}^{x_2} a_1xdx + \int_{x_1}^{x_2} a_1x^2dx + \ldots + \int_{x_1}^{x_2} a_nx^ndx + \ldots.$$

Особый интерес представляет интегрирование степенного ряда (5.45) с переменным верхним пределом по отрезку $[0, x]$, где $|x| < R$:

$$\int_0^x f(t)dt = a_0x + \frac{a_1x^2}{2} + \frac{a_2x^3}{3} + \ldots + \frac{a_nx^{n+1}}{n+1} + \ldots. \qquad (5.47)$$

В этом случае получаем степенной ряд (5.47), который имеет тот же радиус сходимости, что и ряд (5.45).

§ 5.5. Разложение функций в степенные ряды

1. Ряд Маклорена

Для приложений важно уметь данную функцию $f(x)$ *разлагать в степенной ряд*. Для этого надо ответить на два вопроса:

1) может ли эта функция на данном отрезке быть представлена в виде суммы некоторого степенного ряда?

2) если да, то как найти этот ряд?

В этом пункте мы дадим ответ на второй из поставленных вопросов; ответ на первый вопрос будет дан в следующем пункте.

Предположим, что данная функция $f(x)$ на некотором отрезке $[-r, r]$ может быть разложена в степенной ряд

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots . \tag{5.48}$$

Найдем коэффициенты $a_0, a_1, a_2, a_3, \ldots$ этого ряда.

Из предыдущего параграфа мы знаем, что степенной ряд (5.49) можно дифференцировать почленно любое число раз. Поэтому для любого x из интервала $(-r, r)$ имеем

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + \ldots ,$$
$$f''(x) = 2\cdot1a_2 + 3\cdot2a_3x + 4\cdot3a_4x^2 + \ldots ,$$
$$f'''(x) = 3\cdot2\cdot1a_3 + 4\cdot3\cdot2a_4x + \ldots$$

и т.д.; вообще,

$$f^{(n)}(x) = n(n-1)\ldots3\cdot2\cdot1a_n + \ldots$$

(в разложении $f^{(n)}(x)$ указан только свободный член). Полагая в этих равенствах, а также в разложении (5.49) $x = 0$, получаем $f(0) = a_0$, $f'(0) = a_1$, $f''(0) = 2\cdot1a_2$, $f'''(0) = 3\cdot2\cdot1a_3$ и т.д.; вообще, $f^{(n)}(0) = n!a_n$. Отсюда имеем формулу

$$a_n = \frac{f^{(n)}(0)}{n!} \quad (n = 0, 1, 2, \ldots).$$

Определение. *Пусть функция $f(x)$ определена в некоторой окрестности точки $x = 0$ и имеет в этой точке производные всех порядков. Степенной ряд*

$$f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + \ldots \tag{5.49}$$

называют ***рядом Маклорена*** *для функции $f(x)$.*

Ответ на поставленный выше вопрос 2 можно дать в виде следующей теоремы.

Теорема 5.16. *Если функция $f(x)$ разлагается в некоторой окрестности точки $x = 0$ в степенной ряд, то этот ряд является ее рядом Маклорена.*

Теорема 5.16 показывает, что разложение функции в степенной ряд однозначно. Следующий по важности вопрос: к какой функции сходится ее ряд Маклорена?

2. Достаточное условие разложимости функции в ряд Маклорена

Пусть $f(x)$ – произвольная бесконечно дифференцируемая функция. Для нее можно составить ряд Маклорена (5.49). Установим, при каких условиях сумма ряда (5.49) совпадает с функцией $f(x)$. Ответ на этот вопрос дает формула Маклорена. В § 2.18 было показано, что для любой бесконечно дифференцируемой функции справедлива формула Маклорена

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \ldots + \frac{f^{(n)}(0)}{n!}x^n + R_n(x),$$

где остаточный член

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}x^{n+1}, \quad 0 < c < x. \tag{5.50}$$

Если обозначить через $S_n(x)$ частичную сумму ряда Маклорена, то формулу Маклорена можно записать следующим образом:

$$f(x) = S_n(x) + R_n(x). \tag{5.51}$$

Ясно, что вопрос о сходимости ряда Маклорена к функции $f(x)$ сводится к исследованию поведения остаточного члена $R_n(x)$ при $n \to \infty$. Во многих случаях проблема решается следующей теоремой.

Теорема 5.17 (***достаточное условие разложимости функции в ряд Маклорена***). *Пусть функция $f(x)$ определена и бесконечно дифференцируема в интервале $(-r, r)$. Если существует такая константа M, что во всех точках указанного интервала выполняются неравенства*

$$|f^{(n)}(x)| < M \ (n = 0, 1, 2, \ldots), \tag{5.52}$$

то в этом интервале ряд Маклорена сходится к функции $f(x)$.

Доказательство. Из формулы (5.51) следует, что достаточно доказать равенство $\lim\limits_{n\to\infty} R_n(x) = 0.$ Из условия теоремы и формулы (5.50) следует неравенство

$$|R_n(x)| < \frac{Mr^{n+1}}{(n+1)!}.$$

Правая часть этого неравенства только постоянным множителем M отличается от члена сходящегося ряда $\sum\limits_{n=1}^{\infty} \frac{r^n}{n!}$ (см. пример 5.9). По необходимому признаку сходимости $\lim\limits_{n\to\infty} \frac{r^n}{n!} = 0,$ откуда и предел остаточного члена тоже равен нулю, что заканчивает доказательство теоремы.

3. Разложение функции e^x

Пусть $f(x) = e^x$. В любом интервале $(-r, r)$ имеем $|f^{(n)}(x)| = e^x < e^r$. В силу теоремы 5.17 отсюда следует, что функция e^x равна сумме своего ряда Маклорена при $x \in (-r, r)$, а значит, и для любого x ввиду произвольности r. Поскольку $f^{(n)}(0) = e^0 = 1$ при любом n, получаем разложение

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + \ldots\ , \tag{5.53}$$

справедливое для всех x.

4. Разложение функций sin$x$ и cos$x$

Каждая из функций sin$x$, cos$x$ удовлетворяет условиям теоремы 5.17 (при любом r): действительно, производная любого порядка от этих функций есть одна из функций ±sin$x$, ±cos$x$. Поэтому модуль производной любого порядка от этих функций не больше единицы. Таким образом, для любого x каждая функция sin$x$ и cos$x$ равна сумме своего ряда Маклорена.

Имеем

$$(\sin x)' = \cos x,\ (\sin x)'' = -\sin x,\ (\sin x)''' = -\cos x,\ (\sin x)^{(4)} = \sin x.$$

Отсюда видно, что последовательность производных функции sin$x$ периодична с периодом 4. При $x = 0$ получаем

$$\sin 0 = 0,\ \sin'(0) = 1,\ \sin''(0) = 0,\ \sin'''(0) = -1.$$

В общем случае все производные четного порядка равны нулю, а нечетного

$$\sin^{(2n+1)}(0) = (-1)^n.$$

Отсюда ряд Маклорена для sin$x$:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \ldots \ . \qquad (5.54)$$

Ряд для cos$x$ получается почленным дифференцированием ряда (5.54):

$$(\sin x)' = (x)' - \left(\frac{x^3}{3!}\right)' + \left(\frac{x^5}{5!}\right)' - \ldots + \left((-1)^n \frac{x^{2n+1}}{(2n+1)!}\right)' + \ldots,$$

откуда

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots \ . \qquad (5.55)$$

Заметим, что в разложение функции $\sin x$ входят только нечетные степени x, соответственно у $\cos x$ только четные степени. Это согласуется с тем, что $\sin x$ – нечетная функция, а $\cos x$ – четная.

5. Разложение функций ln(1 + *x*) и arctg*x*

При нахождении разложения функции $\cos x$ мы использовали свойство почленной дифференцируемости степенных рядов. Аналогично можно использовать и свойство почленной интегрируемости.

Рассмотрим ряд $1 + x + x^2 + x^3 + ...+ x^n + ...$. Данный ряд выражает сумму геометрической прогрессии с первым членом 1 и знаменателем x. Известно, что при $|x| < 1$ данный ряд сходится и его сумма равна $\frac{1}{1-x}$. Следовательно,

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots + x^n + \ldots \ . \tag{5.56}$$

Равенство (5.56) является разложением функции $f(x) = \frac{1}{1-x}$ в степенной ряд при $|x| < 1$. Выполним замену переменной $x = -t$:

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \ldots + (-1)^n t^n + \ldots \ .$$

Проинтегрируем почленно это равенство в пределах от 0 до x ($|x| < 1$):

$$\int_0^x \frac{dt}{1+t} = \ln(1+t)\Big|_0^x = \ln(1+x) =$$

$$= \int_0^x \left(1 - t + t^2 - t^3 + \ldots + (-1)^n t^n + \ldots\right) dt =$$

$$= \left(t - \frac{t^2}{2} + \frac{t^3}{3} - \frac{t^4}{4} + \ldots + (-1)^n \frac{t^{n+1}}{n+1} + \ldots \right)_0^x =$$

$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^n \frac{x^{n+1}}{n+1} + \ldots \ .$$

Отсюда получаем разложение функции $\ln(1+x)$ в степенной ряд:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^n \frac{x^{n+1}}{n+1} + \ldots \ . \qquad (5.57)$$

Оно справедливо для $|x| < 1$. Можно показать, что оно остается верным и для $x = 1$, т.е. выполняется равенство

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots,$$

о котором мы уже упоминали при рассмотрении знакочередующихся рядов.

Найдем теперь разложение функции $\operatorname{arctg} x$. Для этого подставим в равенство (5.56) $x = -t^2$ и проинтегрируем по t от 0 до x. Получим разложение

$$\operatorname{arctg} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \ldots, \qquad (5.58)$$

верное при $|x| < 1$. Так же как для логарифма, можно показать, что оно остается верным и при $x = \pm 1$. В частности, при $x = 1$ слева получаем $\operatorname{arctg} 1 = \frac{\pi}{4}$, так что имеем интересное числовое равенство

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \ldots \ .$$

6. Разложение функции $(1 + x)^a$

Пусть $f(x) = (1 + x)^a$, где a – произвольное число. Тогда имеем:

$$f'(x) = a(1 + x)^{a-1},$$
$$f''(x) = a(a - 1)(1 + x)^{a-2},$$
$$f'''(x) = a(a - 1)(a - 2)(1 + x)^{a-3},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f^{(n)}(x) = a(a - 1)(a - 2)\ldots(a - n + 1)\ (1 + x)^{a-n}.$$

Полагая $x = 0$ во всех этих формулах, получим

$$f(0) = 1,\ f'(0) = a,\ f''(0) = a(a-1),\ \ f'''(0) = a(a-1)(a-2),\ \ldots,$$
$$f^{(n)}(0) = a(a-1)(a-2)\ldots(a-n+1).$$

Подставляя выражения для производных в ряд Маклорена, получим

$$(1+x)^a = 1 + \frac{a}{1}x + \frac{a(a-1)}{1\cdot 2}x^2 + \frac{a(a-1)(a-2)}{1\cdot 2\cdot 3}x^3 + \ldots$$

$$+ \frac{a(a-1)\cdots(a-n+1)}{n!}x^n + \ldots\ . \qquad (5.59)$$

Нетрудно проверить, что радиус сходимости степенного ряда в правой части равенства равен 1, так что ряд сходится при $|x| < 1$.

Для натуральных $a = m$ правая часть равенства (5.59) превращается в многочлен, а само равенство (5.59) – в формулу бинома Ньютона:

$$(1+x)^m = 1 + \frac{m}{1}x + \frac{m(m-1)}{1\cdot 2}x^2 + \ldots$$

(последний член правой части получается при $n = m$ и равен $\dfrac{m(m-1)\cdots(m-m+1)}{m!}x^m = x^m$).

Докажем, что равенство (5.59) выполняется в общем случае при $|x| < 1$.

Во-первых, проверим, что функция $y = (1 + x)^a$ удовлетворяет следующему уравнению:

$$(1 + x)y' = ay. \qquad (5.60)$$

Действительно,

$$y' = a(1 + x)^{a-1} \text{ и } (1 + x)y' = a(1 + x)^a = ay.$$

Во-вторых, проверим, что равенство (5.60) выполняется и для степенного ряда, стоящего в правой части (5.59). Дифференцируя почленно, имеем

$$y' = a + \frac{a(a-1)}{1!}x + \frac{a(a-1)(a-2)}{2!}x^2 + \ldots$$
$$\ldots + \frac{a(a-1)(a-2)\ldots(a-n+1)}{(n-1)!}x^{n-1} + \ldots .$$

Отсюда

$$(1+x)y' = a + \frac{a(a-1)}{1!}x + \frac{a(a-1)(a-2)}{2!}x^2 + \ldots +$$
$$+ ax + \frac{a(a-1)}{1!}x^2 + \ldots =$$
$$= a + \frac{a}{1!}(a-1+1)x + \frac{a(a-1)}{2!}(a-2+2)x^2 + \ldots =$$
$$= a\left(1 + \frac{a}{1!}x + \frac{a(a-1)}{2!}x^2 + \ldots\right) = ay.$$

Мы видим, что и левая, и правая части равенства (5.59) удовлетворяют одному дифференциальному уравнению (5.60). Поскольку очевидно, что они удовлетворяют одному и тому же начальному условию $y(0) = 1$, то по теореме единственности для дифференциальных уравнений они равны.

§ 5.6. Степенные ряды с произвольным центром

1. Интервал сходимости

Выполнив в степенном ряде замену переменной x на $x - x_0$, можем рассмотреть *степенные ряды с центром в* x_0. В общем случае они имеют вид:

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \ldots + a_n(x - x_0)^n + \ldots . \qquad (5.61)$$

До этого мы систематически рассматривали случай, когда $x_0 = 0$. Общий случай ничем существенным не отличается от частного случая, все сводится к переносу начала в точку $x = x_0$. Чтобы это увидеть, выполним замену переменной $X = x - x_0$. После такой замены ряд (5.61) принимает вид

$$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n + \ldots, \tag{5.62}$$

т.е. превращается в степенной ряд с центром $x_0 = 0$. Из предыдущего изложения следует, что ряд (5.62) имеет интервал сходимости $(-R, R)$, т.е. сходится при $|X| < R$ и расходится при $|X| > R$. Для ряда (5.61) это означает, что он сходится при $|x - x_0| < R$ и расходится при $|x - x_0| > R$. Следовательно, *интервал сходимости ряда* (5.61) *есть* $(x_0 - R, x_0 + R)$ (рис. 5.2).

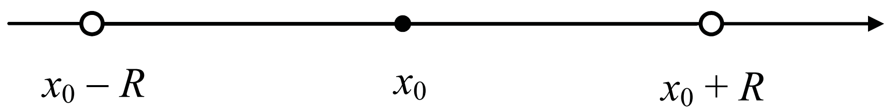

Рис. 5.2

Все свойства степенных рядов, которыми обладает ряд (5.62), переносятся на ряд (5.61) в интервале $(x_0 - R, x_0 + R)$. К ним относится возможность почленного дифференцирования и интегрирования ряда (5.61).

2. Ряд Тейлора

Аналогично задаче о разложении функции в ряд по степеням x ставится и задача о разложении по степеням $x - x_0$.

Определение. *Пусть функция $f(x)$ определена в некоторой окрестности точки x_0 и имеет в этой точке производные всех порядков. Степенной ряд*

$$f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \ldots$$

называют ***рядом Тейлора*** *с центром x_0 для функции $f(x)$.*

Справедливо утверждение, обобщающее теорему 5.16.

Теорема 5.18. *Если функция разлагается в некоторой окрестности точки x_0 в ряд по степеням $x - x_0$, то он является ее рядом Тейлора с центром x_0.*

Доказательство. Положим $X = x - x_0$ и $F(X) = f(x)$. Если функция $f(x)$ в окрестности точки x_0 разлагается в ряд (5.61), то функция $F(X)$ в окрестности нуля разлагается в ряд (5.62). Тогда согласно теореме 5.16 из § 5.5 последний ряд должен быть рядом Маклорена для $F(X)$, т.е. имеет место разложение

$$F(X) = F(0) + \frac{F'(0)}{1!}X + \frac{F''(0)}{2!}X^2 + \ldots + \frac{F^{(n)}(0)}{n!}X^n + \ldots\ .$$

Возвращаясь к переменной x, получим

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \ldots,$$

что и требовалось доказать.

§ 5.7. Приложения степенных рядов к приближенным вычислениям

1. Вычисление значений показательной функции

Для показательной функции справедливо разложение в ряд

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + \ldots\ (-\infty < x < \infty). \qquad (5.63)$$

При больших по модулю значениях x ряд (5.63) малопригоден для вычислений. В этих случаях обычно поступают так: представляют x в виде суммы

$$x = E(x) + q,$$

где $E(x)$ – целая часть числа x (наибольшее целое число, не превосходящее x) и q – дробная его часть, $0 \le q < 1$. Тогда

$$e^x = e^{E(x)} \cdot e^q.$$

Первый множитель $e^{E(x)}$ находят с помощью возведения в степень числа e (считается, что это число нам известно с произвольной точностью). Второй множитель вычисляют с помощью разложения (5.63). При $0 \le x < 1$ этот ряд быстро сходится, поскольку остаток ряда $R_n(x)$ оценивается следующим образом:

$$0 \le R_n(x) < \frac{x^{n+1}}{n!\,n}.$$

Пример 5.21. Найти $\sqrt{e}$ с точностью до 10^{-5}.

Решение. Пользуемся формулой

$$e^{\frac{1}{2}} = \sum_{k=0}^{n} u_k + R_n\left(\frac{1}{2}\right), \tag{5.64}$$

где $u_0 = 1$, $u_k = \frac{u_{k-1}}{2k}$ ($k = 1, 2, ..., n$). Слагаемые подсчитываем с двумя запасными десятичными знаками. Последовательно имеем:

$$u_0 = 1, \qquad u_1 = \frac{u_0}{2} = 0{,}50000000,$$

$$u_2 = \frac{u_1}{4} = 0{,}1250000, \qquad u_3 = \frac{u_2}{6} = 0{,}0208333,$$

$$u_4 = \frac{u_3}{8} = 0{,}0026042, \qquad u_5 = \frac{u_4}{10} = 0{,}0002604,$$

$$u_6 = \frac{u_5}{12} = 0{,}0000217, \qquad u_7 = \frac{u_6}{14} = 0{,}0000016;$$

$$S_7 = 1{,}6487212.$$

Округляя сумму до пяти десятичных знаков после запятой, получим

$$\sqrt{e} = 1{,}64872.$$

2. Вычисление значений логарифмической функции

Непосредственное применение разложения логарифмической функции в степенной ряд

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots + (-1)^n \frac{x^{n+1}}{n+1} + \ldots \qquad (5.65)$$

затруднительно из-за ограничения $-1 < x \le 1$ и медленной сходимости. Поэтому заменим x на $-x$ и получим

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \ldots - \frac{x^{n+1}}{n+1} - \ldots \ .$$

Вычитая из первого равенства второе, находим

$$\ln\frac{1+x}{1-x} = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \ldots\right), \qquad (5.66)$$

где $|x| < 1$.

Формула (5.66) очень удобна для вычисления логарифмов по двум причинам: 1) аргумент логарифмической функции для указанных значений x принимает произвольные положительные значения; 2) ряд (5.66) сходится довольно быстро. Оценим остаток ряда (5.66). Так как отношение последующего члена ряда к предыдущему равно

$$\frac{x^{2n+1}}{2n+1} : \frac{x^{2n-1}}{2n-1} = \frac{2n-1}{2n+1} x^2$$

и потому меньше x^2, то остаток меньше, чем сумма геометрической прогрессии с первым членом $2\frac{x^{2n+1}}{2n+1}$ и знаменателем x^2. Таким образом, выполняется неравенство

$$R_n < \frac{2x^{2n+1}}{(2n+1)\left(1-x^2\right)}. \qquad (5.67)$$

Пример 5.22. Найти ln5 с точностью до 10^{-6}.

Решение. Чтобы ускорить сходимость ряда (5.66), представим число 5 в виде произведения $5 = e\cdot\frac{5}{e}$, тогда $\ln 5=\ln e+\ln\frac{5}{e}=$ $=1+\ln\frac{5}{e}$. Полагая $\frac{1+x}{1-x}=\frac{5}{e}$, находим $x=\frac{5-e}{5+e}\approx 0{,}29562514$ (здесь предполагается, что число e мы можем вычислять с требуемой точностью, как в п. 1). Из оценки (5.67) следует, что остаток приблизительно равен первому отброшенному члену. Будем вести все вычисления с двумя запасными знаками

$$u_1=2x=0{,}59125029,\qquad u_2=2\frac{x^3}{3}=0{,}01722395,$$

$$u_3=2\frac{x^5}{5}=0{,}00090316,\qquad u_4=2\frac{x^7}{7}=0{,}00005638,$$

$$u_5=2\frac{x^9}{9}=0{,}00000383,\qquad u_6=2\frac{x^{11}}{11}=0{,}00000027.$$

Требуемая точность достигнута. Складывая члены ряда, получим

$$\ln 5 = 1{,}609438.$$

3. Вычисление значений синуса и косинуса

Для вычислений значений функций $\sin x$ и $\cos x$ используем разложения в ряд

$$\sin x = x-\frac{x^3}{3!}+\frac{x^5}{5!}-\ldots+(-1)^n\frac{x^{2n+1}}{2n+1}+\ldots\ (-\infty<x<\infty), \quad (5.68)$$

$$\cos x = 1-\frac{x^2}{2!}+\frac{x^4}{4!}-\ldots+(-1)^n\frac{x^{2n}}{(2n)!}+\ldots\ (-\infty<x<\infty). \quad (5.69)$$

Ряды (5.68) и (5.69) при больших x сходятся медленно. Однако, учитывая периодичность функций $\sin x$ и $\cos x$ и формулы приведения тригонометрических функций, достаточно уметь вычислять

$\sin x$ и $\cos x$ для промежутка $0 \le x \le \frac{\pi}{4}$. Так как в этом промежутке ряды (5.68) и (5.69) являются знакочередующимися с убывающими по модулю членами, то остаток ряда в обоих случаях не превышает модуля первого отброшенного члена.

Пример 5.23. Вычислить $\sin\frac{\pi}{4}$ с точностью до 10^{-4}.

Решение. Значение аргумента с двумя запасными знаками: $x = 0{,}785398$. Применяя (5.68), получим:

$$u_1 = x = 0{,}785398, \qquad u_2 = -\frac{x^2 u_1}{2\cdot 3} = -0{,}080746,$$

$$u_3 = -\frac{x^2 u_2}{4\cdot 5} = 0{,}002490, \qquad u_4 = -\frac{x^2 u_3}{6\cdot 7} = -0{,}000037.$$

Требуемая точность достигнута, поэтому

$$\sin\frac{\pi}{4} = 0{,}7071.$$

Разумеется, этот результат можно было получить, используя значение $\sin\frac{\pi}{4}$, равное $\frac{\sqrt{2}}{2}$.

4. Приближенное нахождение интегралов

Пример 5.24. Вычислить интеграл

$$I = \int_0^1 e^{-x^2}\,dx,$$

разлагая подынтегральную функцию в степенной ряд и используя семь членов этого разложения. Оценить погрешность.

Решение. Имеем

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \ldots + (-1)^n\frac{x^{2n}}{n!} + \ldots\ .$$

Этот ряд сходится при любом x; проинтегрировав почленно первые семь членов, получим

$$I \approx \left(x - \frac{x^3}{3} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \frac{x^9}{9 \cdot 4!} - \frac{x^{11}}{11 \cdot 5!} + \frac{x^{13}}{13 \cdot 6!} \right) \Bigg|_0^1 =$$

$$= 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42} + \frac{1}{216} - \frac{1}{1320} + \frac{1}{9360}. \qquad (5.70)$$

Оценим остаток ряда:

$$|R_7| \le \frac{1}{15 \cdot 7!} = \frac{1}{75600} < 1{,}5 \cdot 10^{-5}.$$

Учитывая это, вычислим сумму (5.70) с пятью знаками после запятой (с одним запасным знаком). Окончательно получим

$$I = \int_0^1 e^{-x^2} dx = 0{,}7468.$$

Пример 5.25. Вычислить интеграл

$$I = \int_0^{0,5} \frac{dx}{1 - x^5}$$

с точностью до 10^{-4}.

Решение. Разложим подынтегральную функцию в степенной ряд

$$\frac{1}{1 - x^5} = 1 + x^5 + x^{10} + x^{15} + \ldots \ .$$

Областью сходимости этого ряда является интервал $(-1,\ 1)$. Отрезок интегрирования входит в этот интервал, следовательно, данный ряд можно почленно интегрировать.

Интегрируя по отрезку $[0;\ 0{,}5]$, получим

$$I = \left(x + \frac{x^6}{6} + \frac{x^{11}}{11} + \frac{x^{16}}{16} + \ldots \right) \Bigg|_0^{1/2} = \frac{1}{2} + \frac{1}{6 \cdot 2^6} + \frac{1}{11 \cdot 2^{11}} + \frac{1}{16 \cdot 2^{16}} + \ldots .$$

Так как уже третий член меньше, чем 10^{-4}, то попробуем взять в качестве приближенного значения интеграла сумму первых двух членов.

Оценим остаток R_2:

$$R_2 = \frac{1}{11 \cdot 2^{11}} + \frac{1}{16 \cdot 2^{16}} + \ldots \ .$$

Для этого заменим первые множители, стоящие в знаменателях, на число 11. Тогда получим сумму бесконечной убывающей геометрической прогрессии

$$R_2 < \frac{1}{11 \cdot 2^{11}} + \frac{1}{11 \cdot 2^{16}} + \ldots = \frac{1}{11 \cdot 2^{11}} \cdot \frac{1}{1 - \frac{1}{2^5}} = \frac{1}{11 \cdot 2^6 \cdot 31} < 0{,}5 \cdot 10^{-4}.$$

Таким образом, сумма первых двух членов дает значение интеграла с требуемой точностью. Окончательно получим

$$\int_0^{0,5} \frac{dx}{1 - x^5} = \frac{1}{2} + \frac{1}{6 \cdot 2^6} = 0{,}5026.$$

В заключение заметим, что разложение функций в степенные ряды используется также при решении дифференциальных уравнений, о чем речь пойдет в следующих главах.

§ 5.8. Ряды из матриц

Рассмотрим $M(k, l)$ – множество матриц размером $k \times l$ с действительными элементами.

1. Последовательности из матриц

Пусть для каждого натурального n дана матрица $A_n \in M(k, l)$. Последовательность $A_1, A_2, \ldots, A_n, \ldots$, или $\{A_n\}$, $n \in \mathbb{N}$, называют *последовательностью матриц*.

Ясно, что всякая последовательность матриц из $M(k, l)$ порождает $k \cdot l$ обычных числовых последовательностей, которые получаются выписыванием элементов матриц A_n с фиксированными индексами i и j.

Пример 5.26. Рассмотрим последовательность 2×2 матриц

$$A_n = \left(a_{ij}^{(n)}\right),\ a_{ij}^{(n)} = \frac{1}{(ij)^n},\ i = 1, 2, j = 1, 2.$$

В развернутом виде имеем:

$$A_1 = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} \end{pmatrix},\ A_2 = \begin{pmatrix} 1 & \left(\frac{1}{2}\right)^2 \\ \left(\frac{1}{2}\right)^2 & \left(\frac{1}{4}\right)^2 \end{pmatrix}, \ldots, A_n = \begin{pmatrix} 1 & \left(\frac{1}{2}\right)^n \\ \left(\frac{1}{2}\right)^n & \left(\frac{1}{4}\right)^n \end{pmatrix}, \ldots,$$

так что получим четыре числовые последовательности:

1, 1, ..., 1, ... в первой строке и первом столбце;

$\frac{1}{2}, \left(\frac{1}{2}\right)^2, \ldots, \left(\frac{1}{2}\right)^n, \ldots$ в первой строке и втором столбце;

$\frac{1}{2}, \left(\frac{1}{2}\right)^2, \ldots, \left(\frac{1}{2}\right)^n, \ldots$ во второй строке и первом столбце;

$\frac{1}{4}, \left(\frac{1}{4}\right)^2, \ldots, \left(\frac{1}{4}\right)^n, \ldots$ во второй строке и во втором столбце.

Определение. *Матрицу $A = (a_{ij})$ называют* ***пределом*** *последовательности матриц $\{A_n\}$, $n \in \mathbb{N}$, если для каждой пары i и j выполняются равенства*

$$a_{ij} = \lim_{n \to \infty} a_{ij}^{(n)},\ i = 1, \ldots, k,\ j = 1, \ldots, l.$$

Иными словами, сходимость должна быть поэлементной.

Пример 5.27. Найти предел последовательности матриц $\{A_n\}$, $n \in \mathbb{N}$, из предыдущего примера 5.26.

Решение. Очевидно, $\lim\limits_{n \to \infty} A_n = A$, где $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

2. Ряды из матриц. Определения и примеры

Определение. ***Рядом из матриц*** $\{A_n \in M(k, l)\}$, $n \in \mathbb{N}$, *называют выражение вида*

$$A_1 + A_2 + \ldots + A_n + \ldots \tag{5.71}$$

Ряд из матриц – это непосредственное обобщение понятия числового ряда. Определение его сходимости копирует определение сходимости числового ряда (5.1). А именно, обозначим через

$$S_1 = A_1,\, S_2 = A_1 + A_2,\, \ldots \tag{5.72}$$

частичные суммы ряда (5.71).

n-я частичная сумма S_n – это, по сути, просто набор из $k \cdot l$ частичных сумм обычных числовых рядов, которые получаются на фиксированном месте соответствующих матриц с индексами i и j, $i = 1,\ldots, k$; $j = 1,\ldots, l$.

Определение. *Ряд* (5.71) ***сходится***, *если последовательность* $\{S_n\}$ *его частичных сумм сходится.*

В случае сходимости ряда (5.71) обозначим предел частичных сумм через S. Тогда имеем

$$S = \lim_{n\to\infty} S_n. \tag{5.73}$$

Определение. *Для сходящегося ряда матрицу S называют* ***суммой ряда*** (5.71).

В этом случае сумму ряда записывают в виде символического равенства $S = \sum\limits_{n=1}^{\infty} A_n$.

Пример 5.28. Найти сумму ряда из матриц

$$A_1 + A_2 + \ldots + A_n + \ldots, \text{ где } A_n = \begin{pmatrix} \left(\frac{1}{2}\right)^n & \left(\frac{1}{3}\right)^n \\ \left(\frac{1}{4}\right)^n & \left(\frac{1}{5}\right)^n \end{pmatrix}.$$

Р е ш е н и е . Ясно, что в качестве частичной суммы S_n мы получим матрицу, элементами которой будут частичные суммы геометрических прогрессий со знаменателями, меньшими единицы. Поэтому данный ряд сходится, и его суммой будет

$$S = \begin{pmatrix} \dfrac{\frac{1}{2}}{1-\frac{1}{2}} & \dfrac{\frac{1}{3}}{1-\frac{1}{3}} \\ \dfrac{\frac{1}{4}}{1-\frac{1}{4}} & \dfrac{\frac{1}{5}}{1-\frac{1}{5}} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{4} \end{pmatrix}.$$

Более интересные примеры возникают, если мы по аналогии с геометрической прогрессией рассмотрим ряд, составленный из степеней квадратной матрицы A размером k:

$$E + A + A^2 + \ldots + A^n + \ldots . \qquad (5.74)$$

Такие ряды рассматривались в первой части учебника при изучении балансовых моделей, где предполагалось, что $A \geq 0$. Оказывается, что сходимость такого ряда эквивалентна *продуктивности матрицы Леонтьева A*. Чтобы дать исчерпывающий ответ на вопрос о сходимости ряда (5.74), нам предстоит рассмотреть *степенные матричные ряды*.

3. Степенные матричные ряды

Определение. *Пусть даны квадратная матрица A порядка k и степенной ряд* $\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_n x^n + \ldots$. ***Степенным матричным рядом*** *называют ряд, полученный заменой в степенном ряде переменной x на A*:

$$\alpha_0 + \alpha_1 A + \alpha_2 A^2 + \ldots + \alpha_n A^n + \ldots = \sum_{n=0}^{\infty} \alpha_n A^n. \qquad (5.75)$$

Оказывается, что определение сходимости степенного матричного ряда (5.75) полностью сводится к вопросу о сходимости обычного степенного ряда. Напомним, что число λ называют *соб-*

ственным значением матрицы A, если найдется ненулевой *собственный вектор* $\vec{x}$, для которого выполняется равенство

$$A\vec{x} = \lambda\vec{x}.$$

Вместе с матричным степенным рядом (5.75) рассмотрим степенной ряд

$$\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \alpha_n\lambda^n + \ldots = \sum_{n=0}^{\infty}\alpha_n\lambda^n. \tag{5.76}$$

Теорема 5.19. *Матричный степенной ряд* (5.75) *сходится тогда и только тогда, когда сходится степенной ряд* (5.76) *для каждого собственного значения* λ *матрицы* A.

Доказательство. Пусть матричный ряд (5.75) сходится и λ – собственное значение матрицы A с собственным вектором $\vec{x}$. Обозначим $B = \sum_{n=0}^{\infty}\alpha_n A^n$. Рассмотрим вектор $B\vec{x}$. Поскольку для любого натурального n выполняется равенство $A^n\vec{x} = \lambda^n\vec{x}$, то справедливо равенство

$$B\vec{x} = \sum_{n=0}^{\infty}\alpha_n\lambda^n\vec{x}. \tag{5.77}$$

Отсюда следует сходимость ряда (5.76).

При доказательстве достаточности ограничимся случаем, когда *собственные векторы матрицы A образуют базис пространства* $\mathbb{R}^k$. Заметим, что для проверки сходимости ряда (5.75) достаточно проверить, что для любого вектора x пространства $\mathbb{R}^k$ сходится ряд из векторов

$$\alpha_0 x + \alpha_1 Ax + \alpha_2 A^2 x + \ldots + \alpha_n A^n x + \ldots\ . \tag{5.78}$$

Если x – собственный вектор матрицы A, то ряд (5.78) сходится по условию. В общем случае вектор x представляется в виде линейной комбинации собственных векторов. Поэтому ряд (5.78) также представляется в виде линейной комбинации рядов такого же типа для собственных векторов, каждый из которых сходится. Следовательно, сходится и ряд (5.78), что заканчивает доказательство теоремы.

Следствие 1. *Матрица Леонтьева A продуктивна тогда и только тогда, когда число Фробениуса матрицы A меньше единицы.*

Доказательство. Мы воспользуемся критерием продуктивности из гл. 3 первой части учебника: неотрицательная матрица A продуктивна тогда и только тогда, когда сходится матричный ряд (5.74):

$$E + A + A^2 + \ldots + A^n + \ldots .$$

Согласно теореме 5.19 ряд (5.74) сходится, если для любого собственного значения λ матрицы A сходится числовой ряд

$$1 + \lambda + \lambda^2 + \ldots + \lambda^n + \ldots ,$$

т.е. когда $|\lambda| < 1$. Поскольку матрица A неотрицательна, то по теореме Перрона – Фробениуса максимальное по модулю собственное значение λ_A матрицы A действительно и неотрицательно. Поэтому условие сходимости матричного ряда (5.74) эквивалентно условию $\lambda_A < 1$, что и требовалось доказать.

Следствие 2. *Для любой квадратной матрицы A сходится ряд*

$$e^A = E + A + \frac{A^2}{2!} + \ldots + \frac{A^n}{n!} + \ldots . \tag{5.79}$$

Согласно теореме 5.19 достаточно проверить, что для любого собственного значения λ сходится ряд для обычной показательной функции

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \ldots + \frac{\lambda^n}{n!} + \ldots ,$$

а это хорошо известно.

Разложение в ряд (5.79) используют при решении систем линейных дифференциальных уравнений.

ГЛАВА 6

КРАТНЫЕ ИНТЕГРАЛЫ

§ 6.1. Двойной интеграл и его свойства

Подобно тому как задача вычисления площади криволинейной трапеции приводит к определенному интегралу от функции одной переменной, аналогичная задача вычисления объема тела приводит к понятию *двойного интеграла*. Определение последнего мало отличается от определения определенного интеграла для функции одной переменной.

1. Определение двойного интеграла

Пусть функция $f(x, y)$ двух переменных задана в прямоугольной области P вида $\{a \le x \le b, c \le y \le d\}$ и сама является ограниченной функцией в P. Разобьем область P несколькими прямыми, параллельными оси x, а также несколькими прямыми, параллельными оси y, на части (рис. 6.1), которые будем называть *ячейками*; само же разбиение такого рода будем называть *декартовым*. Обозначим разбиение буквой T. Будем считать, что ячейки пронумерованы числами 1, 2, …, n. Пусть σ_i – площадь i-й ячейки и m_i, M_i – соответственно точная нижняя и точная верхняя границы множества значений функции f на i-й ячейке. Составим две суммы Дарбу:

нижнюю сумму $s_T = s_T(f) = \sum\limits_{i=1}^{n} m_i \sigma_i$

и верхнюю сумму $S_T = S_T(f) = \sum\limits_{i=1}^{n} M_i \sigma_i$.

Очевидно, что $s_T \le S_T$. Докажем, как и в главе 4, что справедливо более общее неравенство

$$s_{T_1} \le S_{T_2}, \tag{6.1}$$

где T_1 и T_2 – любые два декартовых разбиения области P.

Для этой цели рассмотрим разбиение T, которое будем называть *совмещением разбиений* T_1 и T_2. Чтобы получить разбиение T, нужно взять все прямые (параллельные осям координат), задающие разбиения T_1 и T_2; эти прямые и будут порождать разбиение T. Ясно, что разбиение T является «измельчением» как разбиения T_1, так и разбиения T_2: каждая ячейка A разбиения T целиком содержится в некоторой ячейке B разбиения T_1 и некоторой ячейке C разбиения T_2.

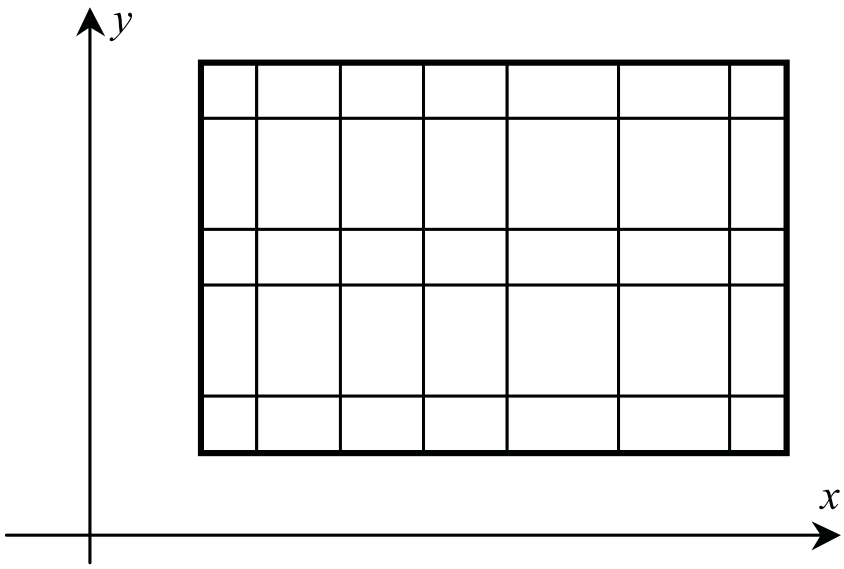

Рис. 6.1

Точную нижнюю (соответственно верхнюю) границу множества значений функции f на произвольном множестве D обозначим $m(D)$ (соответственно $M(D)$). В силу вышесказанного имеем

$$m(A) \geq m(B), \quad m(A) \geq m(C),$$
$$M(A) \leq M(B), \quad M(A) \leq M(C).$$

Из этих неравенств (точнее, из первого и последнего) очевидным образом следует $s_{T_1} \leq s_T$, $S_T \leq S_{T_2}$, так что имеем

$$s_{T_1} \leq s_T \leq S_T \leq S_{T_2}.$$

Отсюда и следует (6.1).

Определение. *Если для ограниченной функции $f(x, y)$, заданной на прямоугольнике $P = \{a \le x \le b, c \le y \le d\}$, существует* ***единственное*** *число J, разделяющее верхние и нижние суммы Дарбу для всевозможных декартовых разбиений прямоугольника P,*

$$s_{T_1} \le J \le S_{T_2}$$

(T_1 и T_2 – любые два разбиения), то это число называется ***двойным интегралом от функции f по P*** *и обозначается*

$$\iint_P f(x, y)dxdy.$$

При этом функция f называется ***подынтегральной функцией***.

Может случиться, что подынтегральная функция f равна нулю вне некоторого меньшего прямоугольника $P' \subset P$. Если для такой функции требуется найти двойной интеграл по P, то область интегрирования может быть сужена до P':

$$\iint_P f(x, y)dxdy = \iint_{P'} f(x, y)dxdy. \tag{6.2}$$

Действительно, любое разбиение T прямоугольника P можно продолжить, добавив к набору разбивающих прямых продолжения сторон P'. Из свойств сумм Дарбу следует, что при определении двойного интеграла от f по P можно ограничиться только разбиениями такого вида. Далее, если T' – разбиение P' прямыми, входящими в разбиение T, то очевидно, что суммы Дарбу $s_T(f)$, $S_T(f)$ отличаются от сумм Дарбу $s_{T'}(f)$, $S_{T'}(f)$ только наличием нескольких дополнительных слагаемых, равных 0. Поэтому из равенств

$$s_T(f) = s_{T'}(f), \;\; S_T(f) = S_{T'}(f)$$

вытекает соотношение (6.2).

В дальнейшем для сокращения формулировок, будем под “прямоугольниками” понимать только прямоугольники интересующего нас вида $\{a \le x \le b,\ c \le y \le d\}$.

Определение. *Пусть ограниченная функция* $f(x, y)$ *задана на ограниченном множестве* G. *Расширим область определения* $f(x, y)$, *положив* $f(x, y) = 0$ *вне* G. *Функция* $f(x, y)$ *называется* ***интегрируемой по G***, *если она интегрируема по минимальному прямоугольнику* P', *содержащему* G. *Двойной интеграл от* f *по* G *определяется формулой*

$$\iint\limits_{G} f(x,y)dxdy = \iint\limits_{P'} f(x,y)dxdy, \tag{6.3}$$

множество G *при этом называется* ***областью интегрирования***.

Рассмотрим прямоугольник P, содержащий область интегрирования G. Очевидно, что P содержит и минимальный прямоугольник P', а функция $f(x, y) = 0$ вне P'. Отсюда с учетом (6.2) вытекает важный вывод:

в определении двойного интеграла вместо минимального прямоугольника P' *в формуле* (6.3) *можно использовать любой прямоугольник* P, *содержащий область интегрирования* G.

2. Геометрический смысл двойного интеграла

Пусть $f(x,y) \ge 0$ в области $P = \{a \le x \le b,\ c \le y \le d\}$. Для любой области $G \subseteq P$ определим тело

$$C(f,G) = \{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in G,\ 0 \le z \le f(x,y)\}.$$

С геометрической точки зрения $C(f,G)$ – столбик с основанием G, ограниченный сверху поверхностью $z = f(x,y)$ (рис. 6.2). Объем тела $C(f,G)$ обозначим $V(f,G)$.

Рассмотрим декартово разбиение T прямоугольной области P на ячейки $P_1, P_2, \ldots, P_n$ и соответствующие суммы Дарбу $\sum m_i\sigma_i$ и $\sum M_i\sigma_i$, где $\sigma_i = \sigma(P_i)$ (площадь ячейки P_i).

Слагаемое $m_i\sigma_i$ в нижней сумме Дарбу равно объему параллелепипеда, в основании которого лежит P_i, а высота равна m_i.

Аналогично слагаемое $M_i\sigma_i$ в верхней сумме Дарбу равно объему параллелепипеда, в основании которого лежит P_i, а высота равна M_i.

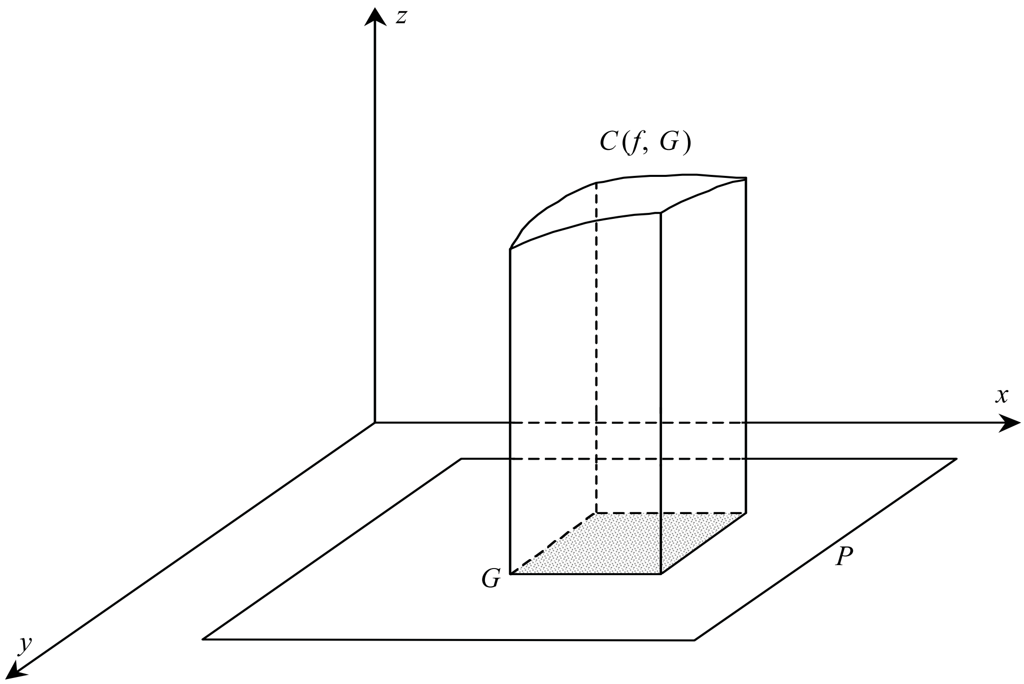

Рис. 6.2

Ясно, что тело $C(f, P_i)$ заключено между двумя этими параллелепипедами. Отсюда следует, что $m_i\sigma_i \le V(f, P_i) \le M_i\sigma_i$. Суммируя по всем ячейкам разбиения T, получаем двойное неравенство

$$s_T(f) \le V(f, P) \le S_T(f).$$

Отсюда с учетом того, что интеграл $\iint\limits_P f(x, y)dxdy$ – единственное число, разделяющее верхние и нижние суммы Дарбу, вытекает соотношение

$$\iint\limits_P f(x, y)dxdy = V(f, P).$$

Аналогичное равенство справедливо и для интеграла по области G произвольной формы.

Действительно, продолжая f нулем на дополнении G до объемлющего прямоугольника $P \supset G$, имеем

$$\iint\limits_G f(x,y)dxdy = \iint\limits_P f(x,y)dxdy = V(f,P) = V(f,G).$$

Таким образом, мы пришли к следующему заключению: *для неотрицательной функции f интеграл от f по области G равен объему тела, ограниченного снизу областью G, а сверху – поверхностью $z = f(x, y)$ (графиком функции).*

3. Свойства двойного интеграла

Рассмотрим основные свойства двойного интеграла.

1. *Если существует интеграл от функции $f \equiv 1$ по G, то его значение совпадает с площадью области интегрирования*:

$$\iint\limits_G 1dxdy = \sigma(G), \tag{6.4}$$

где $\sigma(G)$ – площадь G.

Доказательство. Фиксируем прямоугольник $P \supset G$ и рассмотрим какое-либо его разбиение T. Нижняя сумма Дарбу постоянной функции $f \equiv 1$ имеет вид $s_T = \sum\limits_i m_i\sigma_i$, где $m_i = 1$, если i-я ячейка P_i целиком содержится в G, и $m_i = 0$, если P_i содержит точки, не принадлежащие G. Отсюда следует, что $s_T = \sum\limits_{P_i \subset G} \sigma_i$. Аналогично $S_T = \sum\limits_i M_i\sigma_i$, где $M_i = 1$, если $P_i \cap G \neq \varnothing$, и $M_i = 0$, если $P_i \cap G = \varnothing$. Таким образом, $S_T = \sum\limits_{P_i \cap G \neq \varnothing} \sigma_i$. Рассмотрим фигуры, составленные из ячеек разбиения, соответствующие нижней и верхней суммам Дарбу: $A_T = \bigcup\limits_{P_i \subset G} P_i$ и $B_T = \bigcup\limits_{P_i \cap G \neq \varnothing} P_i$. Поскольку $A_T \subset G \subset B_T$, имеем

$$s_T = \sigma(A_T) \leq \sigma(G) \leq \sigma(B_T) = S_T. \tag{6.5}$$

Завершая доказательство, заметим, что интеграл $\iint\limits_P f(x,y)dxdy$ – единственная точка, разделяющая множества верхних и нижних сумм Дарбу, поэтому $\iint\limits_G 1dxdy = \sigma(G)$.

З а м е ч а н и е. Ввиду некоторой проблематичности самого понятия площади множества мы должны здесь заметить следующее. Простейшим эталоном для измерения площади являются, конечно, прямоугольники, а вслед за ними – фигуры, составленные из конечного числа прямоугольников. Временно назовем такие фигуры *составными многоугольниками*. По определению, множество G на плоскости имеет площадь, равную $\sigma(G)$, если число $\sigma(G)$ является единственным разделяющим числом для площадей составных многоугольников, содержащих G, и площадей составных многоугольников, содержащихся в G. Другими словами (см. выше неравенства (6.5)), площадь $\sigma(G)$ есть интеграл $\iint_G 1dxdy$, если только этот интеграл существует. При таком подходе к понятию площади равенство (6.4) превращается по существу в тавтологию. Отметим также полезный для дальнейшего факт: *утверждение, что* $\sigma(G)=0$, *означает, что существуют составные многоугольники как угодно малой площади, содержащие* G.

2. *Справедливо равенство*

$$\iint\limits_G Cfdxdy = C\iint\limits_G fdxdy, \tag{6.6}$$

где C – const.

Д о к а з а т е л ь с т в о очевидно.

3. *Если функции f_1 и f_2 интегрируемы по области G, то справедливо равенство*

$$\iint\limits_G (f_1+f_2)dxdy = \iint\limits_G f_1dxdy + \iint\limits_G f_2dxdy. \tag{6.7}$$

Д о к а з а т е л ь с т в о. Пусть $P \supset G$ – прямоугольник, объемлющий область интегрирования. Как обычно, считаем, что f_1 и f_2 продолжены нулем на дополнении G до P. Каково бы ни было разбиение T прямоугольника P, для любой его i-й ячейки имеем два очевидных неравенства:

$$m_i(f) \geq m_i(f_1) + m_i(f_2),$$
$$M_i(f) \leq M_i(f_1) + M_i(f_2).$$

Отсюда следует, что известное нам неравенство $s_T(f) \leq S_T(f)$ для функции $f = f_1 + f_2$ можно продолжить:

$$s_T(f_1) + s_T(f_2) \leq s_T(f) \leq S_T(f) \leq S_T(f_1) + S_T(f_2). \qquad (6.8)$$

Зададим теперь любое $\varepsilon > 0$. Ввиду интегрируемости f_1 в области G найдем такое разбиение T_1 прямоугольника P, что

$$S_{T_1}(f_1) - s_{T_1}(f_1) < \frac{\varepsilon}{2}, \qquad (6.9)$$

а ввиду интегрируемости f_2 – такое разбиение T_2, что

$$S_{T_2}(f_2) - s_{T_2}(f_2) < \frac{\varepsilon}{2}. \qquad (6.10)$$

Рассмотрим теперь разбиение T, являющееся совмещением разбиений T_1 и T_2 (см. с.397). Поскольку T есть измельчение T_1 и T_2, имеем вслед за (6.9) и (6.10)

$$S_T(f_1) - s_T(f_1) < \frac{\varepsilon}{2},\ S_T(f_2) - s_T(f_2) < \frac{\varepsilon}{2},$$

поэтому

$$(S_T(f_1) + S_T(f_2)) - (s_T(f_1) + s_T(f_2)) < \varepsilon;$$

отсюда и из (6.8) следует, что числа $s_T(f_1 + f_2)$ и $S_T(f_1 + f_2)$ отличаются меньше, чем на ε, что означает существование интеграла $\iint\limits_P (f_1+f_2)dxdy.$ Последний заключен между $s_T(f_1+f_2)$ и $S_T(f_1+f_2)$, а, значит, ввиду (6.8) – между $s_T(f_1)+s_T(f_2)$ и $S_T(f_1)+S_T(f_2)$. Но в тех же пределах заключена, очевидно, и сумма $\iint\limits_G f_1 dxdy+\iint\limits_G f_2 dxdy.$ Следовательно,

$$\left|\iint\limits_G (f_1+f_2)dxdy-\left(\iint\limits_G f_1 dxdy+\iint\limits_G f_2 dxdy\right)\right|<\varepsilon.$$

Ввиду того, что число ε может быть выбрано как угодно малым, отсюда следует равенство (6.7).

4. *Если* f_1 *и* f_2 *– две функции, интегрируемые в* G, *и* $f_1\le f_2$, *то*

$$\iint\limits_G f_1 dxdy\le\iint\limits_G f_2 dxdy.$$

Доказательство очевидно.

5. *Если* $f\le C$ (const) *во всей области* G, *то*

$$\iint\limits_G f dxdy\le C\sigma(G).$$

Доказательство. Это свойство вытекает из предыдущего свойства.

6. *Если функция f интегрируема в областях* G, H *и* $G\cap H$, *то f интегрируема в* $G\cup H$ *и*

$$\iint\limits_{G\cup H} f dxdy=\iint\limits_G f dxdy+\iint\limits_H f dxdy-\iint\limits_{G\cap H} f dxdy.$$

Доказательство. Для произвольного множества M на плоскости Oxy назовем *индикатором M* функцию I_M, равную 1 на

M и равную 0 вне M. Для индикаторов областей G, H и $G \cap H$ справедливо, очевидно, равенство

$$I_{G \cup H} = I_G + I_H - I_{G \cap H}.$$

Считая, что $f = 0$ вне $G \cup H$, отсюда получаем

$$f = f \cdot I_G + f \cdot I_H - f \cdot I_{G \cap H}.$$

Интегрируя f по $G \cup H$, с учетом свойства 3) имеем

$$\iint_{G \cup H} f dxdy = \iint_{G \cup H} f \cdot I_G dxdy + \iint_{G \cup H} f \cdot I_H dxdy - \iint_{G \cup H} f \cdot I_{G \cap H} dxdy.$$

Для завершения доказательства осталось заметить, что

$$\iint_{G \cup H} f \cdot I_G dxdy = \iint_G f dxdy, \quad \iint_{G \cup H} f \cdot I_H dxdy = \iint_H f dxdy,$$
$$\iint_{G \cup H} f \cdot I_{G \cap H} dxdy = \iint_{G \cap H} f dxdy,$$

поскольку $f \cdot I_G = 0$ вне G, $f \cdot I_H = 0$ вне H, а $f \cdot I_{G \cap H} = 0$ вне $G \cap H$.

7. *Если функция f ограничена на множестве G нулевой площади, то*

$$\iint_G f dxdy = 0.$$

Доказательство. Пусть $P \supset G$ – минимальный прямоугольник, содержащий G; T – разбиение P; B_T – фигура, составленная из ячеек разбиения T, имеющих непустое пересечение с G. Поскольку площадь $\sigma(G) = 0$, для любого $\varepsilon > 0$ можно выбрать разбиение так, чтобы площадь $\sigma(B_T)$ была меньше ε. В силу ограниченности f на G найдется константа C, для которой выполняется двойное неравенство $-C \le f \le C$ во всем G. Отсюда вытекают неравенства для сумм Дарбу

$$-C\varepsilon \le s_T(f) \le S_T(f) \le C\varepsilon.$$

Следовательно, при $\varepsilon \to 0$ верхняя и нижняя суммы Дарбу также стремятся к нулю, что и означает равенство 0 интеграла от f по G.

8. *Свойство аддитивности интеграла. Если область интегрирования G функции f разбита на две части* $G = A \cup B$ *таким образом, что площадь* $A \cap B$ *равна нулю, а f интегрируема по A и по B, то*

$$\iint\limits_G f dx dy = \iint\limits_A f dx dy + \iint\limits_B f dx dy. \tag{6.11}$$

Доказательство немедленно следует из предыдущих свойств 6 и 7.

§ 6.2. Для каких функций существует двойной интеграл?

Теорема 6.1. *Если функция* $f(x, y)$ *определена и непрерывна в прямоугольнике* $P = \{a \le x \le b,\ c \le y \le d\}$, *то существует двойной интеграл* $\iint\limits_P f dx dy$.

Доказательство не приводим, так как оно полностью аналогично доказательству соответствующей теоремы для определенного интеграла.

Теорема 6.2. *Пусть G – ограниченная область, f – ограниченная функция на G,* Γ *– объединение границы G и множества точек разрыва f на G. Предположим, что площадь* Γ *равна нулю. Тогда существует интеграл* $\iint\limits_G f dx dy$.

Доказательство. Пусть P – прямоугольник, содержащий G. Зададим $\varepsilon > 0$ и выберем разбиение прямоугольника P так, чтобы общая площадь ячеек, покрывающих Γ, была меньше ε. Обозначим часть прямоугольника P, занятую этими ячейками, через F; остальную часть P обозначим Q (рис. 6.3).

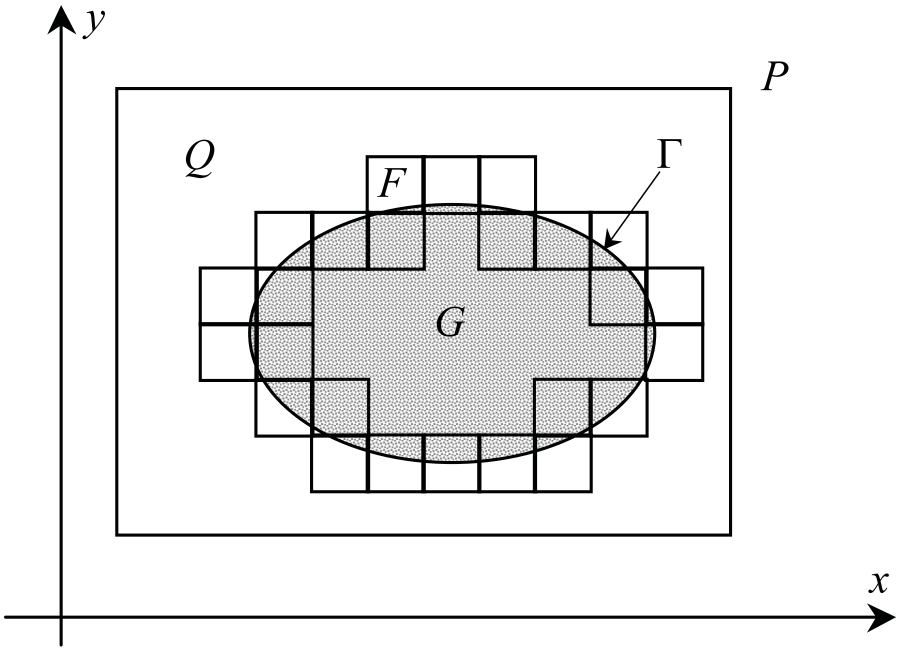

Рис. 6.3

Очевидно, Q есть объединение некоторого числа прямоугольников, поэтому в силу теоремы 6.1 и свойства аддитивности двойного интеграла существует интеграл $\iint\limits_G f dx dy$. Это означает, что можно найти разбиение T_2 области Q, для которого

$$\sum_{T_2} (M_i - m_i)\sigma_i < \varepsilon. \tag{6.12}$$

Заметим, что в прямоугольнике P справедливо $|f| < C$ (где C – const), следовательно, каждая разность $M_i - m_i < 2C$.

Составим новое разбиение прямоугольника P, включив в число разбивающих прямых все прямые, задающие T_1 и T_2. Для нового разбиения (обозначим его T) разность $S_T - s_T = \sum\limits_T (M_i - m_i)\sigma_i$ распадается на две части: сумму по прямоугольникам, входящим в Q, и сумму по ячейкам, составляющим F. Первая сумма согласно

(6.12) меньше ε; вторая сумма не превосходит $2C\varepsilon$, поскольку площадь F меньше ε. В итоге получим, что для разбиения T выполняется неравенство

$$S_T - s_T < \varepsilon + 2C\varepsilon.$$

Мы видим, что его левая часть стремится к 0 при $\varepsilon \to 0$, что означает интегрируемость f по G.

Можно доказать, что класс интегрируемых функций замкнут относительно сложения и умножения функций. Следующая теорема утверждает, что класс интегрируемых функций замкнут и относительно замены функции ее модулем.

Теорема 6.3. *Если функция $f(x,y)$ интегрируема по области G, то функция $|f(x,y)|$ также интегрируема по G.*

Доказательство. Пусть P – прямоугольник, содержащий G, T – некоторое разбиение P. Как обычно, считаем, что $f(x,y)$ равна 0 вне G. Рассмотрим точные нижние и верхние границы множеств значений функций $f(x,y)$ и $|f(x,y)|$ на i-й ячейке разбиения T:

$$m_i(f), M_i(f) \text{ и } m_i(|f|), M_i(|f|).$$

Если числа $m_i(f), M_i(f)$ одного знака или одно из них равно 0, то

$$M_i(|f|) - m_i(|f|) = M_i(f) - m_i(f). \qquad (6.13)$$

Если числа $m_i(f), M_i(f)$ имеют разные знаки, то, как нетрудно видеть,

$$M_i(|f|) - m_i(|f|) < M_i(|f|) < M_i(f) - m_i(f). \qquad (6.14)$$

Из (6.13) и (6.14) следует, что для любой ячейки выполняется неравенство

$$M_i(|f|) - m_i(|f|) \le M_i(f) - m_i(f). \qquad (6.15)$$

Умножив для каждого i неравенство (6.15) на площадь i-й ячейки и просуммировав по i, получим

$$S(|f|)-s(|f|)\le S(f)-s(f). \qquad (6.16)$$

Итак, если функция $f(x,y)$ интегрируема, то правая часть (6.16) стремится к 0 при измельчении разбиения T. Но тогда и левая часть (6.16) стремится к 0. Таким образом, функция $|f(x,y)|$ также интегрируема.

§ 6.3. Сведение двойного интеграла к повторному

Пусть область G – криволинейная трапеция, ограниченная слева и справа прямыми $x=a$ и $x=b$ ($a<b$), снизу – графиком функции $y=g_1(x)$, $x\in[a,b]$, сверху – графиком $y=g_2(x)$, $x\in[a,b]$. Тогда справедлива следующая теорема, сводящая нахождение двойного интеграла к последовательному вычислению двух интегралов от функций одной переменной.

Теорема 6.4. *Если функция $f(x,y)$ интегрируема в области G и при любом фиксированном x из $[a,b]$ существует интеграл $\int\limits_{g_1(x)}^{g_2(x)} f(x,y)dy$, то справедлива формула*

$$\iint\limits_G f(x,y)dxdy=\int\limits_a^b\left\{\int\limits_{g_1(x)}^{g_2(x)} f(x,y)dy\right\}dx. \qquad (6.17)$$

Доказательство начнем со случая, когда область G – прямоугольник вида $a\le x\le b$, $c\le y\le d$. Разобьем отрезки $[a,b]$ и $[c,d]$ на части точками

$$x_0=a<x_1<x_2<\ldots<x_n=b, \qquad (6.18)$$

$$y_0=c<y_1<y_2<\ldots<y_m=d. \qquad (6.19)$$

Данным разбиениям отрезков соответствует декартово разбиение T прямоугольника G прямыми линиями $x = x_i$ и $y = y_j$. Ячейки разбиения T будем нумеровать с помощью двух индексов:

$$G_{ij} = \{x_i \le x \le x_{i+1}, y_j < y < y_{j+1}\}.$$

Нетрудно видеть, что площадь ячейки G_{ij} равна $\Delta x_i \Delta y_j$, где $\Delta x_i = x_{i+1} - x_i$, $\Delta y_j = y_{j+1} - y_j$.

В пределах ячейки G_{ij} имеем

$$m_{ij} \le f(x, y) \le M_{ij}, \tag{6.20}$$

где m_{ij} и M_{ij} – точная нижняя и, соответственно, точная верхняя границы множества значений f в ячейке G_{ij}.

Интегрируя $f(x, y)$ по y при фиксированном значении $x \in [x_i, x_{i+1}]$, приходим к соотношению

$$m_{ij}\Delta y_j \le \int\limits_{y_j}^{y_{j+1}} f(x, y)dy \le M_{ij}\Delta y_j.$$

Суммируя подобные неравенства по $j = 0, 1, \ldots, n$, получим

$$\sum_j m_{ij}\Delta y_j \le \int\limits_a^b f(x, y)dy \le \sum_j M_{ij}\Delta y_j. \tag{6.21}$$

Рассмотрим функцию $F(x) = \int\limits_c^d f(x, y)dy$. Обозначив точную нижнюю границу ее значений на отрезке $[x_i, x_{i+1}]$ через m_i, а точную верхнюю границу – M_i, получим из (6.21)

$$\sum_j m_{ij}\Delta y_j \le m_i \le \sum_j M_{ij}\Delta y_j,$$

$$\sum_j m_{ij}\Delta y_j \le M_i \le \sum_j M_{ij}\Delta y_j.$$

Умножая все части обоих неравенств на Δx_i и суммируя по i ($i = 0$, $1, \ldots, n$), получим

$$s_T(f) \le \sum_i m_i \Delta x_i \le S_T(f), \tag{6.22}$$

$$s_T(f) \le \sum_i M_i \Delta x_i \le S_T(f), \tag{6.23}$$

где $s_T(f)$, $S_T(f)$ – нижняя и соответственно верхняя суммы Дарбу для функции $f(x, y)$ относительно разбиения T.

Неравенства (6.22) и (6.23) означают, что нижняя и верхняя суммы Дарбу для функции $F(x)$ относительно разбиения (6.18) отрезка $[a, b]$ заключены между $s_T(f)$ и $S_T(f)$. Выбирая разбиение T так, чтобы суммы $s_T(T)$ и $S_T(f)$ неограниченно сближались, получим, что тогда нижняя и верхняя суммы Дарбу для функции $F(x)$ будут неограниченно сближаться, стремясь к общему пределу, равному $\iint\limits_G f dx$. Это доказывает существование правой части формулы (6.17), а также и само равенство (6.17).

Пусть теперь G – криволинейная трапеция. Заключим ее в прямоугольник H вида $H = \{a \le x \le b, c \le y \le d\}$ (рис. 6.4) и продолжим f на весь прямоугольник H, полагая $f(x, y) = 0$, если $(x, y) \notin G$.

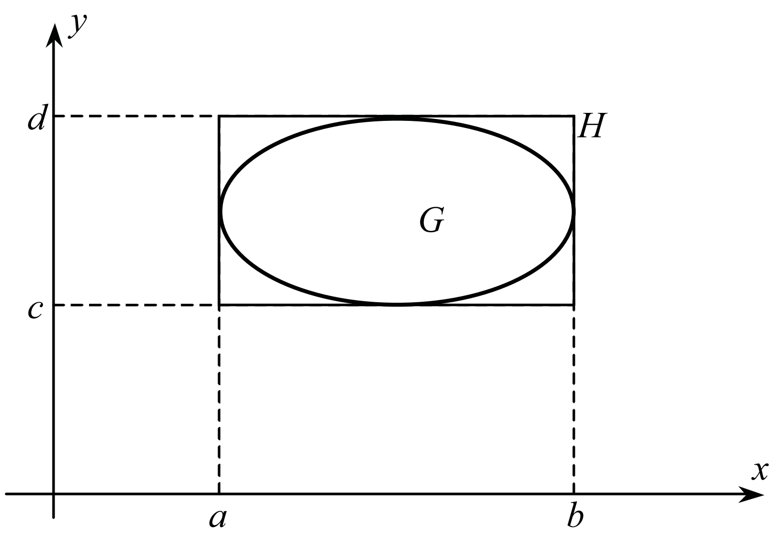

Рис. 6.4

Применив уже рассмотренный случай теоремы к интегралу от f по H, получим

$$\iint\limits_H f(x,y)dxdy = \int\limits_a^b \left\{ \int\limits_{g_1(x)}^{g_2(x)} f(x,y)dy \right\} dx.$$

Отсюда с учетом того, что $f(x,y)=0$ вне G, вытекает (6.17).

Замечание. Выражение $\int\limits_a^b \left\{ \int\limits_{g_1(x)}^{g_2(x)} f(x,y)dy \right\} dx$ называется *повторным интегралом* и часто записывается в виде

$$\int\limits_a^b dx \int\limits_{g_1(x)}^{g_2(x)} f(x,y)dy.$$

Пример 6.1. Вычислить интеграл от функции $f(x,y)=\dfrac{x^2}{y^2}$ по области G, ограниченной прямыми $x=2$, $y=x$ и гиперболой $xy=1$ (рис. 6.5).

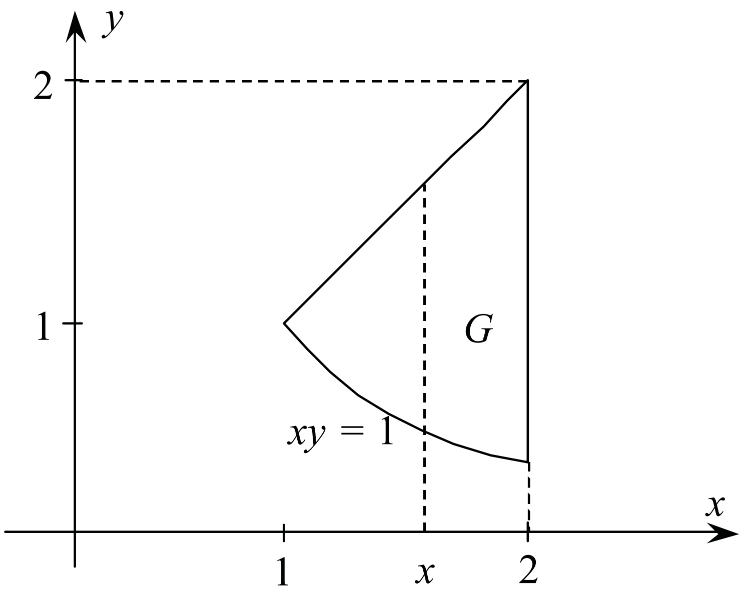

Рис. 6.5

Р е ш е н и е. Проекция области G на ось Ox есть отрезок $[1;2]$. При фиксированном значении x из этого отрезка соответствующая прямая, параллельная оси Oy, пересекает область G по отрезку $\frac{1}{x} \le y \le x$. Поэтому согласно формуле (6.17)

$$\iint_G f(x,y)dxdy = \int_1^2 dx \int_{1/x}^{x} \frac{x^2}{y^2} dy = \int_1^2 x^2 \left(-\frac{1}{y}\right)\Bigg|_{\frac{1}{x}}^{x} dx =$$

$$= \int_1^2 x^2 \left(x - \frac{1}{x}\right) dx = \int_1^2 (x^3 - x) dx = \left(\frac{x^4}{4} - \frac{x^2}{2}\right)\Bigg|_1^2 = \frac{9}{4}.$$

§ 6.4. Другой подход к понятию двойного интеграла

Пусть снова $f(x,y)$ – ограниченная функция двух переменных, заданная в ограниченной области G. Разобьем G с помощью некоторого набора кривых линий на части $G_1, G_2, ..., G_n$, называемые *ячейками*. В каждой ячейке G_i выберем точку (x_i, y_i) и составим сумму

$$\sum_{i=1}^{n} f(x_i, y_i)\sigma_i, \tag{6.24}$$

где σ_i – площадь ячейки G_i. Эта сумма зависит как от выбора разбиения области G, так и от выбора точек (x_i, y_i) в ячейках. Она называется *интегральной суммой*.

Обозначим само разбиение через T. Пусть d_i – максимум расстояния между точками ячейки G_i; число d_i называется *диаметром ячейки*. Наибольшее из чисел $d_1, d_2, ..., d_n$ назовем *диаметром разбиения* T и обозначим $d(T)$. На рис. 6.6 показано разбиение прямоугольной области на 4 ячейки; максимальное расстояние между точками ячейки G_1 есть PQ.

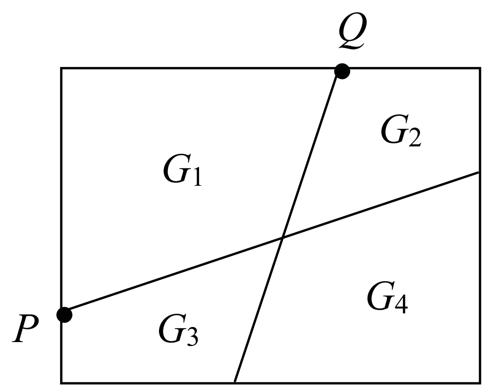

Рис. 6.6

Число $d(T)$ характеризует как бы “степень мелкости” разбиения T: чем меньше это число, тем мельче разбиение T. Если разбиения $T_1, T_2, \ldots$ выбираются так, что $d(T_i) \to 0$ при $i \to \infty$, то будем говорить о *неограниченном измельчении* разбиения области G.

Определение. *Если при* $d(T) \to 0$, *т.е. при любом неограниченном измельчении разбиения области* G, *интегральная сумма* (6.24) *стремится к определенному пределу, то этот предел называется* ***двойным интегралом*** *от функции* $f(x, y)$ *по области* G *и обозначается* $\iint\limits_G f(x, y)dxdy$.

Справедлива следующая теорема.

Теорема 6.5. *Если существует двойной интеграл в смысле определения, данного в* § 6.1, *то существует и двойной интеграл в смысле последнего определения и обратно*; *при этом оба интеграла равны*.

Оставим эту теорему без доказательства.

§ 6.5. Замена переменных в двойном интеграле

С самого начала заметим, что в этом параграфе излагаются только основные идеи замены переменных; строгих доказательств мы здесь не приводим.

1. Предварительная формула замены переменных

Пусть имеется область G в плоскости Oxy, по которой мы можем интегрировать различные функции $f(x,y)$ (например, ограниченные и непрерывные). Пусть, далее, имеется другая область H в плоскости Ouv, по которой мы также можем интегрировать различные функции $g(u,v)$. Пусть, наконец, по некоторому правилу установлено взаимно однозначное соответствие между точками $P \in G$ и точками $Q \in H$:

$$P = \varphi(Q), \quad Q = \psi(P),$$

или в координатах:

$$\begin{cases} x = \varphi_1(u,v), \\ y = \varphi_2(u,v), \end{cases} \qquad \begin{cases} u = \psi_1(x,y), \\ v = \psi_2(x,y). \end{cases} \tag{6.25}$$

Мы хотели бы связать интеграл от функции $f(P)$ по области G с интегралом от функции $g(Q) = f(\varphi(Q))$ по области H.

Каждое разбиение области G на ячейки порождает соответствующее разбиение области H. При этом интегральная сумма для функции $f(P)$ может быть преобразована к виду

$$\sum_i f(P_i)\sigma(G_i) = \sum_i f(\varphi(Q_i))\frac{\sigma(G_i)}{\sigma(H_i)}\sigma(H_i), \tag{6.26}$$

где точка Q_i соответствует P_i, а $\sigma(G_i)$ и $\sigma(H_i)$ – площади соответствующих ячеек (в плоскостях Oxy и Ouv).

Отношение $\sigma(G_i)/\sigma(H_i)$ показывает, во сколько раз изменилась площадь ячейки H_i при переходе к соответствующей ячейке G_i. Допустим, что при неограниченном уменьшении размеров ячейки H_i, точнее, когда H_i стягивается в какую-то точку Q, отношение $\sigma(G_i)/\sigma(H_i)$ имеет предел. Мы обозначим его через $\tau(Q)$ и назовем *коэффициентом искажения площади* в точке Q при отображении φ.

Допустим, что $\tau(Q)$ – непрерывная функция точки Q в области H. Тогда можно записать

$$\frac{\sigma(G_i)}{\sigma(H_i)} = \tau(Q_i) + \varepsilon_i,$$

где ε_i – как угодно малая величина, когда размеры ячейки H_i достаточно малы. Таким образом,

$$\sum_i f(P_i)\sigma(G_i) = \sum_i f(\varphi(Q_i))\frac{\sigma(G_i)}{\sigma(H_i)}\sigma(H_i) = \sum_i f(\varphi(Q_i))[\tau(Q_i)+\varepsilon_i]\sigma(H_i) =$$

$$= \sum_i f(\varphi(Q_i))\tau(Q_i)\sigma(H_i) + \sum_i \varepsilon_i f(\varphi(Q_i))\sigma(H_i). \qquad (6.27)$$

Если мы обозначим наибольшее из чисел $|\varepsilon_1|,\ldots,|\varepsilon_n|$ через ε, то вторая сумма в правой части (6.27) по модулю не будет превосходить

$$\varepsilon C \sum_{i=1}^{n} \sigma(H_i) = \varepsilon C\, \sigma(H),$$

где $|f| < C$ (функция f предполагается ограниченной), $\sigma(H)$ – площадь H. Если принять, что $\varepsilon \to 0$ при неограниченном измельчении ячеек G_i (а значит, и H_i), то из (6.27) получим в пределе следующее равенство

$$\iint_G f(P)dxdy = \iint_H f(\varphi(Q))\tau(Q)dudv. \qquad (6.28)$$

Эту формулу будем называть предварительной формулой замены переменных в кратном интеграле. Точная формулировка условий, при которых справедлива формула (6.28), будет дана позже, в пункте 3.

Для практического применения формулы (6.28) необходимо уметь вычислять коэффициент $\tau(Q)$ искажения площади при отображении φ. Рассмотрению этого вопроса и посвящен следующий пункт.

2. Вычисление $\tau(Q)$

Рассмотрим вначале случай, когда отображение φ, задающее взаимно однозначное соответствие между плоскостями Oxy и Ouv, является линейным:

$$\begin{cases} x = \varphi_1(u,v) = a_{11}u + a_{12}v, \\ y = \varphi_2(u,v) = a_{21}u + a_{22}v. \end{cases} \tag{6.29}$$

В плоскости Ouv рассмотрим ячейку ΔH (рис. 6.7) – прямоугольник $Q_0Q_1Q_2Q_3$, построенный на векторах $\overrightarrow{Q_0Q_1} = (\Delta u, 0)$ и $\overrightarrow{Q_0Q_2} = (0, \Delta v)$.

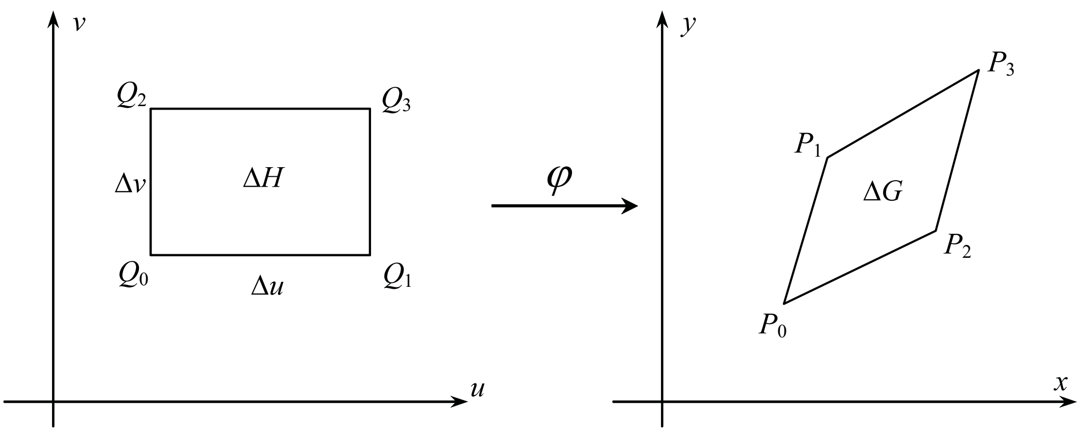

Рис. 6.7

Из линейности φ следует, что образ ячейки ΔH при этом отображении $\Delta G = \varphi(\Delta H)$ является параллелограммом с вершинами $P_i(x_i, y_i)$, соответствующими точкам $Q_i(u_i, v_i)$. При этом векторам $\overrightarrow{Q_0Q_1}$ и $\overrightarrow{Q_0Q_2}$ соответствуют векторы

$$\begin{aligned}
&\overrightarrow{P_0P_1} = (x_1 - x_0,\ y_1 - y_0) = \\
&= (\varphi_1(u_1, v_1) - \varphi_1(u_0, v_0),\ \varphi_2(u_1, v_1) - \varphi_2(u_0, v_0)) = \\
&= (\varphi_1(u_1 - u_0, v_1 - v_0),\ \varphi_2(u_1 - u_0, v_1 - v_0)) = (\varphi_1(\Delta u, 0),\ \varphi_2(\Delta u, 0)) = \\
&= (a_{11}\Delta u,\ a_{21}\Delta u), \\
&\overrightarrow{P_0P_2} = (\varphi_1(0, \Delta v),\ \varphi_2(0, \Delta v)) = (a_{12}\Delta v,\ a_{22}\Delta v).
\end{aligned}$$

Площадь $\sigma(\Delta H)$ ячейки ΔH равна $\Delta u \cdot \Delta v$. Площадь ячейки ΔG по известным формулам аналитической геометрии равна абсолютной величине определителя

$$\begin{vmatrix} a_{11}\Delta u & a_{12}\Delta v \\ a_{21}\Delta u & a_{22}\Delta v \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \Delta u \Delta v.$$

Таким образом,

$$\sigma(\Delta G) = \left| \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right| \cdot \sigma(\Delta H)$$

и, следовательно,

$$\tau(Q) = \lim \frac{\sigma(\Delta G)}{\sigma(\Delta H)} = \left| \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right|. \qquad (6.30)$$

Мы видим, что при линейном отображении коэффициент искажения $\tau(Q)$ есть постоянная величина (6.30).

Нетрудно видеть, что в случае линейного отображения (6.29) определитель $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ может быть записан следующим образом:

$$\begin{vmatrix} \dfrac{\partial \varphi_1}{\partial u} & \dfrac{\partial \varphi_1}{\partial v} \\ \dfrac{\partial \varphi_2}{\partial u} & \dfrac{\partial \varphi_2}{\partial v} \end{vmatrix}.$$

Разумеется, такой определитель можно составить и в случае нелинейного отображения φ, если только для функций φ_1 и φ_2 существуют частные производные. Этот определитель называется *якобианом* отображения φ и обозначается

$$\frac{\partial(\varphi_1, \varphi_2)}{\partial(u, v)} \quad \text{или} \quad \frac{\partial(x, y)}{\partial(u, v)}.$$

Итак, для случая линейного отображения φ мы получили формулу

$$\tau(Q) = \left| \frac{\partial(\varphi_1, \varphi_2)}{\partial(u, v)} \right|. \qquad (6.31)$$

Оказывается, формула (6.31) носит универсальный характер, т.е. она справедлива и в случае нелинейного отображения: *коэффициент* $\tau(Q)$ *искажения площади равен абсолютной величине якобиана в точке* Q.

Обоснование формулы (6.31) для общего случая мы опускаем.

Впрочем, сама идея доказательства формулы (6.31) достаточно проста. Фиксируем точку $Q_0(x_0, y_0)$ и рассмотрим наряду с отображением φ, "касательное отображение" $\Phi: (u,v) \to (x,y)$, определив его формулами

$$x = \varphi_1(Q_0) + \frac{\partial \varphi_1(Q_0)}{\partial u}(u - u_0) + \frac{\partial \varphi_1(Q_0)}{\partial v}(v - v_0),$$

$$y = \varphi_2(Q_0) + \frac{\partial \varphi_2(Q_0)}{\partial u}(u - u_0) + \frac{\partial \varphi_2(Q_0)}{\partial v}(v - v_0).$$

Главное для понимания связи между φ и Φ – это то, что по мере приближения точки Q к Q_0 образы этой точки при отображениях φ и Φ, т.е. точки $P = \varphi(Q)$ и $P' = \Phi(Q)$, стремятся друг к другу быстрее, чем Q стремится к Q_0:

$$\frac{\rho(P,P')}{\rho(Q_0,Q)} \to 0 \text{ при } Q \to Q_0 \text{ (} \rho \text{ – расстояние).} \tag{6.32}$$

Коэффициент искажения площади для отображения Φ в точке Q_0, очевидно, равен якобиану отображения φ, вычисленному в точке Q_0. Используя (6.32), можно показать, что коэффициент искажения площади для отображения φ в точке Q_0 тот же, что и для Φ.

3. Окончательный вид формулы замены переменной в кратном интеграле

С учетом (6.31) формула (6.28) приобретает теперь вид

$$\iint_G f(P)dxdy = \iint_H f(\varphi(Q)) \left| \frac{\partial(\varphi_1, \varphi_2)}{\partial(u,v)} \right| dudv. \tag{6.33}$$

Здесь P – переменная точка области G, а Q – переменная точка области H; сами области соответствуют друг другу посредством

(6.25). Формула замены переменной (6.33) в полностью координатной форме записывается следующим образом:

$$\iint_G f(x,y)dxdy = \iint_H f(\varphi_1(u,v),\varphi_2(u,v))\left|\frac{\partial(\varphi_1,\varphi_2)}{\partial(u,v)}\right|dudv. \qquad (6.34)$$

При выводе формулы (6.34) мы ограничились лишь основными идеями. Дадим точную формулировку теоремы о замене переменных.

Теорема 6.6. *Пусть H – ограниченная область в плоскости Ouv; $\varphi_1(u,v)$ и $\varphi_2(u,v)$ – функции, определенные в области H и имеющие в ней непрерывные частные производные. Предположим, что равенства*

$$x = \varphi_1(u,v), \quad y = \varphi_2(u,v)$$

задают взаимно однозначное соответствие между точками области H и точками некоторой области G в плоскости Oxy. Тогда для любой интегрируемой в области G функции $f(x,y)$ справедлива формула (6.34).

Наиболее типичные примеры применения замены переменной для вычисления двойного интеграла связаны с переходом от прямоугольных координат x, y на плоскости к полярным координатам r, φ. В этом случае

$$x = r\cos\varphi, \quad y = r\sin\varphi,$$

$$\frac{\partial(x,y)}{\partial(r,\varphi)} = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \varphi} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r,$$

следовательно, формула преобразования интеграла приобретает вид

$$\iint_G f(x,y)dxdy = \iint_H f(r\cos\varphi, r\sin\varphi)\, rdrd\varphi. \qquad (6.35)$$

Например, если G – круг $x^2 + y^2 \le R^2$, $\sigma(G)$ – его площадь, то

$$\sigma(G) = \iint\limits_G dxdy = \iint\limits_H rdrd\varphi,$$

где область H в плоскости r, φ задается неравенствами $0 \le r \le R$, $0 \le \varphi \le 2\pi$ (рис. 6.8).

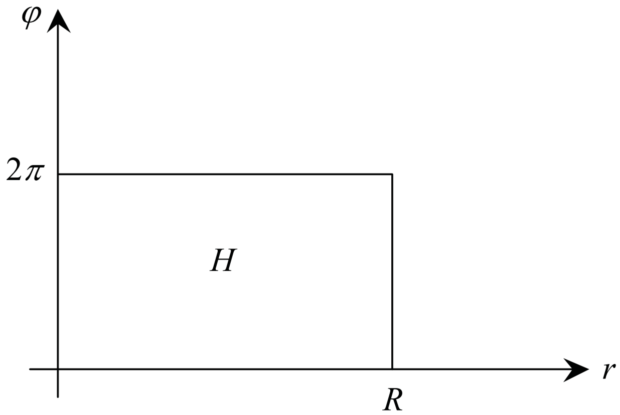

Рис. 6.8

Последний интеграл сводится к повторному

$$\int\limits_0^R \left\{ \int\limits_0^{2\pi} d\varphi \right\} rdr = \int\limits_0^R 2\pi rdr = \pi R^2,$$

что дает нам знакомую формулу для площади круга.

§ 6.6. Тройной интеграл

Определение тройного интеграла ничем принципиально не отличается от определения двойного интеграла. Тройной интеграл

$$\iiint\limits_D f(x, y, z) dxdydz$$

задается областью D в трехмерном пространстве $Oxyz$ и функцией f, определенной и ограниченной в D. Остаются в силе два подхода к понятию интеграла: как единственное разделяющее число для множества нижних и множества верхних сумм Дарбу и как предел интегральных сумм.

Почти дословно сохраняются и все свойства двойного интеграла, изложенные в § 6.1, необходимо лишь заменить термин “площадь” на “объем”. Например, свойство 1) для тройного интеграла формулируется так:

Если существует тройной интеграл от функции $f \equiv 1$ *по D, то его значение совпадает с объемом области интегрирования*:

$$\iiint\limits_D 1 dxdydz = V(D), \tag{6.36}$$

где $V(D)$ *– объем D.*

С учетом аналогичной замены сохраняется и достаточное условие интегрируемости (теорема 6.2):

Пусть D – ограниченная область в $\mathbb{R}^3$. *Тогда для интегрируемости ограниченной функции f по D достаточно, чтобы граница D и множество точек разрыва f в D имели нулевой объем.*

Заслуживает отдельного рассмотрения вопрос о сведении тройного интеграла к повторному.

Пусть ограниченная функция *f* задана в ограниченной области *D* трехмерного пространства *Oxyz*. Обозначим через *G* проекцию области *D* на координатную плоскость *Oxy*. Предположим, что существуют определенные в *D* функции $z_1(x,y)$ и $z_2(x,y)$, такие, что принадлежность точки $P(x,y,z)$ области *D* эквивалентна выполнению двух условий:

1) $(x,y) \in G$;

2) $z_1(x,y) \le z \le z_2(x,y)$.

Другими словами, область *D* ограничена снизу графиком функции $z_1(x,y)$, а сверху – графиком функции $z_2(x,y)$. В этом случае справедлива формула

$$\iiint\limits_D f(x,y,z)dxdydz = \iint\limits_G \left\{ \int\limits_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)dz \right\} dxdy. \tag{6.37}$$

Нетрудно видеть, что интеграл в фигурных скобках есть интеграл от функции $f(x,y,z)$ по отрезку – пересечению области D с прямой, параллельной оси Oz и пересекающей область G в точке (x, y). Таким образом, интеграл по трехмерной области приводится к интегралу по двумерной области от определенного интеграла с переменными пределами.

Пример 6.2. Вычислить интеграл

$$\iiint\limits_D \frac{dxdydz}{(1+x+y+z)^3},$$

где D – треугольная пирамида, ограниченная плоскостями $x=0, y=0, z=0$ и $x+y+z=1$ (рис. 6.9).

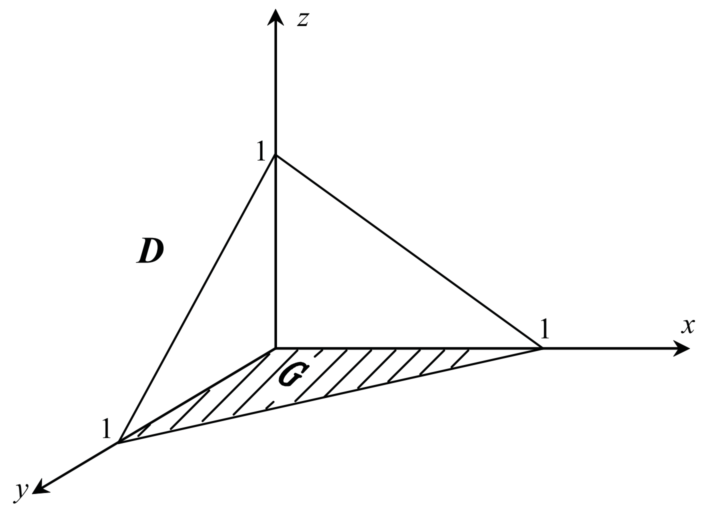

Рис. 6.9

Решение. Проекция G области D на плоскость Oxy – треугольник, выделенный на рис. 6.9 штриховкой. При фиксированной точке $(x,y)\in G$ соответствующая прямая, параллельная оси Oz, пересекает область D по отрезку $0 \le z \le 1-x-y$. Поэтому согласно формуле (6.37)

$$\iiint\limits_D \frac{dxdydz}{(1+x+y+z)^3} = \iint\limits_G \left\{ \int\limits_0^{1-x-y} \frac{dz}{(1+x+y)^3} \right\} dxdy.$$

Интеграл, заключенный в фигурных скобках, равен

$$-\frac{1}{2}\frac{1}{(1+x+y+z)^2}\Bigg|_{z=0}^{z=1-x-y}=\frac{1}{2(1+x+y)^2}-\frac{1}{8}.$$

Полученную функцию от x и y мы должны проинтегрировать по треугольной области G (рис. 6.10).

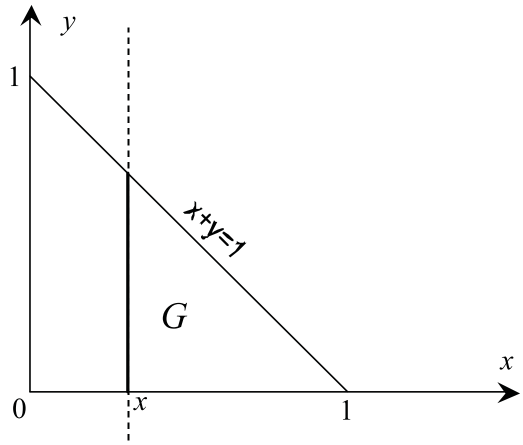

Рис. 6.10

Проекция этой области на ось Ox есть отрезок $0 \le x \le 1$. Фиксируем x на этом отрезке; соответствующая вертикальная прямая пересекает G по отрезку $0 \le y \le 1-x$. Отсюда

$$\iint_G\left(\frac{1}{2(1+x+y)^2}-\frac{1}{8}\right)dxdy=\int_0^1\left\{\int_0^{1-x}\left[\frac{1}{2(1+x+y)^2}-\frac{1}{8}\right]dy\right\}dx=$$

$$=\int_0^1\left(-\frac{1}{2(1+x+y)}-\frac{y}{8}\right)\Bigg|_{y=0}^{y=1-x}dx=\int_0^1\left(-\frac{1}{4}-\frac{1-x}{8}+\frac{1}{2(1+x)}\right)dx=$$

$$=\left(-\frac{x}{4}+\frac{(1-x)^2}{16}+\frac{1}{2}\ln(1+x)\right)\Bigg|_0^1=-\frac{1}{4}+\frac{1}{2}\ln 2-\frac{1}{16}=\frac{1}{2}\ln 2-\frac{5}{16}.$$

Формула замены переменных в тройном интеграле также аналогична формуле замены переменных в двойном интеграле. Если область D в пространстве $Oxyz$ находится во взаимно однозначном соответствии с областью Ω в пространстве $Ouvw$, причем это соответствие осуществляется с помощью формул $x = \varphi_1(u, v, w)$, $y = \varphi_2(u, v, w)$, $z = \varphi_3(u, v, w)$, то справедлива формула

$$\iiint_D f(x, y, z)dxdydz =$$

$$= \iiint_\Omega f(\varphi_1(u, v, w), \varphi_2(u, v, w), \varphi_3(u, v, w)) \left| \frac{\partial(\varphi_1, \varphi_2, \varphi_3)}{\partial(u, v, w)} \right| dudvdw, \quad (6.38)$$

где

$$\frac{\partial(\varphi_1, \varphi_2, \varphi_3)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial\varphi_1}{\partial u} & \frac{\partial\varphi_1}{\partial v} & \frac{\partial\varphi_1}{\partial w} \\ \frac{\partial\varphi_2}{\partial u} & \frac{\partial\varphi_2}{\partial v} & \frac{\partial\varphi_2}{\partial w} \\ \frac{\partial\varphi_3}{\partial u} & \frac{\partial\varphi_3}{\partial v} & \frac{\partial\varphi_3}{\partial w} \end{vmatrix}$$

– якобиан отображения φ.

Разумеется, вслед за тройным интегралом мы можем рассмотреть интегралы любой кратности; ввиду полной аналогии с двойными и тройными интегралами останавливаться на этом вопросе не будем.

§ 6.7. Несобственные кратные интегралы

В предыдущих параграфах мы изучали кратные *собственные* интегралы, т.е. интегралы от ограниченных функций по ограниченным областям. Определяемые ниже двойные несобственные интегралы вполне аналогичны несобственным однократным интегралам, рассмотренным в § 4.7, поскольку в обоих случаях условие ограниченности области и функции преодолевается одинаковым образом – за счет предельного перехода по расширяющейся области интегрирования.

По аналогии с однократными несобственными интегралами можно ввести два рода кратных несобственных интегралов:

1) от ограниченной функции по неограниченной области;

2) от неограниченной функции по ограниченной области.

Однако для упрощения изложения ограничимся рассмотрением лишь двойных несобственных интегралов первого рода.

Напомним, что множество $G \subset \mathbb{R}^2$ называется ограниченным, если существует число r, превосходящее расстояние от любой точки $M \in G$ до начала координат. Круги и прямоугольники являются, очевидно, примерами ограниченных множеств. Далее, если некоторое множество H содержится в ограниченном множестве G, то и H также является ограниченным. Поэтому для того, чтобы некоторое множество было ограниченным, необходимо и достаточно, чтобы существовал прямоугольник (круг), покрывающий данное множество.

Определение. *Пусть* $T = (a, +\infty)$ *– некоторый бесконечный интервал. Семейство* $\{D_t\}$ *ограниченных множеств* $D_t \subset \mathbb{R}^2 (t \in T)$ *называется* ***допустимым***, *если*:

1) *для любого* $t \in T$ *площадь границы* D_t *равна* 0;

2) $D_{t_1} \subset D_{t_2}$ *при* $t_1 < t_2$ (*условие монотонности*);

3) всякое ограниченное множество содержится в некотором D_t (условие исчерпывания $\mathbb{R}^2$).

Пример 6.3. Для любого $t > 0$ зададим круг C_t неравенством $x^2 + y^2 \le t^2$. Семейство $\{C_t\}$ является допустимым, так как условия 1), 2) и 3) выполняются очевидным образом.

Пример 6.4. Для любого $t > 0$ неравенствами $-t \le x \le t$, $-t \le y \le t$ зададим квадрат K_t. Пусть G – некоторое ограниченное множество. Согласно определению, найдется $r > 0$ такое, что $G \subset C_r$. Но в этом случае $G \subset K_r$. Следовательно, семейство квадратов $\{K_t\}$ также является допустимым.

Определение. *Пусть* $G \subset \mathbb{R}^2$ *– неограниченное множество,* $f(x, y)$ *– функция, интегрируемая по всякому подмножеству в* G

вида $G \cap D$, *где D – ограниченное множество с границей нулевой площади. Если для любого допустимого семейства* $\{D_t\}$ *предел*

$$\lim_{t\to\infty} \iint\limits_{G\cap D_t} f(x,y)dxdy \tag{6.39}$$

существует и не зависит от выбора семейства $\{D_t\}$, *то данный предел обозначается* $\iint\limits_{G} f(x,y)dxdy$ *и называется* ***несобственным двойным интегралом*** *от f по G.*

Таким образом, несобственный интеграл является пределом собственного интеграла и, как и всякий предел, может быть равным $+\infty$ или $-\infty$, или же быть числом. В последнем случае говорят, что несобственный интеграл *сходится*. Если же несобственный интеграл не является числом, то говорят, что он *расходится*. Несобственный интеграл *расходится* и в тех случаях, когда, несмотря на интегрируемость $f(x,y)$ по любой конечной части *G*, предел (6.39) не существует для какого-либо допустимого семейства $\{D_t\}$ или существует, но зависит от $\{D_t\}$.

Теорема 6.7. *Пусть* $G \subset \mathbb{R}^2$ *– неограниченное множество,* $f(x,y)$ *– неотрицательная функция, интегрируемая по всякому подмножеству в G вида* $G \cap D$, *где D – ограниченное множество с границей нулевой площади. Тогда для любого допустимого семейства* $\{D_t\}$ *имеем*

$$\iint\limits_{G} f(x,y)dxdy = \lim_{t\to\infty} \iint\limits_{G\cap D_t} f(x,y)dxdy. \tag{6.40}$$

Доказательство. Согласно определению несобственного интеграла достаточно доказать, что предел в правой части (6.40) существует и не зависит от $\{D_t\}$. Рассмотрим с этой целью функцию

$$I(t) = \iint\limits_{G\cap D_t} f(x,y)dxdy.$$

Поскольку при $t_1 < t_2$ выполняется неравенство

$$I(t_2) - I(t_1) = \iint\limits_{G\cap D_{t_2}\setminus D_{t_1}} f(x,y)dxdy \geq 0,$$

$I(t)$ – неубывающая функция и, следовательно, предел

$$\lim_{t\to\infty} I(t) = \lim_{t\to\infty} \iint\limits_{G\cap D_t} f(x,y)dxdy \tag{6.41}$$

существует для любого допустимого семейства $\{D_t\}$.

Рассмотрим другое допустимое семейство $\{\widetilde{D}_t\}$. Так как $\widetilde{D}_t$ – ограниченное множество, то для него найдется множество $D_{\tilde{t}}$ из семейства $\{D_t\}$, такое что $\widetilde{D}_t \subset D_{\tilde{t}}$. Из неотрицательности $f(x,y)$ вытекает неравенство

$$\iint\limits_{G\cap \widetilde{D}_t} f(x,y)dxdy \le \iint\limits_{G\cap D_{\tilde{t}}} f(x,y)dxdy. \tag{6.42}$$

Переходя к пределу в (6.42), получим

$$\lim_{t\to\infty} \iint\limits_{G\cap \widetilde{D}_t} f(x,y)dxdy \le \lim_{t\to\infty} \iint\limits_{G\cap D_{\tilde{t}}} f(x,y)dxdy. \tag{6.43}$$

Меняя ролями $\{\widetilde{D}_t\}$ и $\{D_t\}$, из (6.43) получаем обратное неравенство. Таким образом, предел правой части (6.40) не зависит от $\{D_t\}$, что и завершает доказательство.

П р и м е р 6.5. Найти несобственный интеграл

$$\iint\limits_{|y|\le x} \frac{dxdy}{(1+x^2)(1+y^2)}.$$

Р е ш е н и е. Пусть $\{K_t\}$ – семейство квадратов из примера 6.4. Нетрудно видеть, что пересечение области интегрирования с квадратом K_t задается неравенствами $|y| \le x \le t$. Имеем

$$\iint\limits_{|y|\le x} \frac{dxdy}{(1+x^2)(1+y^2)} = \lim_{t\to\infty} \iint\limits_{|y|\le x\le t} \frac{dxdy}{(1+x^2)(1+y^2)} = \lim_{t\to\infty} \int_0^t dx \int_{-x}^{x} \frac{dy}{(1+x^2)(1+y^2)} =$$

$$= \lim_{t\to\infty} \int_0^t \left.\frac{\operatorname{arctg} y}{(1+x^2)}\right|_{y=-x}^{y=x} dx = \lim_{t\to\infty} \int_0^t \frac{2\operatorname{arctg} x}{(1+x^2)} dx = \lim_{t\to\infty} \operatorname{arctg}^2 x\Big|_0^t = \frac{\pi^2}{4}.$$

Теорема 6.8. *Пусть G – неограниченное множество; $f(x,y)$ и $h(x,y)$ – функции, такие, что $0 \le f(x,y) \le h(x,y)$ на G. Предположим, что $f(x,y)$ и $h(x,y)$ интегрируемы по всякому подмножеству в G вида $G \cap D$, где D – ограниченное множество с границей нулевой площади. Тогда*:

1) *сходимость* $\iint\limits_G h(x,y)dxdy$ *влечет сходимость* $\iint\limits_G f(x,y)dxdy$;

2) *расходимость* $\iint\limits_G f(x,y)dxdy$ *влечет расходимость* $\iint\limits_G h(x,y)dxdy$.

Доказательство. Поскольку f и h – неотрицательные функции, интегралы

$$I_f = \iint\limits_G f(x,y)dxdy \text{ и } I_h = \iint\limits_G h(x,y)dxdy$$

равны неотрицательным числам или $+\infty$. Следовательно, расходимость I_f, I_h равносильна равенству этих интегралов $+\infty$.

Пусть $\{D_t\}$ – какое-либо допустимое семейство. Рассмотрим функции

$$I_f(t) = \iint\limits_{G \cap D_t} f(x,y)dxdy \text{ и } I_h(t) = \iint\limits_{G \cap D_t} h(x,y)dxdy.$$

Имеем $I_f(t) \to I_f, I_h(t) \to I_h$ при $t \to \infty$. Кроме того, для любого t выполняется неравенство $I_f(t) \le I_h(t)$. Переходя в этом неравенстве к пределу при $t \to \infty$, получим $I_f \le I_h$. Отсюда одновременно вытекают оба утверждения теоремы.

Теорема 6.9. *Пусть $G \subset \mathbb{R}^2$ – неограниченное множество, $f(x,y)$ – функция, определенная на G и интегрируемая по всякому подмножеству в G вида $G \cap D$, где D – ограниченное множество с границей нулевой площади. Тогда из сходимости интеграла* $\iint\limits_G |f(x,y)|dxdy$ *вытекает сходимость интеграла* $\iint\limits_G f(x,y)dxdy$.

Доказательство. Определим функции

$$f^{+}=\frac{|f|+f}{2} \text{ и } f^{-}=\frac{|f|-f}{2}.$$

По теореме 6.3 функция $|f(x,y)|$ интегрируема по тем же ограниченным областям, что и $f(x,y)$. Вследствие (6.6), (6.7) то же самое верно для f^{+} и f^{-}. Поскольку $0 \le f^{+} \le |f|$ и $0 \le f^{-} \le |f|$, по теореме 6.8 сходятся интегралы

$$\iint\limits_G f^{+}dxdy \text{ и } \iint\limits_G f^{-}dxdy.$$

Но тогда сходится и интеграл от функции $f = f^{+} - f^{-}$.

Замечание. Если сходимость несобственного интеграла $\iint\limits_G f(x,y)dxdy$ устанавливается как следствие сходимости интеграла $\iint\limits_G |f(x,y)|dxdy$, то говорят, что интеграл $\iint\limits_G f(x,y)dxdy$ *сходится абсолютно*. При этом в отличие от теории несобственного интеграла от функции одной переменной нового понятия не возникает, поскольку можно доказать, что сходимость несобственного кратного интеграла эквивалентна его абсолютной сходимости.

Пример 6.6. Покажем сходимость несобственного интеграла

$$\iint\limits_{|y|\le x} \frac{\cos x \cos y}{\left(1+x^2\right)\left(1+y^2\right)}dxdy. \tag{6.44}$$

Действительно, функция

$$\frac{|\cos x \cos y|}{\left(1+x^2\right)\left(1+y^2\right)}$$

непрерывна и, следовательно, интегрируема по любому ограниченному множеству с границей нулевой площади. С учетом сходимости интеграла

$$\iint\limits_{|y|\le x}\frac{dxdy}{\left(1+x^2\right)\left(1+y^2\right)}\ \text{(пример 6.5)}$$

из неравенств

$$0\le\frac{|\cos x\cos y|}{\left(1+x^2\right)\left(1+y^2\right)}\le\frac{1}{\left(1+x^2\right)\left(1+y^2\right)}$$

по теореме 6.8 получаем сходимость интеграла

$$\iint\limits_{|y|\le x}\frac{|\cos x\cos y|}{\left(1+x^2\right)\left(1+y^2\right)}dxdy.$$

Отсюда по теореме 6.9 вытекает сходимость интеграла (6.44).

Для вычисления несобственных интегралов, так же как и для собственных, применяются два основных приема:

1) сведение кратного интеграла к повторному интегралу;

2) замена переменных.

При этом имеются два возможных пути применения данных приемов. Первый путь – сведение кратного интеграла к повторному или замена переменных в собственном интеграле с последующим предельным переходом по расширяющейся области интегрирования; по такому пути мы пошли в примере 6.5. Второй, более короткий путь – применение несобственных аналогов формул (6.17), (6.34). Несобственным аналогом (6.17), является формула

$$\iint\limits_{\mathbb{R}^2}f\left(x,y\right)dxdy=\int\limits_{-\infty}^{\infty}\left\{\int\limits_{-\infty}^{\infty}f\left(x,y\right)dy\right\}dx,$$

записываемая обычно без внутренних фигурных скобок в виде

$$\iint\limits_{\mathbb{R}^2}f\left(x,y\right)dxdy=\int\limits_{-\infty}^{\infty}dx\int\limits_{-\infty}^{\infty}f\left(x,y\right)dy. \tag{6.45}$$

Равенство (6.45) выполняется при условии абсолютной сходимости повторного интеграла в правой части:

$$\int\limits_{-\infty}^{\infty}dx\int\limits_{-\infty}^{\infty}|f(x,y)|\,dy<+\infty. \tag{6.46}$$

Мы не будем доказывать (6.45), однако, заметим, что дополнительное условие (6.46) связано с тем, что сходимость несобственного однократного интеграла в отличие от двойного не означает его абсолютной сходимости.

Что же касается формулы замены переменных в несобственном двойном интеграле, то эта формула по виду совпадает с формулой (6.34). Отличие только в том, что область интегрирования не ограничена.

Пример 6.7. Найти несобственный интеграл

$$\iint\limits_{\mathbb{R}_+^2} \frac{dxdy}{(1+x+y)^3},$$

где, как обычно, $\mathbb{R}_+^2$ задается неравенствами $x \geq 0, y \geq 0$.

Решение. Мы не можем непосредственно применить (6.45), поскольку область интегрирования отлична от $\mathbb{R}^2$. Тем не менее можно воспользоваться стандартным приемом – расширить область интегрирования до $\mathbb{R}^2$, положив подынтегральную функцию равной 0 вне исходной области интегрирования. В результате получим

$$\iint\limits_{\mathbb{R}_+^2} \frac{dxdy}{(1+x+y)^3} = \int\limits_0^\infty dx \int\limits_0^\infty \frac{dy}{(1+x+y)^3} = \int\limits_0^\infty \left(-\frac{1}{2(1+x+y)^2} \right) \Bigg|_{y=0}^{y=\infty} dx =$$

$$= \int\limits_0^\infty \frac{dx}{2(1+x)^2} = \left(-\frac{1}{2(1+x)} \right) \Bigg|_0^\infty = \frac{1}{2}.$$

Пример 6.8. Найти несобственный интеграл

$$\iint\limits_{\mathbb{R}_+^2} e^{-(x+y)} \cos x \, dxdy. \tag{6.47}$$

Решение. Данный интеграл легко сводится к повторному

$$\iint\limits_{\mathbb{R}_+^2} e^{-(x+y)} \cos y \, dxdy = \int\limits_0^\infty dx \int\limits_0^\infty e^{-(x+y)} \cos y dy = \int\limits_0^\infty e^{-x} dx \cdot \int\limits_0^\infty e^{-y} \cos y dy.$$

Вычислив предварительно неопределенные интегралы

$$\int e^{-x}dx = -e^{-x} + C, \quad \int e^{-y}\cos y dy = \frac{1}{2}e^{-y}(\sin y - \cos y) + C,$$

найдем несобственный интеграл (6.47)

$$\int_0^\infty e^{-x}dx \int_0^\infty e^{-y}\cos y dy = (-e^{-x})|_0^\infty \cdot \left(\frac{1}{2}e^{-y}(\sin y - \cos y)\right)\Big|_0^\infty = \frac{1}{2}.$$

Поскольку подынтегральная функция знакопеременная, решение будет неполным, если не проверить условие (6.46) абсолютной сходимости повторного интеграла:

$$\int_0^\infty dx \int_0^\infty |e^{-(x+y)}\cos y| dy \le \int_0^\infty dx \int_0^\infty e^{-(x+y)}dy = \int_0^\infty e^{-x}dx \int_0^\infty e^{-y}dy = 1 < +\infty.$$

Пример 6.9. Найти несобственный интеграл

$$\iint_{\mathbb{R}^2} \frac{dxdy}{\left(1+(x+y)^2\right)\left(1+(x-y)^2\right)}.$$

Решение. Перейдем к новым переменным $u = x + y, v = x - y$. Так как

$$x = \frac{u+v}{2}, \; y = \frac{u-v}{2},$$

то якобиан

$$\begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Кроме того, при переходе к новым переменным область интегрирования остается прежней, поэтому

$$\iint_{\mathbb{R}^2} \frac{dxdy}{\left(1+(x+y)^2\right)\left(1+(x-y)^2\right)} = \frac{1}{2}\iint_{\mathbb{R}^2} \frac{dudv}{\left(1+u^2\right)\left(1+v^2\right)} =$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} \frac{du}{\left(1+u^2\right)} \int_{-\infty}^{\infty} \frac{dv}{\left(1+v^2\right)} = \frac{\pi^2}{2}.$$

Завершая главу о кратных интегралах, вычислим интеграл

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

имеющий приложения в различных разделах математики:

$$\int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = \left(\int\limits_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx \int\limits_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy\right)^{1/2} = \left(\iint\limits_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}}dxdy\right)^{1/2} =$$

$$= \left(\iint\limits_{\substack{0\le r\le\infty \\ 0\le\varphi\le 2\pi}} re^{-\frac{r^2}{2}}d\varphi dr\right)^{1/2} = \left(\int\limits_0^{2\pi} d\varphi \int\limits_0^{\infty} re^{-\frac{r^2}{2}}dr\right)^{1/2} = \left(2\pi\int\limits_0^{\infty} e^{-u}du\right)^{1/2} = \sqrt{2\pi}.$$

Так, в теории вероятностей применяется функция Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int\limits_0^x e^{-\frac{t^2}{2}}dt.$$

С учетом найденного значения интеграла, получим

$$\Phi(+\infty) = \frac{1}{\sqrt{2\pi}}\int\limits_0^{+\infty} e^{-\frac{x^2}{2}}dx = \frac{1}{2\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}dx = \frac{1}{2}.$$

Здесь мы использовали четность функции $e^{-x^2/2}$.

ГЛАВА 7

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

§ 7.1. Общие понятия и примеры

Различные вопросы математики, естествознания, экономики приводят к необходимости решения уравнений, содержащих в качестве неизвестной некоторую функцию $y(x)$ и ее производные до некоторого порядка n. С одним из наиболее простых таких уравнений, уравнением вида $y' = f(x)$, мы уже встречались в интегральном исчислении. Его решением является неопределенный интеграл от $f(x)$. Приведем другие примеры таких уравнений:

$$y' + 2y = x^2; \quad y''' + y' = 0; \quad y'' = xy.$$

Определение. *Уравнение, связывающее независимую переменную x с неизвестной функцией $y(x)$ и ее производными до некоторого порядка n включительно, называется **дифференциальным уравнением n-го порядка.***

Таким образом, приведенные выше уравнения являются примерами дифференциальных уравнений соответственно первого, третьего и второго порядков.

Любое дифференциальное уравнение может быть записано в виде

$$F(x, y, y', \ldots, y^{(n)}) = 0, \tag{7.1}$$

где F – некоторая заданная функция, x – независимая переменная, $y(x)$ – искомая функция, а $y'(x), \ldots, y^{(n)}(x)$ – ее производные.

Определение. ***Решением** дифференциального уравнения* (7.1) *называется функция $y(x)$, имеющая производные до n-го порядка включительно, и такая, что ее подстановка в уравнение* (7.1) *обращает его в тождество.*

Так, например, решением уравнения $y' = 2y$ является функция $y = e^{2x}$. Легко видеть, что решением этого уравнения будет также любая функция вида

$$y(x) = Ce^{2x}, \tag{7.2}$$

где C – произвольная постоянная. В дальнейшем мы покажем, что формулой (7.2) определяются все решения данного уравнения, или, как говорят, задается *общее решение* уравнения. Отметим, что в нашем случае общее решение зависит от произвольной постоянной C. Придавая ей определенные числовые значения, мы будем получать конкретные решения, или, как принято говорить, *частные решения*.

Рассмотрим в качестве еще одного примера уравнение

$$y'' = x. \tag{7.3}$$

Имеем

$$y' = \frac{x^2}{2} + C$$

и далее

$$y = \frac{x^3}{6} + C_1 x + C_2, \tag{7.4}$$

где C_1, C_2 – постоянные. Очевидно, что при любых значениях C_1 и C_2 полученная функция y будет решением уравнения (7.3), т.е. формула (7.4) задает общее решение уравнения (7.3). Оно, как видно, зависит от двух произвольных постоянных C_1 и C_2. Придавая им конкретные значения, мы получаем частные решения.

В дальнейшем понятия общего и частного решения будут уточнены. Однако ряд важных обстоятельств мы можем отметить уже сейчас, исходя из рассмотренных выше примеров:

1) дифференциальное уравнение имеет бесконечно много решений;

2) общее решение дифференциального уравнения зависит от произвольных постоянных, число которых равно порядку дифференциального уравнения;

3) частные решения получаются из общего путем придания конкретных значений указанным постоянным.

Отметим также, что процесс нахождения решения дифференциального уравнения принято называть *интегрированием* этого уравнения, а график решения – *интегральной кривой* данного уравнения.

§ 7.2. Дифференциальные уравнения первого порядка

1. Общий вид дифференциального уравнения первого порядка есть

$$F(x, y, y') = 0.$$

Если это уравнение можно разрешить относительно y', т.е. записать в виде

$$y' = f(x, y), \tag{7.5}$$

то говорят, что уравнение записано в *нормальной форме* (или в *форме Коши*).

Замечание 1. Уравнение (7.5) является частным случаем уравнения вида

$$N(x, y)dx + M(x, y)dy = 0, \tag{7.6}$$

называемого *уравнением в дифференциалах*. Уравнения (7.5) и (7.6) равносильны (т.е. имеют одинаковые решения) в области, где выполнено условие $M(x,y) \neq 0$. Уравнение (7.6) также называют *симметричной формой* записи дифференциального уравнения первого порядка. Это название обусловлено тем, что в этом уравнении не обозначено явным образом, какая из переменных является искомой функцией. Поэтому в зависимости от ситуации в качестве таковой может быть выбрана любая из переменных.

2. Рассмотрим *геометрическую трактовку* нахождения решений уравнения (7.5). Возьмем некоторую точку (x_0, y_0) из области определения D функции $f(x, y)$. Пусть $y = \varphi(x)$ – интегральная кри-

вая, проходящая через эту точку (т.е. $y_0 = \varphi(x_0)$). Из уравнения (7.5) вытекает, что

$$\varphi'(x_0) = f(x_0, y_0).$$

Таким образом, угловой коэффициент касательной к интегральной кривой, проходящей через точку (x_0, y_0), равен (при $x = x_0$) числу $f(x_0, y_0)$.

Построим теперь для каждой точки (x_0, y_0) из области определения D прямую, проходящую через эту точку и имеющую угловой коэффициент, равный $f(x_0, y_0)$. В этом случае принято говорить, что эта прямая определяет *направление* в точке (x_0, y_0), а на множестве D задано *поле направлений*. Таким образом, с геометрической точки зрения решить уравнение (7.5) означает найти кривую, касательная к которой в каждой точке совпадает с направлением поля в этой точке.

Пример 7.1. Рассмотрим дифференциальное уравнение

$$y' = \frac{y}{x}.$$

Правая часть уравнения определена на множестве D, состоящем из точек (x, y), где $x \neq 0$. Следовательно, поле направлений для данного уравнения можно построить на всей плоскости, кроме оси Oy. В каждой точке (x, y) угловой коэффициент y' касательной совпадает с угловым коэффициентом прямой, проходящей через данную точку и начало координат. Вдоль этих прямых угловой коэффициент постоянен, т.е. $\frac{y}{x} = C = \text{const}.$ Отсюда следует, что интегральными кривыми этого уравнения являются прямые $y = Cx$, где C – произвольная постоянная. На рис. 7.1 изображено поле направлений данного уравнения.

Заметим также, что решения этого уравнения удовлетворяют (при $y \neq 0$) условию

$$\frac{y'x}{y} = 1,$$

означающему, что эти кривые, и только они, имеют эластичность, равную 1.

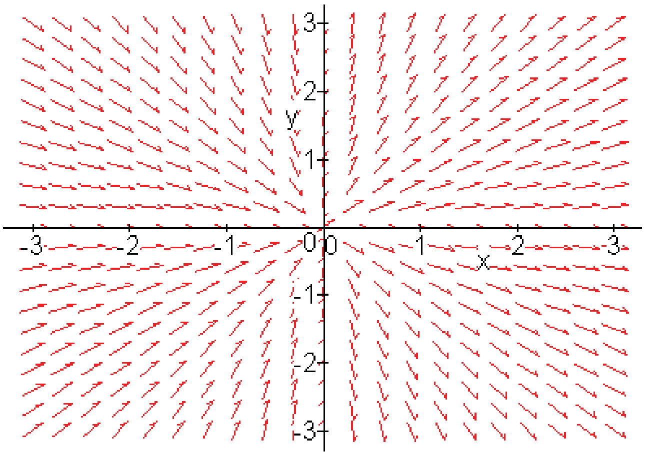

Рис. 7.1

3. Как мы уже отмечали выше, дифференциальное уравнение имеет бесконечно много решений. Чтобы из этого множества выделить какое-то конкретное решение, необходимо указать дополнительное условие; чаще всего оно задается в виде начального условия

$$y(x_0) = y_0. \tag{7.7}$$

Задача о нахождении решений дифференциального уравнения (7.5), удовлетворяющих начальному условию (7.7), называется *задачей Коши*.

В общем случае задача Коши может иметь единственное решение, бесконечно много решений либо вообще не иметь решений. Условия, при которых решение задачи Коши существует и единственно, формулируются в следующей теореме Коши.

Теорема 7.1 (***о существовании и единственности решения задачи Коши***). *Если в некоторой окрестности точки* (x_0, y_0) *функция* $f(x, y)$ *определена, непрерывна и имеет непрерывную частную производную* f_y', *то существует такая окрестность точки* (x_0, y_0), *в которой задача Коши* (7.5), (7.7) *имеет решение, притом единственное.*

Эту теорему мы примем без доказательства.

На основании теоремы Коши мы можем теперь уточнить понятия общего и частного решений.

Определение. *Если задача Коши* (7.5), (7.7) *имеет единственное решение, то это решение называется* ***частным решением*** *уравнения* (7.5).

Множество всех частных решений называется *общим решением* дифференциального уравнения.

Отметим, что в некоторых случаях процесс решения дифференциального уравнения приводит не к явному выражению $y = \varphi(x, C)$ для общего решения, а к некоторому соотношению вида

$$\Phi(x, y, C) = 0, \tag{7.8}$$

определяющему решение y как неявную функцию. Это соотношение называется *общим интегралом* дифференциального уравнения. Из соотношения (7.8) при помощи выбора постоянных C может быть получено уравнение любой интегральной кривой, т.е. может быть получено в неявном виде любое частное решение.

Процесс нахождения решения дифференциального уравнения называют *интегрированием* этого уравнения; обычно под этим понимают отыскание общего решения или общего интеграла.

Замечание 2. В теории дифференциальных уравнений под выражением вида $\int f(x)dx$ принято понимать не множество всех первообразных функции $f(x)$, а какую-либо одну фиксированную первообразную. Само выражение $\int f(x)dx$ часто называют *квадратурой* функции $f(x)$, а решить дифференциальное уравнение *в квадратурах* означает выразить его общее решение (или общий интеграл) в виде конечного числа квадратур от элементарных функций или их первообразных.

4. Выше нами была сформулирована теорема 7.1 о существовании и единственности решения задачи Коши для уравнения $y' = f(x, y)$. Условием, гарантирующим как существование решения, так и его единственность, является дифференцируемость функции $f(x, y)$. В отдельных точках это условие может нарушаться; через такие точки может не проходить ни одной интегральной кривой или проходить несколько интегральных кривых.

Точки, через которые не проходит ни одна интегральная кривая или проходит более одной интегральной кривой, называются *особыми точками* данного дифференциального уравнения.

Может случиться, что некоторая интегральная кривая уравнения состоит только из особых точек. Такая кривая называется *особым решением* уравнения.

Например, для уравнения

$$y' = 3\sqrt[3]{y^2}$$

функция $f(x, y) = 3y^{2/3}$ определена и непрерывна на всей плоскости Oxy. Ее частная производная $f'_y = \dfrac{2}{\sqrt[3]{y}}$ существует и непрерывна во всех точках, где $y \neq 0$, т.е. во всех точках, не принадлежащих оси Ox. Таким образом, через любую точку, не лежащую на оси Ox, проходит единственная интегральная кривая уравнения. Чтобы выяснить, как обстоит дело с точками оси Ox, проинтегрируем данное уравнение. Имеем

$$\frac{dy}{3\sqrt[3]{y^2}} = dx,$$

откуда следует

$$y^{1/3} + C = x,$$

или

$$y = (x - C)^3.$$

Итак, общее решение представляет собой семейство кубических парабол. Однако имеется еще одно решение $y(x) = 0$. Следовательно, через любую точку $(C, 0)$ оси Ox проходят по крайней мере две интегральные кривые: ось Ox и парабола $y = (x - C)^3$. Это показывает, что точки оси Ox являются особыми точками уравнения $y' = 3\sqrt[3]{y^2}$, а функция $y(x) = 0$ – особым решением.

Нетрудно установить, что через любую точку вида $(C, 0)$ проходит в действительности бесчисленное множество интегральных кривых. Любую из них можно составить из трех кусков: "нижней" половины параболы $y = (x - C_1)^3$, где C_1 – число, меньшее или равное C; отрезка C_1C_2 оси Ox, где $C_2 > C_1$, и "верхней" половины параболы $y = (x - C_2)^3$ (рис. 7.2).

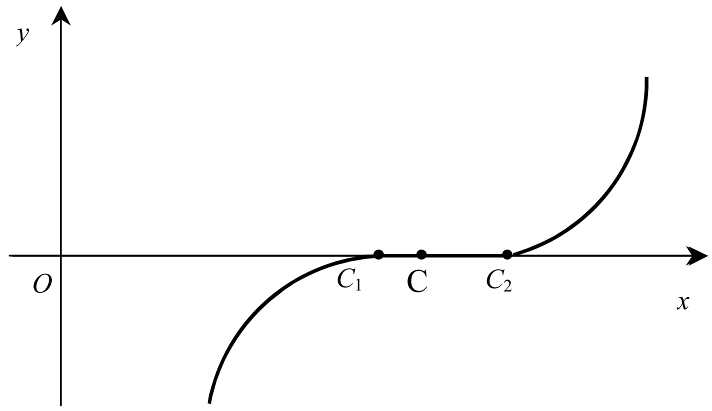

Рис. 7.2

Из этого примера можно понять, почему в формулировке теоремы Коши мы были вынуждены говорить о существовании и единственности решения лишь в некоторой окрестности начальной точки (x_0, y_0), а не во всей области существования функции $f(x, y)$.

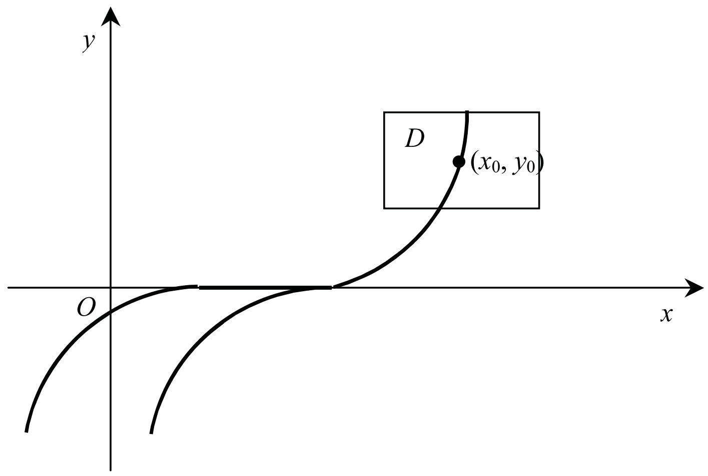

Рис. 7.3

В самом деле, пусть точка (x_0, y_0) не является особой для уравнения $y' = 3\sqrt[3]{y^2}$, т.е. $y \neq 0$. Если взять столь малую окрестность D

точки (x_0, y_0), чтобы она не пересекала ось Ox, то внутри такой окрестности через точку (x_0, y_0) будет проходить единственная кубическая парабола вида $y = (x - C)^3$. Однако если взять достаточно большую окрестность (например, всю плоскость Oxy), то окажется, что внутри такой окрестности через точку (x_0, y_0) проходит бесчисленное множество интегральных кривых. Любую из них можно составить указанным выше способом из трех кусков, один из которых проходит через точку (x_0, y_0) (рис. 7.3).

§ 7.3. Уравнения с разделяющимися переменными. Автономные уравнения

1. Одним из наиболее простых, но весьма важных (с точки зрения приложений) типов дифференциальных уравнений являются *уравнения с разделяющимися переменными*. Это дифференциальные уравнения вида

$$y' = f(x)g(y), \tag{7.9}$$

где $f(x)$ и $g(y)$ – непрерывные функции.

Запишем уравнение (7.9) в форме

$$\frac{dy}{dx} = f(x)g(y).$$

Для отыскания решения этого уравнения необходимо, как говорят, *разделить* в нем *переменные*, т.е. переписать уравнение следующим образом:

$$\frac{dy}{g(y)} = f(x)dx$$

в предположении, что в рассматриваемой области $g(y) \neq 0$. Теперь левая часть уравнения содержит только переменную y, а правая – только x. Интегрируя обе части этого уравнения, получим

$$\int \frac{dy}{g(y)} = \int f(x)dx + C.$$

Таким образом, найден общий интеграл уравнения (7.9).

Пример 7.2. Найти функцию, имеющую постоянную эластичность, равную k.

Решение. По условию задачи имеем

$$\frac{y'x}{y} = k,$$

$$\frac{dy}{dx} \cdot \frac{x}{y} = k.$$

Отсюда при естественном предположении $x \neq 0$ получим

$$\frac{dy}{y} = \frac{dx}{x} \cdot k.$$

Интегрируя обе части полученного равенства, находим

$$\ln|y| = k\ln|x| + \ln C,$$

откуда следует, что

$$y = C \cdot x^k.$$

Замечание 1. Обобщением уравнения (7.9) является уравнение в дифференциалах вида

$$f_1(x)g_1(y)dx + f_2(x)g_2(y)dy = 0,$$

где $f_1(x)$, $f_2(x)$, $g_1(y)$ и $g_2(y)$ – непрерывные функции. Его также называют *уравнением с разделяющимися переменными,* и к нему полностью применим метод интегрирования, описанный выше для уравнения (7.9).

2. Одним из важных частных случаев дифференциальных уравнений с разделяющимися переменными являются так называемые *автономные уравнения*. Это уравнения вида

$$y' = g(y). \tag{7.10}$$

Такие уравнения часто встречаются в различных вопросах экономической динамики. Обычно в качестве независимой переменной рассматривается время; его отсутствие в правой части уравнения (7.10) можно трактовать как неизменность законов, по которым развивается экономическая система в рассматриваемый промежуток времени.

Замечание 2. Если y^* – корень уравнения $g(y) = 0$, то $y = y^*$ (y – const) является решением уравнения (7.10). Такое решение называется *стационарным*.

Отметим еще одно интересное свойство, которым обладают решения автономного уравнения.

Теорема 7.2. *Если $y = \varphi(x)$ – решение автономного дифференциального уравнения, то $y = \varphi(x + C)$ также является решением этого уравнения.*

Доказательство. Пусть $y = \varphi(x)$ – решение уравнения (7.10), т.е.

$$\varphi'(x) = g(\varphi(x)).$$

Это равенство выполняется для любого x из области определения, поэтому мы можем заменить в нем x на $x + C$, в результате получим

$$\varphi'(x + C) = g(\varphi(x + C)). \tag{7.11}$$

Положим $\bar{y} = \varphi(x + C)$. Принимая во внимание равенство (7.11) и правило дифференцирования сложной функции, находим

$$\bar{y}' = \varphi'(x + C) \cdot (x + C)' = g(\varphi(x + C)) \cdot 1 = g(\bar{y}).$$

Это говорит о том, что функция $\bar{y} = \varphi(x + C)$ также является решением. Теорема доказана.

Замечание 3. Геометрическая трактовка данной теоремы заключается в том, что при параллельном переносе вдоль оси Ox интегральные кривые автономного уравнения переходят друг в друга.

Замечание 4. Если $g(y) \neq 0$, то общее решение автономного уравнения задается формулой $y = \varphi(x + C)$, где $\varphi(x)$ – произвольное частное решение.

3. К автономным уравнениям приводятся дифференциальные уравнения вида

$$y' = f(ky + lx + p), \tag{7.12}$$

где k,l,p – некоторые постоянные. Естественно предполагать, что k и l отличны от нуля. Тогда, применяя подстановку $z=ky+lx$, получаем, что $z' = ky' + l$. Следовательно, уравнение (7.12) равносильно уравнению $z' = kf(z+p)+l$, которое, в свою очередь, совпадает с (7.10), если положить $g(z) = kf(z+p)+l$. Нетрудно видеть, что к данному результату также можно было прийти, применив замену $z = ky + lx + p$.

Аналогичным образом к однородным уравнениям приводятся и уравнения вида

$$y' = f\left(\frac{ky + lx + p}{my + nx + q}\right), \tag{7.13}$$

где $k,\ l,\ p,\ m,\ n,\ q$ – некоторые действительные числа, удовлетворяющие условиям: $kn - lm = 0$, $k \neq 0$, $l \neq 0$. Из этих условий вытекает, что

$$\frac{m}{k} = \frac{n}{l} = a.$$

Поэтому после замены $z = ky + lx$ уравнение (7.13) преобразуется к виду

$$z' = kf\left(\frac{z+p}{az+q}\right) + l,$$

то есть к уравнению вида (7.10).

Пример 7.3. Рассмотрим дифференциальное уравнение

$$y' = \cos(2y + 2x + 6).$$

Р е ш е н и е. Делая замену $z = 2y + 2x + 6$, находим $z' = 2y' + 2$. Следовательно,

$$z' = 2\cos z + 2,$$

или

$$z' = 4\cos^2\frac{z}{2}. \tag{7.14}$$

Мы получили автономное дифференциальное уравнение. Возможны два случая.

1) $\cos\frac{z}{2} \neq 0$. Тогда, разделяя в (7.14) переменные, получаем

$$\frac{dz}{2\cos^2\frac{z}{2}} = 2dx.$$

Интегрируя это уравнение, находим

$$\operatorname{tg}\frac{z}{2} = 2x + C,$$

откуда

$$z = 2\operatorname{arctg}(2x + C) + 2\pi k. \tag{7.15}$$

2) $\cos\frac{z}{2} = 0$. Этот случай дает нам стационарные решения уравнения (7.14)

$$z = (2k + 1)\pi, \quad k \in \mathbb{Z}. \tag{7.16}$$

Теперь, возвращаясь в (7.15) и (7.16) к переменной $y = z/2 - x - 3$, получаем все решения исходного уравнения

$$y = \operatorname{arctg}(2x + C) + \pi k - x - 3,$$

$$y = \frac{\pi}{2} + \pi k - x - 3, \quad k \in \mathbb{Z}.$$

Пример 7.4. Решить дифференциальное уравнение

$$y' = \left(\frac{y - x}{2y - 2x + 1}\right)^2 + 1.$$

Решение. Положив $z = y - x$, имеем $z' = y' - 1$. Следовательно,

$$z' = \left(\frac{z}{2z + 1}\right)^2.$$

Очевидно, что $z = 0$ является решением этого уравнения. Если же $z \neq 0$, то, разделяя переменные, приходим к уравнению

$$\left(2 + \frac{1}{z}\right)^2 dz = dx.$$

Интегрируя его, находим

$$4z + 4\ln|z| - \frac{1}{z} = x + C.$$

Сделав обратную замену, получаем общий интеграл исходного уравнения

$$4y - 4\ln|y - x| - \frac{1}{(y - x)} = 5x + C.$$

Кроме того, случай $z = 0$ дает еще одно решение $y = x$.

§ 7.4. Математические модели экономической динамики с непрерывным временем

Рассматривая ниже некоторые математические модели роста, мы будем предполагать, что в качестве аргумента неизвестной функции выбрано время t. Данные модели, базирующиеся на дифференциальных уравнениях, называются моделями роста с непрерывным временем. Отметим, что существуют также дискретные аналоги этих моделей, они будут рассмотрены ниже.

1. Модель естественного роста (рост при постоянном темпе)

Пусть $y(t)$ – интенсивность (т.е. величина выпуска на единицу времени) выпуска продукции некоторого предприятия. Будем предполагать, что имеет место *аксиома о ненасыщаемости* потребителя, т.е. что весь выпущенный предприятием товар будет продан, а также что объем продаж не является столь высоким, чтобы существенно повлиять на цену товара p, которую ввиду этого мы будем считать фиксированной[1].

Для увеличения интенсивности выпуска $y(t)$ необходимо, чтобы *чистые инвестиции* $I(t)$ (т. е. разность между общим объемом инвестиций и амортизационными затратами) были больше нуля.

В случае $I(t) = 0$ общие инвестиции только лишь покрывают затраты на амортизацию, и уровень выпуска продукции остается неизменным.

Случай $I < 0$ приводит к уменьшению основных фондов и, как следствие, к уменьшению уровня выпуска продукции.

Таким образом, мы видим, что скорость увеличения интенсивности выпуска продукции является возрастающей функцией от I.

Пусть эта зависимость выражается прямой пропорциональностью, т.е. имеет место так называемый *принцип акселерации*:

$$my' = I \quad (m = \text{const}), \tag{7.17}$$

где m – *норма акселерации*. Пусть α – *норма чистых инвестиций*, т.е. часть дохода py, которая тратится на чистые инвестиции, тогда

$$I = \alpha py.$$

Подставляя выражение для I в (7.17), получаем

$$y' = \frac{\alpha p}{m} y,$$

или

$$y' = ky, \tag{7.18}$$

где $k = \alpha p/m$.

[1] Данное предположение соответствует модели конкурентной фирмы, рассмотренной в § 2.15.

Разделяя переменные в уравнении (7.18), имеем

$$\frac{dy}{y} = kdt.$$

Отсюда после интегрирования обеих частей находим

$$\ln|y| = kt + \ln|C|,$$

или, что то же самое,

$$y = Ce^{kt}. \quad (7.19)$$

Интегральная кривая уравнения (7.18) имеет вид, показанный на рис. 7.4.

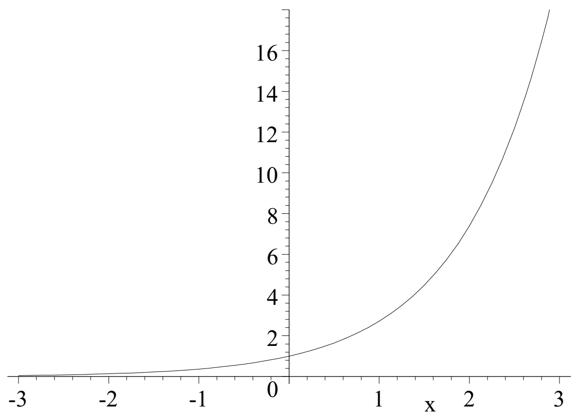

Рис. 7.4

Если $y(t_0) = y_0$, то из (7.19) следует, что $C = y_0e^{-kt_0}$; т.е.

$$y = y_0e^{k(t-t_0)}. \quad (7.20)$$

Уравнение (7.18) называется *уравнением естественного роста*. Этим уравнением описываются также динамика роста цен при постоянном темпе инфляции, процессы радиоактивного распада и процессы размножения бактерий. Данную модель целесообразно применять для исследования начальных этапов развития экономической системы и в течение ограниченного промежутка времени, поскольку, как это следует из уравнения (7.20), с течением времени y может принимать как угодно большие значения, что не может не сказаться на изменении цены (в данной модели мы ее предполагали постоянной).

2. Логистический рост

1. Рассмотрим более сложный случай по сравнению с предыдущим пунктом. Пусть $p = p(y)$ – убывающая функция $\left(\dfrac{dp}{dy} < 0\right)$, т.е. с увеличением выпуска будет происходить насыщение рынка, и цена будет падать[1]. Проведя аналогичные рассуждения, получим уравнение:

$$y' = kp(y)\cdot y, \tag{7.21}$$

где $k = \alpha/m$. Уравнение (7.21) представляет собой автономное дифференциальное уравнение. Так как $k > 0$, $p > 0$, $y > 0$, то из (7.21) следует, что $y(t)$ есть возрастающая функция ($y' > 0$). Исследуем $y(t)$ на выпуклость. Дифференцируя уравнение (7.21) по t, имеем:

$$y'' = ky'\left(\frac{dp}{dy}\cdot y + p\right) = ky'p\left(\frac{dp}{dy}\cdot\frac{y}{p} + 1\right),$$

$$y'' = ky'p\left(1 - \frac{1}{|e_y|}\right),$$

где $e_y = \dfrac{dy}{dp}\cdot\dfrac{p}{y} < 0$ – эластичность спроса[2]. Из последнего соотношения вытекает, что если спрос эластичен ($|e_y| > 1$), то $y'' > 0$ и интенсивность выпуска $y(t)$ – выпуклая функция, а если спрос неэластичен ($|e_y| < 1$), то $y'' < 0$ и $y(t)$ – вогнутая функция.

---

[1] Данное предположение выполняется, например, в случае монополии.

[2] Функция $y(p)$, обратная к $p(y)$, является функцией спроса на продукцию данного предприятия (отрасли), см. § 2.12.

2. Рассмотрим наиболее простой вид зависимости цены от выпуска – линейный, т.е. предположим, что

$$p(y) = b - ay \ (a, b > 0),$$

тогда уравнение (7.21) принимает вид

$$y' = k(b - ay)y. \qquad (7.22)$$

Из (7.22) легко получить, что $y' = 0$, если $y = 0$, или $y = \frac{b}{a}$, а также что $y'' > 0$ при $y < \frac{b}{2a}$ и $y'' < 0$ при $y > \frac{b}{2a}$. График $y(t)$ изображен на рис. 7.5.

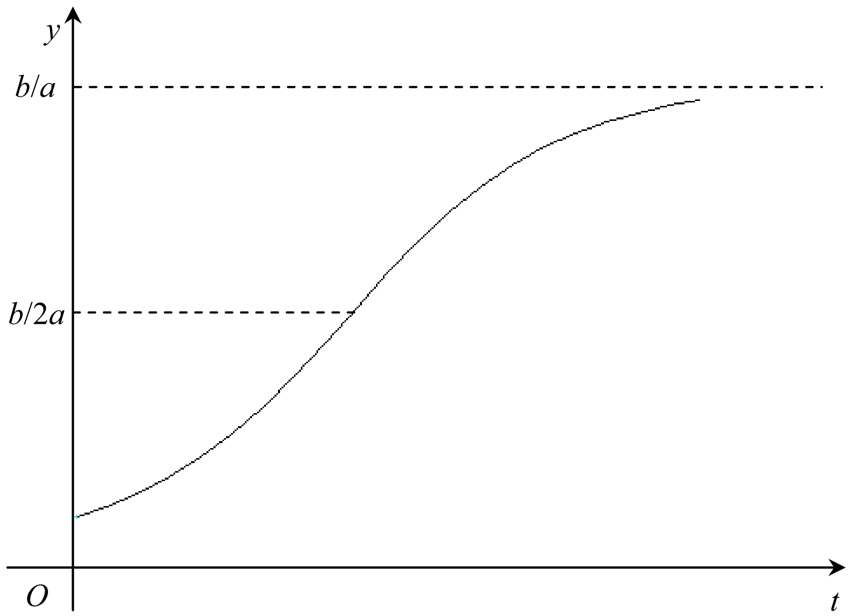

Рис. 7.5

В данном случае нетрудно получить и явное выражение для $y(t)$. Разделяя переменные в уравнении (7.22), находим

$$\frac{dy}{y(b - ay)} = kdt, \quad \text{или} \quad \frac{dy}{b}\left(\frac{1}{y} + \frac{a}{b - ay}\right) = kdt.$$

Проинтегрировав это соотношение, имеем:

$$\ln|y| - \ln|b - ay| = kbt + \ln C,$$

$$\frac{y}{b - ay} = Ce^{kbt}.$$

Отсюда получим, что

$$y = \frac{Cbe^{bkt}}{1 + Cae^{bkt}}.$$

Интегральная кривая уравнения (7.22) называется *логистической кривой*. Она также описывает некоторые модели распространения информации (рекламы), динамику эпидемий, жизненные циклы технологических укладов, процессы размножения бактерий в ограниченной среде обитания и др.

Замечание 1. Из графика логистической кривой видно, что при малых t логистический рост схож с естественным ростом, однако при больших t характер роста меняется, темпы роста замедляются и кривая асимптотически приближается к прямой $y = \frac{b}{a}$. Эта прямая является стационарным решением уравнения (7.22) и соответствует случаю $p(y) = 0$. Для уравнения (7.22) существуют решения и при $y > \frac{b}{a}$. График одного из таких решений показан на рис. 7.6. Но так как в этом случае $p(y) < 0$, то они не имеют экономической интерпретации.

Замечание 2. Интегральная кривая дифференциального уравнения

$$y' = k(t)(y - y_1)(y_2 - y), \qquad (7.23)$$

где $k(t) > 0$, $y_2 > y_1 > 0$, называется *обобщенной логистической кривой.*

Уравнение (7.23) используется для описания моделей роста в макроэкономике, в частности роста в условиях инфляции или роста с учетом технологического прогресса. При этом множитель $k(t)$ носит соответственно название *мультипликатор инфляции*, или *мультипликатор технологического прогресса*. Проделав выкладки, аналогичные тем, что были проведены выше для уравнения (7.22), нетрудно получить решения:

$$y = y_1, \quad y = y_2, \;\; y = \frac{y_1 + y_2 Ce^{P(t)}}{1 + Ce^{P(t)}},$$

где $P(x) = (y_2 - y_1)\int k(t)dt.$

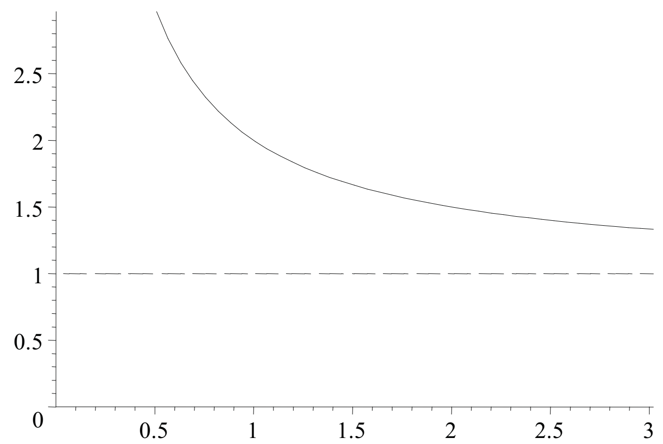

Рис. 7.6

3. Более реалистичной является модель, в которой скорость роста зависит не от дохода, а от прибыли. Пусть $S(y) = \alpha y + \beta$ – функция издержек, где αy – переменные издержки, а β – постоянные (α, β – const), тогда

$$y' = k(p(y)\cdot y - \alpha y - \beta). \tag{7.24}$$

Если теперь предположить, что $p(y) = b - ay$, то правая часть уравнения (7.24) представляет собой квадратный трехчлен относительно y с отрицательным коэффициентом перед y^2. В этом случае возможны три варианта.

а) $D < 0$. Следовательно, $y' < 0$. Издержки настолько велики, что это приводит к постоянному падению уровня производства и в конце концов – к банкротству (рис.7.7).

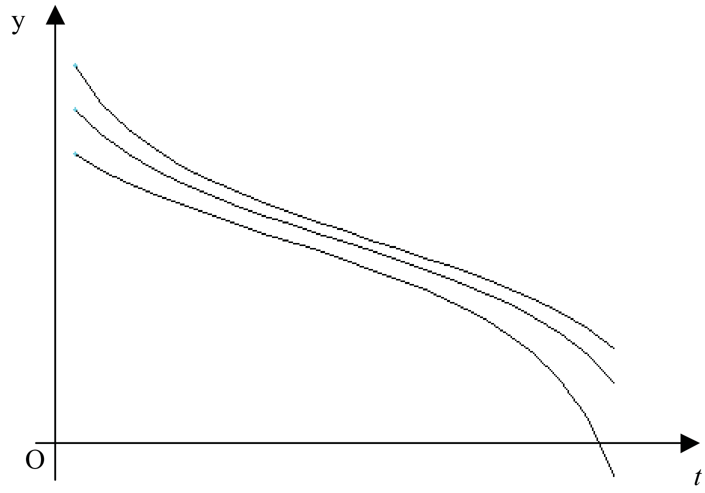

Рис. 7.7

б) $D = 0$. В этом случае $y' \geq 0$ и имеется одна стационарная кривая

$$y = y^* < \frac{b}{a}.$$

При этом интегральные кривые, удовлетворяющие начальному условию $y(t_0) = y_0 > y^*$ (тогда $y_0 < b/a$), будут асимптотически приближаться к y^* на $+\infty$, а удовлетворяющие условию $y(t_0) < y^*$, будут асимптотически приближаться к y^* на $-\infty$ (рис. 7.8).

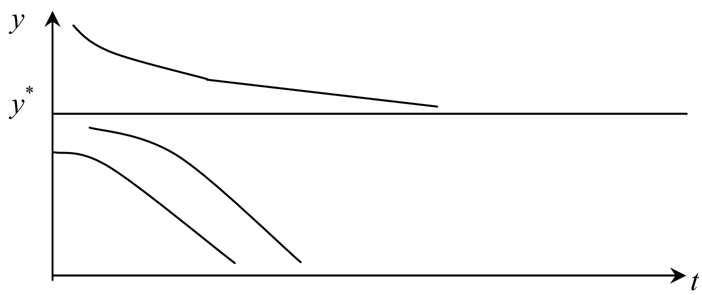

Рис. 7.8

в) $D > 0$. В этом случае существуют два стационарных решения: $y = y_1$, $y = y_2$.($0 < y_1 < y_2 < b/a$). При этом $y' > 0$ при $y_1 < y < y_2$ и $y' < 0$ при $y < y_1$ или $y > y_2$ (рис. 7.9).

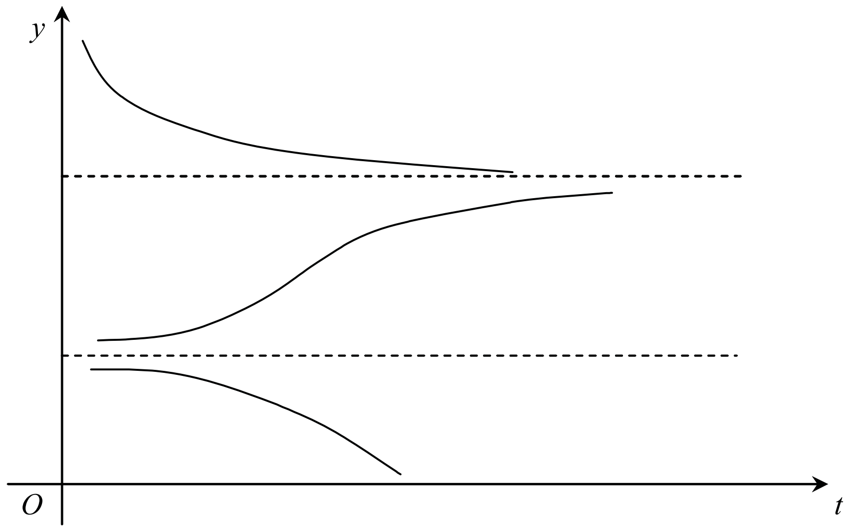

Рис. 7.9

Предлагаем читателю самостоятельно проинтегрировать уравнение (7.24) (в предположении $p(y) = b - ay$) и получить явные выражения для y.

3. Неоклассическая модель роста

Пусть $Y = F(K, L)$ – национальный доход, где K – объем капиталовложений (фондов), L – величина затрат труда, $F(K, L)$ – линейно-однородная производственная функция

$$F(tK, tL) = tF(K, L).$$

Обозначим через $f(k)$ производительность труда:

$$f(k) = \frac{F(K, L)}{L} = F\left(\frac{K}{L}, 1\right) = F(k, 1),$$

где $k = K/L$ – *фондовооруженность*. Производственная функция называется *неоклассической*, если

$$f'(k) > 0, \ \ f''(k) < 0.$$

Будем предполагать, что:

1) происходит естественный прирост трудовых ресурсов, т.е.

$$L' = \alpha L \ (\alpha = \text{const}); \tag{7.25}$$

2) инвестиции направлены как на увеличение производственных фондов, так и на амортизацию, т.е.

$$I = K' + \beta K \tag{7.26}$$

(β – норма амортизации).

Пусть v – норма инвестиций (т.е. $I = vY$), тогда в силу (7.26) имеем:

$$vY = K' + \beta K,$$

$$K' = vY - \beta K. \tag{7.27}$$

Из определения фондовооруженности вытекает, что

$$k' = \left(\frac{K}{L}\right)' = \frac{K'L - KL'}{L^2}.$$

Подставляя сюда значения для L' и K' из (7.25) и (7.27), находим

$$k' = \frac{(vY - \beta K)L - \alpha LK}{L^2},$$

т.е. $k' = \dfrac{vY}{L} - (\beta + \alpha)k.$

Учитывая, что $f = \dfrac{Y}{L}$, получаем

$$k' = vf(k) - (\alpha + \beta)k. \tag{7.28}$$

Уравнение (7.28) называется *уравнением неоклассического роста*.

Замечание 3. У автономного дифференциального уравнения (7.28) существует стационарное решение $k = k^*$. Действительно, так как $f''(k) < 0$, то графики $vf(k)$ и $(\alpha + \beta)k$ обязательно пересекутся (рис. 7.10).

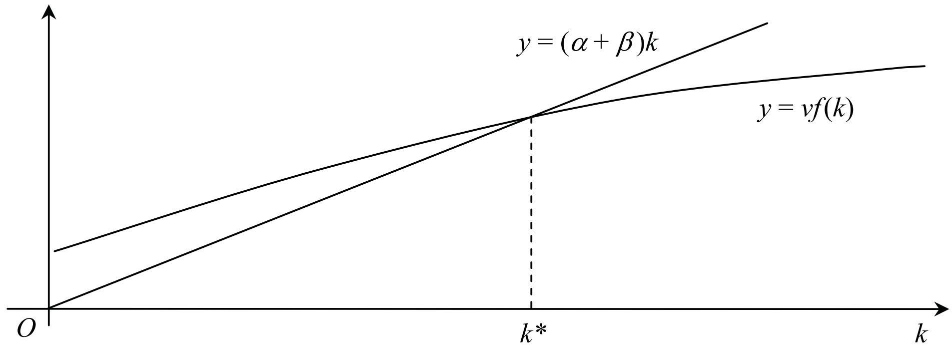

Рис. 7.10

Кроме того, так как функция $f'(k)$ непрерывная монотонно убывающая, то существует такое k_1, что $f'(k_1) = (\alpha + \beta)/m$ (т.е. у $k(t)$ существует точка перегиба). Итак, при $k > k^*$ имеем $k' < 0$; при $k < k^*$ будет $k' > 0$. При $k < k_1$ имеем $k'' > 0$, а при $k > k_1$ – $k'' < 0$.

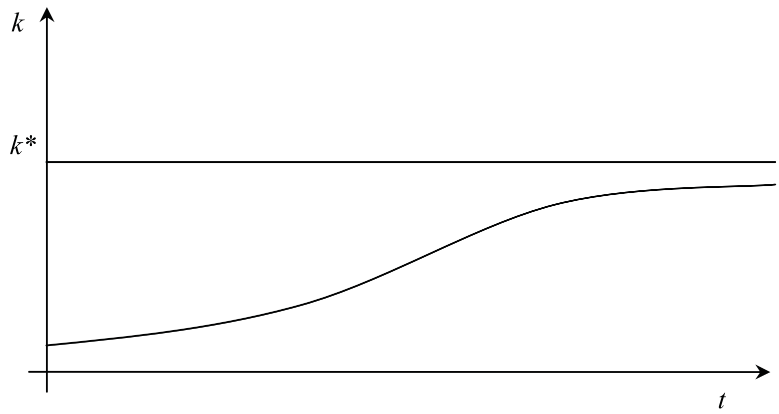

Рис. 7.11

Ввиду этого интегральная кривая уравнения (7.28) очень напоминает логистическую кривую (рис. 7.11).

§ 7.5. Однородные уравнения

1. **Определение.** *Дифференциальное уравнение первого порядка*

$$N(x, y)dx + M(x, y)dy = 0, \qquad (7.29)$$

где $N(x, y)$ и $M(x, y)$ – однородные функции одной и той же степени, называется ***однородным уравнением***.

Рассмотрим другие формы записи однородного уравнения. Дифференциальное уравнение, заданное в нормальной форме

$$y' = F(x, y), \qquad (7.30)$$

является однородным тогда и только тогда, когда функция $F(x, y)$ есть однородная функция нулевой степени. Кроме того, если в рассматриваемой области выполнено условие $x \neq 0$, то на основании определения однородной функции нулевой степени правая часть уравнения (7.30) может быть преобразована следующим образом:

$$F(x, y) = F\left(x \cdot 1, x \cdot \frac{y}{x}\right) = x^0 F\left(1, \frac{y}{x}\right) = F\left(1, \frac{y}{x}\right).$$

В силу этого уравнение (7.30) принимает вид

$$y' = f\left(\frac{y}{x}\right), \qquad (7.31)$$

где $f\left(\frac{y}{x}\right) = F\left(1, \frac{y}{x}\right)$.

Пример 7.5. Рассмотрим дифференциальное уравнение

$$(x^2 + y^2) \cdot dy - xy \cdot dx = 0. \qquad (7.32)$$

Оно является однородным, так как обе функции $N(x, y) = x^2 + y^2$ и $M(x, y) = -xy$ являются однородными второй степени. В области, не включающей в себя начало координат, данное уравнение может быть записано в нормальной форме:

$$y' = \frac{xy}{x^2 + y^2}\,.$$

Если дополнительно потребовать выполнение условия $x \neq 0$, то данное уравнение принимает вид

$$y' = \frac{\dfrac{y}{x}}{1 + \left(\dfrac{y}{x}\right)^2}\,. \tag{7.33}$$

2. Остановимся теперь на методе интегрирования однородного уравнения. Подстановкой

$$y(x) = xu(x)$$

оно приводится к уравнению с разделяющимися переменными. Продемонстрируем это на примере уравнения (7.31). Имеем

$$y' = u(x) + xu'(x).$$

Ввиду этого соотношение (7.31) принимает вид

$$u + xu' = f(u).$$

Разделяя переменные, получим

$$\frac{du}{f(u) - u} = \frac{dx}{x}\,.$$

Отсюда находим общий интеграл уравнения:

$$\int \frac{du}{f(u) - u} = \ln|x| + C.$$

После нахождения $u(x)$ необходимо вернуться к функции $y(x) = xu(x)$.

Замечание. Если существуют корни уравнения $f(u) = u$, то к найденным решениям добавляются стационарные решения $u(x) = u^*$, где u^* – любой корень уравнения $f(u) = u$.

Пример 7.6. Решить уравнение

$$y' = 3\left(\frac{y}{x}\right) - \left(\frac{y}{x}\right)^2 - 1.$$

Решение. Это уравнение является однородным. Выполнив замену $y(x) = xu(x)$, приходим к уравнению

$$u' = \frac{2u - u^2 - 1}{x}. \qquad (7.34)$$

Разделяя в нем переменные, получим при $u \neq 1$

$$\int \frac{du}{2u - u^2 - 1} = \int \frac{dx}{x},$$

или

$$\frac{1}{u-1} = \ln|x| + C,$$

$$u = (\ln|x| + C)^{-1} + 1.$$

Кроме этого у уравнения (7.34) имеется стационарное решение $u(x) = 1$. Таким образом, решением исходного уравнения являются функции

$$y(x) = \frac{x}{\ln|x| + C} + x; \;\; y(x) = x.$$

3. К однородным уравнениям приводятся дифференциальные уравнения вида

$$y' = f\left(\frac{ky + lx + p}{my + nx + q}\right),\ kn–lm \neq 0, \tag{7.35}$$

где $k,\ l,\ p,\ m,\ n,\ q$ – некоторые постоянные. Этого можно добиться, сделав замену переменных

$$t = x - \alpha; \quad z = y - \beta \tag{7.36}$$

и выбрав постоянные α и β таким образом, чтобы свободные члены в правой части выражения (7.35) стали равными нулю. Действительно, на основании (7.35), (7.36) имеем

$$\frac{dz}{dt} = \frac{dz}{dx} \cdot \frac{dx}{dt} = \left(\frac{dy}{dx} - 0\right) \cdot 1 = f\left(\frac{kz + lt + k\beta + l\alpha + p}{mz + nt + m\beta + n\alpha + q}\right).$$

Приравнивая к нулю свободные члены в числителе и знаменателе дроби, получаем следующую систему линейных алгебраических уравнений

$$\begin{cases} k\beta + l\alpha = -p, \\ m\beta + n\alpha = -q. \end{cases} \tag{7.37}$$

Определитель этой системы равен $kn - ml$ и в силу (7.35) отличен от нуля. Следовательно, у этой системы существует единственное решение и пара (β, α) определена однозначно. Таким образом, в результате проведенной замены уравнение (7.35) преобразуется к однородному уравнению

$$\frac{dz}{dt} = f\left(\frac{kz + lt}{mz + nt}\right).$$

Пример 7.7. Решить уравнение

$$y' = \frac{1}{2}\left(\frac{y + x + 3}{x + 2}\right)^2. \tag{7.38}$$

Р е ш е н и е. В данном случае замена очевидна и можно обойтись без решения системы (7.37). Положим $t = x + 2$, $z = y + 1$. Тогда уравнение (7.37) принимает вид

$$\frac{dz}{dt} = \frac{1}{2}\left(\frac{z+t}{t}\right)^2,$$

или

$$\frac{dz}{dt} = \frac{1}{2}\left(\frac{z}{t}+1\right)^2.$$

Применяя в полученном однородном уравнении подстановку $z = ut$, находим

$$\frac{du}{dt} = \frac{u^2+1}{2t}.$$

Разделяя переменные, имеем

$$\frac{du}{u^2+1} = \frac{dt}{2t}.$$

Интегрируя это уравнение, получаем

$$\operatorname{arctg} u = \frac{1}{2}\ln|t| + C.$$

Возвращаясь сначала к переменной z, а затем к переменным x и y, находим общий интеграл уравнения (7.38)

$$\operatorname{arctg}\frac{y+1}{x+2} = \frac{1}{2}\ln|x+2| + C.$$

§ 7.6. Линейные дифференциальные уравнения первого порядка

1. **Определение.** *Дифференциальное уравнение вида*

$$\alpha(x)y' + \beta(x)y + \gamma(x) = 0 \qquad (7.39)$$

называется ***линейным дифференциальным уравнением первого порядка***.

Предполагается, что, $\alpha(x)$, $\beta(x)$ и $\gamma(x)$ – непрерывные на некотором промежутке функции. Если $\alpha(x) \neq 0$, то уравнение (7.39) можно преобразовать следующим образом:

$$y' + p(x)y = f(x), \qquad (7.40)$$

где $p(x) = \dfrac{\beta(x)}{\alpha(x)}$; $f(x) = -\dfrac{\gamma(x)}{\alpha(x)}$.

Дифференциальное уравнение

$$y' + p(x)y = 0 \qquad (7.41)$$

называется *линейным однородным уравнением*, соответствующим уравнению (7.40).

Для отыскания общего решения неоднородного уравнения (7.40) применим *метод Лагранжа* (другое его название – *метод вариации постоянной*). Сначала найдем решение уравнения (7.41), которое представляет собой уравнение с разделяющимися переменными. Очевидно, что $y = 0$ является решением уравнения (7.41). При $y \neq 0$ имеем

$$\frac{dy}{y} = -p(x)dx.$$

Интегрируя это уравнение, получим

$$\ln|y| = -P(x) + \ln|C|,$$

где $P(x) = \int p(x)dx,$ а C – отличная от нуля постоянная. Из последнего уравнения находим общее решение уравнения (7.41):

$$y = Ce^{-P(x)}. \tag{7.42}$$

Здесь C – уже произвольная постоянная, так как решение $y = 0$ входит в (7.42) при $C = 0$.

Теперь заменим в формуле (7.42) постоянную C на некоторую (искомую) функцию $C(x)$, т.е. общее решение уравнения (7.40) будем искать в виде

$$y = C(x)e^{-P(x)}. \tag{7.43}$$

Из (7.43) следует

$$y' = C'e^{-P(x)} - Ce^{-P(x)}p(x). \tag{7.44}$$

Подставляя выражения для y и y' из (7.43) и (7.44) в (7.40), находим

$$C'e^{-P(x)} - Ce^{-P(x)}p(x) + p(x)Ce^{-P(x)} = f(x).$$

Отсюда получим, что

$$C' = f(x)e^{P(x)}.$$

Следовательно,

$$C(x) = \int f(x)e^{P(x)}dx. \tag{7.45}$$

Подставив в (7.43) выражение для $C(x)$, получим общее решение уравнения (7.40).

Итак, сформулируем алгоритм решения линейного неоднородного дифференциального уравнения *методом Лагранжа* (*методом вариации постоянной*):

1) для заданного неоднородного уравнения (7.40) выписать соответствующее ему однородное уравнение (формула (7.41));

2) методом разделения переменных найти общее решение однородного уравнения (формула (7.42));

3) в общем решении однородного уравнения заменить постоянную C на функцию $C(x)$ (формула (7.43));

4) подставить полученное в пункте 3 выражение в исходное неоднородное уравнение и найти $C(x)$ (формула (7.45));

5) выписать общее решение неоднородного уравнения, подставив выражение для $C(x)$ в (7.43).

Пример 7.8. Решить уравнение

$$y' + \frac{2y}{x} = 5x^2.$$

Решение. Это линейное дифференциальное уравнение. Соответствующее однородное уравнение имеет вид

$$y' + 2\frac{y}{x} = 0.$$

Разделяя переменные, имеем

$$\frac{dy}{y} = -\frac{2dx}{x},$$

откуда $\ln|y| = -2\ln|x| + \ln C$, или $y = C/x^2$.

Полагая $C = C(x)$, находим

$$y' = \frac{C'(x)}{x^2} - \frac{2C(x)}{x^3}.$$

Подставив выражение для y и y' в исходное уравнение, получим

$$\frac{C'(x)}{x^2}-\frac{2C(x)}{x^3}+\frac{2C(x)}{x^3}=5x^2.$$

Отсюда следует, что

$$C'(x) = 5x^4.$$

Значит,

$$C(x) = x^5 + C_1 \ (C_1 = \text{const}).$$

Таким образом,

$$y(x)=\frac{x^5+C_1}{x^2}, \text{ или } y(x)=x^3+\frac{C_1}{x^2}.$$

Замечание 1. В некоторых случаях дифференциальное уравнение может быть приведено к линейному, если поменять ролями переменные y и x – искомую функцию и ее аргумент.

Пример 7.9. Решить задачу Коши

$$y'=\frac{y}{x+2y^3}, \quad y(8) = 2. \tag{7.46}$$

Решение. Данное дифференциальное уравнение не является линейным относительно неизвестной функции $y(x)$. Однако если предположить, что неизвестной является функция $x(y)$, то уравнение станет линейным. Действительно, из (7.46) следует

$$\frac{dx}{dy}=\frac{x+2y^3}{y}, \tag{7.47}$$

или

$$\frac{dx}{dy}=\frac{x}{y}+2y^2.$$

Решением соответствующего линейного однородного уравнения является

$$x = Cy.$$

В соответствии с методом вариации постоянной будем искать решение неоднородного уравнения в виде $x = C(y)y$. Подставляя выражение для x в (7.47), после приведения подобных получаем

$$C'y = 2y^2.$$

Отсюда находим

$$C(y) = y^2 + C_1 \quad (C_1 = \text{const}).$$

Таким образом, решением уравнения (7.47) является

$$x = y^3 + C_1 y. \tag{7.48}$$

Теперь найдем решение задачи Коши (7.46). Полагая в (7.48) $x = 8$, $y = 2$, получаем, что $C_1 = 0$. Следовательно, $x = y^3$, а значит,

$$y = \sqrt[3]{x}.$$

2. Для решения линейного дифференциального уравнения может быть также применен *метод Бернулли*, который заключается в следующем:

1) решение уравнения (7.40) ищется в виде

$$y = u(x)v(x), \tag{7.49}$$

где $u(x)$ и $v(x)$ – некоторые функции, которые необходимо определить;

2) выражение для y подставляется в исходное уравнение, и после группировки слагаемых оно принимает вид

$$u'v + u(v' + p(x)v) = f(x); \tag{7.50}$$

3) функция $v(x)$ находится из уравнения

$$v' + p(x)v = 0; \tag{7.51}$$

4) после подстановки найденной функции $v(x)$ в уравнение (7.50) получается следующее уравнение для определения функции $u(x)$

$$u' = -\frac{f(x)}{v(x)}; \tag{7.52}$$

5) искомое общее решение выписывается в виде произведения функций $u(x)$ и $v(x)$.

Пример 7.10. Решить уравнение

$$y' + y\operatorname{ctg} x = 2\cos x.$$

Решение. Полагая в уравнении $y = u(x)v(x)$, имеем

$$u'v + u(v' + v\operatorname{ctg} x) = 2\cos x. \tag{7.53}$$

Приравнивая выражение в скобках к нулю, получаем уравнение с разделяющимися переменными

$$\frac{v'}{v} = -\operatorname{ctg} x.$$

Его общее решение имеет вид

$$v(x) = \frac{C}{\sin x}.$$

В качестве $v(x)$ возьмем частное решение

$$v(x) = \frac{1}{\sin x}, \tag{7.54}$$

получающееся из общего при $C = 1$. Подставляя выражение для $v(x)$ из (7.54) в (7.53), находим

$$u' = 2\sin x\cos x.$$

Отсюда имеем

$$u(x) = \sin^2 x + C.$$

Тогда

$$y = uv = \frac{\sin^2 x + C}{\sin x}.$$

Таким образом, общим решением исходного уравнения будет

$$y = \sin x + \frac{C}{\sin x}.$$

§ 7.7. Уравнения Бернулли и Риккати

Определение. *Дифференциальное уравнение вида*

$$y' + p(x)y = f(x)y^n \quad (n \neq 0,\ n \neq 1) \qquad (7.55)$$

называется ***уравнением Бернулли***.

Заменой $z = y^{1-n}$ уравнение Бернулли приводится к линейному уравнению

$$z' + (1 - n)\,p(x)z = (1 - n)f(x).$$

Пример 7.11. Решить уравнение

$$y' - y = \frac{e^{6x}}{y^2}.$$

Р е ш е н и е. Это уравнение Бернулли (n= –2). Выполнив замену $z = y^3$, получим $z' = 3y^2y'$. Умножая обе части исходного уравнения на $3y^2$ ($3y^2 \neq 0$), с учетом выражений для z и z' находим

$$z' - 3z = 3e^{6x}.$$

Соответствующее линейное однородное уравнение $z' - 3z = 0$ имеет решение $z = Ce^{3x}$. Применяя метод вариации постоянной, получаем

$$C'(x)e^{3x} + 3C(x)e^{3x} - 3C(x)e^{3x} = 3e^{6x},$$

т.е. $C'(x) = 3e^{3x}$, $C(x) = e^{3x} + C_1$, где C_1 – произвольная постоянная. Таким образом, $z = e^{6x} + C_1e^{3x}$. Значит, $y = \sqrt[3]{e^{6x} + C_1e^{3x}}$.

З а м е ч а н и е 1. Уравнение Бернулли может быть решено непосредственно при помощи метода Лагранжа или метода Бернулли без предварительного приведения к линейному уравнению.

П р и м е р 7.12. Решить задачу Коши

$$y' - \frac{y}{x} = \frac{y^2}{x^4}, \quad y(1) = \frac{1}{2}.$$

Р е ш е н и е. Данное уравнение является уравнением Бернулли ($n = 2$). Решим его методом Лагранжа. Для этого рассмотрим сначала однородное линейное уравнение

$$y' - \frac{y}{x} = 0.$$

Разделяя переменные, находим

$$y=Cx.$$

В соответствии с методом вариации постоянной решения неоднородного уравнения будем искать в виде $y=C(x)x$. Подставляя выражения для y и y' в исходное уравнение, имеем

$$C'x + C - C = \frac{C^2x^2}{x^3}, \qquad \frac{C'}{C^2} = \frac{1}{x^2}.$$

Интегрируя это уравнение, получаем

$$\frac{-1}{C} = \frac{-1}{x} - C_1,$$

где C_1 – произвольная постоянная. Отсюда

$$C(x) = \frac{x}{C_1 x + 1}.$$

Следовательно, общее решение исходного уравнения имеет вид

$$y = \frac{x^2}{C_1 x + 1}. \tag{7.56}$$

Полагая в (7.56) $x = 1$, $y = \frac{1}{2}$, находим

$$C_1 + 1 = 2.$$

Таким образом, $C_1 = 1$ и решением задачи Коши будет

$$y = \frac{x^2}{x + 1}.$$

Пример 7.13. Решить уравнение

$$y' - \frac{2xy}{1 + x^2} = \frac{2 \operatorname{arctg} x}{3(1 + x^2)^4} y^4.$$

Решение. Данное уравнение представляет собой уравнение Бернулли ($n = 4$). Для его решения воспользуемся методом Бернулли. Положим $y = u(x)v(x)$. В силу этого уравнение принимает вид

$$u'v + u\left(v' - \frac{2xv}{1 + x^2}\right) = \frac{2 \operatorname{arctg} x}{3(1 + x^2)^4}(uv)^4. \tag{7.57}$$

Функцию $v(x)$ найдем как решение уравнения с разделяющимися переменными:

$$v' - \frac{2xv}{1+x^2} = 0.$$

Интегрируя его, получаем частное решение

$$v(x) = 1 + x^2.$$

Подставляя выражение для $v(x)$ в уравнение (7.57), получаем

$$3u'(1+x^2) = 2u^4 \operatorname{arctg} x.$$

Разделяя переменные, находим

$$\frac{3du}{u^4} = \frac{2dx}{1+x^2} \operatorname{arctg} x,$$

$$\frac{-1}{u^3} = \operatorname{arctg}^2 x - C,$$

$$u = \frac{1}{\sqrt[3]{C - \operatorname{arctg}^2 x}}.$$

Перемножая u и v, получаем общее решение исходного уравнения

$$u = \frac{1+x^2}{\sqrt[3]{C - \operatorname{arctg}^2 x}}.$$

В § 7.6 при рассмотрении некоторых моделей роста мы встречались с дифференциальными уравнениями первого порядка, в которых производная искомой функции была равна квадратному трехчлену (с постоянными коэффициентами) от этой функции. С точки зрения экономической динамики также представляют интерес и уравнения того же типа с коэффициентами, зависящими от времени. На них мы и остановимся.

Определение. *Дифференциальное уравнение вида*

$$y' = a(x)y^2 + b(x)y + c(x) \tag{7.58}$$

называется ***уравнением Риккати.***

В общем случае уравнение Риккати не удается проинтегрировать в квадратурах. Однако если известно частное решение $y_1(x)$ этого уравнения, то можно найти и его общее решение. Действительно, делая замену $y(x) = y_1(x) + z(x)$, где $z(x)$ – новая неизвестная функция, приходим к уравнению

$$z' - (2a(x)y_1 + b(x))z = a(x)z^2,$$

являющемуся уравнением Бернулли ($n = 2$).

Пример 7.14. Решить уравнение

$$y' = y^2 - 2xy + x^2 + 1.$$

Нетрудно убедиться в том, что решением является $y_1(x) = x$. Тогда в результате замены

$$y(x) = x + z(x)$$

получаем автономное уравнение

$$z' = z^2.$$

Это уравнение имеет стационарное решение $z = 0$. Если же $z \neq 0$, то, разделяя переменные, находим

$$z = -\frac{1}{x + C}.$$

Возвращаясь к переменной y, находим решение исходного уравнения

$$y = x - \frac{1}{x + C}, \quad y = x.$$

§ 7.8. Уравнения в полных дифференциалах. Интегрирующий множитель

Определение. *Дифференциальное уравнение в симметрической форме*

$$N(x, y)dx + M(x, y)dy = 0, \tag{7.59}$$

где $N(x, y)$, $M(x, y)$ – непрерывные в некоторой области D функции, называется ***уравнением в полных дифференциалах****, если существует такая непрерывно дифференцируемая функция $U(x, y)$, что*

$$dU = N(x, y)dx + M(x, y)dy. \tag{7.60}$$

Из (7.60) следует, что общий интеграл уравнения в полных дифференциалах задается соотношением

$$U(x, y) = C.$$

Возникает естественный вопрос: каким образом по коэффициентам $N(x,y)$ и $M(x,y)$ можно определить, что данное уравнение является уравнением в полных дифференциалах? В случае, когда $U(x,y)$ – дважды непрерывно дифференцируемая функция, необходимое условие установить несложно. Действительно, из (7.60) вытекает, что

$$N = \frac{\partial U}{\partial x}, \quad M = \frac{\partial U}{\partial y}. \tag{7.61}$$

Так как $\frac{\partial^2 U}{\partial x \partial y} = \frac{\partial^2 U}{\partial y \partial x}$, из (7.61) получаем

$$\frac{\partial M}{\partial x} = \frac{\partial N}{\partial y}. \tag{7.62}$$

Соотношения (7.62) будут также и достаточными, если, например, рассматриваемая область является выпуклой. Однако доказательство этого факта выходит за рамки нашего курса.

Пример 7.15. Решить уравнение

$$(3x^2 - \sin(x + y^2) + \ln y) \cdot dx + \left(\frac{x}{y} - 2y \sin(x + y^2) + 2y\right) \cdot dy = 0.$$

Решение. В данном случае функции

$$N(x, y) = 3x^2 - \sin(x + y^2) + \ln y \text{ и } M(x, y) = \frac{x}{y} - 2y \sin(x + y^2) + 2y$$

непрерывно дифференцируемы в выпуклой области D, задаваемой условием $y > 0$. Кроме того,

$$\frac{\partial M}{\partial x} = \frac{\partial N}{\partial y} = \frac{1}{y} - 2y\cos(x + y^2),$$

а значит, условие (7.62) выполнено и данное уравнение является уравнением в полных дифференциалах. Найдем функцию $U(x, y)$ из системы уравнений

$$\frac{\partial U}{\partial x} = 3x^2 - \sin(x + y^2) + \ln y,$$
$$\frac{\partial U}{\partial y} = \frac{x}{y} - 2y \sin(x + y^2) + 2y.$$

Из первого уравнения получаем

$$U = x^3 - \cos(x + y^2) + x \ln y + g(y),$$

где $g(y)$ – некоторая непрерывно дифференцируемая функция. Чтобы ее определить, подставим выражение для U во второе уравнение системы. Имеем

$$\frac{x}{y} - 2y \sin(x + y^2) + g'(y) = \frac{x}{y} - 2y \sin(x + y^2) + 2y,$$

$$g'(y) = 2y, \qquad g(y) = y^2.$$

Таким образом, общим интегралом исходного уравнения будет

$$x^3 - \cos\ (x + y^2)\ + x\ \ln y + y^2 = C.$$

В некоторых случаях уравнения, записанные в симметричной форме, могут быть сведены к уравнению в полных дифференциалах умножением на некоторый множитель.

Пример 7.16. Решить уравнение

$$x(\sin y^2 - 2\sin x^2 + \cos x^2)dx + y(\sin y^2 + 2\cos y^2 + \cos x^2)dy = 0.$$

Решение. Непосредственной проверкой убеждаемся, что условия (7.62) не выполнены и данное уравнение не является уравнением в полных дифференциалах. Однако оно станет таковым, если обе части этого уравнения умножить на e^{xy}. Действительно, имеем

$$e^{xy}x(\sin y^2 - 2\sin x^2 + \cos x^2)dx + e^{xy}y(\sin y^2 + 2\cos y^2 + \cos x^2)dy = 0.$$

Теперь, как нетрудно убедиться, условия (7.62) выполнены. Решая систему

$$\frac{\partial U}{\partial x} = e^{xy}x(\sin y^2 - 2\sin x^2 + \cos x^2),$$
$$\frac{\partial U}{\partial y} = e^{xy}x(\sin y^2 + 2\cos y^2 + \cos x^2),$$

находим

$$U(x, y) = e^{xy}(\sin y^2 + \cos x^2).$$

Следовательно, уравнение

$$e^{xy}(\sin\ y^2 + \cos\ x^2) = C$$

определяет общий интеграл исходного уравнения.

Определение. *Непрерывно дифференцируемая в области G функция* $\mu(x, y) \neq 0$ *называется* ***интегрирующим множителем*** *уравнения* (7.59), *если уравнение*

$$\mu(x, y)(N(x, y)dx + M(x, y)dy) = 0 \tag{7.63}$$

является уравнением в полных дифференциалах.

Если для данного уравнения существует интегрирующий множитель, то он на основании (7.62) должен удовлетворять условию

$$\frac{\partial(\mu M)}{\partial x} = \frac{\partial(\mu N)}{\partial y},$$

или

$$M\frac{\partial\mu}{\partial x} - N\frac{\partial\mu}{\partial y} + \mu\left(\frac{\partial M}{\partial x} - \frac{\partial N}{\partial y}\right) = 0. \tag{7.64}$$

Равенство (7.64) представляет собой так называемое *дифференциальное уравнение в частных производных*. Решить это уравнение ничуть не проще, чем уравнение (7.59). Однако поскольку нас интересует только лишь одно частное решение уравнения (7.64), то иногда его можно найти, используя особенности коэффициентов N и M. В частности, в некоторых случаях это удается сделать в предположении, что $\mu = \mu(x)$ или $\mu = \mu(y)$.

Пример 7.17. Решить уравнение

$$(x + y^2)\, dx + xydy = 0. \tag{7.65}$$

Решение. Будем искать интегрирующий множитель этого уравнения в виде $\mu = \mu(x)$. Тогда в силу (7.65) уравнение (7.64) примет вид

$$xy\mu' = (2y - y)\mu,$$

или

$$\frac{d\mu}{\mu} = \frac{dx}{x}.$$

Из этого равенства находим $\mu = Cx$. В качестве интегрирующего множителя для уравнения (7.65) выберем $\mu = 6x$. В результате получаем уравнение в полных дифференциалах

$$6(x^2 + xy^2)\,dx + 6x^2 y dy = 0.$$

Решая систему

$$\frac{\partial U}{\partial x} = 6(x^2 + xy^2), \qquad \frac{\partial U}{\partial y} = 6x^2 y,$$

из первого уравнения находим

$$U = 2x^3 + 3x^2 y^2 + g(y),$$

где $g(y)$ – некоторая непрерывно дифференцируемая функция. Подставляя выражение для U во второе уравнение системы, имеем

$$6xy^2 + g'(y) = 6xy^2,$$
$$g'(y) = 0, \qquad g(y) = C.$$

Таким образом, общим интегралом исходного уравнения будет

$$2x^3 + 3x^2 y^2 = C.$$

§ 7.9. Дифференциальные уравнения высших порядков

1. Общие сведения

В данном параграфе мы остановимся на свойствах дифференциальных уравнений высших порядков. Напомним, что в общем виде дифференциальное уравнение n-го порядка задается соотношением (7.1). Если это уравнение разрешено относительно старшей производной, т.е.

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}), \tag{7.66}$$

то говорят, что оно записано в *нормальной форме* (или в *форме Коши*).

Начальные условия для уравнения n-го порядка имеют вид

$$y(x_0) = y_0, y'(x_0) = y'_0, \dots, y^{(n-1)}(x_0) = y_0^{(n-1)}. \tag{7.67}$$

Аналогичным образом задача об отыскании решений уравнения (7.66), удовлетворяющих начальному условию (7.67), называется *задачей Коши* уравнения n-го порядка. Так же, как и в случае дифференциального уравнения первого порядка, имеет место следующая теорема.

Теорема 7.3 (***о существовании и единственности решения задачи Коши для дифференциального уравнения n-го порядка***). *Если в некоторой окрестности точки* $(x_0, y'_0, \dots, y_0^{(n-1)})$ *функция* $f(x, y, y', \dots, y^{(n-1)})$ *определена, непрерывна и имеет непрерывные частные производные по* $y, y', \dots, y^{(n-1)}$*, то существует такая окрестность этой точки, в которой задача Коши* (7.66), (7.67) *имеет решение, притом единственное.*

Для уравнения второго порядка геометрический смысл этой теоремы заключается в том, что через данную точку в заданном направлении проходит единственная интегральная кривая.

2. Уравнения, допускающие понижение порядка

В некоторых случаях удается понизить порядок дифференциального уравнения. Не ограничивая общности, продемонстрируем это на примере уравнения второго порядка.

Если уравнение не содержит в явном виде аргумента, т.е.

$$F(y, y', y'') = 0,$$

то порядок уравнения может быть понижен при помощи замены $y' = z(y)$, где $z(y)$ – новая неизвестная функция от аргумента y.

Пример 7.18. Решить уравнение

$$y'' = -6y(y')^3.$$

Решение. В данном уравнении отсутствует в явном виде аргумент x. Положим $y'=z(y)$, тогда на основании правила дифференцирования сложной функции получаем

$$y'' = \frac{dz}{dy}y' = \frac{dz}{dy}z.$$

В силу этого данное уравнение преобразуется к виду

$$\frac{dz}{dy}z = -6yz^3.$$

Отсюда следует, что, либо $z = 0$, а значит,

$$y = C, \tag{7.68}$$

либо

$$\frac{dz}{dy} = -6yz^2.$$

Разделяя переменные и интегрируя полученное уравнение, находим

$$3y^2 + C_1 = \frac{1}{z}.$$

Следовательно,

$$z = \frac{1}{3y^2 + C_1}.$$

Возвращаясь к переменной y, имеем

$$y' = \frac{1}{3y^2 + C_1}.$$

Интегрируя это уравнение путем разделения переменных, получаем общий интеграл уравнения

$$y^3 + C_1 y = x + C_2. \tag{7.69}$$

Таким образом, множество всех решений исходного уравнения задается соотношениями (7.68), (7.69).

Если уравнение не содержит в явном виде искомую функцию, т.е.

$$F(y, y', y'') = 0,$$

то порядок уравнения может быть понижен при помощи замены $y' = z(x)$.

Пример 7.19. Решить уравнение

$$y'' = (y')^2 + 1.$$

Решение. В данном уравнении отсутствует в явном виде искомая функция. Сделав замену $y' = z(x)$, получаем

$$z' = z^2 + 1.$$

Разделяя в полученном уравнении переменные и затем интегрируя его, находим

$$z = \operatorname{tg}(x + C_1),$$

$$y' = \operatorname{tg}(x + C_1),$$

$$y = -\ln|\cos(x + C_1)| + C_2.$$

Порядок дифференциального уравнения, однородного относительно неизвестной функции и ее производных, т.е. удовлетворяющего условию

$$F(x, ty, ty', ty'') = t^m F(x, y, y', y''),$$

может быть понижен при помощи замены $y' = yz(x)$.

Пример 7.20. Решить уравнение

$$y''y-(y')^2-\frac{3yy'}{x}=(6x-3x^2)y^2.$$

Решение. Данное уравнение является однородным второй степени относительно y, y' и y''. Сделаем замену $y' = z(x)y$, имеем

$$y''=z'y+zy'=(z'+z^2)y.$$

На основании этого исходное уравнение принимает вид

$$\left(z'-\frac{3z}{x}-6x+3x^2\right)y^2=0.$$

Отсюда вытекает, что либо $y = 0$, либо

$$z'-\frac{3z}{x}-6x+3x^2=0.$$

Решая полученное линейное уравнение, находим

$$z=-6x^2-3x^3\ln|x|+C_1x^3,$$

$$\frac{y'}{y}=-6x^2-3x^3\ln|x|+C_1x^3,$$

$$\ln|y|=-2x^3-\frac{3}{4}x^4\ln|x|+\frac{3}{16}x^4+C_2x^4+\ln|C_3|,$$

$$y=C_3e^{C_2x^4-2x^3-\frac{3}{4}x^4\ln|x|+\frac{3}{16}x^4}. \qquad (7.70)$$

Заметим, что соотношение (7.70) задает общее решение исходного уравнения, так как полученное выше частное решение $y = 0$ входит в (7.70) при $C_3 = 0$.

§ 7.10. Линейные дифференциальные уравнения высших порядков

1. В § 7.6 мы научились решать линейные дифференциальные уравнения первого порядка. Данный параграф посвящен линейным уравнениям более высоких порядков. Этот класс уравнений является очень важным, т.к. многие практические задачи либо описываются линейными уравнениями, либо могут быть приближенно решены при помощи линейных уравнений.

Определение. *Дифференциальное уравнение n-го порядка называется* ***линейным****, если оно имеет вид*

$$y^{(n)}+a_1(x)y^{(n-1)}+a_2(x)y^{(n-2)}+\ldots+a_n(x)y=f(x), \qquad (7.71)$$

где $a_1(x)$, $a_2(x)$, ... , $a_n(x)$, $f(x)$ – непрерывные функции.

Имеет место следующая теорема.

Теорема 7.4 (***существования и единственности***). *Пусть функции* $a_1(x)$, ... , $a_n(x)$, $f(x)$ *непрерывны на отрезке* $[a, b]$. *Тогда для любой точки* $x_0 \in (a, b)$ *существует, причем единственное, решение* $y(x)$ *уравнения* (7.71), *определенное на отрезке* $[a, b]$ *и удовлетворяющее начальным условиям*

$$y(x_0) = y_0,\ y'(x_0) = y'_0,\ \ldots,\ y^{(n-1)}(x_0) = y_0^{(n-1)}.$$

Примем эту теорему без доказательства. В то же время отметим, что если в теореме 7.4 утверждалось существование и единственность решения в некоторой окрестности точки x_0, то в данной теореме утверждается существование и единственность решения линейного уравнения на всем промежутке $[a, b]$, т.е. утверждение теоремы носит уже не локальный, а глобальный характер.

Введем следующее обозначение

$$L(y) = y^{(n)} + a_1(x)y^{(n-1)} + \ldots + a_n(x)y. \qquad (7.72)$$

Заметим, что $L(y)$ часто называют *линейным дифференциальным оператором n-го порядка*. Отметим свойство оператора $L(y)$, необходимое нам в дальнейшем.

Лемма 1. *Пусть $y_1(x)$ и $y_2(x)$ – произвольные функции, имеющие производные до n-го порядка включительно, C_1 и C_2 – произвольные постоянные, тогда*

$$L(C_1y_1 + C_2y_2) = C_1L(y_1) + C_2L(y_2). \tag{7.73}$$

Справедливость этого утверждения легко установить непосредственной проверкой.

Используя принятые выше обозначения, уравнение (7.71) можно записать в виде

$$L(y) = f(x). \tag{7.74}$$

Уравнение

$$L(y) = 0 \tag{7.75}$$

называется *линейным однородным уравнением*, соответствующим уравнению (7.75). В противоположность этому уравнение (7.74) (при $f(x) \not\equiv 0$) называется *неоднородным.* Следующее утверждение связывает решения уравнений (7.74) и (7.75).

Теорема 7.5. *Общее решение неоднородного уравнения* (7.74) *есть сумма частного решения $\overline{y}(x)$ этого уравнения и общего решения соответствующего ему однородного уравнения* (7.75).

Доказательство. Покажем сначала, что сумма $y(x)$ частного решения уравнения неоднородного уравнения $\overline{y}(x)$ и произвольного решения $y_0(x)$ однородного уравнения также является решением неоднородного уравнения. Действительно, в силу леммы имеем

$$L(\overline{y} + y_0) = L(\overline{y}) + L(y_0) = f(x) + 0 = f(x),$$

что и требовалось доказать. Теперь нам остается доказать, что всякое решение $y(x)$ неоднородного уравнения есть сумма $\overline{y}(x)$ и некоторого частного решения $y_0(x)$ уравнения (7.74). Имеем

$$L(y - \overline{y}) = L(y) - L(\overline{y}) = f(x) - f(x) = 0.$$

Следовательно, $y_0(x) = y(x) - \bar{y}(x)$ – решение уравнения (7.75), значит, $y(x) = y_0(x) + \bar{y}(x)$, что и завершает доказательство.

2. Отметим одно важное свойство линейных уравнений, часто используемое при нахождении решений.

Теорема 7.6. *Пусть $y_1(x)$ и $y_2(x)$ – соответственно решения уравнений $L(y) = f_1(x)$ и $L(y) = f_2(x)$, тогда $y_1(x) + y_2(x)$ есть решение уравнения $L(y) = f_1(x) + f_2(x)$.*

Действительно, $L(y_1 + y_2) = L(y_1) + L(y_2) = f_1(x) + f_2(x)$, что и утверждалось в теореме.

3. Рассмотрим *уравнение Самуэльсона – Солоу*

$$p' = k(D(p) - S(p)), \tag{7.76}$$

моделирующее связь между изменением цены p и неудовлетворенным спросом $D(p) - S(p)$ (здесь $D(p)$ и $S(p)$ – соответственно величины спроса и предложения при цене p, $k > 0$). Предположим, что спрос и предложение задаются линейными функциями

$$D(p) = a - bp, S(p) = m + np, \tag{7.77}$$

где a, b, m, n – некоторые положительные числа. С учетом (7.77) уравнение (7.76) примет вид

$$p' = k(a - m) - k(n + b)p. \tag{7.78}$$

Уравнение (7.78) является линейным дифференциальным уравнением. Найдем решение соответствующего ему однородного уравнения. Имеем:

$$\frac{dp}{dt} = k(n+b)p,$$

$$\frac{dp}{p} = k(n+b)dt,$$

$$\ln p = k(n + b)t + \ln C,$$

$$p(t) = C\,e^{k(n + b)t}. \tag{7.79}$$

В качестве частного решения уравнения (7.78) можно использовать стационарное равновесное решение

$$p(t) = \overline{p} = \text{const},$$

где $\overline{p}$ – корень уравнения $D(p) = S(p)$ (в этом случае обе части уравнения (7.78) будут равны нулю).

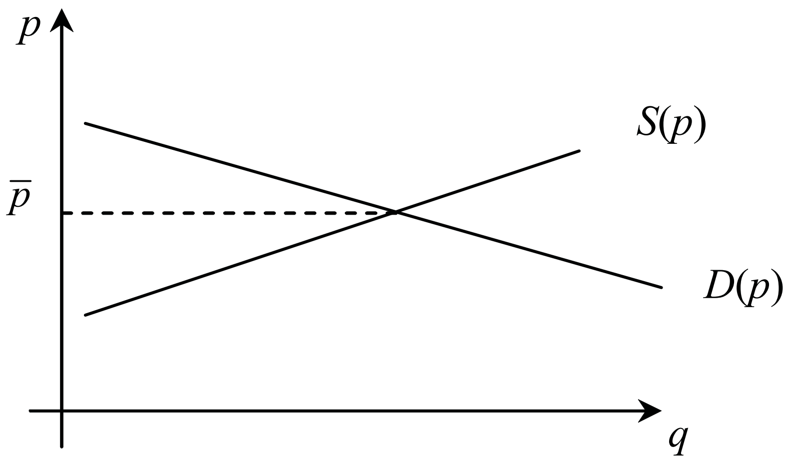

Рис. 7.12

Из (7.78) нетрудно найти (рис. 7.12), что

$$\overline{p} = \frac{a-m}{b+n}.$$

Таким образом, общее решение уравнения (7.78) имеет вид

$$p(t) = \frac{a-m}{b+n} + Ce^{k(n+b)t}. \qquad (7.80)$$

Из (7.80), в частности, вытекает, что если $n > b$, то с течением времени интегральные кривые будут отдаляться от состояния равновесия $\overline{p}$. В случае $n = b$ величина $p(t)$ – постоянна. Если же $n < b$, то с течением времени интегральные кривые будут асимптотически приближаться к состоянию равновесия $\overline{p}$. Данную модель можно рассматривать как непрерывный аналог паутинной модели рынка.

§ 7.11. Линейные однородные уравнения. Фундаментальный набор решений

В данном параграфе мы остановимся на свойствах частного и общего решения линейного однородного дифференциального уравнения n-го порядка:

$$L(y) = 0.$$

1. **Лемма 2.** *Пусть* $y_1(x), y_2(x), ..., y_k(x)$ *– произвольные решения линейного однородного дифференциального уравнения и* $C_1, C_2, \ldots, C_k$ *– произвольные постоянные, тогда линейная комбинация* $C_1y_1(x) + C_2y_2(x) + ... + C_ky_k(x)$ *также является решением этого уравнения.*

Действительно, на основании леммы 1 из предыдущего параграфа имеем

$$L(C_1y_1 + C_2y_2 + \ldots + C_ky_k) = C_1L(y_1) + C_2L(y_2) + \ldots + C_kL(y_k) = 0,$$

что и требовалось доказать.

Замечание 1. Из леммы 2 следует, что множество решений линейного однородного уравнения образует *линейное пространство*. Естественно, возникают вопросы: какова размерность этого пространства и как устроен его базис? Чтобы ответить на них, необходимо иметь критерий, дающий возможность определить, является ли данная система решений $y_1(x), y_2(x), ..., y_k(x)$ *линейно зависимой* или нет? Один из способов, позволяющих это сделать, связан с так называемым определителем Вронского.

Пусть $y_1(x), ..., y_k(x)$ – система, состоящая из k функций, тогда *определитель Вронского* этой системы имеет вид:

$$W(y_1, y_2, \ldots, y_k) = \begin{vmatrix} y_1 & y_2 & \cdots & y_k \\ y_1' & y_2' & \cdots & y_k' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(k-1)} & y_2^{(k-1)} & \cdots & y_k^{(k-1)} \end{vmatrix}. \qquad (7.81)$$

Справедливы следующие утверждения.

Теорема 7.7. *Если функции y_1, ..., y_k линейно зависимы, то их определитель Вронского тождественно равен нулю.*

Доказательство. Пусть y_1, ..., y_k – линейно зависимы. Тогда одна из этих функций является линейной комбинацией остальных. Для определенности положим, что таковой является y_1. В этом случае первый столбец определителя Вронского будет представлять собой линейную комбинацию остальных столбцов, а следовательно, определитель тождественно равен нулю. Теорема доказана.

Теорема 7.8. *Если $y_1(x)$, ..., $y_k(x)$ – линейно независимые решения уравнения* (7.75), *то их определитель Вронского ни при одном значении x не обращается в нуль.*

Доказательство теоремы проведем от противного. Пусть существует точка x_0, в которой определитель Вронского равен нулю. Рассмотрим однородную систему линейных алгебраических уравнений

$$\begin{cases} C_1 y_1(x_0) + C_2 y_2(x_0) + \ldots + C_k y_k(x_0) = 0, \\ C_1 y_1'(x_0) + C_2 y_2'(x_0) + \ldots + C_k y_k'(x_0) = 0, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ C_1 y_1^{(k-1)}(x_0) + C_2 y_2^{(k-1)}(x_0) + \ldots + C_k y_k^{(k-1)}(x_0) = 0 \end{cases} \tag{7.82}$$

относительно искомых неизвестных чисел C_1, C_2, ..., C_k. Определителем этой системы является определитель Вронского в точке x_0. Так как по предположению он равен нулю, то система имеет ненулевое решение $\overline{C}_1, \overline{C}_2, \ldots, \overline{C}_k$. Рассмотрим функцию

$$\varphi(x) = \overline{C}_1 y_1(x) + \overline{C}_2 y_2(x) + \ldots + \overline{C}_k y_k(x),$$

которая в силу леммы 1 является решением уравнения (7.75). При этом $\varphi(x) \neq 0$, так как в противном случае функции y_1, ..., y_k были бы линейно зависимы: одну из них, имеющую ненулевой коэффициент, можно было бы выразить в виде линейной комбинации остальных. С другой стороны, равенства (7.82) означают, что

$$\varphi(x_0) = 0, \varphi'(x_0) = 0, \ldots, \varphi^{(k-1)}(x_0) = 0. \tag{7.83}$$

Но этим же начальным условиям удовлетворяет и другое решение уравнения (7.75), а именно $y(x) = 0$. Это противоречит единственности решения задачи Коши для уравнения (7.75). Следовательно, $W(y_1, ..., y_k) \neq 0$ для всех x. Теорема доказана.

Определение. *Систему функций* $y_1(x), ..., y_n(x)$, *состоящую из* n *линейно независимых решений уравнения* (7.75), *будем называть* ***фундаментальным набором решений*** *этого уравнения.*

2. Ранее мы уже отмечали, что общее решение дифференциального уравнения n-го порядка зависит от n произвольных постоянных. С другой стороны, если $y_1(x), ..., y_n(x)$ – решения уравнения (7.75), то функция

$$y(x) = C_1 y_1(x) + C_2 y_2(x) + \ldots C_n y_n(x) \tag{7.84}$$

также является решением этого уравнения. Напрашивается вопрос: при каких условиях формулой (7.84) определяется общее решение линейного однородного уравнения? Ответ на этот вопрос дает основная теорема теории линейных однородных дифференциальных уравнений.

Теорема 7.9 (*об общем решении линейного однородного уравнения*). *Пусть* $y_1(x), ..., y_n(x)$ – *фундаментальный набор решений уравнения* (7.75), *тогда общее решение этого уравнения задается формулой*

$$y = C_1 y_1 + ... + C_n y_n. \tag{7.85}$$

Доказательство. То, что функция $y(x)$, определяемая формулой (7.82), является решением уравнения (7.75), следует из леммы 2. Покажем теперь, что любое решение $\varphi(x)$ уравнения (7.75) представимо в виде линейной комбинации функций $y_1, ..., y_n$. Зафиксируем некоторую точку x_0. Введем следующие обозначения:

$$\varphi(x_0) = y_0, \varphi'(x_0) = y_0', \ldots, \varphi^{(n-1)}(x_0) = y_0^{(n-1)}.$$

Рассмотрим систему линейных алгебраических уравнений:

$$\begin{cases} C_1 y_1(x_0) + C_2 y_2(x_0) + \ldots + C_n y_n(x_0) = y_0, \\ C_1 y_1'(x_0) + C_2 y_2'(x_0) + \ldots + C_n y_n'(x_0) = y_0', \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ C_1 y_1^{(n-1)}(x_0) + C_2 y_2^{(n-1)}(x_0) + \ldots + C_n y_n^{(n-1)}(x_0) = y_0^{(n-1)}. \end{cases} \tag{7.86}$$

Определителем этой системы является определитель Вронского для функций y_1, ..., y_n в точке x_0. Ввиду линейной независимости этих функций данный определитель не равен нулю. Следовательно, у системы (7.86) существует решение $\left(\overline{C}_1, \overline{C}_2, \ldots, \overline{C}_n\right)$.

Тогда функция

$$y(x) = \overline{C}_1 y_1(x) + \overline{C}_2 y_2(x) + \ldots + \overline{C}_n y_n(x),$$

как это вытекает из (7.86), удовлетворяет тем же начальным условиям. В силу единственности решения задачи Коши имеем

$$\varphi(x) = y(x),$$

т.е. $\varphi(x)$ есть линейная комбинация функций y_1, ..., y_n. Теорема доказана.

Замечание 2. Из теоремы 7.9 следует, что множество решений линейного однородного уравнения образует n-мерное линейное пространство, а фундаментальный набор решений является его *базисом*.

Пример 7.21. Для уравнения

$$y'' - 4y = 0$$

функции $y_1(x) = e^{2x}$ и $y_2(x) = e^{-2x}$ являются частными решениями. Эти решения линейно независимы, так как их определитель Вронского

$$W(y_1, y_2) = \begin{vmatrix} e^{2x} & e^{-2x} \\ 2e^{2x} & -2e^{-2x} \end{vmatrix} = -4$$

не равен нулю. Значит, y_1, y_2 образуют фундаментальный набор и общее решение уравнения имеет вид

$$y = C_1 e^{2x} + C_2 e^{-2x}.$$

§ 7.12. Линейные дифференциальные уравнения с постоянными коэффициентами

Данный параграф посвящен нахождению решений линейных дифференциальных уравнений с постоянными коэффициентами.

1. Линейные однородные уравнения

Нахождение общего решения линейного однородного уравнения требует знания какого-либо фундаментального набора решений. Последнее в общем случае является довольно сложной задачей. Однако эта задача намного упрощается, если коэффициенты уравнения постоянны. Рассмотрим уравнение

$$y^{(n)} + a_1 y^{(n-1)} + a_2 y^{(n-2)} + \ldots + a_n y = 0, \tag{7.87}$$

где $a_1, \ldots , a_n$ – некоторые постоянные.

Будем искать решение уравнения (7.87) в виде функции

$$y = e^{\lambda x}.$$

Тогда

$$y' = \lambda e^{\lambda x}, \cdots, y^{(n)} = \lambda^n e^{\lambda x}.$$

Подставляя выражения для y и ее производных в уравнение (7.87), имеем

$$\lambda^n e^{\lambda x} + a_1 \lambda^{n-1} e^{\lambda x} + \ldots + a_{n-1} y e^{\lambda x} + a_n e^{\lambda x} = 0.$$

Так как $e^{\lambda x} \neq 0,$ то это соотношение эквивалентно уравнению

$$\lambda^n + a_1 \lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n = 0. \tag{7.88}$$

Определение. *Алгебраическое уравнение* (7.88) *называется* ***характеристическим уравнением*** *линейного однородного дифференциального уравнения* (7.87).

Замечание. Нами установлено соответствие между корнями характеристического уравнения и решениями уравнения (7.87). Так как корни характеристического уравнения могут быть комплексными, то необходимо внести ясность, что мы будем понимать в этом случае под $e^{\lambda x}$. Для любого комплексного числа $z = \alpha + i\beta$ положим по определению

$$e^z = e^\alpha (\cos\beta + i\sin\beta). \tag{7.89}$$

Соотношение (7.89) носит название *формулы Эйлера*. Принимая во внимание соответствующие свойства комплексных чисел, нетрудно убедиться, что определенная по формуле Эйлера функция комплексного аргумента обладает всем набором свойств, характерных для экспоненциальной функции. В частности, для произвольных комплексных чисел v и u

$$e^v e^u = e^{u+v}, \quad \frac{e^v}{e^u} = e^{u-v}.$$

В дальнейшем для простоты изложения будем оперировать с уравнениями второго порядка, т.е. с уравнениями вида

$$y'' + py' + qy = 0. \tag{7.90}$$

Однако заметим, что полученные для этих уравнений результаты могут быть без труда перенесены на случай уравнений более высоких порядков.

При решении характеристического уравнения

$$\lambda^2 + p\lambda + q = 0$$

могут возникнуть три случая.

Случай 1. Дискриминант D больше нуля. Тогда существуют два действительных и различных решения λ_1 и λ_2 характеристического уравнения. Соответствующие им решения уравнения (7.90)

$$y_1 = e^{\lambda_1 x}, \ y_2 = e^{\lambda_2 x}$$

линейно независимы. Действительно, их определитель Вронского

$$W(y_1, y_2) = \begin{vmatrix} e^{\lambda_1 x} & e^{\lambda_2 x} \\ \lambda_1 e^{\lambda_1 x} & \lambda_2 e^{\lambda_2 x} \end{vmatrix} = (\lambda_2 - \lambda_1) e^{(\lambda_1 + \lambda_2)x}$$

ввиду того, что $\lambda_1 \neq \lambda_2$, отличен от нуля. Следовательно, y_1 и y_2 образуют фундаментальный набор и общее решение уравнения (7.89) имеет вид

$$y = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x}.$$

Пример 7.22. Решить уравнение

$$y'' + y' - 2 = 0.$$

Решение. Корнями характеристического уравнения

$$\lambda^2 + \lambda - 2 = 0$$

являются числа $\lambda_1 = 1$ и $\lambda_2 = -2$. Следовательно, общее решение уравнения имеет вид

$$y = C_1 e^{x} + C_2 e^{-2x}.$$

Случай 2. Дискриминант $D < 0$. Характеристическое уравнение имеет два комплексно сопряженных корня $\lambda_{1,2} = \alpha \pm \beta i.$ Уравнение (7.90) имеет два линейно независимых комплексно сопряженных решения

$$y_1 = e^{(\alpha + i\beta)x}, \ y_2 = e^{(\alpha - i\beta)x},$$

или в соответствии с формулой Эйлера –

$$y_1 = e^{\alpha x}(\cos\beta x + i\sin\beta x), \quad y_2 = e^{\alpha x}(\cos\beta x - i\sin\beta x).$$

Чтобы получить действительные решения, заменим y_1 и y_2 их линейными комбинациями:

$$y_1^* = \frac{1}{2}(y_1 + y_2), \quad y_2^* = \frac{1}{2i}(y_1 - y_2),$$

представляющими собой соответственно действительную и мнимую часть y_1.

Таким образом, мы получили два линейно независимых действительных решения:

$$y_1^* = e^{\alpha x}\cos\beta x, \quad y_2^* = e^{\alpha x}\sin\beta x. \tag{7.91}$$

Следовательно, общее решение в этом случае имеет вид

$$y = e^{\alpha x}(C_1\cos\beta x + C_2\sin\beta x). \tag{7.92}$$

Пример 7.23. Решить уравнение

$$y'' - 2y' + 2y = 0.$$

Решение. Корнями характеристического уравнения

$$\lambda^2 - 2\lambda + 2 = 0$$

будут $\lambda_{1,2} = 1 \pm i.$ В данном случае $\alpha = 1$, $\beta = 1$, так что общим решением данного дифференциального уравнения будет

$$y = e^x(C_1\cos x + C_2\sin x).$$

Случай 3. Дискриминант $D = 0$. У характеристического уравнения существует единственный корень $\lambda = -\frac{p}{2}$, которому соответствует решение $y_1 = e^{-\lambda x}$. Но для построения общего решения нам необходимо еще одно линейно независимое с y_1 решение уравнения (7.90). Покажем, что таковым является функция

$$y_2 = xe^{\lambda x}.$$

Действительно,

$$y_2' = \lambda x e^{\lambda x} + e^{\lambda x}, \quad y_2'' = \lambda^2 x e^{\lambda x} + 2e^{\lambda x}.$$

Подставляя выражения для y, y' и y'' в уравнение (7.90), имеем

$$e^{\lambda x}\left[x(\lambda^2 + p\lambda + q) + 2\lambda + p\right] = 0.$$

Учитывая, что λ является корнем характеристического уравнения и при этом $\lambda = -\frac{p}{2}$, видим, что последнее равенство выполняется тождественно, т.е. y_2 – решение уравнения (7.90). Покажем теперь, что y_1 и y_2 линейно независимы. Имеем

$$W(y_1, y_2) = \begin{vmatrix} e^{\lambda x} & xe^{\lambda x} \\ \lambda e^{\lambda x} & (1 + \lambda x)e^{\lambda x} \end{vmatrix} = e^{2\lambda x} \neq 0.$$

Следовательно, y_1 и y_2 линейно независимы и образуют фундаментальный набор. Таким образом, общим решением в этом случае будет

$$y = e^{\lambda x}(C_1 + C_2 x).$$

Пример 7.24. Найти общее решение уравнения

$$y'' + 6y + 9 = 0.$$

Р е ш е н и е . Характеристическое уравнение

$$\lambda^2 + 6\lambda + 9 = 0$$

имеет единственный корень $\lambda = -3$. Поэтому, общее решение имеет вид

$$y = e^{-3x}(C_1 + C_2x).$$

Сформулируем теперь *алгоритм решения линейного однородного дифференциального уравнения n-го порядка.*

1) составить характеристическое уравнение и найти его корни;

2) построить фундаментальную систему решений, поставив в соответствие каждому действительному корню характеристического уравнения λ кратности k совокупность из k линейно независимых частных решений

$$y_1 = e^{\lambda x},\ y_2 = xe^{\lambda x},\ \dots\ ,\ y_k = x^{k-1}e^{\lambda x} \tag{7.93}$$

и каждой паре комплексно сопряженных корней $\alpha \pm i\beta$ кратности k – совокупность из $2k$ линейно независимых решений

$$y_1 = e^{\alpha x}\cos\beta x,\ y_2 = xe^{\alpha x}\cos\beta x,\ \dots\ ,\ y_k = x^{k-1}e^{\alpha x}\cos\beta x,$$

$$y_{k+1} = e^{\alpha x}\sin\beta x,\ y_{k+2} = xe^{\alpha x}\sin\beta x,\ \dots\ ,\ y_{2k} = x^{k-1}e^{\alpha x}\sin\beta x; \tag{7.94}$$

3) по формуле (7.85) выписать общее решение уравнения.

П р и м е р 7.25. Решить уравнение

$$y^{\text{IV}} - 2y''' + 2y' - y = 0.$$

Р е ш е н и е . Характеристическое уравнение имеет вид

$$\lambda^4 - 2\lambda^3 + 2\lambda - 1 = 0.$$

После несложных тождественных преобразований получаем

$$(\lambda + 1)(\lambda - 1)^3 = 0.$$

Данное уравнение имеет два корня: $\lambda_1 = -1$ и $\lambda_2 = 1$. Причем кратность λ_2 равна 3. Следовательно, общим решением исходного уравнения будет

$$y = e^x\left(C_1 + C_2x + C_3x^2\right) + C_4e^{-x}.$$

П р и м е р 7.26. Найти общее решение уравнения

$$y''' - 2y'' + y' - 2y = 0.$$

Решая характеристическое уравнение

$$\lambda^3 - 2\lambda^2 + \lambda - 2 = 0,$$

находим $\lambda_1 = 2$; $\lambda_{2,3} = \pm i$. Таким образом, общее решение исходного уравнения имеет вид

$$y = C_1e^{2x} + C_2\cos x + C_3\sin x.$$

2. Линейные неоднородные уравнения

Для отыскания решения неоднородного уравнения

$$y^{(n)} + a_1y^{(n-1)} + a_2y^{(n-2)} + \ldots + a_ny = f(x), \qquad (7.95)$$

так же, как и в § 7.6, мы можем применить *метод вариации постоянных.* Продемонстрируем это на примере уравнения второго порядка

$$y'' + py' + qy = f(x). \qquad (7.96)$$

Будем искать решение в виде

$$y = C_1(x)y_1 + C_2(x)y_2, \tag{7.97}$$

где $y_1(x)$, $y_2(x)$ – фундаментальный набор решений соответствующего однородного уравнения, а $C_1(x)$, $C_2(x)$ – искомые функции. Из (7.97) находим

$$y' = C_1'y_1 + C_2'y_2 + C_1y_1' + C_2y_2'.$$

Потребуем дополнительно, чтобы

$$C_1'y_1 + C_2'y_2 = 0. \tag{7.98}$$

Тогда получим

$$y' = C_1y_1' + C_2y_2'.$$

Следовательно,

$$y'' = C_1'y_1' + C_2'y_2' + C_1y_1'' + C_2y_2''.$$

Подставляя найденные значения для y, y' и y'' в уравнение (7.96), имеем

$$C_1(y_1'' + py_1' + qy_1) + C_2(y_2'' + py_2' + qy_2) + C_1'y_1' + C_2'y_2' = f(x).$$

Оба выражения, заключенные в скобки, равны нулю, так как y_1 и y_2 являются решениями однородного уравнения (7.90). Следовательно, мы получим уравнение

$$C_1'y_1' + C_2'y_2' = f(x).$$

Таким образом, неизвестные функции C_1 и C_2 должны удовлетворять системе уравнений

$$\begin{cases} C_1'y_1 + C_2'y_2 = 0, \\ C_1'y_1' + C_2'y_2' = f(x). \end{cases} \tag{7.99}$$

Определителем этой системы является определитель Вронского, который отличен от нуля в силу линейной независимости y_1 и y_2. Следовательно, из системы (7.99) мы можем однозначно определить функции $C_1'(x)$ и $C_2'(x)$. Проинтегрировав их, найдем $C_1(x)$ и $C_2(x)$, а затем по формуле (7.96) – и решение уравнения (7.97).

Пример 7.27. Решить уравнение

$$y'' + 3y' - 4y = 30e^{2x}.$$

Решение. Соответствующее однородное уравнение имеет линейно независимые решения $y_1 = e^x$ и $y_2 = e^{-4x}$ ($\lambda_1 = 1$ и $\lambda_2 = -4$ – корни характеристического уравнения). Потому решение этого неоднородного уравнения будем искать в виде

$$y = C_1(x)e^x + C_2(x)e^{-4x}.$$

Для нахождения $C_1(x)$ и $C_2(x)$ решим систему (7.99), которая в нашем случае примет вид

$$\begin{cases} e^x C_1'(x) + e^{-4x} C_2'(x) = 0, \\ e^x C_1'(x) - 4e^{-4x} C_2'(x) = 30e^{2x}. \end{cases}$$

Получим

$$C_1(x) = 6e^x + \overline{C}_1, \quad C_2(x) = -e^{6x} + \overline{C}_2,$$

где $\overline{C}_1$ и $\overline{C}_2$ – произвольные постоянные. Таким образом, общим решением искомого уравнения будет

$$y = 5e^{2x} + \overline{C}_1 e^x + \overline{C}_2 e^{-4x}.$$

При специальном виде правой части $f(x)$ уравнения (7.95) методика отыскания частного решения неоднородного уравнения, изложенная выше, может быть заменена более простыми приемами. Опишем их вкратце.

В случае, когда коэффициенты левой части уравнения (7.95) постоянны, а правая часть имеет специальный вид

$$f(x) = e^{ax}\,(P_n(x)\cos bx + Q_k(x)\sin bx), \tag{7.100}$$

где $P_n(x)$ и $Q_k(x)$ – многочлены от x степени соответственно n и k, для отыскания частного решения уравнения (7.95) удобнее всего воспользоваться *методом неопределенных коэффициентов*. Он заключается в следующем. Если число $\gamma = a + bi$ не является корнем характеристического уравнения, то частное решение ищется в виде

$$y = e^{ax}\left(\overline{P}_m(x)\cos bx + \overline{Q}_m(x)\sin bx\right), \tag{7.101}$$

где $\overline{P}_m(x)$ и $\overline{Q}_m(x)$ – многочлены степени $m = \max\{k, n\}$, коэффициенты которых необходимо определить путем подстановки неизвестной функции y в уравнение (7.95). Если же γ является корнем характеристического уравнения кратности l, то частное решение ищется в форме

$$y = x^l e^{ax}\left(\overline{P}_m(x)\cos bx + \overline{Q}_m(x)\sin bx\right). \tag{7.102}$$

Этот случай часто называют *резонансным*.

Пример 7.28. Решить уравнение

$$y'' + 5y' + 6y = 56e^{5x}.$$

Решение. В соответствии с теоремой 7.6 общее решение данного линейного неоднородного уравнения представляет собой сумму некоторого частного решения этого уравнения и общего решения соответствующего ему однородного уравнения. Характеристическое уравнение

$$\lambda^2 + 5\lambda + 6 = 0$$

имеет корни $\lambda_1 = -2$ и $\lambda_2 = -3$. Следовательно, общим решением соответствующего однородного уравнения будет

$$y = C_1e^{-2x} + C_2e^{-3x}.$$

Теперь найдем частное решение. Так как правая часть уравнения имеет вид (7.100), то мы можем воспользоваться методом неопределенных коэффициентов. В нашем случае $a = 5$, $b = 0$, $P_n(x) = 1$, $Q_n(x) = 0$. Учитывая, что $P_n(x)$ и $Q_n(x)$ – многочлены нулевой степени и что число $\gamma = 5$ не является корнем характеристического уравнения, частное решение будем искать в виде

$$y = Ae^{5x} \quad (A = \text{const}).$$

Здесь $\overline{P}_0(x) = A$ – общий вид многочлена нулевой степени.

Находим

$$y' = 5e^{5x}; \quad y'' = 25e^{5x}.$$

Подставляя y, y' и y'' в исходное уравнение, имеем

$$56Ae^{5x} = 56e^{5x}.$$

Отсюда $A = 1$. Таким образом, частным решением исходного уравнения будет

$$y = e^{5x}.$$

Следовательно, общее решение уравнения имеет вид

$$y = e^{5x} + C_1e^{-2x} + C_2e^{-3x}.$$

Пример 7.29. Найти общее решение уравнения

$$y''' + y'' - 4y' - 4y = 10 \sin x.$$

Решение. Имеем $a = 0$, $b = 1$, $P_n(x) = 0$, $Q_n(x) = 10$. Число $\gamma = i$ не совпадает с корнями $\lambda_{1,2} = \pm 2$ и $\lambda_3 = -1$ характеристического уравнения, поэтому частное решение будем искать в виде

$$y = A \sin x + B \cos x \quad (A, B - \text{const}).$$

Отсюда получаем

$$y' = A\cos x - B\sin x,$$
$$y'' = -A\sin x - B\cos x,$$
$$y''' = -A\cos x + B\sin x.$$

Подставляя выражения для искомой функции и ее производных в исходное уравнение, находим

$$(5B - 5A)\sin x - (5B + 5A)\cos x = 10\sin x.$$

Так как это равенство выполняется тождественно, то

$$\begin{cases} 5B - 5A = 10, \\ 5B + 5A = \ \ 0; \end{cases} \qquad \begin{cases} A = -1, \\ B = \ \ 1. \end{cases}$$

Таким образом, общим решением данного уравнения будет

$$y = \cos x - \sin x + C_1 e^{-2x} + C_2 e^{2x} + C_3 e^{-x}.$$

Пример 7.30. Решить уравнение

$$y'' + y = x^2 + 2xe^x.$$

Решение. Так как правая часть данного линейного неоднородного уравнения представляет собой сумму двух выражений вида (7.100), то в соответствии с теоремой 7.6 его частное решение будем искать в виде суммы $y_1 + y_2$ частного решения y_1 уравнения

$$y'' + y = x^2 \tag{7.103}$$

и частного решения y_2 уравнения

$$y'' + y = 2xe^x. \tag{7.104}$$

Начнем с уравнения (7.103). В этом случае $a = 0$, $b = 0$, $P_2(x) = x^2$, $Q(x) = 0$. Число $\gamma = 0$ не совпадает с корнями $\lambda_{1,2} = \pm i$ характеристического уравнения, поэтому частное решение мы будем искать в форме

$$y = Ax^2 + Bx + C \quad (A, B, C - \text{const}).$$

Имеем $y' = 2Ax + B$, $y'' = 2A$. Подставляя в уравнение (7.103), получим

$$2A + Ax^2 + Bx + C = x^2,$$

или

$$(A - 1)\,x^2 + Bx + C + 2A = 0.$$

Следовательно,

$$\begin{cases} A - 1 = 0, \\ B = 0, \\ C + 2A = 0, \end{cases} \quad \text{то есть} \quad \begin{cases} A = 1, \\ B = 0, \\ C = -2. \end{cases}$$

Отсюда находим частное решение уравнения (7.103)

$$y = x^2 - 2. \tag{7.105}$$

Теперь рассмотрим уравнение (7.104). Имеем $a = 1$, $b = 0$, $P_2(x) = 2x$, $Q(x) = 0$. Число $\gamma = 1$ не совпадает с корнями $\lambda_{1,2} = \pm i$ характеристического уравнения, поэтому частное решение ищем в форме

$$y = (Mx + K)e^x \quad (M, K - \text{const}).$$

Находим

$$y' = (Mx + K + M)e^x,$$

$$y'' = (Mx + K + 2M)e^x.$$

Подставляя в уравнение (7.104), получим

$$(2Mx + 2K + 2M)e^{x} = 2x\,e^{x}.$$

$$\begin{cases} 2M = 2, \\ 2M + 2K = 0; \end{cases} \qquad \begin{cases} M = 1, \\ K = -\,1. \end{cases}$$

Отсюда получаем частное решение уравнения (7.104)

$$y = (x - 1\,)e^{x},$$

Следовательно, общим решением исходного уравнения будет

$$y = x^{2} - 2 + (x - 1\,)e^{x} + C_{1}\cos x + C_{2}\sin x.$$

Пример 7.31. Решить уравнение

$$y'' - 2y' + y = e^{x}.$$

Решение. В данном случае $a = 1$, $b = 0$, $P_n(x) = 1$, $Q_k(x) = 0$. Число $\gamma = 1$ совпадает с корнем характеристического уравнения $\lambda = 1$, который имеет кратность 2 (резонансный случай). Частное решение исходного уравнения будем искать в виде

$$y = Ax^{2}e^{x}.$$

Имеем

$$y' = 2Axe^{x} + Ax^{2}e^{x}; \quad y'' = (2A + 4Ax + Ax^{2})e^{x}.$$

Подставляя полученные выражения в исходное уравнение, находим

$$2Ae^{x} = e^{x}, \quad A = 0{,}5.$$

Таким образом, общим решением данного уравнения является

$$y = (0{,}5x^{2} + C_{1}x + C_{2})e^{x}.$$

Пример 7.32. Решить уравнение

$$y^{(4)} + 3y'' - 4y = e^{2x}(-40\cos x - 70\sin x).$$

Решение. Имеем $a = 2$, $b = 1$, $P_n(x) = 64$, $Q_k(x) = 42$. Характеристическое уравнение

$$\lambda^4 + 3\lambda^2 - 4 = 0$$

имеет корни $\lambda_1 = -1$, $\lambda_2 = 1$, $\lambda_3 = 2i$, $\lambda_4 = -2i$, не совпадающие с числом $\gamma = 2+i$, поэтому частное решение будем искать в виде

$$y = e^{2x}(A\cos x + B\sin x) \quad (A, B - \text{const}).$$

Находим

$$y' = e^{2x}((B + 2A)\cos x + (2B - A)\sin x),$$

$$y'' = e^{2x}((4B + 3A)\cos x + (3B - 4A)\sin x),$$

$$y''' = e^{2x}((11B + 2A)\cos x + (2B - 11A)\sin x),$$

$$y^{(4)} = e^{2x}((24B - 7A)\cos x + (-7B - 24A)\sin x).$$

Подставляя выражения для y, y'' и $y^{(4)}$ в исходное уравнение, получаем

$$e^{2x}((36B - 2A)\cos x + (-2B - 36A)\sin x) = e^{2x}(-40\cos x - 70\sin x).$$

Следовательно,

$$\begin{cases} 36B - 2A = -40, \\ -2B - 36A = -70. \end{cases}$$

Решением этой системы является пара $B = -1$, $A = 2$. Таким образом, общим решением исходного уравнения будет

$$y = e^{2x}(2\cos x - \sin x) + C_1 e^{x} + C_2 e^{-x} + C_3 \sin 2x + C_4 \cos 2x.$$

§ 7.13. Системы дифференциальных уравнений

1. Решение той или иной задачи может потребовать нахождения не одной, а сразу нескольких неизвестных функций. Если каждое из этих уравнений является дифференциальным, т.е. имеет вид соотношения, связывающего неизвестные функции $y_1(x)$, $y_2(x)$, … , $y_n(x)$ и их производные, то говорят о *системе дифференциальных уравнений*. Ограничимся рассмотрением систем дифференциальных уравнений первого порядка, у которых число неизвестных функций совпадает с числом уравнений. Это системы вида

$$\begin{cases} F_1(x, y_1, ..., y_n, y_1', ..., y_n') = 0, \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ F_n(x, y_1, ..., y_n, y_1', ..., y_n') = 0. \end{cases} \tag{7.106}$$

Замечание. Часто при рассмотрении систем дифференциальных уравнений в качестве аргумента используют время t. В этом случае производные неизвестных функций принято обозначать через $\dot{y}_1(t), \cdots, \dot{y}_n(t)$.

Отметим некоторые специальные формы записи систем дифференциальных уравнений. Говорят, что система записана в *нормальной форме* (или в *форме Коши*), если она приведена к виду

$$\begin{cases} y_1' = f_1(x, y_1, ..., y_n), \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ y_n' = f_n(x, y_1, ..., y_n). \end{cases} \tag{7.107}$$

Если же система преобразована к виду

$$\frac{dy_1}{f_1(x, y_1, ..., y_n)} = \frac{dy_2}{f_2(x, y_1, ..., y_n)} = \ldots =$$

$$= \frac{dy_n}{f_n(x, y_1, ..., y_n)} = \frac{dx}{g(x, y_1, ..., y_n)}, \tag{7.108}$$

то мы будем говорить, что она записана в *симметричной форме*.

Например, система

$$\begin{cases} \dfrac{y_1'}{4} + \sqrt{y_1} + 5(y_2 x)^2 = 0, \\ y_2' - 3y_1\sqrt{y_2} - x^4 = 0 \end{cases}$$

в нормальной и симметрической формах соответственно выглядит так:

$$\begin{cases} y_1' = -4\sqrt{y_1} - 20(y_2 x)^2, \\ y_2' = 3y_1\sqrt{y_2} + x^4, \end{cases}$$

$$\left\{ \frac{dy_1}{-4\sqrt{y_1} - 20(y_2 x)^2} = \frac{dy_2}{3y_1\sqrt{y_2} + x^4} = \frac{dx}{1}. \right.$$

На системы дифференциальных уравнений естественным образом обобщается постановка задачи Коши для одного уравнения. Так, в случае системы (7.107) задача Коши состоит в нахождении решения, удовлетворяющего начальным условиям

$$y_1(x_0) = y_1^0,\ \ldots,\ y_n(x_0) = y_n^0, \tag{7.109}$$

где $x_0, y_1^0, \ldots, y_n^0$ – заданные числа. Для случая системы может быть доказана теорема существования и единственности решения задачи Коши, аналогичная теореме 7.1.

2. К системам дифференциальных уравнений первого порядка в известном смысле сводятся уравнения (и системы уравнений) любого порядка. Проиллюстрируем это на примере уравнения третьего порядка. Пусть дано уравнение

$$y''' = f(x, y, y', y'').$$

Если обозначить функции y' и y'' соответственно через u и v, то уравнение можно заменить системой

$$\begin{cases} y' = u, \\ u' = v, \\ v' = f(x, y, u, v), \end{cases}$$

состоящей из трех уравнений первого порядка с тремя неизвестными функциями $y(x)$, $u(x)$, $v(x)$. Аналогичное представление допускает любое другое дифференциальное уравнение (или система уравнений).

§ 7.14. Системы линейных дифференциальных уравнений первого порядка

1. Общие сведения о линейных системах

В дальнейшем ограничимся рассмотрением *линейных систем.* Это системы дифференциальных уравнений вида

$$\begin{cases} y_1' = a_{11}(x)y_1 + a_{12}(x)y_2 + \ldots + a_{1n}(x)y_n + f_1(x), \\ y_2' = a_{21}(x)y_1 + a_{22}(x)y_2 + \ldots + a_{2n}(x)y_n + f_2(x), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n' = a_{n1}(x)y_1 + a_{n2}(x)y_2 + \ldots + a_{nn}(x)y_n + f_n(x), \end{cases} \tag{7.110}$$

где коэффициенты a_{ij} и f_i – некоторые функции независимой переменной x. Будем считать их непрерывными; тогда для данной системы заведомо выполняются условия теоремы о существовании и единственности решения задачи Коши. Если все $f_i \equiv 0$, то система называется *однородной*, в противном случае она называется *неоднородной*. Система

$$\begin{cases} y_1' = a_{11}(x)y_1 + a_{12}(x)y_2 + \ldots + a_{1n}(x)y_n, \\ y_2' = a_{21}(x)y_1 + a_{22}(x)y_2 + \ldots + a_{2n}(x)y_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_n' = a_{n1}(x)y_1 + a_{n2}(x)y_2 + \ldots + a_{nn}(x)y_n, \end{cases} \tag{7.111}$$

называется *однородной системой, соответствующей неоднородной системе* (7.110).

При изучении линейных систем удобно использовать матричные обозначения

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \ Y = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix}, \ Y' = \begin{pmatrix} y'_1(x) \\ y'_2(x) \\ \vdots \\ y'_n(x) \end{pmatrix}, \ F = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix},$$

позволяющие записать систему (7.110) в виде одного матричного уравнения

$$Y' = AY+F. \tag{7.112}$$

Например, система

$$\begin{cases} y_1' = 3y_1 - 5y_2 - x^2, \\ y_2' = 4y_1 - 2y_2 + 3x^4 \end{cases}$$

в матричной записи выглядит следующим образом:

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 3 & -5 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} -x^2 \\ 3x^4 \end{pmatrix}.$$

Нетрудно показать, что так же, как и в случае линейных уравнений, общее решение неоднородной системы представляет собой сумму частного решения этой системы и общего решения соответствующей ей однородной системы. В свою очередь, общее решение однородной системы имеет вид

$$Y = C_1 Y_1 + C_2 Y_2 + \ldots + C_n Y_n, \tag{7.113}$$

где C_1, …, C_n– произвольные постоянные, а

$$Y_1 = \begin{pmatrix} y_{11}(x) \\ y_{21}(x) \\ \vdots \\ y_{n1}(x) \end{pmatrix}, \ldots, \; Y_n = \begin{pmatrix} y_{1n}(x) \\ y_{2n}(x) \\ \vdots \\ y_{nn}(x) \end{pmatrix}$$

– произвольные линейно независимые решения, называемые *фундаментальным набором решений* этой системы. Критерием линейной независимости этих решений является неравенство нулю определителя Вронского

$$W(Y_1, \ldots, Y_n) = \begin{vmatrix} y_{11} & \cdots & y_{1n} \\ \ldots & \ldots & \ldots \\ y_{n1} & \cdots & y_{nn} \end{vmatrix}. \tag{7.114}$$

2. Метод сведения линейной системы к одному уравнению более высокого порядка

Один из методов интегрирования линейной системы заключается в сведении системы к одному уравнению n-го порядка с одной неизвестной функцией (сведение одного уравнения произвольного порядка к системе уравнений первого порядка рассматривалось выше). Продемонстрируем это на примере системы двух уравнений.

$$\begin{cases} y_1' = a_{11}y_1 + a_{12}y_2 + q_1(x), \\ y_2' = a_{21}y_1 + a_{22}y_2 + q_2(x). \end{cases} \tag{7.115}$$

Дифференцируя (по x) обе части первого уравнения системы (7.115), находим

$$y_1'' = a_{11}y_1' + a_{12}y_2' + a_{11}'y_1 + a_{12}'y_2 + q_1',$$

откуда, заменяя производные y_1', y_2' их выражениями из самой системы, имеем

$$y_1'' = a_{11}(a_{11}y_1 + a_{12}y_2 + q_1) + a_{12}(\alpha_{21}y_1 + a_{22}y_2 + q_2) + a_{11}'y_1 + a_{12}'y_2 + q_1'.$$

Группируя в правой части, получим уравнение вида

$$y_1'' = b_1y_1 + b_2y_2 + d_1\,, \tag{7.116}$$

где коэффициенты b_1, b_2 и d_1 определенным образом выражаются через коэффициенты a_{ij} и q_1 и их производные (записывать эти выражения не будем). Комбинируя уравнение (7.116) с первым уравнением системы (7.115), получим

$$\begin{cases} y_1' - q_1 = a_{11}y_1 + a_{12}y_2, \\ y_1'' - d_1 = b_1y_1 + b_2y_2. \end{cases} \tag{7.117}$$

Предположим, что в рассматриваемой области изменения x определитель $\begin{vmatrix} a_{11} & a_{12} \\ b_1 & b_2 \end{vmatrix}$ отличен от нуля. Тогда систему (7.117) можно решить относительно y_1 и y_2, т.е. выразить y_1 и y_2 через y_1' и y_1''.

В результате приходим к уравнениям вида

$$y_1 = ky_1' + ly_1'' + p, \tag{7.118}$$

$$y_2 = my_1' + sy_1'' + v \tag{7.119}$$

(выражения для k, l, m, s, p, v приводить не будем). Первое из них представляет собой линейное дифференциальное уравнение второго порядка с одной неизвестной функцией $y_1(t)$. К нему приложима вся теория, изложенная в § 7.10 и § 7.11. Заметим, что если в исходной системе (7.115) все коэффициенты a_{ij} постоянны, то уравнение (7.118) также является уравнением с постоянными коэффициентами; для решения таких уравнений имеется эффективный метод, изложенный в § 7.12.

Пример 7.33. Решить систему

$$\begin{cases} y_1' = 3y_1 - 2y_2 + 2, \\ y_2' = 2y_1 - y_2. \end{cases}$$

Решение. Дифференцируя обе части первого уравнения, имеем

$$y_1'' = 3y_1' - 2y_2' = 3(3y_1 - 2y_2 + 2) - 2(2y_1 - y_2) = 5y_1 - 4y_2 + 6.$$

В комбинации с первым уравнением данной системы это приводит к системе

$$\begin{cases} y_1' - 2 = 3y_1 - 2y_2, \\ y_1'' - 6 = 5y_1 - 4y_2. \end{cases}$$

Отсюда находим выражения для y_1 и y_2 через y_1' и y_1'':

$$y_1 = 2y_1' - y_1'' + 2, \tag{7.120}$$

$$y_2 = \frac{5}{2}y_1' - \frac{3}{2}y_1'' + 4. \tag{7.121}$$

В результате приходим к уравнению второго порядка для неизвестной функции $y_1(t)$:

$$y_1'' - 2y_1' + y_1 = 2.$$

Решая это уравнение известным способом, получим

$$y_1 = (C_1 + C_2 x)e^x + 2,$$

после чего из (7.121) находим

$$y_2 = \left(C_1 - \frac{C_2}{2} + C_2 x\right)e^x + 4.$$

§ 7.15. Линейные системы дифференциальных уравнений с постоянными коэффициентами

1.Однородные линейные системы

Рассмотрим однородную линейную систему (7.111). В матричном виде она записывается следующим образом:

$$Y' = AY. \tag{7.122}$$

Если все элементы a_{ij} матрицы A постоянны, то для решения системы можно использовать методы линейной алгебры. Прежде всего заметим, что однородная система имеет очевидное частное решение

$$y_1(x) \equiv 0, \ \dots \ , y_n(x) \equiv 0.$$

Это решение называется *тривиальным* (*нулевым*). Интерес представляют, конечно, нетривиальные решения. Будем искать такие решения в виде

$$y_1 = p_1 e^{\lambda t}, \dots, \ y_n = p_n e^{\lambda t}, \tag{7.123}$$

или, используя матричную запись, в виде

$$Y = Pe^{\lambda t}, \tag{7.124}$$

где $P=(p_1, p_2, \ldots, p_n)^T$ – ненулевая матрица (вектор-столбец) с постоянными элементами.

Имеем

$$Y' = \lambda P e^{\lambda t}.$$

Подставляя выражения для Y и Y' в уравнение (7.122), получим

$$\lambda P e^{\lambda t} = A P e^{\lambda t},$$

откуда после сокращения на $e^{\lambda t}$ находим

$$AP = \lambda P. \tag{7.125}$$

Это уравнение говорит о том, что λ является собственным значением матрицы A, а P – собственным вектором, соответствующим λ.

Определение. *Характеристическое уравнение матрицы* A

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0 \tag{7.126}$$

называется ***характеристическим уравнением*** *однородной линейной системы* (7.122) *с постоянными коэффициентами.*

Замечание 1. Напомним, что для нахождения собственного вектора P, соответствующего собственному значению λ матрицы A, необходимо найти решение следующей алгебраической системы уравнений:

$$\begin{pmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{7.127}$$

Так же, как и в случае линейных уравнений, имея корни характеристического уравнения, мы можем построить общее решение однородной системы. Для простоты изложения продемонстрируем это на примере систем двух уравнений с двумя неизвестными, т.е. систем вида

$$\begin{cases} y_1' = a_{11}y_1 + a_{12}y_2, \\ y_2' = a_{21}y_1 + a_{22}y_2. \end{cases} \tag{7.128}$$

Однако заметим, что полученные в этом случае результаты могут быть без труда перенесены на случай систем большего числа уравнений.

Характеристическое уравнение

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0$$

системы (7.128) является алгебраическим уравнением второго порядка. При его решении могут возникнуть три случая.

Случай 1. Собственные значения λ_1 и λ_2 действительные и различные. Тогда соответствующие им собственные векторы P_1 и P_2 будут действительными и линейно независимыми. Определяемые ими два частных решения уравнения (7.122)

$$Y_1 = P_1 e^{\lambda_1 t}, \qquad Y_2 = P_2 e^{\lambda_2 t}$$

также будут линейно независимыми, так как определитель Вронского

$$W(Y_1, Y_2) = e^{(\lambda_1 + \lambda_2)x} \begin{vmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{vmatrix}$$

отличен от нуля. Общее же решение, как это следует из (7.113), имеет вид

$$Y = C_1 Y_1 + C_2 Y_2, \tag{7.129}$$

где C_1, C_2 – произвольные постоянные.

Пример 7.34. Решить систему

$$\begin{cases} y_1' = 7y_1 + 6y_2, \\ y_2' = -y_1 + 2y_2. \end{cases}$$

Решение. Здесь

$$A = \begin{pmatrix} 7 & 6 \\ -1 & 2 \end{pmatrix}.$$

Характеристическое уравнение матрицы A имеет вид

$$\begin{vmatrix} 7-\lambda & 6 \\ -1 & 2-\lambda \end{vmatrix} = 0,$$

или

$$\lambda^2 - 9\lambda + 20 = 0.$$

Его корни $\lambda_1 = 5$, $\lambda_2 = 4$. Найдем теперь собственные векторы, отвечающие найденным собственным значениям.

Для $\lambda = 5$ получаем следующую систему

$$\begin{pmatrix} 2 & 6 \\ -1 & -3 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Ее решением будет

$$p_1 = -3p_2.$$

Полагая $p_2 = 1$, находим $p_1 = -3$. Таким образом,

$$P_1 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Аналогично в случае $\lambda = 4$ получаем

$$P_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Следовательно, частными решениями системы являются

$$Y_1 = \begin{pmatrix} -3 \\ 1 \end{pmatrix} e^{5x}, \quad Y_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{4x}.$$

Общее решение в матричной записи имеет вид

$$Y = C_1 \begin{pmatrix} -3 \\ 1 \end{pmatrix} e^{5x} + C_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix} e^{4x},$$

а в развернутой форме оно запишется следующим образом:

$$y_1(x) = -3C_1 e^{5x} - 2C_2 e^{4x}, \quad y_2(x) = C_1 e^{5x} + C_2 e^{4x}.$$

Случай 2. Собственные значения λ_1 и λ_2 комплексно сопряжены: $\lambda_{1,2} = \alpha \pm \beta i$, где $\beta \neq 0$. Покажем, что соответствующие им собственные векторы также будут комплексно сопряжены. Действительно, пусть P – собственный вектор, отвечающий λ_1 (разумеется, комплексный). Тогда из (7.125) имеем

$$\overline{A}\,\overline{P} = \overline{\lambda}_1 \overline{P}$$

(черта обозначает операцию комплексного сопряжения) или, учитывая, что $\overline{A} = A$ (матрица A состоит из действительных чисел) и $\overline{\lambda}_1 = \lambda_2$,

$$A\overline{P} = \lambda_2 \overline{P}.$$

Таким образом, вектор $\overline{P}$ является собственным, отвечающим собственному значению λ_2 и, следовательно, частные решения $Y_1 = Pe^{\lambda_1 t}$, $Y_2 = \overline{P} e^{\overline{\lambda}_1 t}$ комплексно сопряжены. Чтобы получить действительные решения, заменим Y_1 и Y_2 их линейными комбинациями

$$Y_1^* = \frac{1}{2}(Y_1 + Y_2), \qquad Y_2^* = \frac{1}{2i}(Y_1 - Y_2),$$

которые, как нетрудно видеть, представляют собой соответственно действительную и мнимую части Y_1. Итак, общее решение в этом случае имеет вид

$$Y = C_1 Y_1^* + C_2 Y_2^*.$$

Пример 7.35. Решить систему

$$\begin{cases} y_1' = y_1 + y_2, \\ y_2' = -3y_1 + y_2. \end{cases}$$

Решение. Решая характеристическое уравнение

$$\begin{vmatrix} 1-\lambda & 1 \\ -3 & 1-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda^2 - 2\lambda + 4 = 0,$$

находим $\lambda_1 = 1 + 2i$, $\lambda_2 = 1 - 2i$. Соответствующими им собственными векторами будут

$$P_1 = \begin{pmatrix} -i \\ 2 \end{pmatrix}, \quad P_2 = \begin{pmatrix} i \\ 2 \end{pmatrix}.$$

Рассмотрим частное решение, соответствующее λ_1:

$$Y_1 = e^{(1+2i)x} \begin{pmatrix} -i \\ 2 \end{pmatrix}.$$

На основании формулы Эйлера получаем

$$Y_1 = e^x (\cos 2x + i \sin 2x) \begin{pmatrix} -i \\ 2 \end{pmatrix}.$$

$$Y_1 = e^x \begin{pmatrix} \sin 2x - i\cos 2x \\ 2\cos 2x + 2i \sin 2x \end{pmatrix} = e^x \begin{pmatrix} \sin 2x \\ 2\cos 2x \end{pmatrix} + ie^x \begin{pmatrix} -\cos 2x \\ 2\sin 2x \end{pmatrix}.$$

Вычислив действительную и мнимую части Y_1, находим два линейно независимых решения исходного уравнения

$$Y_1^* = e^x \begin{pmatrix} \sin 2x \\ 2\cos 2x \end{pmatrix}, \qquad Y_2^* = e^x \begin{pmatrix} -\cos 2x \\ 2\sin 2x \end{pmatrix}.$$

Таким образом, общее решение в матричной записи имеет вид

$$Y = C_1 e^x \begin{pmatrix} \sin 2x \\ 2\cos 2x \end{pmatrix} + C_2 e^x \begin{pmatrix} -\cos 2x \\ 2\sin 2x \end{pmatrix}.$$

Отсюда

$$y_1(x) = C_1 e^x \sin 2x - C_2 e^x \cos 2x, \quad y_2(x) = 2C_1 e^x \cos 2x + 2C_2 e^x \sin 2x.$$

Случай 3. Характеристическое уравнение имеет единственный корень λ (кратности 2), которому соответствуют два линейно независимых собственных вектора P_1 и P_2 (т.е. кратность корня совпадает с числом линейно независимых собственных векторов). Векторы P_1 и P_2 порождают два линейно независимых решения

$$Y_1 = e^{\lambda t} P_1, \qquad Y_2 = e^{\lambda t} P_2,$$

и общее решение, так же как и в случае 1, находится по формуле (7.129).

Замечание 2. Предоставляем читателю возможность доказать самостоятельно, что для системы двух уравнений данный случай может иметь место только лишь для систем вида

$$\begin{cases} y_1' = \lambda y_1, \\ y_2' = \lambda y_2. \end{cases}$$

Случай 4. Характеристическое уравнение имеет единственный корень λ (кратности 2), которому с точностью до постоянного множителя соответствует один собственный вектор P_1 (т.е. кратность корня больше числа линейно независимых собственных векторов). В этом случае для отыскания решения целесообразно применить *метод неопределенных коэффициентов*, который мы уже использовали при решении линейных дифференциальных уравнений. Согласно этому методу общее решение необходимо искать в форме

$$\begin{cases} y_1 = e^{\lambda x}\left(c_{11} + c_{12}x\right), \\ y_2 = e^{\lambda x}\left(c_{21} + c_{22}x\right), \end{cases} \tag{7.130}$$

где постоянные c_{ij} требуют определения путем подстановки этих выражений в исходную однородную систему.

Пример 7.36. Решить систему

$$\begin{cases} y_1' = y_1 + y_2, \\ y_2' = -y_1 + 3y_2. \end{cases}$$

Решение. Характеристическое уравнение

$$\lambda^2 - 4\lambda + 4 = 0$$

имеет единственный корень $\lambda = 2$ (кратности 2). Ему соответствует единственный собственный вектор

$$P_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Поэтому решение в этом случае будем искать в виде

$$\begin{cases} y_1 = e^{2x}\left(c_{11} + c_{12}x\right), \\ y_2 = e^{2x}\left(c_{21} + c_{22}x\right). \end{cases}$$

Подставляя выражения для y_1 и y_2 в исходную систему, находим

$$\begin{cases} e^{2x}(2c_{11} + c_{12} + 2c_{12}x) = e^{2x}(c_{11} + c_{21} + (c_{12} + c_{22})x), \\ e^{2x}(2c_{21} + c_{22} + 2c_{22}x) = e^{2x}(3c_{21} - c_{11} + (3c_{22} - c_{12})x). \end{cases}$$

Отсюда получаем систему

$$\begin{cases} 2c_{11} + c_{12} = c_{11} + c_{21}, \\ 2c_{12} = c_{12} + c_{22}, \\ 2c_{21} + c_{22} = 3c_{21} - c_{11}, \\ 2c_{22} = 3c_{22} - c_{12}. \end{cases}$$

Решая ее, находим

$$\begin{cases} c_{11} = C_1, \\ c_{12} = C_2, \\ c_{21} = C_1 + C_2, \\ c_{22} = C_2, \end{cases}$$

где C_1, C_2 – произвольные постоянные. Таким образом, общее решение системы имеет вид

$$\begin{cases} y_1 = e^{2x}(C_1 + C_2 x), \\ y_2 = e^{2x}(C_1 + C_2 + C_2 x). \end{cases} \tag{7.131}$$

Замечание 3. Для решения однородных систем в случае, когда корень характеристического уравнения λ кратный и ему соответствует единственный собственный вектор P_1, может быть также применен *метод присоединенных векторов*. Суть его такова. Пусть P_2 – вектор-столбец, являющийся решением уравнения

$$AP_2 = \lambda P_2 + P_1, \tag{7.132}$$

тогда однородная система (7.128) имеет два линейно независимых решения

$$Y_1 = e^{\lambda x} P_1, \quad Y_2 = e^{\lambda x}(xP_1 + P_2).$$

Покажем, что Y_2 является решением. Имеем

$$Y_2' = e^{\lambda x}(\lambda x P_1 + \lambda P_2 + P_1).$$

Учитывая, что P_1 – собственный вектор, а P_2 удовлетворяет условию (7.132), получаем

$$Y_2' = e^{\lambda x}(xAP_1 + AP_2) = e^{\lambda x}A(xP_1 + P_2) = AY_2.$$

Нетрудно также убедиться, что Y_1 и Y_2 линейно независимы. Следовательно, они образуют фундаментальный набор решений, и общее решение может быть найдено по формуле (7.129).

В общем случае корню характеристического уравнения λ кратности $k > 1$, имеющему один собственный вектор P_1, соответствует k линейно независимых решений

$$Y_1 = e^{\lambda x}P_1,\ Y_2 = e^{\lambda x}(xP_1 + P_2),\ Y_3 = e^{\lambda x}\left(\frac{x^2}{2!}P_1 + \frac{x}{1!}P_2 + P_3\right), \ldots$$

$$Y_k = e^{\lambda x}\left(\frac{x^k}{k!}P_1 + \frac{x^{k-1}}{(k-1)!}P_2 + \ldots + \frac{x}{1!}P_{k-1} + P_k\right), \tag{7.133}$$

где присоединенные векторы $P_2, P_3, \ldots, P_k$ являются последовательными решениями следующих алгебраических систем

$$\begin{gathered} AP_2 = \lambda P_2 + P_1, \\ \ldots\ldots\ldots\ldots\ldots \\ AP_k = \lambda P_k + P_{k-1}. \end{gathered} \tag{7.134}$$

Пример 7.37. Решить методом присоединенных векторов систему

$$\begin{cases} y_1' = y_1 + y_2, \\ y_2' = -y_1 + 3y_2. \end{cases}$$

Решение. Данная система рассматривалась в примере 7.36, где было установлено, что характеристическое уравнение имеет единственный корень $\lambda = 2$ кратности 2, которому соответствует единственный собственный вектор

$$P_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

С учетом этого, соотношения (7.131) примут вид

$$\begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} p_{21} \\ p_{22} \end{pmatrix} = 2\begin{pmatrix} p_{21} \\ p_{22} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

где p_{21} и p_{22} – элементы вектора P_2. Они равносильны уравнению

$$p_{22} = 1 - p_{21},$$

где в качестве свободной переменной выбрана p_{21}. Таким образом, у нас имеется некоторый произвол в выборе вектора P_2. Полагая $p_{21} = 0$, получаем $p_{22} = 1$. Следовательно,

$$P_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

По формулам (7.133) находим фундаментальный набор решений системы

$$Y_1 = e^{2x}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad Y_2 = e^{2x}\begin{pmatrix} x \\ x+1 \end{pmatrix}.$$

Общее же решение в матричной форме имеет вид

$$Y = C_1\begin{pmatrix} 1 \\ 1 \end{pmatrix}e^{2x} + C_2\begin{pmatrix} x \\ x+1 \end{pmatrix}e^{2x}.$$

Нетрудно убедиться, что оно совпадает с решением, полученным в примере 7.36 (см. соотношение (7.131)).

Как отмечалось выше, методы, разработанные для решения однородных систем из двух уравнений, могут быть перенесены на случай систем большего числа уравнений. Продемонстрируем это на примерах.

Пример 7.38. Решить систему

$$\begin{cases} y_1' = 5y_1 - 2y_2 - 2y_3, \\ y_2' = 2y_1 + y_2 - 2y_3, \\ y_3' = 2y_1 - 2y_2 + y_3. \end{cases}$$

Р е ш е н и е . Составим характеристическое уравнение системы

$$\begin{vmatrix} 5-\lambda & -2 & -2 \\ 2 & 1-\lambda & -2 \\ 2 & -2 & 1-\lambda \end{vmatrix} = 0 .$$

Раскрывая определитель, получаем

$$\lambda^3 - 7\lambda^2 + 15\lambda - 9 = 0.$$

Данное уравнение после несложных преобразований принимает вид

$$(\lambda - 1)(\lambda^2 - 6\lambda + 9) = 0.$$

Отсюда находим: $\lambda_1 = 1$ (простой корень), ему соответствует собственный вектор

$$P_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

и $\lambda_2 = 3$ (корень кратности 2), которому соответствуют два линейно независимых собственных вектора

$$P_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Следовательно, общее решение системы имеет вид

$$Y = C_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^{x} + C_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} e^{3x} + C_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^{3x}.$$

П р и м е р 7.39. Решить задачу Коши

$$\begin{cases} y_1' = 3y_1 + 2y_2 - 2y_3, \\ y_2' = 4y_1 + 3y_2 - 4y_3, \\ y_3' = 4y_1 + 2y_2 - 3y_3, \end{cases}$$

$$y_1(0) = 2,\ y_2(0) = 2,\ y_3(0) = 3.$$

Р е ш е н и е . Характеристическое уравнение системы имеет вид

$$\lambda^3 - 3\lambda^2 - \lambda + 3 = 0.$$

Его корнями будут $\lambda_1 = -1$, $\lambda_2 = 1$, $\lambda_3 = 3$. Им соответствуют следующие собственные векторы:

$$P_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},\ P_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},\ P_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Следовательно, общее решение системы имеет вид

$$Y = C_1 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} e^{-x} + C_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^{x} + C_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^{3x}. \qquad (7.135)$$

Полагая $x = 0$ в (7.135) и принимая во внимание начальные условия, мы приходим к следующей системе:

$$\begin{cases} C_2 + C_3 = 2, \\ C_1 + C_3 = 2, \\ C_1 + C_2 + C_3 = 3. \end{cases}$$

Решая ее, находим

$$\begin{cases} C_1 = 1, \\ C_2 = 1, \\ C_3 = 1. \end{cases}$$

Таким образом, решением задачи Коши будет

$$Y = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} e^{-x} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^{x} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^{3x}.$$

Пример 7.40. Решить систему

$$\begin{cases} y_1' = y_1 + 2y_2 - y_3, \\ y_2' = 2y_1 + 3y_2 - 3y_3, \\ y_3' = y_1 + 3y_2 - 2y_3. \end{cases}$$

Решение. Решая характеристическое уравнение

$$\lambda^3 - 2\lambda^2 + \lambda - 2 = 0,$$

находим $\lambda_1 = i$, $\lambda_2 = -i$, $\lambda_3 = 2$. Соответствующими им собственными векторами будут

$$P_1 = \begin{pmatrix} 1-i \\ 1+i \\ 2 \end{pmatrix}, \ P_2 = \begin{pmatrix} 1+i \\ 1-i \\ 2 \end{pmatrix}, \ P_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Рассмотрим частное решение, соответствующее λ_1:

$$Y_1 = e^{ix} \begin{pmatrix} 1-i \\ 1+i \\ 2 \end{pmatrix}.$$

Оно является комплексным. Для нахождения действительных решений вычислим действительную и мнимую части Y_1. На основании формулы Эйлера

$$Y_1 = (\cos x + i\sin x)\begin{pmatrix} 1-i \\ 1+i \\ 2 \end{pmatrix} = \begin{pmatrix} \sin x + \cos x + i(\sin x - \cos x) \\ -\sin x + \cos x + i(\sin x + \cos x) \\ 2\cos x + 2i\sin x \end{pmatrix} =$$

$$= \begin{pmatrix} \sin x + \cos x \\ -\sin x + \cos x \\ 2\cos x \end{pmatrix} + i\begin{pmatrix} \sin x - \cos x \\ \sin x + \cos x \\ 2\sin x \end{pmatrix}.$$

Отсюда получаем два линейно независимых решения исходного уравнения

$$Y_1^* = \begin{pmatrix} \sin x + \cos x \\ -\sin x + \cos x \\ 2\cos x \end{pmatrix}, \qquad Y_2^* = \begin{pmatrix} \sin x - \cos x \\ \sin x + \cos x \\ 2\sin x \end{pmatrix}.$$

Вместе с частным решением

$$Y_3 = e^{2x}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

соответствующим вектору P_3, они образуют фундаментальный набор решений системы. Таким образом, общее решение имеет вид

$$Y = C_1\begin{pmatrix} \sin x + \cos x \\ -\sin x + \cos x \\ 2\cos x \end{pmatrix} + C_2\begin{pmatrix} \sin x - \cos x \\ \sin x + \cos x \\ 2\sin x \end{pmatrix} + C_3 e^{2x}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Пример 7.41. Решить систему

$$\begin{cases} y_1' = 3y_1 + y_2 - y_3, \\ y_2' = y_1 + y_2 + y_3, \\ y_3' = 2y_1 + 2y_3. \end{cases}$$

Решение. Характеристическое уравнение системы

$$\lambda^3 - 6\lambda^2 + 12\lambda - 8 = 0$$

имеет единственный корень $\lambda = 2$ кратности 3, которому соответствует единственный собственный вектор

$$P_1 = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

В этом случае целесообразно применить метод присоединенных векторов (см. замечание 3). В соответствии с этим методом первоначально мы должны найти решение системы матричных уравнений (7.134). Решением первого уравнения системы (7.134) является собственный вектор P_1. Второе уравнение системы $AP_2 = \lambda P_2 + P_1$ равносильно следующей алгебраической системе:

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

Одним из решений этой системы является вектор

$$P_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

В свою очередь, решая систему

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},$$

соответствующую матричному уравнению $AP_3 = 2P_3 + P_2$, находим

$$P_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Тогда, как это следует из (7.133), фундаментальный набор решений данной системы имеет вид

$$Y_1 = e^{2x} \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, \quad Y_2 = e^{2x} \begin{pmatrix} 1 \\ 2x+1 \\ 2x+2 \end{pmatrix}, \quad Y_3 = e^{2x} \begin{pmatrix} 1 \\ x^2+x \\ x^2+2x \end{pmatrix}.$$

А общим решением исходной системы дифференциальных уравнений будет

$$Y = C_1 e^{2x} \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} + C_2 e^{2x} \begin{pmatrix} 1 \\ 2x+1 \\ 2x+2 \end{pmatrix} + C_3 e^{2x} \begin{pmatrix} 1 \\ x^2+x \\ x^2+2x \end{pmatrix}.$$

2. Неоднородные линейные системы

Для решения неоднородных линейных систем применяются методы, аналогичные методам, используемым для решения неоднородных линейных уравнений. Одним из таких методов является хорошо известный нам *метод вариации постоянных.* Продемонстрируем его суть на следующем примере.

Пример 7.42. Решить систему

$$\begin{cases} y_1' = y_1 + 2y_2 + 8 + e^{-x}\ln x, \\ y_2' = 3y_1 + 2y_2 + 12 - e^{-x}\ln x. \end{cases}$$

Решение. Решая характеристическое уравнение

$$\lambda^2 - 3\lambda - 4 = 0,$$

находим корни $\lambda_1 = -1$, $\lambda_2 = 4$. Собственными векторами, отвечающими найденным собственным значениям, будут следующие векторы:

$$P_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Следовательно, общее решение соответствующей однородной системы имеет вид

$$\widetilde{Y} = C_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-x} + C_2 \begin{pmatrix} 2 \\ 3 \end{pmatrix} e^{4x}.$$

Решение неоднородного уравнения методом вариации постоянной будем искать в форме

$$Y = C_1(x) \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-x} + C_2(x) \begin{pmatrix} 2 \\ 3 \end{pmatrix} e^{4x}.$$

Для нахождения $C_1(x)$ и $C_2(x)$, подставив выражение для Y в исходную систему, получим

$$\begin{cases} C_1' e^{-x} + 2C_2' e^{4x} = 8 + e^{-x}\ln x, \\ -C_1' e^{-x} + 3C_2' e^{4x} = 12 - e^{-x}\ln x. \end{cases}$$

Отсюда находим:

$$\begin{cases} C_1' = \ln x, \\ C_2' = 4e^{-4x}; \end{cases}$$

$$\begin{cases} C_1(x) = x\ln x - x + \tilde{C}_1, \\ C_2(x) = -e^{-4x} + \tilde{C}_2, \end{cases}$$

где $\tilde{C}_1, \tilde{C}_2$ – произвольные постоянные. Таким образом, решение исходной системы будет

$$Y = \begin{pmatrix} x\ln x - x + \tilde{C}_1 \\ -x\ln x + x - \tilde{C}_1 \end{pmatrix} e^{-x} + \begin{pmatrix} 2(-e^{-4x} + \tilde{C}_2) \\ 3(-e^{-4x} + \tilde{C}_2) \end{pmatrix} e^{4x}.$$

В случае, когда столбец свободных членов системы имеет специальный вид

$$F = e^{ax}\,(P_m(x) \cos bx + Q_k(x) \sin bx),$$

где $P_m(x)$ и $Q_k(x)$ – вектор-столбцы, элементами которых являются многочлены от x степени, не превышающей соответственно n и k, для отыскания частного решения уравнения целесообразно воспользоваться *методом неопределенных коэффициентов*. Для систем он имеет определенную специфику. Суть метода такова.

Если число $\gamma = a + bi$ не является корнем характеристического уравнения, то частное решение ищется в виде

$$Y = e^{ax}(\tilde{P}_m(x)\cos bx + \tilde{Q}_m(x)\sin bx),$$

где $\tilde{P}_m(x)$ и $\tilde{Q}_m(x)$ – вектор-столбцы, элементами которых являются многочлены от x степени $m = \max\{k, n\}$.

Если же γ является корнем характеристического уравнения кратности l (*резонансный случай*), то частное решение ищется в форме

$$Y = e^{ax}\left(\tilde{P}_{m+l}(x)\cos bx + \tilde{Q}_{m+l}(x)\sin bx\right).$$

Пример 7.43. Решить систему

$$\begin{cases} y_1' = y_1 + 3y_2 + 2x, \\ y_2' = 2y_1 + 2y_2 + 1. \end{cases}$$

Решение. Решая характеристическое уравнение системы

$$\lambda^2 - 3\lambda - 4 = 0,$$

находим его корни $\lambda_1 = -1$, $\lambda_2 = 4$. Собственными векторами, отвечающими найденным собственным значениям, будут соответственно

$$P_1 = \begin{pmatrix} -3 \\ 2 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Следовательно, общее решение однородной системы, соответствующей данной неоднородной, имеет вид

$$\widetilde{Y} = C_1 \begin{pmatrix} -3 \\ 2 \end{pmatrix} e^{-x} + C_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{4x}.$$

Теперь найдем частное решение. В рассматриваемом случае элементы столбца свободных членов представляют собой многочлены степени, не превышающей 1, и так как число $\gamma = 0$ не совпадает с корнями характеристического уравнения, то частное решение неоднородной системы будем искать в виде

$$\hat{Y} = \begin{pmatrix} px + c \\ qx + d \end{pmatrix},$$

где p, q, c и d – некоторые постоянные. Для их определения подставим выражение для $\hat{Y}$ в исходную систему. Получим

$$\begin{cases} p = px + c + 3qx + 3d + 2x, \\ q = 2px + 2c + 2qx + 2d + 1. \end{cases}$$

Отсюда

$$\begin{cases} p = c + 3d, \\ 0 = p + 3q + 2, \\ q = 2c + 2d + 1, \\ 0 = 2p + 2q. \end{cases}$$

Решив эту систему, находим $p = 1$, $q = -1$, $c = -2$ и $d = 1$. Следовательно,

$$\widehat{Y} = \begin{pmatrix} x - 2 \\ -x + 1 \end{pmatrix}.$$

Так как общее решение неоднородной системы Y представляет собой сумму частного решения $\widehat{Y}$ и общего решения $\widetilde{Y}$ соответствующей однородной системы, то окончательно получаем

$$Y = \begin{pmatrix} x - 2 \\ -x + 1 \end{pmatrix} + C_1 \begin{pmatrix} -3 \\ 2 \end{pmatrix} e^{-x} + C_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{4x}.$$

Пример 7.44. Решить систему

$$\begin{cases} y_1' = -4y_1 + y_2 + e^x, \\ y_2' = -5y_1 + 2y_2 + 5e^x. \end{cases}$$

Решение. Решим характеристическое уравнение системы

$$\lambda^2 + 2\lambda - 3 = 0.$$

Его корнями будут $\lambda_1 = 1$, $\lambda_2 = -3$. Им соответствуют собственные векторы

$$P_1 = \begin{pmatrix} 1 \\ 5 \end{pmatrix}, \qquad P_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Следовательно, общее решение соответствующей однородной системы имеет вид

$$\widetilde{Y} = C_1 \begin{pmatrix} 1 \\ 5 \end{pmatrix} e^x + C_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{-3x}.$$

Теперь найдем частное решение. В данном случае число $\gamma = 1$ совпадает с корнем λ_1 характеристического уравнения (резонансный случай). Так как элементы столбца свободных членов представляют собой многочлены нулевой степени, частное решение неоднородной системы будем искать в виде

$$\widehat{Y} = \begin{pmatrix} px + c \\ qx + d \end{pmatrix} e^x,$$

где p, q, c и d – некоторые постоянные. Подставим выражение для $\widehat{Y}$ в исходную систему. Имеем

$$\begin{cases} (px + c + p)e^x = (-4px - 4c + qx + d + 1)e^x, \\ (qx + d + q)e^x = (-5px - 5c + 2qx + 2d + 5)e^x. \end{cases}$$

Отсюда

$$\begin{cases} p = -4p + q, \\ c + p = -4c + d + 1, \\ q = -5p + 2q, \\ d + q = -5c + 2d + 5. \end{cases}$$

Решая эту систему, находим

$$\begin{cases} p = 1, \\ q = 5, \\ d = 5c. \end{cases}$$

Полагая $c = 1$, получаем $d = 5$. Следовательно,

$$\widehat{Y} = \begin{pmatrix} x + 1 \\ 5x + 5 \end{pmatrix} e^x.$$

Таким образом, общее решение системы имеет вид

$$Y=\begin{pmatrix} x+1 \\ 5x+5 \end{pmatrix}e^{x}+C_1\begin{pmatrix} 1 \\ 5 \end{pmatrix}e^{x}+C_2\begin{pmatrix} 1 \\ 1 \end{pmatrix}e^{-3x}.$$

§ 7.16. Разностные уравнения

1. Общие понятия и примеры

Определение. *Уравнение вида*

$$F(n, y_n, y_{n+1},\ldots, y_{n+k}) = 0, \tag{7.136}$$

где k – фиксированное, а n – произвольное натуральное число, y_n, y_{n+1}, … , y_{n+k} – члены некоторой числовой последовательности, называется ***разностным уравнением k-го порядка.***

Решить разностное уравнение означает найти его *общее решение*, т.е. все последовательности $y_n = y(n)$, удовлетворяющие уравнению (7.136).

Разностные уравнения используются в моделях экономической динамики с дискретным временем, а также для приближенного решения дифференциальных уравнений.

Некоторые типы разностных уравнений нам знакомы еще из школьной математики. Так, например, разностное уравнение второго порядка

$$2y_{n+1} = y_n + y_{n+2}$$

задает арифметическую прогрессию. Его решением является последовательность $y_n = a_1 + d(n-1)$, где a_1 и d – произвольные действительные числа. Аналогично уравнение

$$y_{n+1}^2 = y_n \cdot y_{n+2}$$

определяет геометрическую прогрессию, и его решением является последовательность $y_n = b_1 q^{n-1}$, где b_1 и q – произвольные действительные числа.

Между теориями разностных и дифференциальных уравнений прослеживается определенная аналогия. Если в уравнении (7.136) произвести формальную замену:

$$n \mapsto x,\ \ y_n \mapsto y(x),\ \ y_{n+1} \mapsto y'(x),\ \ \ldots,\ \ y_{n+k} \mapsto y^{(k)}(x), \quad (7.137)$$

то определение разностного уравнения трансформируется в общее определение обыкновенного дифференциального уравнения порядка k. Проведя формальную замену (7.137), нетрудно получить *нормальную форму* записи разностного уравнения

$$y_{n+k} = f(n, y_n, y_{n+1}, \ldots, y_{n+k-1}). \quad (7.138)$$

Аналогичным образом определяется и *задача Коши* – как задача нахождения решения уравнения (7.138), удовлетворяющего начальным условиям

$$y(n_0) = a_0,\ \ y(n_0+1) = a_1,\ \ldots,\ y(n_0+k-1) = a_{k-1}. \quad (7.139)$$

Ниже будет показано, что решения соответствующих классов дифференциальных и разностных уравнений (например, линейных) осуществляются схожими методами.

Имеет место следующая теорема.

Теорема 7.10. *Решение* $y_n = y(n)$ *задачи Коши* (7.138), (7.139) *при* $n \ge n_0$ *определено однозначно.*

Доказательство. Подставляя значения для $y(n_0)$, $y(n_0+1)$, … , $y(n_0+k-1)$ из (7.139) в (7.138), мы находим $y(n_0+k)$. Это, в свою очередь, дает нам возможность определить $y(n_0+k+1)$. Продолжая этот процесс, можно рекуррентным способом по формуле (7.138) найти любое значение $y(n)$ при $n \ge n_0$. Теорема доказана.

Следует, однако, заметить, что нахождение общего решения разностного уравнения, в отличие от задачи Коши, является гораздо более сложной задачей.

2. Линейные разностные уравнения

Определение. *Разностное уравнение вида*

$$a_k(n)\, y_{n+k} + \ldots + a_1(n)\, y_{n+1} + a_0(n)\, y_n = f(n), \qquad (7.140)$$

где $a_0, a_1, \ldots, a_k, f$ *– некоторые функции от* n $(a_0 \neq 0,\ a_k \neq 0)$, *называется* ***линейным разностным уравнением k-го порядка.***

Введем следующее обозначение

$$L(y) = a_k y_{n+k} + \ldots + a_1 y_{n+1} + a_0 y_n. \qquad (7.141)$$

Это выражение называется *линейным разностным оператором k-го порядка*.

С учетом этих обозначений уравнение (7.141) может быть записано в виде

$$L(y) = f(n). \qquad (7.142)$$

Уравнение

$$L(y) = 0 \qquad (7.143)$$

называется *линейным однородным разностным уравнением*, соответствующим уравнению (7.142). Само же уравнение (7.142) (при $f(n) \not\equiv 0$) называется *неоднородным.* Проведя рассуждения, аналогичные использованным в § 7.10 и § 7.11, нетрудно убедиться в справедливости следующих утверждений.

Теорема 7.11. (***об общем решении линейного неоднородного уравнения***). *Общее решение линейного неоднородного разностного уравнения* (7.142) *есть сумма частного решения* $\overline{y}(n)$ *этого уравнения и общего решения соответствующего ему однородного уравнения* (7.143).

Теорема 7.12 (*об общем решении линейного однородного уравнения*). *Пусть* $y^{(1)}(n)$, ..., $y^{(k)}(n)$ *– система, состоящая из k линейно независимых решений линейного однородного разностного уравнения, тогда общее решение этого уравнения задается формулой*

$$y(n) = C_1 y^{(1)}(n) + \ldots + C_k y^{(k)}(n). \tag{7.144}$$

Множество решений линейного однородного разностного уравнения k-го порядка образует k-мерное линейное пространство, а любой набор $y^{(1)}(n)$, ..., $y^{(k)}(n)$ из k линейно независимых решений (называемый *фундаментальным набором*) будет его базисом. Признаком линейной независимости решений $y^{(1)}(n)$, ..., $y^{(k)}(n)$ однородного уравнения является неравенство нулю *определителя Казоратти*

$$\Delta = \begin{vmatrix} y_n^{(1)} & y_n^{(2)} & \cdots & y_n^{(k)} \\ y_{n+1}^{(1)} & y_{n+1}^{(2)} & \cdots & y_{n+1}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n+k}^{(1)} & y_{n+k}^{(k)} & \cdots & y_{n+k}^{(k)} \end{vmatrix}, \tag{7.145}$$

который представляет собой аналог определителя Вронского в теории обыкновенных дифференциальных уравнений.

В случае, когда коэффициенты a_0, a_1, … , a_k являются постоянными, методы решения линейного однородного разностного уравнения

$$a_k y_{n+k} + \ldots + a_1 y_{n+1} + a_0 y_n = 0 \tag{7.146}$$

во многом аналогичны методам решения дифференциальных уравнений с постоянными коэффициентами. Действительно, будем искать решения уравнения в виде

$$y_n = \lambda^n, \tag{7.147}$$

где $\lambda \neq 0$ – некоторая постоянная. Подставляя выражение для y_n из (7.147) в (7.146), находим

$$a_k \lambda^{n+k} + a_{k-1}\lambda^{n+k-1} + \ldots + a_1 \lambda^{n+1} + a_0 \lambda^n = 0.$$

Разделив обе части этого уравнения на λ^n, получим

$$a_k\lambda^k + a_{k-1}\lambda^{k-1} + \ldots + a_1\lambda + a_0 = 0. \tag{7.148}$$

Уравнение (7.148) называется *характеристическим уравнением* однородного линейного разностного уравнения.

Так же как и в случае линейных дифференциальных уравнений с постоянными коэффициентами, знание корней характеристического уравнения позволяет построить общее решение однородного разностного уравнения. Для простоты изложения продемонстрируем это на примере уравнения второго порядка

$$ay_{n+2} + by_{n+1} + cy_n = 0, \quad a \neq 0, c \neq 0. \tag{7.149}$$

Однако заметим, что полученные в этом случае результаты могут быть без труда перенесены на случай уравнений более высокого порядка.

В зависимости от значения дискриминанта $D = b^2 - 4ac$ характеристического уравнения

$$a\lambda^2 + b\lambda + c = 0 \tag{7.150}$$

возможны следующие случаи.

Случай 1. Дискриминант $D > 0$. Характеристическое уравнение имеет два действительных и различных корня λ_1 и λ_2. Тогда $\lambda_1 \neq 0$ и $\lambda_2 \neq 0$. Действительно, если хотя бы один корень равен нулю, то коэффициент $c = \lambda_1\lambda_2$ также будет равен нулю, что противоречит определению линейного разностного уравнения второго порядка. Корням λ_1 и λ_2 характеристического уравнения соответствуют два решения:

$$y_n^{(1)} = \lambda_1^n, \quad y_n^{(2)} = \lambda_2^n$$

уравнения (7.149). Покажем, что эти решения линейно независимы. Для этого вычислим определитель Казоратти

$$\Delta = \begin{vmatrix} \lambda_1^n & \lambda_2^n \\ \lambda_1^{n+1} & \lambda_2^{n+1} \end{vmatrix} = \lambda_1^n \lambda_2^{n+1} - \lambda_1^{n+1} \lambda_2^n = \lambda_1^n \lambda_2^n (\lambda_2 - \lambda_1).$$

Так как корни λ_1 и λ_2 различны, то $\lambda_2 - \lambda_1 \neq 0,$ следовательно, $\Delta \neq 0,$ а значит, решения линейно независимы. В этом случае общее решение уравнения имеет вид

$$y(n) = C_1 \lambda_1^n + C_2 \lambda_2^n, \qquad (7.151)$$

где C_1, C_2 – произвольные постоянные.

Пример 7.45. Решить уравнение

$$y_{n+2} + 4y_{n+1} - 5y_n = 0.$$

Решение. Составим характеристическое уравнение

$$\lambda^2 + 4\lambda - 5 = 0.$$

Оно имеет два действительных различных корня: $\lambda_1 = 1,$ $\lambda_2 = -5.$ Поэтому общее решение исходного уравнения имеет вид

$$y_n = C_1 + C_2(-5)^n.$$

Случай 2. Дискриминант $D < 0.$ Характеристическое уравнение имеет два комплексно сопряженных корня λ_1 и $\lambda_2,$ которые, используя тригонометрическую форму записи, могут быть представлены следующим образом:

$$\lambda_1 = r(\cos\varphi + i\sin\varphi), \qquad \lambda_2 = r(\cos\varphi - i\sin\varphi),$$

где r – модуль $\lambda_1,$ а φ – его аргумент. Соответствующие решения разностного уравнения также комплексно сопряжены и на основании формулы Муавра имеют вид

$$y_n^{(1)} = r^n(\cos n\varphi + i\sin n\varphi), \qquad y_n^{(2)} = r^n(\cos n\varphi - i\sin n\varphi).$$

Чтобы получить действительные решения, заменим $y_n^{(1)}$ и $y_n^{(2)}$ их линейными комбинациями

$$z_n^{(1)} = \frac{1}{2}\left(y_n^{(1)} + y_n^{(2)}\right), \quad z_n^{(2)} = \frac{1}{2i}\left(y_n^{(1)} - y_n^{(2)}\right).$$

Итак, мы получили два линейно независимых действительных решения

$$z_n^{(1)} = r^n \cos n\varphi, \qquad z_n^{(2)} = r^n \sin n\varphi.$$

Таким образом, в данном случае общее решение имеет вид

$$y_n = r^n (C_1 \cos n\varphi + C_2 \sin n\varphi).$$

Пример 7.46. Решить уравнение

$$y_{n+2} - 2y_{n+1} + 4y_n = 0.$$

Решение. Характеристическое уравнение

$$\lambda^2 - 2\lambda + 4 = 0$$

имеет два комплексно сопряженных корня $\lambda_1 = 1 + \sqrt{3}i$ и $\lambda_2 = 1 - \sqrt{3}i,$ которые могут быть записаны в виде:

$$\lambda_1 = 2\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right), \qquad \lambda_2 = 2\left(\cos\frac{\pi}{3} - i\sin\frac{\pi}{3}\right).$$

Следовательно, общим решением исходного уравнения будет

$$y_n = 2^n\left(C_1 \cos\frac{n\pi}{3} + C_2 \sin\frac{n\pi}{3}\right).$$

Случай 3. Дискриминант $D=0$. Оба корня характеристического уравнения действительны и равны $(\lambda_1=\lambda_2=\lambda)$. Покажем, что в этом случае, кроме решения

$$y_n^{(1)}=\lambda^n,$$

существует еще одно решение, линейно независимое с $y_n^{(1)}$. Действительно, нетрудно убедиться, что таковым является

$$y_n^{(2)}=n\lambda^n.$$

Сначала докажем, что $y_n^{(2)}$ будет решением уравнения (7.150). В самом деле, подставляя выражение для $y_n^{(2)}$ в уравнение (7.150), получим

$$a(n+2)\lambda^{n+2}+b(n+1)\lambda^{n+1}+cn\lambda^n=\lambda^n\left(a(n+2)\lambda^2+b\lambda(n+1)+cn\right)=$$
$$=\lambda^n(n(a\lambda^2+b\lambda+c)+2a\lambda^2+b\lambda)=0$$

в силу того, что $a\lambda^2+b\lambda+c=0$ и $\lambda=-\dfrac{b}{2a}$. Вычислим теперь определитель Казоратти. Имеем

$$\Delta=\begin{vmatrix}\lambda^n & n\lambda^n\\ \lambda^{n+1} & (n+1)\lambda^{n+1}\end{vmatrix}=\lambda^{2n+1}\neq 0,$$

так как $\lambda\neq 0$. Следовательно, частные решения $y_n^{(1)}$ и $y_n^{(2)}$ линейно независимы, и общее решение уравнения (7.150) имеет вид

$$y_n=\lambda^n(C_1+nC_2).$$

Пример 7.47. Решить уравнение

$$y_{n+2}+6y_{n+1}+9y_n=0.$$

Решение. Характеристическое уравнение

$$\lambda^2+6\lambda+9=0$$

имеет единственный действительный корень $\lambda = -3$. Следовательно, общее решение исходного уравнения таково:

$$y_n = (-3)^n (C_1 + nC_2).$$

Для нахождения решения неоднородного линейного разностного уравнения, так же как и в случае линейных дифференциальных уравнений, используется *метод неопределенных коэффициентов*, основанный на подборе частного решения неоднородного уравнения по виду правой части $f(n)$.

Пример 7.48. Решить уравнение

$$y_{n+2} + 2y_{n+1} - 3y_n = 64 \cdot 5^n.$$

Решение. Будем искать частное решение в виде

$$y(n) = p5^n.$$

Подставляя это выражение в наше уравнение, получим

$$p(25 + 10 - 3)5^n = 64 \cdot 5^n.$$

Следовательно, $p = 2$, а значит,

$$y(n) = 2 \cdot 5^n.$$

Решая характеристическое уравнение

$$\lambda^2 + 2\lambda - 3 = 0,$$

находим $\lambda_1 = 1$, $\lambda_2 = -3$. Таким образом, общее решение уравнения имеет вид

$$y_n = 2 \cdot 5^n + C_1 + C_2(-3)^n.$$

§ 7.17. Модели экономической динамики с дискретным временем

1. Модель Самуэльсона – Хикса

В данном параграфе мы рассмотрим несколько примеров, иллюстрирующих применение разностных уравнений. Начнем с *модели делового цикла Самуэльсона-Хикса* (*динамического варианта модели Кейнса*). Эта модель основывается на уже упоминавшемся *принципе акселерации* (см. § 7.4), то есть на предположении, что объемы инвестирования прямо пропорциональны приросту национального дохода. Данное предположение характеризуется уравнением

$$I(t) = V(Y(t-1) - Y(t-2)), \tag{7.152}$$

где коэффициент $V > 0$ – фактор акселерации, $I(t)$ – величина инвестиций в период t, $Y(t-1), Y(t-2)$ – величины национального дохода соответственно в (t – 1)-м и (t – 2)-м периодах. Предполагается также, что спрос на данном этапе зависит от величины национального дохода на предыдущем этапе, т.е.

$$C(t) = aY(t-1) + b. \tag{7.153}$$

Условие равенства спроса и предложения имеет вид

$$Y(t) = I(t) + C(t). \tag{7.154}$$

Подставляя в (7.154) выражение для $I(t)$ из (7.152) и выражение для $C(t)$ из (7.153), находим

$$Y(t) = (a+V)Y(t-1) - VY(t-2) + b. \tag{7.155}$$

Уравнение (7.155) известно как *уравнение Хикса.* Оно представляет собой неоднородное линейное разностное уравнение второго порядка с постоянными коэффициентами (если предположить, что на протяжении рассматриваемых периодов величины a и V постоянны).

Замечание 1. Мы можем легко найти частное решение уравнения (7.155), если положим

$$Y(t) = Y(t-1) = Y(t-2) = Y^*, \tag{7.156}$$

т.е. взяв в качестве частного решения, равновесное решение $Y^* = \text{const}$. Из (7.155) в силу (7.156) имеем

$$Y^* = (a+V)Y^* - VY^* + b,$$

откуда

$$Y^* = b(1-a)^{-1}. \tag{7.157}$$

Заметим также, что выражение $(1-a)^{-1}$ в формуле (7.157) носит название *мультипликатора Кейнса* и является одномерным аналогом *матрицы полных затрат*.

Пример 7.49. Рассмотрим модель Самуэльсона – Хикса при условии, что $a = 0{,}5$; $V = 0{,}5$; $b = 4$. В этом случае уравнение (7.155) принимает вид:

$$Y(t) - Y(t-1) + 0{,}5Y(t-2) = 4.$$

Его частным решением будет

$$y(t) = \frac{4}{1-0{,}5} = 8.$$

Найдем корни характеристического уравнения

$$\lambda^2 - \lambda + 0{,}5 = 0.$$

Имеем

$$\lambda_{1,2} = \frac{1 \pm i}{2} = \frac{\sqrt{2}}{2}\left(\cos\frac{\pi}{4} \pm i\sin\frac{\pi}{4}\right).$$

Таким образом, общим решением соответствующего однородного уравнения является

$$\tilde{Y}(t)=\left(\frac{\sqrt{2}}{2}\right)^t\left(C_1\cos\frac{t\pi}{4}+C_2\sin\frac{t\pi}{4}\right).$$

Следовательно, общим решением уравнения будет

$$Y(t)=8+\left(\frac{\sqrt{2}}{2}\right)^t\left(C_1\cos\frac{t\pi}{4}+C_2\sin\frac{t\pi}{4}\right).$$

Замечание 2. В зависимости от значений a и V возможны четыре типа динамики. Она может быть растущей или затухающей и при этом иметь или не иметь колебательный характер. Так, в рассмотренном выше примере динамика носила колебательный характер с убывающей амплитудой. Мы рекомендуем читателю самостоятельно определить виды динамики в зависимости от a и V.

2. Паутинная модель рынка

При помощи разностных уравнений можно дать трактовку процессов сходимости и расходимости в паутинных моделях рынка. Для упрощения выкладок предположим, что спрос и предложение задаются линейными функциями, но при этом спрос зависит от цены в данный момент времени, а предложение зависит от цены на предыдущем этапе, т.е.

$$d_t=a-bp_t,\quad s_t=m+np_{t-1}, \tag{7.158}$$

где a, b, m, n – положительные действительные числа.

Таким образом, если $s_t=d_t$, то

$$a-m=bp_t+np_{t-1}. \tag{7.159}$$

Уравнение (7.159) представляет собой линейное разностное уравнение первого порядка с постоянными коэффициентами.

В качестве частного решения можно использовать равновесное решение

$$p_t = \overline{p} = \text{const}. \qquad (7.160)$$

Действительно, подставив выражение для p_t из формулы (7.160) в (7.159), легко получить

$$\overline{p} = \frac{a-m}{b+n}. \qquad (7.161)$$

Решая характеристическое уравнение

$$b\lambda + n = 0,$$

находим $\lambda = -\frac{n}{b}$. Следовательно,

$$p_t = C_1\left(-\frac{n}{b}\right)^t + \frac{a-m}{b+n}. \qquad (7.162)$$

Таким образом, из (7.162) вытекает, что динамика цен носит колебательный характер. При этом если $n < b,$ то последовательность цен $\{p_t\}$ будет сходиться к равновесному состоянию (рис. 1.16), если $n > b,$ то с течением времени последовательность $\{p_t\}$ будет удаляться от равновесного состояния (рис. 1.17), если же $n = b,$ то будут иметь место циклические колебания цены относительно равновесного состояния.

3. Задача об определении текущей стоимости купонной облигации

Пусть F – номинальная стоимость купонной облигации (денежная сумма, выплачиваемая эмитентом в момент погашения, совпадающий с концом последнего купонного периода), K – величина купона (денежная сумма, выплачиваемая в конце каждого купонного периода), $P(n)$ – текущая стоимость облигации в конце n-го купонного периода, k – число купонных периодов. Пусть также r – процентная ставка за один купонный период, выраженная в частях. Предполагается, что она неизменна в течение всего срока обращения облигации. Перечисленные величины связаны между собой следующими соотношениями:

$$P(k) = F, \tag{7.163}$$

$$P(n+1) + K = (1+r)P(n). \tag{7.164}$$

Таким образом, задача об определении текущей стоимости купонной облигации сводится к решению задачи Коши (7.163), (7.164) для неоднородного линейного разностного уравнения первого порядка с постоянными коэффициентами. В качестве частного решения выберем равновесное решение

$$P(n+1) = P(n) = P^*. \tag{7.165}$$

Подставив выражение для $P(n)$ из формулы (7.165) в (7.164), получаем

$$P^* = \frac{K}{r}. \tag{7.166}$$

Заметим, что величина K/r есть не что иное, как текущая стоимость бесконечной ренты, т.е. сумма, которую необходимо уплатить в настоящий момент, чтобы в течение бесконечно длительного времени получать сумму K через каждый промежуток времени t при процентной ставке r. Действительно,

$$\frac{K}{r} = \frac{K}{1+r} + \frac{K}{(1+r)^2} + \frac{K}{(1+r)^3} + \cdots$$

В справедливости этого равенства легко убедиться, посчитав сумму бесконечно убывающей геометрической прогрессии, находящейся в правой части формулы.

Решив характеристическое уравнение

$$\lambda - (1-r) = 0,$$

находим

$$P(n) = C(1+r)^n + (K/r). \tag{7.167}$$

Полагая в соотношении (7.167) $n = k$ и учитывая (7.162), имеем

$$C = (F - K/r)\,(1+r)^{-k}. \tag{7.168}$$

Из (7.167) в силу (7.168) следует, что последовательность $P(n)$ будет возрастающей, если номинальная стоимость облигации выше, чем стоимость бесконечной ренты, убывающей, если она меньше, и постоянной, если они равны.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

1. *Архипов Г.И.* Лекции по математическому анализу. / *Г.И. Архипов*, *В.А. Садовничий*, *В.Н. Чубариков.* – М.: Высш. шк, 1999.

2. *Ашманов С.А.* Введение в математическую экономику. / *С.А. Ашманов.* – М.: Наука, 1984.

3. *Бугров Я.С.* Высшая математика. Дифференциальное и интегральное исчисление / *Я.С. Бугров*, *С.М. Никольский.* – М.: Наука, 1988.

4. *Ильин В.А.*. Математический анализ. Т.1, 2 / *В.А. Ильин*, *В.А. Садовничий*, *Б.Х. Сендов.* – М.: Изд-во Моск. ун-та, 1985.

5. *Зорич В.А.* Математический анализ. Ч.1. / *В.А. Зорич.* – М.: Фазис, 1997. Ч.2. – М.: Наука, 1990.

6. *Карманов В.Г.* Математическое программирование. / *В.Г. Карманов.* – М.: Наука, 1986.

7. *Кудрявцев Л.Д.* Курс математического анализа. Т. 1–3. / *Л.Д. Кудрявцев* – М.: Высш. шк., 1981.

8. *Маленво Э.* Лекции по микроэкономическому анализу. / *Э. Маленво* – М.: Наука, 1985.

9. *Кубонива М.* Математическая экономика на персональном компьютере / *М. Кубонива* – М.: Финансы и статистика, 1991.

10. *Макконелл К.Р.* Экономикс / *К.Р. Макконелл*, *С.Л. Брю.* – М.: Республика, 1992.

11. *Пиндайк Р.С.* Микроэкономика / *Р.С. Пиндайк*, *Д.Л. Рубинфельд.* – М.: Дело, 2001.

12. *Фихтенгольц Г.М.* Курс математического анализа. Т. 1–3 / *Г.М. Фихтенгольц.* – М.: Физматгиз, 1962.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Н

О

П

Р

С

Т

У

Ф

Учебное издание

Солодовников Александр Самуилович
Бабайцев Владимир Алексеевич
Браилов Андрей Владимирович
Шандра Игорь Георгиевич

МАТЕМАТИКА В ЭКОНОМИКЕ

Часть 2

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

Заведующая редакцией *Л.А. Табакова*
Ведущий редактор *Л.Д. Григорьева*
Младший редактор *О.О. Салтыкова*
Художественный редактор *Г.Г. Семенова*
Технический редактор *Т.С. Маринина*
Компьютерная верстка *А.В. Браилов*

ИБ № 5386

Подписано в печать 26.08.2012. Формат $60x90^1/_{16}$
Гарнитура «Таймс». Печать офсетная
Усл. п.л. 35,0. Уч.-изд. л. 34,2
Тираж 1000 экз. Заказ

Издательство «Финансы и статистика»
101000, Москва, ул. Покровка, 7
Телефон (495) 625-35-02. Факс (495) 625-09-57
E-mail: mail@finstat.ru http://www.finstat.ru