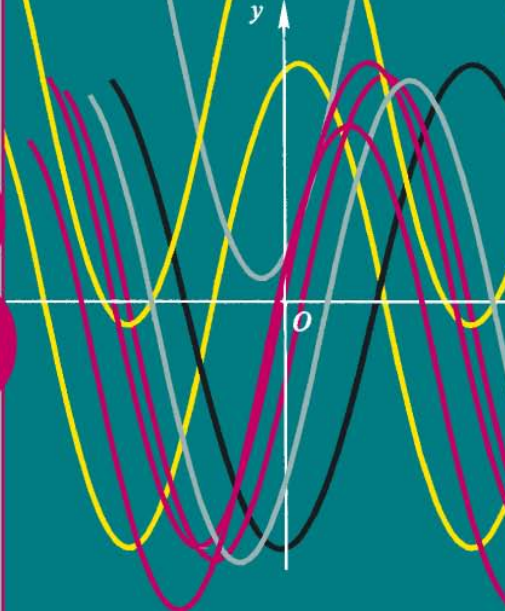


УВ
УЧЕБНИК
ДЛЯ ВУЗОВ



И. И. БАВРИН

КУРС ВЫСШЕЙ МАТЕМАТИКИ

ГУМАНИТАРНЫЙ
ИЗДАТЕЛЬСКИЙ
ЦЕНТР

ВЛАДОС

И.И. БАВРИН

КУРС ВЫСШЕЙ МАТЕМАТИКИ

2-е издание, переработанное и дополненное

Рекомендовано Министерством образования
Российской Федерации в качестве учебника
для студентов высших педагогических учебных
заведений, обучающихся по направлению
«Естествознание», специальности «Физика»

Москва

ГУМАНИТАРНЫЙ
ИЗДАТЕЛЬСКИЙ
ЦЕНТР
ВЛАДОС
2004

УДК 51(075.8)
ББК 21.1я73
Б13

Р е ц е н з е н т ы:

доктор физико-математических наук,
профессор кафедры математического анализа МГУ
им. М. В. Ломоносова **В.И. Гаврилов**;
доктор педагогических наук, профессор,
заведующий кафедрой математического анализа МПУ
Г.Л. Луканкин

Баврин И. И.

Б13 Курс высшей математики: Учеб. для студ. высш. пед.
учеб. заведений / 2-е изд., перераб. и доп. — М.: Гуманит.
изд. центр ВЛАДОС, 2004. — 560 с.
ISBN 5-691-00117-5.
Агентство СІР РГБ.

Учебник соответствует примерной программе дисциплины «Математика» для направления 540100 «Естественнознание», специальности «Физика» педагогических вузов.

Состоит из трех разделов. Первый раздел — аналитическая геометрия и линейная алгебра, второй — математический анализ, третий — специальные главы высшей математики, в том числе теория поля, элементы теории функций комплексной переменной, интеграл Фурье, основные уравнения и задачи математической физики, теория вероятностей, элементы математической статистики, элементы вариационного и операционного исчисления. В приложении приведены таблицы из теории вероятностей и математической статистики, дополнительная таблица интегралов и основные соотношения и формулы из школьной математики. Приведено много разнообразных примеров и задач, иллюстрирующих понятия высшей математики и ее методы.

УДК 51(075.8)
ББК 22.1я73

- © Баврин И. И., 2004
- © ООО «Гуманитарный издательский центр ВЛАДОС», 2004
- © Серия «Учебник для вузов»
и серийное оформление.
ООО «Гуманитарный издательский центр ВЛАДОС», 2004
- © Макет. ООО «Гуманитарный издательский центр ВЛАДОС», 2004

ISBN 5-691-00117-5

Предисловие ко второму изданию

Первое издание учебника «Курс высшей математики» (М., 1992) переработано и дополнено в соответствии с новой примерной программой дисциплины «Математика» для направления 540100 «Естествознание» и программой по математике для специальности 01.40.00 «Физика» педагогических высших учебных заведений.

Укажем наиболее существенные изменения и дополнения книги.

Выделены глава «Линейная алгебра» и раздел «Специальные главы». В этот раздел в качестве нового материала включены элементы теории функций комплексной переменной, математической статистики, операционного исчисления и некоторые численные методы.

И. И. Баврин

Раздел I

АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ И ЛИНЕЙНАЯ АЛГЕБРА

Глава 1. СИСТЕМА КООРДИНАТ НА ПЛОСКОСТИ И ОСНОВНЫЕ ПОНЯТИЯ

§ 1.1. Декартова прямоугольная и полярная системы координат на плоскости

1. Декартовы прямоугольные координаты. Возьмем на плоскости две взаимно перпендикулярные прямые Ox и Oy с указанными на них положительными направлениями (рис. 1). Прямые Ox и Oy называются *координатными осями*, точка их пересечения O — *началом координат*. Обычно полагают, что ось Ox горизонтальна, а ось Oy вертикальна относительно наблюдателя; положительное направление на Ox слева направо, на Oy — снизу вверх.

Выберем единицу масштаба (будем предполагать, что на обеих осях координат выбрана одна и та же единица масштаба). Координатные оси Ox , Oy с выбранной единицей масштаба называются *декартовой прямоугольной* (или кратко *прямоугольной*) *системой координат* на плоскости. (Декартова прямоугольная система координат носит имя французского математика, основателя аналитической геометрии Рене Декарта (1596–1650).)

Произвольной точке M плоскости поставим в соответствие два числа (рис. 1):

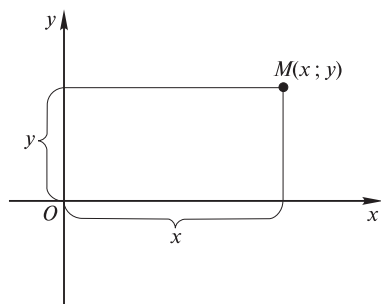


Рис. 1

абсциссу x , равную расстоянию от точки M до оси Oy , взятому со знаком «+», если M лежит правее Oy , и со знаком «-», если M лежит левее Oy ;

ординату y , равную расстоянию от точки M до оси Ox , взятому со знаком «+», если M лежит выше Ox , и со знаком «-», если M лежит ниже Ox .

Абсцисса x и ордината y называются *декартовыми прямоугольными*

(или прямоугольными) координатами точки M . Запись $M(x; y)$ читают: «Точка M с абсциссой, равной x , и ординатой, равной y ».

Отметим, что каждой точке плоскости соответствует одна пара действительных чисел x и y (ее координат). Верно и обратное: каждой паре действительных чисел x и y соответствует одна точка плоскости. Это значит, что положение на плоскости произвольной точки M полностью определяется ее координатами x и y .

Координатные оси Ox и Oy разбивают плоскость на I, II, III, IV квадранты (рис. 2). Знаки координат точек в различных квадрантах указаны в таблице:

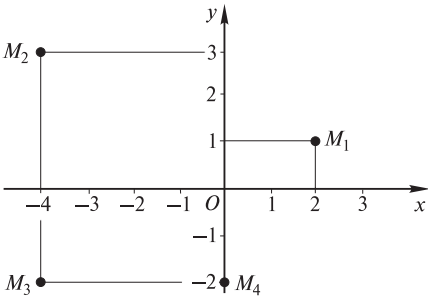


Рис. 2

	I	II	III	IV
x	+	−	−	+
y	+	+	−	−

При этом если точка $M(x; y)$ лежит на оси Oy , то $x=0$; если $M(x; y)$ лежит на оси Ox , то $y=0$.

На рисунке 2 построены четыре точки $M_1(2; 1)$, $M_2(-4; 3)$, $M_3(-4; -2)$ и $M_4(0; -2)$.

2. Полярные координаты. Зафиксируем на плоскости точку O и выходящую из нее полупрямую Op , а также выберем единицу масштаба (рис. 3). Точка O называется *полусом*, полупрямая Op — *полярной осью*.

Произвольной точке M (отличной от O) плоскости поставим в соответствие два числа:

полярный радиус r , равный расстоянию от точки M до полюса O ;

полярный угол φ , равный углу между полярной осью Op и полупрямой OM .

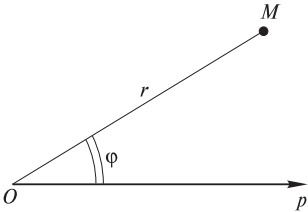


Рис. 3

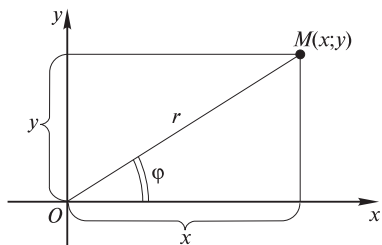


Рис. 4

Запись $M(r; \varphi)$ означает: точка M с полярными координатами r и φ .

Найдем зависимость между прямоугольными и полярными координатами. Будем считать начало координат O прямоугольной системы xOy одновременно полюсом O , а луч Ox примем за полярную ось Op (рис. 4).

Из рисунка 4 видно, что для точки $M(x; y)$ ($M(r; \varphi)$) справедливы соотношения

$$x = r \cos \varphi, \quad y = r \sin \varphi \quad (1)$$

и

$$r = \sqrt{x^2 + y^2}, \quad \operatorname{tg} \varphi = \frac{y}{x}. \quad (2)$$

Формулы (1) выражают прямоугольные координаты точки M через ее полярные координаты. Это можно доказать для любого расположения точки M на координатной плоскости. Формулы (2) выражают полярные координаты точки M через ее прямоугольные координаты и тоже верны при любом положении точки M .

Заметим, что $\operatorname{tg} \varphi = \frac{y}{x}$ дает два значения φ ($-\pi < \varphi \leq \pi$).

Поэтому для вычисления полярного угла φ точки M по ее прямоугольным координатам x и y предварительно выясняют, в каком квадранте лежит точка M .

П р и м е р 1. Даны прямоугольные координаты точки A : $x = 1, y = 1$. Найти ее полярные координаты. По формулам (2) находим: $r = \sqrt{1^2 + 1^2} = \sqrt{2}$, $\operatorname{tg} \varphi = 1$. Из двух значений $\varphi = \frac{\pi}{4}$ и $\varphi = -\frac{3\pi}{4}$ выбираем $\varphi = \frac{\pi}{4}$, так как точка A лежит в первом квадранте. Итак, полярные координаты данной точки $r = \sqrt{2}, \varphi = \frac{\pi}{4}$.

П р и м е р 2. Полярные координаты точки A таковы: $r = 2, \varphi = \frac{\pi}{2}$. Тогда по формулам

(1) прямоугольные координаты этой точки будут $x = 2 \cos \frac{\pi}{2} = 0, y = 2 \sin \frac{\pi}{2} = 2$.

Полярный угол φ измеряется в радианах, отсчет положительных (отрицательных) значений φ ведется от Op против движения (по движению) часовой стрелки. При этом обычно полагают, что $-\pi < \varphi \leq \pi$.

Полюсу O соответствует полярный радиус $r = 0$, полярный угол для него не определен.

§ 1.2. Простейшие задачи на плоскости

1. Расстояние между двумя точками. Найдем расстояние d между двумя данными точками $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$ (рис. 5). Из прямоугольного треугольника M_1NM_2 по теореме Пифагора имеем:

$$d = M_1M_2 = \sqrt{M_1N^2 + M_2N^2}.$$

Известно, что расстояние между точками A и B , расположенными на координатной прямой (оси), вычисляется по формуле $d = AB = |x_B - x_A|$, где x_A и x_B — координаты точек A и B этой прямой. Но $M_1N = A_1A_2 = |x_2 - x_1|$, $M_2N = B_1B_2 = |y_2 - y_1|$. Поэтому

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad (1)$$

П р и м е р. Найти расстояние между точками $A(-1; -2)$ и $B(-4; 2)$. По формуле (1) имеем:

$$AB = \sqrt{(-4+1)^2 + (2+2)^2} = \sqrt{9+16} = 5.$$

2. Деление отрезка в данном отношении. Пусть даны точки $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$. Требуется найти точку $M(x; y)$, лежащую на отрезке M_1M_2 и делящую его в данном отношении:

$$\frac{M_1M}{MM_2} = \lambda. \quad (2)$$

Опустив из точек M_1 , M и M_2 перпендикуляры на ось Ox (рис. 6), получим:

$$\frac{M_1M}{MM_2} = \frac{A_1A}{AA_2}.$$

При выбранном расположении точек имеем:

$$A_1A = x - x_1, \quad AA_2 = x_2 - x.$$

Поэтому заданное отношение (2) принимает вид:

$$\frac{x - x_1}{x_2 - x} = \lambda,$$

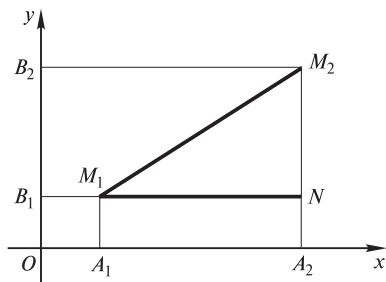


Рис. 5

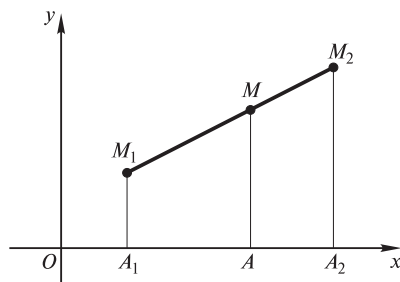


Рис. 6

откуда

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}. \quad (3)$$

Аналогично

$$y = \frac{y_1 + \lambda y_2}{1 + \lambda}. \quad (4)$$

В частности, если $\lambda = 1$, т. е. при делении отрезка $M_1 M_2$ пополам, получаем:

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}.$$

П р и м е ч а н и е. Формулы (3) и (4) верны при любом расположении точек M_1 и M_2 .

П р и м е р. Вычислить координаты точки $M(x; y)$, делящей отрезок $M_1 M_2$, где $M_1(1; 1)$ и $M_2(4; 7)$, в отношении $\frac{M_1 M}{M M_2} = 2$. Согласно формулам (3) и (4) имеем:

$$x = \frac{1 + 2 \cdot 4}{3} = 3, \quad y = \frac{1 + 2 \cdot 7}{3} = 5.$$

§ 1.3. Геометрическое истолкование уравнения с двумя переменными

Прямоугольная и полярная системы координат позволяют задавать различные линии на плоскости их уравнениями.

О п р е д е л е н и е. Уравнением линии на плоскости в прямоугольной системе координат xOy называется уравнение $f(x, y) = 0$, которому удовлетворяют координаты каждой точки данной линии и не удовлетворяют координаты любой точки плоскости, не лежащей на этой линии.

Переменные x и y уравнения линии называются *текущими* координатами.

Покажем, например, что уравнение $x - y = 0$, или

$$x = y, \quad (1)$$

является уравнением биссектрисы I и III координатных углов.

По свойству биссектрисы угла для произвольной точки $M(x; y)$ (лежащей на биссектрисе) имеем $N_2 M = N_1 M$ или $ON_1 = ON_2$ (рис. 7), и поэтому $x = y$, т. е. координаты всех точек биссектрисы удовлетворяют уравнению (1). Очевидно также, что у любой точки, не лежащей на данной биссектрисе, координаты не равны между собой и, значит, не удовлетворяют уравнению (1).

Обратно, если координаты x и y какой-нибудь точки $M(x; y)$ удовлетворяют уравнению (1), то эта точка, очевидно, лежит на биссектрисе I и III координатных углов.

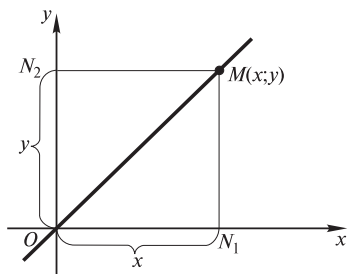


Рис. 7

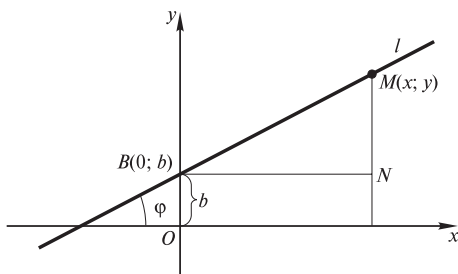


Рис. 8

Однако геометрическим образом данного заранее уравнения не всегда будет линия. Может случиться, что уравнению соответствует лишь несколько точек (уравнению $x^2 + y^2 = 0$, например, на плоскости соответствует только одна точка $(0; 0)$). Встречаются и такие случаи, когда заданному уравнению не соответствует на плоскости ни одна точка (как, например, уравнению $x^2 + y^2 + 1 = 0$).

В связи с изложенным заметим, что всякой линии на плоскости соответствует некоторое уравнение между текущими координатами (x, y) точки этой линии. Наоборот, всякому уравнению между x и y , где x и y — координаты точки на плоскости, соответствует, вообще говоря, некоторая линия, свойства которой вполне определяются данным уравнением.

Отсюда, естественно, возникают две основные задачи аналитической геометрии на плоскости.

1) Дана линия, рассматриваемая как множество точек. Составить уравнение этой линии.

2) Дано уравнение некоторой линии. Изучить по этому уравнению ее геометрические свойства (форму и расположение).

§ 1.4. Прямая линия

1. Уравнение прямой с угловым коэффициентом. Пусть прямая l не параллельна оси Oy (рис. 8). Обозначим точку пересечения l с осью Oy буквой $B(0; b)$, а угол между положительным направлением оси Ox и l обозначим φ . Угол φ , отсчитываемый от оси Ox против часовой стрелки ($0 \leq \varphi < \pi$), называется *углом наклона* прямой l к оси Ox .

Выведем уравнение прямой l .

Пусть $M(x; y)$ — произвольная точка прямой l с текущими координатами x и y . Из прямоугольного треугольника BMN (рис. 8) имеем:

$$\operatorname{tg} \varphi = \frac{y - b}{x}. \quad (1)$$

Эту величину называют *угловым коэффициентом* прямой l и обозначают k : $k = \operatorname{tg} \varphi$. Тогда из равенства (1) получим:

$$k = \frac{y-b}{x},$$

откуда

$$y = kx + b. \quad (2)$$

Уравнение (2) называется *уравнением прямой с угловым коэффициентом*, число b называется *начальной ординатой* (это ордината точки B).

Пример. Если $\varphi = \frac{\pi}{4}$, $b = -3$, то $k = \operatorname{tg} \frac{\pi}{4} = 1$ и уравнение данной прямой имеет вид $y = x - 3$.

Если в уравнении (2) $k = 0$, то имеем уравнение прямой, параллельной оси Ox и проходящей через точку $B(0; b)$:

$$y = b. \quad (3)$$

При $b = 0$ из (3) получаем уравнение координатной оси Ox : $y = 0$.

По аналогии с уравнением (3) уравнение

$$x = a \quad (4)$$

есть уравнение прямой, параллельной оси Oy и проходящей через точку $A(a; 0)$. При $a = 0$ из равенства (4) имеем уравнение координатной оси Oy : $x = 0$.

2. Общее уравнение прямой. Уравнением прямой с угловым коэффициентом может быть задана любая прямая на плоскости, не параллельная оси ординат.

Любую прямую без каких-либо ограничений можно задать уравнением

$$Ax + By + C = 0, \quad (5)$$

где A и B — коэффициенты, одновременно не равные нулю.

Т е о р е м а. Каждая прямая на плоскости с прямоугольной системой координат определяется уравнением первой степени, и, наоборот, каждое уравнение первой степени определяет некоторую прямую на плоскости.

Д о к а з а т е л ь с т в о. 1) Пусть дана прямая, не параллельная оси ординат. В этом случае прямая описывается уравнением прямой с угловым коэффициентом $y = kx + b$, которое является частным случаем уравнения (5) при $A = k$, $B = -1$, $C = b$.

Пусть теперь прямая параллельна оси Oy . Тогда ее уравнение запишется в виде $x = a$. Это уравнение тоже частный случай уравнения (5) при $A = 1$, $B = 0$, $C = -a$. Итак, любая прямая на плоскости определяется уравнением первой степени.

2) Покажем теперь, что произвольному уравнению первой степени (5) (A и B одновременно не равны нулю) соответствует некоторая прямая на плоскости.

Действительно, если $B \neq 0$, то уравнение (5) можно преобразовать в уравнение

$$y = -\frac{A}{B}x - \frac{C}{B},$$

т. е. в уравнение прямой с угловым коэффициентом $k = -\frac{A}{B}$ и начальной ординатой $b = -\frac{C}{B}$. Если $B = 0$, $A \neq 0$, то уравнение (5) преобразуется к виду $x = -\frac{C}{A}$, т. е. в уравнение прямой, параллельной оси Oy . Теорема доказана.

Уравнение первой степени (5) (A и B одновременно не равны нулю), описывающее на плоскости любую прямую, называется *общим уравнением прямой*.

3. Уравнение прямой в отрезках. Предположим, что в общем уравнении прямой $A \neq 0$, $B \neq 0$ и $C \neq 0$. Перенесем C в правую часть и разделим обе части полученного уравнения на $-C$, получим $\frac{A}{-C}x + \frac{B}{-C}y = 1$, или $\frac{x}{-\frac{C}{A}} + \frac{y}{-\frac{C}{B}} = 1$. Обозначив $-\frac{C}{A} = a$, $-\frac{C}{B} = b$, придем к уравнению

$$\frac{x}{a} + \frac{y}{b} = 1. \quad (6)$$

Уравнение (6) называется *уравнением прямой в отрезках*. Это название объясняется тем, что числа a и b определяются отрезками OA и OB , которые прямая отсекает на осях координат (рис. 9). Такой вид уравнения удобен для построения прямой.

Заметим, что прямые, параллельные координатным осям, и прямые, проходящие через начало координат, не могут быть записаны уравнениями в отрезках.

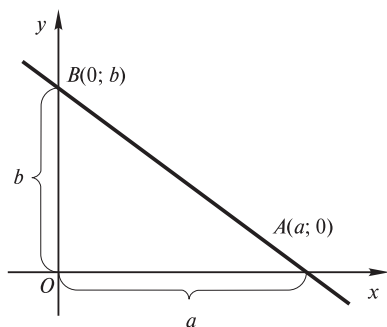


Рис. 9

4. Уравнение прямой, проходящей через данную точку в данном направлении. Выведем уравнение прямой, проходящей через данную точку $M_1(x_1; y_1)$ и имеющей данный угловой коэффициент k . Уравнение этой прямой имеет вид:

$$y = kx + b. \quad (7)$$

Так как искомая прямая проходит через точку M_1 , то

$$y_1 = kx_1 + b. \quad (8)$$

Вычитая из равенства (7) равенство (8), получаем:

$$y - y_1 = k(x - x_1). \quad (9)$$

Это и есть уравнение искомой прямой.

П р и м е р. Уравнение прямой, проходящей через точку $M(-1; 8)$, с угловым коэффициентом $k = 1$ согласно (9) есть $y - 8 = x + 1$, или $x - y + 9 = 0$.

П р и м е ч а н и е 1. В уравнении (9) постоянная k может быть любым действительным числом. Поэтому в форме (9) может быть записано уравнение всякой прямой, проходящей через точку M_1 , не параллельной оси Oy . Уравнение прямой, проходящей через данную точку M_1 параллельно оси Oy , будет иметь вид $x = x_1$ (п. 1).

П р и м е ч а н и е 2. Совокупность всех прямых, проходящих через некоторую точку плоскости, называется *пучком прямых*, а общая их точка — *центром пучка*.

5. Угол между прямыми. Рассмотрим на плоскости две прямые $y = k_1x + b_1$ (l_1) и $y = k_2x + b_2$ (l_2) с углами наклона к оси Ox соответственно φ_1 и φ_2 (рис. 10).

За угол φ ($0 \leq \varphi < \pi$) между прямыми l_1 и l_2 принимают один из двух смежных углов, которые образуют эти прямые. Из рисунка 10 видно, что $\varphi = \varphi_2 - \varphi_1$. Поэтому

$$\operatorname{tg} \varphi = \operatorname{tg}(\varphi_2 - \varphi_1) = \frac{\operatorname{tg} \varphi_2 - \operatorname{tg} \varphi_1}{1 + \operatorname{tg} \varphi_2 \cdot \operatorname{tg} \varphi_1},$$

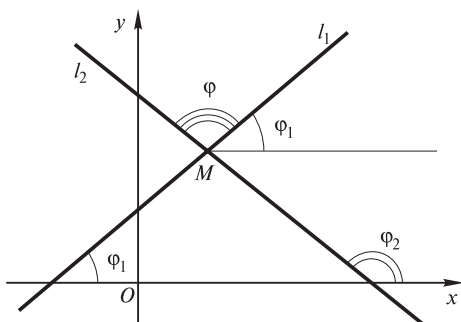


Рис. 10

или

$$\operatorname{tg} \varphi = \frac{k_2 - k_1}{1 + k_1 k_2}. \quad (10)$$

Формула (10) дает выражение тангенса угла между двумя прямыми через угловые коэффициенты этих прямых.

Если прямые l_1 и l_2 параллельны, то $\varphi_1 = \varphi_2$ и, следовательно, $k_1 = k_2$, т. е. *параллельные прямые имеют равные угловые коэффициенты*.

Обратно: если $k_1 = k_2$, то $\operatorname{tg} \varphi_1 = \operatorname{tg} \varphi_2$. Следовательно, $\varphi_1 = \varphi_2$, так как $0 \leq \varphi < \pi$, и прямые l_1 и l_2 параллельны.

Пусть $\varphi = \frac{\pi}{2}$, т. е. l_1 и l_2 взаимно перпендикулярны. В этом случае $\varphi_2 = \varphi_1 + \frac{\pi}{2}$, откуда $\operatorname{tg} \varphi_2 = \operatorname{tg} \left(\varphi_1 + \frac{\pi}{2} \right) = -\operatorname{ctg} \varphi_1 = -\frac{1}{\operatorname{tg} \varphi_1}$, или

$$k_2 = -\frac{1}{k_1}, \quad (11)$$

т. е. *угловые коэффициенты взаимно перпендикулярных прямых обратны по абсолютной величине и противоположны по знаку*.

Справедливо и обратное утверждение.

П р и м е р. Найти две прямые, проходящие через начало координат, одна из которых параллельна, а другая перпендикулярна прямой $y = 2x - 3$.

Так как искомые прямые проходят через точку $(0; 0)$, то их уравнения имеют вид $y = k_1 x$, $y = k_2 x$. Для данной прямой $k = 2$. Отсюда и на основании условий параллельности и перпендикулярности прямых получаем $k_1 = 2$ и $k_2 = -\frac{1}{2}$. Поэтому искомые прямые записываются уравнениями $y = 2x$ и $y = -\frac{1}{2}x$.

6. Взаимное расположение двух прямых на плоскости. Если две прямые l_1 и l_2 лежат на плоскости, то возможны три различных случая их взаимного расположения: 1) пересекаются (т. е. имеют одну общую точку); 2) параллельны и не совпадают; 3) совпадают.

Выясним, как узнать, какой из этих случаев имеет место, если эти прямые заданы своими уравнениями в общем виде:

$$\begin{cases} A_1 x + B_1 y + C_1 = 0, \\ A_2 x + B_2 y + C_2 = 0. \end{cases} \quad (12)$$

Если прямые l_1 и l_2 пересекаются в некоторой точке $M(x; y)$, то координаты этой точки должны удовлетворять обоим уравнениям системы (12). Следовательно, чтобы найти координаты точки пересечения

прямых l_1 и l_2 , надо решить систему уравнений (12). Умножим первое из уравнений (12) на A_2 , а второе на A_1 и вычтем первое из второго. Имеем:

$$(A_1 B_2 - A_2 B_1)y + C_2 A_1 - C_1 A_2 = 0. \quad (13)$$

Аналогично получаем:

$$(A_1 B_2 - A_2 B_1)x + C_1 B_2 - C_2 B_1 = 0. \quad (14)$$

Если $A_1 B_2 - A_2 B_1 \neq 0$, то из (13) и (14) получаем решение системы уравнений (12):

$$x = \frac{B_1 C_2 - B_2 C_1}{A_1 B_2 - A_2 B_1}, y = \frac{C_1 A_2 - C_2 A_1}{A_1 B_2 - A_2 B_1} \quad (15)$$

Это и есть координаты x и y точки пересечения прямых l_1 и l_2 .

Итак, если $A_1 B_2 - A_2 B_1 \neq 0$, то эти прямые пересекаются.

Если

$$A_1 B_2 - A_2 B_1 = 0, \quad (16)$$

то выражения для x и y не имеют смысла. В этом случае прямые l_1 и l_2 параллельны. Действительно, из условия (16) следует, что $-\frac{A_1}{B_1} = -\frac{A_2}{B_2}$, т. е.

$k_1 = k_2$ (если же $B_1 - B_2 = 0$, то прямые l_1 и l_2 параллельны оси Oy и, следовательно, параллельны между собой). Условие (16) можно записать в виде

$$\frac{A_1}{A_2} = \frac{B_1}{B_2}. \quad (17)$$

Здесь один из знаменателей может оказаться равным нулю. Чтобы обойти эту трудность, договоримся всякую пропорцию $\frac{a}{b} = \frac{c}{d}$ понимать в смысле равенства $ad = bc$. Тогда обращение в нуль одного из знаменателей в (17) означает обращение в нуль и соответствующего числителя. В самом деле, если, например, $A_2 = 0$, то, поскольку $B_2 \neq 0$ (A_2 и B_2 не нули одновременно), из равенства $A_1 B_2 = A_2 B_1$ замечаем, что $A_1 = 0$.

В частности, параллельные прямые могут совпадать. Выясним, каков аналитический признак совпадения прямых l_1 и l_2 .

Пусть

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}. \quad (18)$$

Полагая каждое из этих отношений равным q , найдем:

$$A_1 = A_2 q, B_1 = B_2 q, C_1 = C_2 q.$$

Таким образом, первое уравнение (12) получается из второго умножением всех его членов на некоторое число q , т. е. уравнения (12) равносильны. Следовательно, рассматриваемые параллельные прямые совпадают.

Если при выполнении условия (17) хотя бы один из свободных членов уравнений (13) и (14) будет отличен от нуля (или $C_2 A_1 - C_1 A_2 \neq 0$, или $C_1 B_2 - C_2 B_1 \neq 0$), что кратко записывают так:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2}, \quad (19)$$

то уравнения (13) и (14), а значит, и уравнения (12) не будут иметь решений (по крайней мере одно из равенств (13) или (14) будет невозможным). В этом случае параллельные прямые l_1 и l_2 не будут совпадать.

Итак, *условием совпадения двух прямых является пропорциональность соответствующих коэффициентов их уравнений.*

Допустим, что ни одна из прямых l_1 и l_2 не параллельна оси Oy . Тогда их уравнения можно записать в виде

$$y = k_1 x + b_1, y = k_2 x + b_2,$$

где

$$k_1 = -\frac{A_1}{B_1}, k_2 = -\frac{A_2}{B_2}. \quad (20)$$

Условие перпендикулярности таких прямых следует из равенств (11), (20) и имеет вид:

$$A_1 A_2 + B_1 B_2 = 0. \quad (21)$$

Хотя соотношение (21) получено в предположении, что ни одна из прямых l_1 и l_2 не параллельна оси Oy , оно остается верным, если это условие нарушается.

П р и м е р 1. Прямые $3x + 4y - 1 = 0$ и $2x + 3y - 1 = 0$ пересекаются, так как $3 \cdot 3 - 2 \cdot 4 \neq 0$. Координаты точки пересечения согласно (15) $x = -1, y = 1$.

П р и м е р 2. Прямые $2x - y + 2 = 0$ и $4x - 2y - 1 = 0$ параллельны (они не имеют общей точки), так как выполнено условие (19).

П р и м е р 3. Прямые $x + y + 1 = 0$ и $3x + 3y + 3 = 0$ совпадают, так как выполнено условие (18).

Пример 4. Прямые $3x - 4y + 8 = 0$ и $4x + 3y - 9 = 0$ взаимно перпендикулярны, так как выполнено условие (21).

7. Нормальное уравнение прямой.

Пусть дана какая-нибудь прямая, не проходящая через полюс O (рис. 11). Проведем через полюс прямую, перпендикулярную данной, и обозначим P точку ее пересечения с

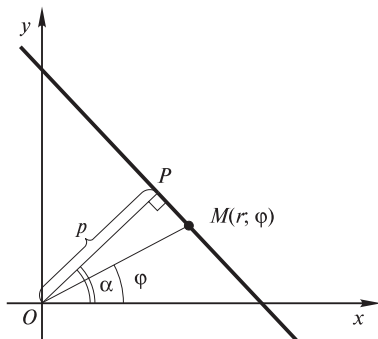


Рис. 11

данной прямой (рис. 11). Пусть α — угол между полярной осью и лучом OP , p — длина отрезка OP . Выведем уравнение данной прямой, считая известными α и $p = OP$. Пусть $M(r; \varphi)$ — произвольная точка данной прямой. Из прямоугольного треугольника OPM имеем:

$$r \cos(\alpha - \varphi) = p. \quad (22)$$

Это уравнение называется *уравнением прямой в полярных координатах*. Перепишем уравнение (22) в виде

$$r \cos \varphi \cos \alpha + r \sin \varphi \sin \alpha - p = 0.$$

Отсюда, учитывая зависимость между полярными и прямоугольными координатами, получим:

$$x \cos \alpha + y \sin \alpha - p = 0. \quad (23)$$

Это уравнение называется *нормальным уравнением прямой*.

Если дано общее уравнение прямой $Ax + By + C = 0$ (A и B не равны нулю одновременно), то его можно привести к виду (23). Действительно, умножив обе части этого общего уравнения на некоторое число $\mu \neq 0$, получим $\mu Ax + \mu By + \mu C = 0$. Выберем μ так, чтобы выполнялись равенства

$$\mu A = \cos \alpha, \mu B = \sin \alpha, \mu C = -p. \quad (24)$$

Возведя обе части двух первых равенств (24) в квадрат и почленно складывая, получим:

$$\mu^2 (A^2 + B^2) = \cos^2 \alpha + \sin^2 \alpha = 1.$$

Так как $A^2 + B^2 \neq 0$ (A и B не равны нулю одновременно), то

$$\mu = \frac{\pm 1}{\sqrt{A^2 + B^2}}.$$

Число μ называется *нормирующим множителем*. Третье из равенств (24) позволяет решить вопрос о выборе знака числа μ . Так как $p > 0$, то $\mu C < 0$, т. е. знак μ выбирается противоположным знаку C ; если $C = 0$, то для μ можно выбрать любой знак.

Итак, общее уравнение прямой приводится к нормальному виду путем умножения его на нормирующий множитель μ .

П р и м е р. Уравнение прямой $3x - 4y - 5 = 0$ привести к нормальному виду. Нормирующий множитель $\mu = \frac{+1}{\sqrt{3^2 + 4^2}} = \frac{1}{5}$. Умножая на него обе части данного уравнения, получим:

$$\frac{3}{5}x - \frac{4}{5}y - 1 = 0.$$

Для данной прямой, следовательно, $p = 1$, $\cos \alpha = \frac{3}{5}$, $\sin \alpha = -\frac{4}{5}$.

8. Расстояние от точки до прямой. Найдем расстояние d от данной точки $M_0(x_0; y_0)$ до прямой l (рис. 12), заданной уравнением (23) (под расстоянием d от точки $M_0(x_0; y_0)$ до прямой l понимается длина перпендикуляра, опущенного из M_0 на l).

Проведем через точку M_0 прямую l_1 , параллельную l . Запишем нормальное уравнение прямой l_1 :

$$x \cos \alpha + y \sin \alpha - (p + d) = 0.$$

Прямая l_1 проходит через точку $M_0(x_0; y_0)$, поэтому

$$x_0 \cos \alpha + y_0 \sin \alpha - (p + d) = 0.$$

Отсюда

$$d = x_0 \cos \alpha + y_0 \sin \alpha - p. \quad (25)$$

Если точка $M_0(x_0; y_0)$ и начало координат лежат по одну сторону от прямой l (рис. 13), то аналогично предыдущему установим, что

$$d = -(x_0 \cos \alpha + y_0 \sin \alpha - p). \quad (26)$$

Из равенств (25) и (26) следует, что

$$d = |x_0 \cos \alpha + y_0 \sin \alpha - p|. \quad (27)$$

Если прямая l задана общим уравнением

$$Ax + By + C = 0,$$

то с учетом вывода, сделанного в конце пункта 7, формула (27) принимает вид:

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}. \quad (28)$$

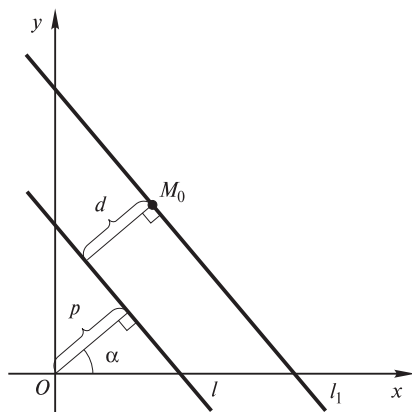


Рис. 12

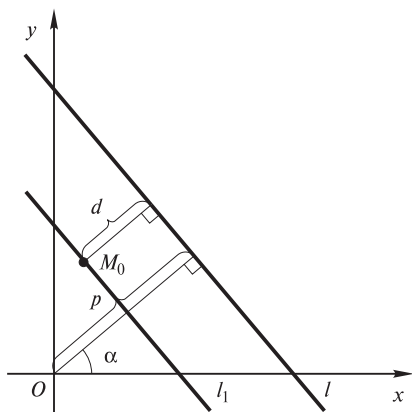


Рис. 13

П р и м е р. Найти расстояние от точки $M_0(-6; 3)$ до прямой $3x - 4y + 15 = 0$. По формуле (28) получаем:

$$d = \frac{|3 \cdot (-6) - 4 \cdot 3 + 15|}{\sqrt{3^2 + (-4)^2}} = \frac{15}{5} = 3.$$

§ 1.5. Основные задачи на прямую

1. Составить уравнение произвольной прямой, проходящей через точку $M_1(x_1; y_1)$. Пусть

$$Ax + By + C = 0 \quad (1)$$

есть уравнение искомой прямой. Значит, M_1 лежит на этой прямой. Поэтому

$$Ax_1 + By_1 + C = 0. \quad (2)$$

Вычитая из (1) почленно (2), получаем:

$$A(x - x_1) + B(y - y_1) = 0. \quad (3)$$

Очевидно, при любых A и B уравнению (3) удовлетворяют координаты точки M_1 .

2. Составить уравнение прямой, проходящей через две данные (различные) точки $M_1(x_1; y_1)$ и $M_2(x_2; y_2)$. Уравнение прямой, проходящей через точку $M_1(x_1; y_1)$, имеет вид (3). Так как прямая проходит и через точку M_2 , то

$$A(x_2 - x_1) + B(y_2 - y_1) = 0,$$

откуда

$$-\frac{A}{B} = \frac{y_2 - y_1}{x_2 - x_1}, \quad (4)$$

и с учетом равенства (3) искомое уравнение запишем в виде:

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}. \quad (5)$$

П р и м е ч а н и е 1. Заметим, что в уравнении (5) один из знаменателей $x_2 - x_1$ или $y_2 - y_1$ может оказаться равным нулю (оба числа $x_2 - x_1$ и $y_2 - y_1$ равняться нулю не могут, ибо точки M_1 и M_2 различные). Так как всякую пропорцию мы условились (см. § 1.4, п. 6) понимать как равенство $ad = bc$, то обращение в нуль одного из знаменателей в равенстве (5) означает обращение в нуль и соответствующего числителя. Если, например, $y_2 - y_1 = 0$, то $y = y_1$.

П р и м е ч а н и е 2. Формула (4) определяет угловой коэффициент прямой, проходящей через точки M_1 и M_2 .

П р и м е р. Написать уравнение прямой, проходящей через точки $M_1(4; -2)$ и $M_2(3; -1)$.

На основании уравнения (5) имеем:

$$\frac{x - 4}{3 - 4} = \frac{y + 2}{-1 + 2}, \quad \text{или} \quad x + y - 2 = 0.$$

§ 1.6. Уравнение линии

1. Уравнение окружности. Как известно, *окружность* называется множество точек плоскости, равноудаленных от данной точки, называемой центром окружности. Выведем уравнение окружности.

Пусть $M_0(x_0; y_0)$ — центр окружности, R — ее радиус, а $M(x; y)$ — произвольная точка окружности с текущими координатами x и y (рис. 14).

По определению окружности $M_0M = R$. Отсюда согласно формуле расстояния между двумя точками $\sqrt{(x - x_0)^2 + (y - y_0)^2} = R$, или

$$(x - x_0)^2 + (y - y_0)^2 = R^2. \quad (1)$$

Формула (1) представляет собой уравнение окружности; это — уравнение второй степени относительно x и y . В главе 5 подробно рассмотрим еще три вида линий (эллипс, гипербола, парабола), уравнения которых в прямоугольной системе координат также являются уравнениями второй степени. Эти четыре вида линий называются *кривыми второго порядка*.

Если центр окружности совпадает с началом координат, то ее уравнение принимает вид:

$$x^2 + y^2 = R^2. \quad (1')$$

В полярных координатах (r ; φ) окружность радиуса R с центром в полюсе изображается уравнением $r = R$.

В тех же координатах (r ; φ) окружность радиуса R с центром в $(R; 0)$ имеет уравнение $r = 2R \cos \varphi$.

2. Параметрические уравнения линий. Иногда бывает удобно вместо уравнения линии, связывающего прямоугольные координаты x и y , рассматривать так называемые *параметрические уравнения* линии, дающие выражения текущих координат x и y в виде функций от некоторой переменной величины t (параметра). Параметрические уравнения играют важную роль, например, в механике, где координаты x и y движущейся точки $M(x; y)$ рассматриваются как функции времени (*уравнения движения*).

З а д а ч а. Определить траекторию полета снаряда, выпущенного из орудия с начальной скоростью v_0 под углом α к горизонту (рис. 15).

Сопротивлением воздуха и размерами снаряда, считая его материальной точкой, пренебрегаем.

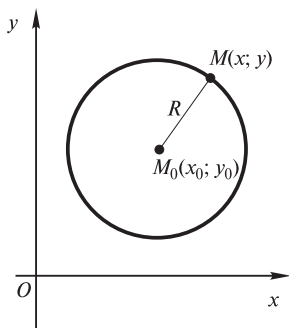


Рис. 14

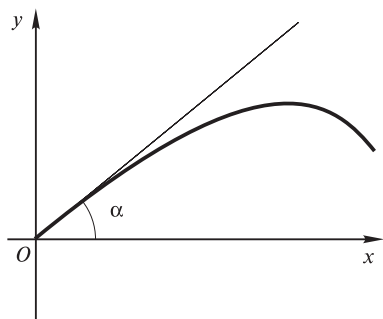


Рис. 15

Выберем систему координат. За начало координат примем точку вылета снаряда из дула. Ox направим горизонтально; а ось Oy — вертикально, расположив их в одной плоскости с дулом орудия. Если бы не было силы земного тяготения, то снаряд двигался бы по прямой, составляющей угол α с осью Ox , и к моменту времени t прошел бы путь $v_0 t$. Координаты снаряда в момент времени t были бы $x = v_0 t \cos \alpha$, $y = v_0 t \sin \alpha$. Вследствие земного тяготения снаряд должен к этому моменту вертикально опуститься на величину $\frac{gt^2}{2}$. Поэтому в действительности в момент времени t координаты снаряда определяются по формулам

$$x = v_0 t \cos \alpha, \quad y = v_0 t \sin \alpha - \frac{gt^2}{2}, \quad (2)$$

где v_0 , α , g — постоянные величины. При изменении t будут изменяться также координаты x и y точки траектории снаряда. Уравнения (2) являются параметрическими уравнениями траектории снаряда, в которых параметр есть время t . Выразив из первого уравнения уравнений (2)

$t = \frac{x}{v_0 \cos \alpha}$ и подставив его во второе уравнение, получим уравнение траектории снаряда в виде

$$y = x \tan \alpha - \frac{gx^2}{2v_0^2 \cos^2 \alpha}.$$

Это — уравнение параболы (см. § 6.2).

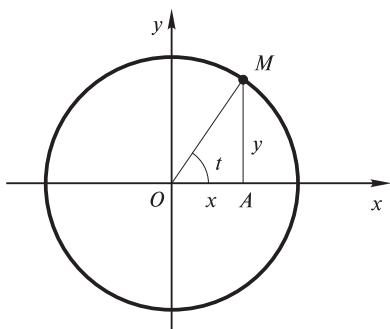


Рис. 16

П р и м е р 1. Установим параметрические уравнения окружности радиуса R с центром в начале координат.

Пусть $M(x; y)$ — любая точка этой окружности, а t — угол AOM (рис. 16). Очевидно,

$$x = R \cos t, \quad y = R \sin t. \quad (3)$$

Это и есть параметрические уравнения окружности. Параметр t может принимать

любые значения, но для того, чтобы точка $M(x; y)$ один раз обошла окружность, следует ограничить изменение параметра t неравенством $0 \leq t < 2\pi$. Заметим, что для исключения параметра t из уравнений (3) достаточно возвести эти уравнения в квадрат и сложить их: мы получим при этом уравнение (1') предыдущего пункта.

П р и м е р 2. Каждой точке $N(X; Y)$, лежащей на окружности $X^2 + Y^2 = a^2$, поставим в соответствие точку $M(x; y)$, где $x = X$, $y = \frac{b}{a}Y$ (a, b — положительные постоянные, причем $b < a$).

Мы как бы «сдавливаем» окружность сверху и снизу (рис. 17). Получающееся при этом множество точек на плоскости называется *эллипсом* (a и b — *полуоси эллипса*). Используя формулы (3), имеем:

$$x = X = a \cos t, \quad y = \frac{b}{a}Y = \frac{b}{a}a \sin t = b \sin t$$

Таким образом, параметрические уравнения эллипса с полуосями a и b есть

$$x = a \cos t, \quad y = b \sin t. \quad (4)$$

Исключив из уравнений (4) параметр t , получим уравнения эллипса

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

П р и м е р 3. Рассмотрим линию, являющуюся траекторией фиксированной точки окружности радиуса R , катящейся без скольжения по прямой. Линия эта называется *циклоидой*. Указанную прямую примем за ось Ox прямоугольной системы координат (рис. 18). Предположим, что фиксированная точка при начальном положении окружности находилась в начале координат, а после того как окружность повернулась на угол t , заняла положение M .

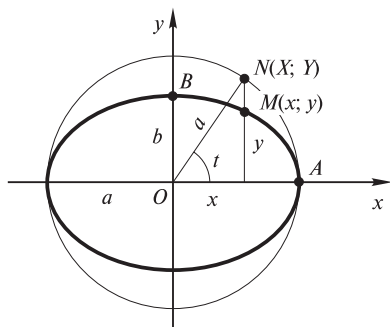


Рис. 17

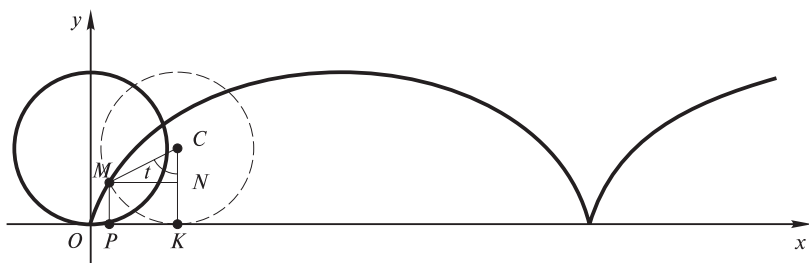


Рис. 18

Поскольку

$$\begin{aligned}x &= OP = OK - PK, \quad y = MP = CK - CN \text{ и } OK = \cup MK = Rt, \\PK &= MN = R \sin t, \quad CK = R, \quad CN = R \cos t, \text{ то } x = Rt - R \sin t, \quad y = R - R \cos t, \text{ или} \\x &= R(t - \sin t), \quad y = R(1 - \cos t).\end{aligned}\quad (5)$$

Уравнения (5) называются *параметрическими уравнениями циклоиды*. При изменении t от 0 до 2π получается одна арка циклоиды.

3. Примеры кривых, заданных уравнениями в полярных координатах.

П р и м е р 1. Рассмотрим уравнение $r = a \varphi$, где a — положительное число, r и φ — полярные координаты. Обозначим через M точку с полярными координатами $(r; \varphi)$. Если $\varphi = 0$, то и $r = 0$; если φ возрастает, начиная с нуля, то r будет возрастать пропорционально φ . Точка $M(r; \varphi)$, таким образом, исходя из полюса, движется вокруг него с ростом φ (в положительном направлении), одновременно удаляясь от него. Множество точек, полярные координаты которых удовлетворяют уравнению $r = a \varphi$, называется *спиралью Архимеда* (рис. 19).

П р и м е р 2. Кривая, задаваемая уравнением

$$r = a(1 + \cos \varphi) \quad (a > 0, 0 \leq \varphi \leq 2\pi),$$

называется *кардиоидой*.

Составляя таблицу значений φ и r , получим:

φ	0	$\pm \frac{\pi}{6}$	$\pm \frac{\pi}{3}$	$\pm \frac{\pi}{2}$	$\pm \frac{2}{3}\pi$	$\pm \frac{5}{6}\pi$	$\pm \pi$
r	$2a$	$\approx 1,9a$	$\frac{3}{2}a$	a	$\frac{a}{2}$	$\approx 0,1a$	0

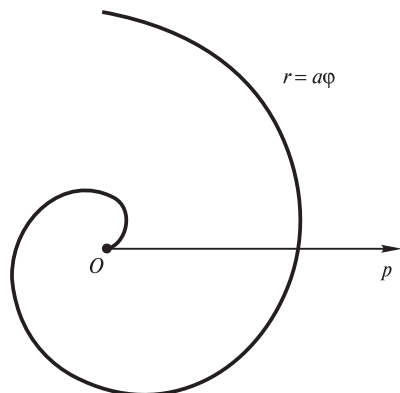


Рис. 19

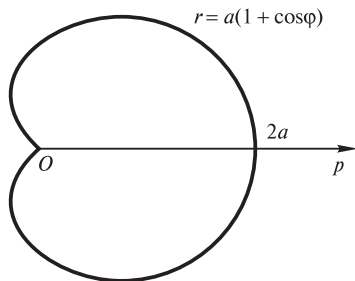


Рис. 20

Построив точки кардиоиды по значениям φ и r из этой таблицы, можно составить приближенное представление о форме этой кривой (рис. 20).

П р и м е р 3. При выводе нормального уравнения прямой (см. § 1.4, п. 7) было получено уравнение прямой в полярных координатах:

$$r \cos(\alpha - \varphi) = p.$$

Глава 2. ВЕКТОРНАЯ АЛГЕБРА

§ 2.1. Понятие вектора и линейные операции над векторами

1. Понятие вектора. При изучении различных разделов физики, механики и технических наук встречаются величины, которые полностью определяются заданием их числовых значений. Такие величины называются *скалярными* или, короче, *скалярами*. Скалярными величинами, например, являются длина, площадь, объем, масса, температура тела и др. Помимо скалярных величин, в различных задачах встречаются величины, для определения которых, кроме числового значения, необходимо знать также их направление. Такие величины называются *векторными*. Физическими примерами векторных величин могут служить смещение материальной точки, движущейся в пространстве, скорость и ускорение этой точки, а также действующая на нее сила.

Векторные величины изображаются с помощью векторов.

Вектором называется направленный отрезок, имеющий определенную длину, т. е. отрезок определенной длины, у которого одна из ограничивающих его точек принимается за начало, а вторая — за конец. Если A — начало вектора и B — его конец, то вектор обозначается символом $\overrightarrow{AB}$. Вектор можно обозначать и одной малой латинской буквой с чертой над ней (например, $\vec{a}$). Изображается вектор отрезком со стрелкой на конце (рис. 21). Начало вектора называют *точкой его приложения*. Если точка A является началом вектора $\vec{a}$, то мы будем говорить, что вектор $\vec{a}$ приложен в точке A .

Длина вектора $\overrightarrow{AB}$ называется его *модулем* и обозначается символом $|\overrightarrow{AB}|$. Модуль вектора $\vec{a}$ обозначается $|\vec{a}|$.

Вектор $\vec{a}$, для которого $|\vec{a}| = 1$, называется *единичным*.

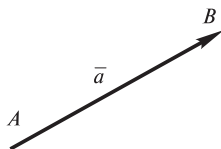


Рис. 21

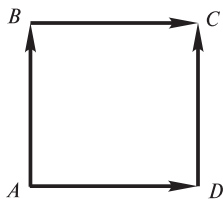


Рис. 22

Вектор называется *нулевым* (обозначается $\vec{0}$), если начало и конец его совпадают. Нулевой вектор не имеет определенного направления и имеет длину, равную нулю.

Векторы $\vec{a}$ и $\vec{b}$, расположенные на одной прямой или на параллельных прямых, называются *коллинеарными*.

Два вектора $\vec{a}$ и $\vec{b}$ называются *равными*, если они коллинеарны, имеют одинаковую длину и одинаковое направление.

В этом случае пишут $\vec{a} = \vec{b}$. Все нулевые векторы считаются равными. Из определения равенства векторов следует, что вектор можно переносить параллельно самому себе, помещая его начало в любую точку пространства (в частности, плоскости). Такой вектор называется *свободным*.

Пример. Рассмотрим квадрат $ABCD$ (рис. 22). На основании определения равенства векторов можем написать $\vec{AD} = \vec{BC}$ и $\vec{AB} = \vec{DC}$, но $\vec{AB} \neq \vec{AD}$, $\vec{BC} \neq \vec{DC}$, хотя $|\vec{AB}| = |\vec{AD}| = |\vec{BC}| = |\vec{DC}|$.

Два коллинеарных вектора (отличные от нулевых векторов), имеющие равные модули, но противоположно направленные, называются *противоположными*.

Вектор, противоположный вектору $\vec{a}$, обозначается $-\vec{a}$. Для вектора $\vec{AB}$ противоположным будет вектор $\vec{BA}$.

2. Линейные операции над векторами. *Линейными операциями* называются операции сложения и вычитания векторов и умножения вектора на число.

Определение 1. Пусть $\vec{a}$ и $\vec{b}$ — два свободных вектора (рис. 23, а). Возьмем произвольную точку O и построим вектор $\vec{OA} = \vec{a}$, затем от точки A отложим вектор $\vec{AB} = \vec{b}$. Вектор $\vec{OB}$, соединяющий начало первого слагаемого вектора с концом второго, называется *суммой* этих векторов и обозначается $\vec{a} + \vec{b}$ (рис. 23, б).

Ту же самую сумму векторов можно получить иным способом. Отложим от точки O векторы $\vec{OA} = \vec{a}$ и $\vec{OC} = \vec{b}$. Построим на этих векторах как на сторонах параллелограмма $OABC$. Вектор $\vec{OB}$, служащий диагональю этого параллелограмма, проведенной из вершины O , является, очевидно, суммой векторов $\vec{a} + \vec{b}$ (рис. 23, в). Из рисунка 23, в непосредственно следует, что сумма двух векторов обладает *переместительным* свойством:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}.$$

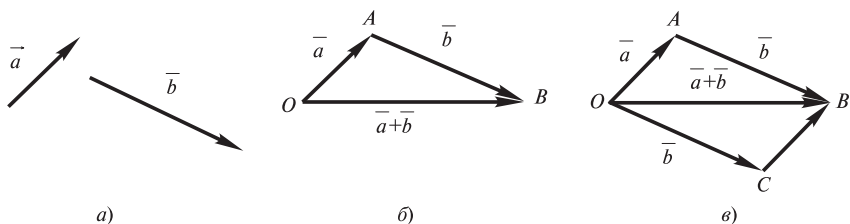


Рис. 23

Действительно, каждый из векторов $\vec{a} + \vec{b}$ и $\vec{b} + \vec{a}$ равен одному и тому же вектору $\vec{OB}$.

Понятие суммы векторов, введенное для двух слагаемых векторов, можно обобщить на случай любого конечного числа слагаемых векторов.

Пусть, например, даны три вектора $\vec{a}$, $\vec{b}$ и $\vec{c}$ (рис. 24, а). Построив сначала сумму векторов $\vec{a} + \vec{b}$, а затем прибавив к этой сумме вектор $\vec{c}$, получим вектор $(\vec{a} + \vec{b}) + \vec{c}$. На рисунке 24, б $\vec{OA} = \vec{a}$, $\vec{AB} = \vec{b}$, $\vec{OB} = \vec{a} + \vec{b}$, $\vec{BC} = \vec{c}$ и $\vec{OC} = \vec{OB} + \vec{BC} = (\vec{a} + \vec{b}) + \vec{c}$.

Из рисунка 24, б видно, что тот же вектор $\vec{OC}$ мы получим, если к вектору $\vec{OA} = \vec{a}$ прибавим вектор $\vec{AC} = \vec{b} + \vec{c}$. Таким образом,

$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}),$$

т. е. сумма векторов обладает *сочетательным* свойством. Поэтому сумму трех векторов $\vec{a}$, $\vec{b}$, $\vec{c}$ записывают просто $\vec{a} + \vec{b} + \vec{c}$.

Итак, сумму трех векторов можно получить следующим образом. Из произвольной точки O откладывается вектор, равный первому слагаемому вектору. К концу первого вектора присоединяется начало второго; к концу второго — начало третьего. Вектор, соединяющий начало первого вектора с концом последнего, является суммой данных векторов. Подобным же образом строится сумма любого конечного числа векторов.

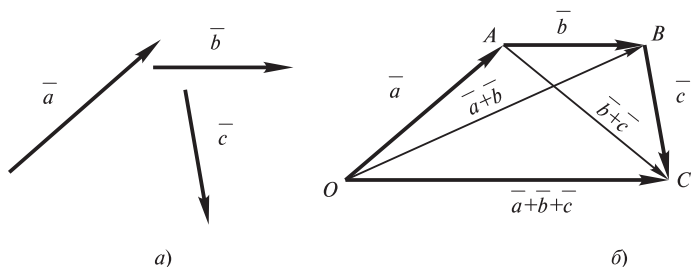


Рис. 24

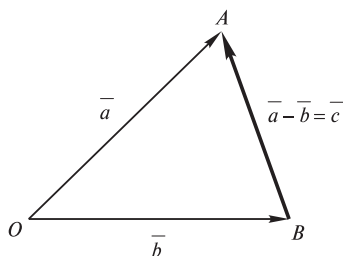


Рис. 25

Если при сложении нескольких векторов конец последнего слагаемого вектора совпадает с началом первого, то сумма векторов равна нулевому вектору. Очевидно, что для любого вектора имеет место равенство $\vec{a} + \vec{0} = \vec{a}$.

О п р е д е л е н и е 2. Разностью двух векторов $\vec{a}$ и $\vec{b}$ называется третий вектор $\vec{c} = \vec{a} - \vec{b}$, сумма которого с вычитаемым вектором $\vec{b}$ дает вектор $\vec{a}$. Таким образом, если $\vec{c} = \vec{a} - \vec{b}$, то $\vec{c} + \vec{b} = \vec{a}$.

Из определения суммы двух векторов вытекает правило построения вектора–разности (рис. 25). Откладываем векторы $\vec{OA} = \vec{a}$ и $\vec{OB} = \vec{b}$ из общей точки O . Вектор $\vec{BA}$, соединяющий концы уменьшаемого вектора $\vec{a}$ и вычитаемого вектора $\vec{b}$ и направленный от вычитаемого к уменьшаемому, является разностью $\vec{c} = \vec{a} - \vec{b}$. Действительно, по правилу сложения векторов

$$\vec{OB} + \vec{BA} = \vec{OA}, \text{ или } \vec{b} + \vec{c} = \vec{a}.$$

О п р е д е л е н и е 3. Произведением $\lambda\vec{a}$ (или $\vec{a}\lambda$) вектора $\vec{a}$ на действительное число λ называется вектор $\vec{b}$, коллинеарный вектору $\vec{a}$, имеющий длину, равную $|\lambda||\vec{a}|$, и то же направление, что и вектор $\vec{a}$, если $\lambda > 0$, и направление, противоположное направлению вектора $\vec{a}$, если $\lambda < 0$. Так, например, $2\vec{a}$ есть вектор, имеющий то же направление, что и вектор $\vec{a}$, а длину, вдвое большую, чем вектор $\vec{a}$. В случае, когда $\lambda = 0$ или $\vec{a} = \vec{0}$, произведение $\lambda\vec{a}$ представляет собой нулевой вектор. Противоположный вектор $-\vec{a}$ можно рассматривать как результат умножения вектора $\vec{a}$ на $\lambda = -1$:

$$-\vec{a} = (-1)\vec{a}.$$

Очевидно, что $\vec{a} + (-\vec{a}) = \vec{0}$. Пусть дан вектор $\vec{a}$. Рассмотрим единичный вектор $\vec{a}_0$, коллинеарный вектору $\vec{a}$ и одинаково с ним направленный. Из определения умножения вектора на число следует, что

$$\vec{a} = |\vec{a}|\vec{a}_0,$$

т. е. каждый вектор равен произведению его модуля на единичный вектор того же направления. Далее из того же определения следует, что если $\vec{b} = \lambda\vec{a}$, где $\vec{a}$ — ненулевой вектор, то векторы $\vec{a}$ и $\vec{b}$ коллинеарны. Очевидно, что и, наоборот, из коллинеарности векторов $\vec{a}$ и $\vec{b}$ следует, что $\vec{b} = \lambda\vec{a}$.

Таким образом, два вектора $\vec{a}$ и $\vec{b}$ коллинеарны тогда и только тогда, когда имеет место равенство

$$\vec{b} = \lambda \vec{a}.$$

Легко убедиться, что умножение вектора на число обладает распределительным свойством:

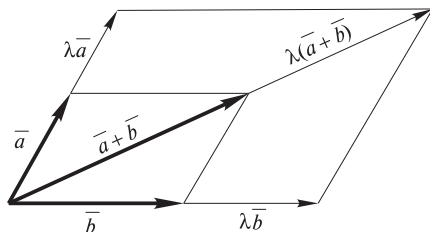


Рис. 26

$$\lambda(\vec{a} + \vec{b}) = \lambda \vec{a} + \lambda \vec{b}, \quad (1)$$

$$(\lambda_1 + \lambda_2)\vec{a} = \lambda_1 \vec{a} + \lambda_2 \vec{a}$$

и сочетательным свойством:

$$\lambda_1 (\lambda_2 \vec{a}) = (\lambda_1 \lambda_2) \vec{a}.$$

Справедливость, например, равенства (1) при $\lambda > 0$, следует из того, что при изменении сторон параллелограмма в λ раз в силу свойств подобия его диагональ также изменяется в λ раз (рис. 26).

3. Понятие линейной зависимости векторов. Векторы $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ называются *линейно зависимыми*, если существуют числа $\lambda_1, \lambda_2, \dots, \lambda_n$, не все равные нулю, для которых имеет место равенство

$$\lambda_1 \vec{a}_1 + \lambda_2 \vec{a}_2 + \dots + \lambda_n \vec{a}_n = \vec{0}. \quad (2)$$

Векторы $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ называются *линейно независимыми*, если равенство (2) имеет место только при $\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$.

Из равенства (2), предполагая, например, что $\lambda_1 \neq 0$, получим

$$\vec{a}_1 = -\frac{\lambda_2}{\lambda_1} \vec{a}_2 - \frac{\lambda_3}{\lambda_1} \vec{a}_3 - \dots - \frac{\lambda_n}{\lambda_1} \vec{a}_n.$$

Полагая

$$-\frac{\lambda_2}{\lambda_1} = \mu_2, -\frac{\lambda_3}{\lambda_1} = \mu_3, \dots, -\frac{\lambda_n}{\lambda_1} = \mu_n,$$

имеем:

$$\vec{a}_1 = \mu_2 \vec{a}_2 + \mu_3 \vec{a}_3 + \dots + \mu_n \vec{a}_n. \quad (3)$$

Выражение

$$\mu_2 \vec{a}_2 + \mu_3 \vec{a}_3 + \dots + \mu_n \vec{a}_n$$

называется *линейной комбинацией* векторов $\vec{a}_2, \vec{a}_3, \dots, \vec{a}_n$.

Таким образом, если несколько векторов линейно зависимы, то хотя бы один из них всегда можно представить в виде линейной комбинации остальных.

Справедливо и обратное утверждение: *если один из векторов представлен в виде линейной комбинации остальных векторов, то все эти векторы линейно зависимы.*

В самом деле, пусть, например, вектор $\bar{a}_1$ является линейной комбинацией векторов $\bar{a}_2, \bar{a}_3, \dots, \bar{a}_n$. Тогда имеет место равенство (3). Перепишав его в виде $-\bar{a}_1 + \mu_2 \bar{a}_2 + \mu_3 \bar{a}_3 + \dots + \mu_n \bar{a}_n = \bar{0}$, убеждаемся в том, что один из коэффициентов (именно коэффициент при $\bar{a}_1$) отличен от нуля. Отсюда в силу определения и вытекает линейная зависимость векторов $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n$.

4. Линейная зависимость векторов на плоскости.

Т е о р е м а 1. *Всякие три вектора $\bar{a}, \bar{b}, \bar{c}$ на плоскости линейно зависимы.*

Д о к а з а т е л ь с т в о. Достаточно убедиться в том, что один из векторов является линейной комбинацией остальных. Возможны два случая:

1. Среди данных векторов имеется пара коллинеарных векторов, например, $\bar{a}$ и $\bar{b}$. Тогда (см. п. 2)

$$\bar{a} = \lambda \bar{b} \quad \text{или} \quad \bar{a} = \lambda \bar{b} + 0 \bar{c},$$

т. е. вектор $\bar{a}$ есть линейная комбинация векторов $\bar{b}$ и $\bar{c}$.

2. Среди данных векторов нет ни одной пары коллинеарных. Допустим, что все три вектора имеют общее начало O (рис. 27). Покажем, что вектор $\bar{a}$ можно представить в виде суммы двух векторов, один из которых коллинеарен вектору $\bar{b}$, а другой — вектору $\bar{c}$.

Для этого через конец M вектора $\bar{a}$ проведем прямые, параллельные векторам $\bar{b}$ и $\bar{c}$, до их пересечения в точках B и C с прямыми, на которых соответственно расположены векторы $\bar{b}$ и $\bar{c}$. Имеем очевидное равенство

$$\overline{OM} = \overline{OB} + \overline{OC}.$$

Так как векторы $\overline{OB}$ и $\overline{OC}$ коллинеарны соответственно векторам $\bar{b}$ и $\bar{c}$, то $\overline{OB} = \lambda_1 \bar{b}$ и $\overline{OC} = \lambda_2 \bar{c}$. Поэтому

$$\bar{a} = \lambda_1 \bar{b} + \lambda_2 \bar{c}, \quad (4)$$

т. е. вектор $\bar{a}$ является линейной комбинацией векторов $\bar{b}$ и $\bar{c}$.

С л е д с т в и е. *Если число данных векторов на плоскости больше трех, то они также линейно зависимы.*

В самом деле, пусть даны n векторов $\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n$ ($n > 3$). Так как три вектора на плоскости всегда линейно зависимы, то для

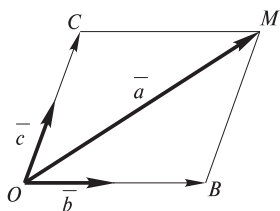


Рис. 27

векторов $\bar{a}_1, \bar{a}_2, \bar{a}_3$ имеем $\bar{a}_1 = \mu_2 \bar{a}_2 + \mu_3 \bar{a}_3$. В таком случае для всех n векторов можно написать:

$$\bar{a}_1 = \mu_2 \bar{a}_2 + \mu_3 \bar{a}_3 + 0 \cdot \bar{a}_4 + \dots + 0 \cdot \bar{a}_n,$$

т. е. вектор $\bar{a}_1$ есть линейная комбинация остальных векторов.

Что касается двух векторов $\bar{a}$ и $\bar{b}$, то, как известно (п. 2), они коллинеарны тогда и только тогда, когда имеет место равенство $\bar{b} = \lambda \bar{a}$, т. е. когда векторы $\bar{a}$ и $\bar{b}$ линейно зависимы. Отсюда непосредственно вытекает следующая теорема:

Т е о р е м а 2. *Для того чтобы два вектора $\bar{a}$ и $\bar{b}$ на плоскости были линейно независимы, необходимо и достаточно, чтобы они были неколлинеарны.*

Из теорем 1 и 2 следует, что *максимальное число линейно независимых векторов на плоскости равно двум.*

5. Линейная зависимость векторов в пространстве.

О п р е д е л е н и е. Векторы называются *компланарными*, если они лежат в одной плоскости или параллельны одной плоскости.

Заметим, что если компланарные векторы имеют общее начало, то они, очевидно, лежат в одной плоскости.

Т е о р е м а 1. *Всякие четыре вектора $\bar{a}, \bar{b}, \bar{c}$ и $\bar{d}$ в пространстве линейно зависимы.*

Д о к а з а т е л ь с т в о. Допустим, что рассматриваемые векторы имеют общее начало. Для того чтобы показать их линейную зависимость, достаточно убедиться в том, что один из векторов является линейной комбинацией остальных. Возможны два случая:

1. Среди данных векторов существует тройка компланарных, например векторы $\bar{a}, \bar{b}$ и $\bar{c}$.

Так как эти векторы лежат в одной плоскости, то по теореме 1 п. 4 один из них, например вектор $\bar{a}$, можно представить в виде линейной комбинации остальных: $\bar{a} = \lambda_1 \bar{b} + \lambda_2 \bar{c}$. В таком случае для всех четырех векторов можно написать равенство $\bar{a} = \lambda_1 \bar{b} + \lambda_2 \bar{c} + 0 \cdot \bar{d}$, а это означает, что вектор $\bar{a}$ есть линейная комбинация векторов $\bar{b}, \bar{c}$ и $\bar{d}$.

2. Среди данных векторов нет ни одной тройки компланарных векторов. В этом случае вектор $\bar{a}$ может быть представлен в виде суммы трех векторов, коллинеарных соответственно векторам $\bar{b}, \bar{c}$ и $\bar{d}$. Для этого, проведя через точку M — конец вектора $\bar{a}$ — плоскости, соответственно параллельные трем плоскостям, определяемым парами векторов $\bar{b}$ и $\bar{c}, \bar{c}$ и $\bar{d}, \bar{d}$ и $\bar{b}$, получаем параллелепипед, диагональю которого является вектор $\bar{a} = \overline{OM}$ (рис. 28). Имеем $\bar{a} = \overline{OM} = \overline{OP} + \overline{OM_3} = \overline{OM_1} + \overline{OM_2} + \overline{OM_3}$. Но $\overline{OM_1} = \lambda_1 \bar{b}, \overline{OM_2} = \lambda_2 \bar{c}, \overline{OM_3} = \lambda_3 \bar{d}$. Следовательно, $\bar{a} = \lambda_1 \bar{b} + \lambda_2 \bar{c} + \lambda_3 \bar{d}$ т. е. векторы $\bar{a}, \bar{b}, \bar{c}$ и $\bar{d}$ линейно зависимы.

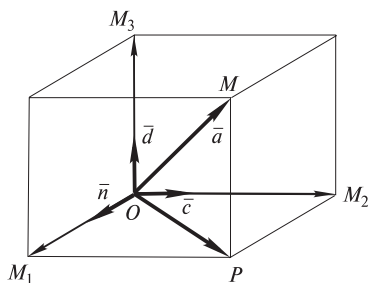


Рис. 28

Из этой теоремы аналогично следствию из пункта 4 получим

С л е д с т в и е. Если число данных векторов в пространстве больше четырех, то они также линейно зависимы.

Аналогично случаю для коллинеарных векторов устанавливается следующее предложение:

Для того чтобы три вектора в пространстве были компланарны, необходимо и достаточно, чтобы они были линейно зависимы.

Отсюда непосредственно вытекает следующая теорема:

Т е о р е м а 2. Для того чтобы три вектора $\vec{a}$, $\vec{b}$ и $\vec{c}$ в пространстве были линейно независимы, необходимо и достаточно, чтобы они были некомпланарны.

Из теорем 1 и 2 следует, что максимальное число линейно независимых векторов в пространстве равно трем.

6. Базис на плоскости и в пространстве.

О п р е д е л е н и е 1. Базисом на плоскости называются два любых линейно независимых вектора.

Из теоремы 2 (п. 4) следует, что два любых неколлинеарных вектора образуют базис. Пусть $\vec{a}$ — любой вектор на плоскости, а векторы $\vec{b}$ и $\vec{c}$ образуют базис. Так как на плоскости всякие три вектора линейно зависимы, то вектор $\vec{a}$ линейно выражается через векторы базиса, т. е. выполняется соотношение

$$\vec{a} = \lambda_1 \vec{b} + \lambda_2 \vec{c}. \quad (5)$$

Если вектор $\vec{a}$ представлен в виде (5), то говорят, что он *разложен по базису*, образованному векторами $\vec{b}$ и $\vec{c}$. Числа λ_1 и λ_2 называют *координатами вектора $\vec{a}$ на плоскости относительно базиса $\vec{b}$ и $\vec{c}$* .

Т е о р е м а 1. Разложение вектора $\vec{a}$ по базису $\vec{b}$ и $\vec{c}$ является единственным.

Д о к а з а т е л ь с т в о. Допустим, что наряду с разложением (5) имеет место разложение

$$\vec{a} = v_1 \vec{b} + v_2 \vec{c}. \quad (6)$$

Покажем, что в этом случае $v_1 = \lambda_1$, $v_2 = \lambda_2$. Действительно, вычитая равенство (6) из равенства (5), получим соотношение

$$\vec{0} = (\lambda_1 - v_1) \vec{b} + (\lambda_2 - v_2) \vec{c}.$$

Возможность почленного вычитания равенств (6) и (5) и производимой группировки членов вытекает из свойств линейных операций над

векторами (см. п. 2). Так как векторы базиса $\bar{b}, \bar{c}$ линейно независимы, то $\lambda_1 - v_1 = 0$ и $\lambda_2 - v_2 = 0$. Отсюда, $\lambda_1 = v_1$ и $\lambda_2 = v_2$, т. е. разложение вектора $\bar{a}$ по базису $\bar{b}, \bar{c}$ единственно.

О п р е д е л е н и е 2. *Базисом в пространстве* называются три любых линейно независимых вектора.

Из теоремы 2 (п. 5) следует, что три любых некомпланарных вектора образуют базис. Как и в случае плоскости, устанавливается, что любой вектор $\bar{a}$ разлагается по векторам $\bar{b}, \bar{c}$ и $\bar{d}$ базиса:

$$\bar{a} = \lambda_1 \bar{b} + \lambda_2 \bar{c} + \lambda_3 \bar{d},$$

причем это разложение единственное.

Числа λ_1, λ_2 и λ_3 называются *координатами вектора $\bar{a}$* в пространстве относительно базиса $\bar{b}, \bar{c}$ и $\bar{d}$.

Основное значение базиса состоит в том, что линейные операции над векторами при задании базиса становятся обычными линейными операциями над числами — координатами этих векторов.

Т е о р е м а 2. *При сложении двух векторов $\bar{a}_1$ и $\bar{a}_2$ их координаты (относительно любого базиса $\bar{b}, \bar{c}$ и $\bar{d}$) складываются. При умножении вектора $\bar{a}_1$ на любое число α все его координаты умножаются на это число.*

Д о к а з а т е л ь с т в о. Пусть

$$\bar{a}_1 = \lambda_1 \bar{b} + \mu_1 \bar{c} + v_1 \bar{d}, \quad \bar{a}_2 = \lambda_2 \bar{b} + \mu_2 \bar{c} + v_2 \bar{d}.$$

Тогда в силу свойств линейных операций (см. п. 2)

$$\begin{aligned} \bar{a}_1 + \bar{a}_2 &= (\lambda_1 + \lambda_2) \bar{b} + (\mu_1 + \mu_2) \bar{c} + (v_1 + v_2) \bar{d}, \\ \alpha \bar{a}_1 &= (\alpha \lambda_1) \bar{b} + (\alpha \mu_1) \bar{c} + (\alpha v_1) \bar{d}. \end{aligned}$$

В силу единственности разложения по базису $\bar{b}, \bar{c}$ и $\bar{d}$ теорема доказана.

7. Проекция вектора на ось и ее свойства.

О п р е д е л е н и е 1. *Углом между векторами $\bar{a}$ и $\bar{b}$* называется наименьший угол φ ($0 \leq \varphi \leq \pi$), на который надо повернуть один из векторов до его совпадения со вторым после приведения этих векторов к общему началу.

Осью называется направленная прямая. Направление прямой на рисунке обычно обозначается стрелкой. Заданное направление оси считается положительным, противоположное — отрицательным.

Рассмотрим ось l , положительное направление которой совпадает с направлением единичного вектора $\bar{l}_0$, расположенного на оси l . Такой вектор называется *ортом* оси l .

О п р е д е л е н и е 2. *Углом между вектором $\bar{a}$ и осью l* называется угол φ между векторами $\bar{a}$ и $\bar{l}_0$ (рис. 29).

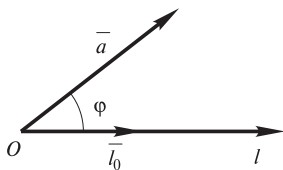


Рис. 29

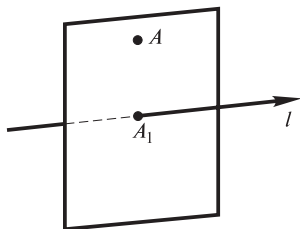


Рис. 30

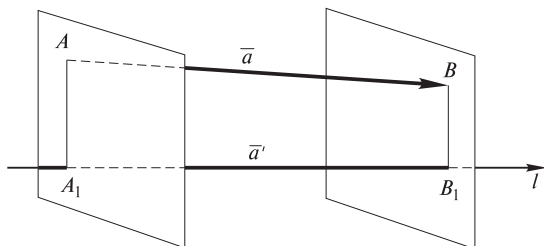


Рис. 31

О п р е д е л е н и е 3. *Проекцией точки A на ось l* (рис. 30) называется точка A_1 , в которой пересекается ось l с плоскостью, перпендикулярной к l , проходящей через точку A .

О п р е д е л е н и е 4. *Компонентой (составляющей) вектора $\vec{a} = \overline{AB}$ на ось l* (рис. 31) называется вектор $\vec{a}' = \overline{A_1B_1}$, где A_1, B_1 соответственно проекции точек A, B на l .

О п р е д е л е н и е 5. *Проекцией вектора $\vec{a}$ на ось l* ($\text{пр}_l \vec{a}$) называется длина его компоненты $\vec{a}'$ на ось l , взятая со знаком «плюс», если направление компоненты совпадает с направлением оси l , и со знаком «минус», если направление компоненты противоположно направлению оси l .

Если $\vec{a} = \vec{0}$, то полагают $\text{пр}_l \vec{a} = 0$.

Т е о р е м а 1. *Проекция вектора $\vec{a}$ на ось l равна произведению его модуля на косинус угла φ между этим вектором и осью l:*

$$\text{пр}_l \vec{a} = |\vec{a}| \cos \varphi.$$

Д о к а з а т е л ь с т в о. Так как вектор $\vec{a} = \overline{OA}$ свободный, то можно предположить, что начало его O лежит на оси l (рис. 32).

Если угол φ острый ($0 \leq \varphi < \frac{\pi}{2}$), то направление компоненты $\vec{a}' = \overline{OA_1}$ вектора $\vec{a}$ совпадает с направлением оси l (рис. 32, а).

В этом случае имеем: $\text{пр}_l \vec{a} = +|\overline{OA_1}| = |\vec{a}| \cos \varphi$.

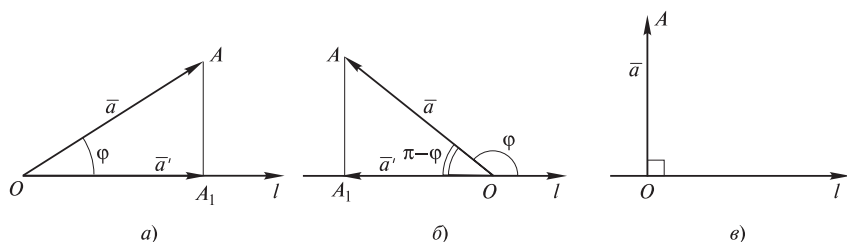


Рис. 32

Если же угол φ тупой $\left(\frac{\pi}{2} < \varphi \leq \pi\right)$ (рис. 32, б), то направление компоненты $\vec{a}' = \overline{OA_1}$ вектора $\vec{a}$ противоположно направлению оси l . Тогда получаем: $\text{пр}_l \vec{a} = -|\overline{OA_1}| = -|\vec{a}| \cos(\pi - \varphi) = |\vec{a}| \cos \varphi$.

Наконец, если $\varphi = \frac{\pi}{2}$ (рис. 32, в), то $\text{пр}_l \vec{a} = 0$ и $\cos \varphi = 0$. Таким образом, снова имеем соотношение $\text{пр}_l \vec{a} = |\vec{a}| \cos \varphi$.

С л е д с т в и е 1. Проекция вектора на ось положительна, если вектор образует с осью острый угол, отрицательна, если этот угол тупой, равна нулю, если этот угол прямой.

С л е д с т в и е 2. Проекции равных векторов на одну и ту же ось равны между собой.

Т е о р е м а 2. Проекции векторов $\vec{a}, \vec{b}$ на данную ось обладают следующими свойствами:

$$\text{пр}_l(\vec{a} + \vec{b}) = \text{пр}_l \vec{a} + \text{пр}_l \vec{b}, \quad (7)$$

$$\text{пр}_l(\lambda \vec{a}) = \lambda \text{пр}_l \vec{a}. \quad (8)$$

Д о к а з а т е л ь с т в о. Свойство (7) иллюстрируется рисунком 33.

Докажем свойство (8). Считая, что угол между вектором $\vec{a} = \overline{AB}$ и направлением l равен φ , имеем:

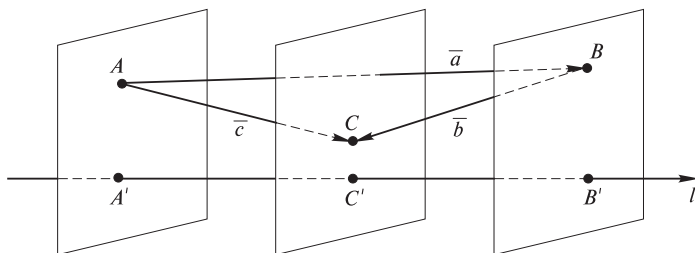


Рис. 33

при $\lambda > 0$ $\text{пр}_l(\lambda \bar{a}) = |\lambda \bar{a}| \cos \varphi = \lambda |\bar{a}| \cos \varphi = \lambda \text{пр}_l \bar{a}$;

при $\lambda < 0$ $\text{пр}_l(\lambda \bar{a}) = |\lambda \bar{a}| \cos(\pi - \varphi) = -\lambda |\bar{a}| \cos(\pi - \varphi) = \lambda |\bar{a}| \cos \varphi =$
 $= \lambda \text{пр}_l \bar{a}$

(при $\lambda < 0$ вектор $\lambda \bar{a}$ направлен в сторону, противоположную направлению $\bar{a}$; если $\bar{a}$ образует с l угол φ , то $\lambda \bar{a}$ образует с l угол $\pi - \varphi$).

При $\lambda = 0$ левая и правая части (8) обращаются в нуль.

8. Декартова прямоугольная система координат в пространстве.

Три взаимно перпендикулярные оси в пространстве (координатные оси) с общим началом O и одинаковой масштабной единицей образуют *декартову прямоугольную* (кратно — *прямоугольную*) *систему координат в пространстве*. Оси упорядочены, т. е. указано, какая из осей считается первой (она называется *осью абсцисс* и обозначается Ox), какая — второй (*ось ординат* Oy) и какая — третьей (*ось аппликат* Oz).

Различают *правую* и *левую* системы декартовых прямоугольных координат (рис. 34, соответственно *а*, *б*). В этой книге принята правая система координат (будем называть ее основной).

Орты осей Ox , Oy , Oz обозначают соответственно через $\bar{i}$, $\bar{j}$, $\bar{k}$. Так как векторы $\bar{i}$, $\bar{j}$, $\bar{k}$ некомпланарны, то они образуют базис (см. п. 6), который называется *декартовым прямоугольным базисом*.

В силу результатов п. 6 каждый вектор $\bar{a}$ может быть, и притом единственным способом, разложен по декартовому прямоугольному базису $\bar{i}$, $\bar{j}$, $\bar{k}$, т. е. для каждого вектора $\bar{a}$ найдется, и притом единственная, тройка чисел a_x , a_y , a_z , такая, что справедливо равенство

$$\bar{a} = a_x \bar{i} + a_y \bar{j} + a_z \bar{k}. \quad (9)$$

Числа a_x , a_y , a_z называются *декартовыми прямоугольными* (или *прямоугольными*) *координатами вектора* $\bar{a}$.

Запись $\bar{a} (a_x; a_y; a_z)$ означает, что вектор $\bar{a}$ имеет декартовы прямоугольные координаты a_x , a_y , a_z .

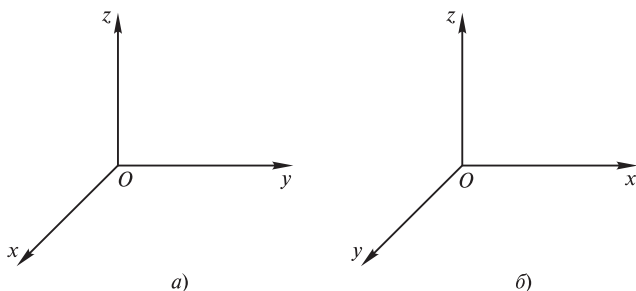


Рис. 34

Выясним геометрический смысл чисел a_x , a_y , a_z . Используя теоремы 2 и 1 о проекциях (см. п. 7), имеем:

$$\text{пр}_{Ox} \bar{a} = a_x \text{пр}_{Ox} \bar{i} + a_y \text{пр}_{Ox} \bar{j} + a_z \text{пр}_{Ox} \bar{k} = a_x.$$

Аналогично устанавливаем $\text{пр}_{Oy} \bar{a} = a_y$, $\text{пр}_{Oz} \bar{a} = a_z$. Следовательно, числа a_x , a_y , a_z в формуле (9) являются проекциями вектора $\bar{a}$ на координатные оси Ox , Oy , Oz соответственно.

Если M — произвольная точка в пространстве, то *радиусом* — *вектором* точки M назовем вектор $\overline{OM}$, имеющий своим началом начало O заданной системы координат, а концом — эту точку.

О п р е д е л е н и е. *Декартовыми прямоугольными координатами точки M* называются проекции ее радиуса—вектора $\overline{OM}$ на соответствующие координатные оси; проекция на первую координатную ось называется *абсциссой* точки M , на вторую — *ординатой*, на третью — *апplikатой*:

$$x = \text{пр}_{Ox} \overline{OM}, \quad y = \text{пр}_{Oy} \overline{OM}, \quad z = \text{пр}_{Oz} \overline{OM}.$$

Символ $M(x; y; z)$ означает, что точка M имеет координаты x , y , z .

Координатные плоскости (плоскости, проходящие через пары координатных осей) делят все пространство на восемь частей, называемых *октантами*, которые нумеруются следующим образом: октант, лежащий над первой четвертью плоскости xOy , — I; лежащий под ней — V; соответственно октанты, лежащие над и под второй четвертью плоскости xOy , — II и VI; над и под третьей четвертью — III и VII; над и под четвертой четвертью — IV и VIII.

Каждому октанту соответствует определенная комбинация знаков координат:

Координаты	Октанты							
	I	II	III	IV	V	VI	VII	VIII
x	+	—	—	+	+	—	—	+
y	+	+	—	—	+	+	—	—
z	+	+	+	+	—	—	—	—

Отметим, что каждой точке пространства соответствует одна упорядоченная тройка действительных чисел $(x; y; z)$ (ее координат). Верно и обратное: каждой упорядоченной тройке действительных чисел $(x; y; z)$ соответствует одна точка пространства. Это значит, что в пространстве

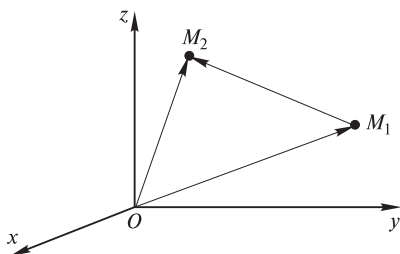


Рис. 35

положение произвольной точки M полностью определяется ее координатами x, y, z .

Пусть задана точка $M(x; y; z)$. Поскольку координаты радиуса-вектора $\overline{OM}$ совпадают с проекциями этого вектора на оси координат, т. е. с координатами точки M , то согласно равенству (9) имеем: $\overline{OM} = x\vec{i} + y\vec{j} + z\vec{k}$. (Если точ-

ка M лежит в плоскости xOy , то $\overline{OM} = x\vec{i} + y\vec{j}$.)

Пусть теперь заданы две точки $M_1(x_1; y_1; z_1)$ и $M_2(x_2; y_2; z_2)$.

Рассмотрим вектор $\overline{M_1M_2}$. Имеем $\overline{M_1M_2} = \overline{OM_2} - \overline{OM_1}$ (рис. 35).

Отсюда в силу теоремы 2 (п. 6) получаем:

$$\overline{M_1M_2} (x_2 - x_1; y_2 - y_1; z_2 - z_1).$$

Итак, чтобы найти координаты некоторого вектора, достаточно из координат его конца вычесть одноименные координаты его начала.

Пусть два ненулевых вектора

$$\vec{a} = a_x\vec{i} + a_y\vec{j} + a_z\vec{k} \quad \text{и} \quad \vec{b} = b_x\vec{i} + b_y\vec{j} + b_z\vec{k}$$

коллинеарны. В этом случае (см. п. 2) $\vec{b} = \lambda\vec{a}$ (λ — скаляр), что в силу следствия 2 из п. 7 равносильно трем равенствам

$$\frac{a_x}{b_x} = \frac{a_y}{b_y} = \frac{a_z}{b_z}. \quad (10)$$

Это есть условие коллинеарности векторов.

Таким образом, векторы коллинеарны тогда и только тогда, когда их одноименные координаты пропорциональны.

П р и м е ч а н и е. В равенстве (10) некоторые из знаменателей могут оказаться равными 0. Напомним, что всякую пропорцию $\frac{a}{b} = \frac{c}{d}$ понимаем в смысле равенства $ad = bc$.

Так, например, равенства $\frac{a_x}{0} = \frac{a_y}{0} = \frac{a_z}{2}$ означают, что $2a_x = 0 \cdot a_z$, $2a_y = 0 \cdot a_z$, $0 \cdot a_x = 0 \cdot a_y$, т. е. что $a_x = 0$, $a_y = 0$.

З а д а ч а. Пусть даны точки $M_1(x_1; y_1; z_1)$ и $M_2(x_2; y_2; z_2)$. Требуется найти точку $M(x; y; z)$, лежащую на отрезке M_1M_2 и делящую его в данном отношении:

$$\frac{M_1M}{MM_2} = \lambda.$$

Очевидно, что $\overline{M_1 M} = \lambda \overline{M M_2}$ или $(x - x_1)\bar{i} + (y - y_1)\bar{j} + (z - z_1)\bar{k} = \lambda(x_2 - x)\bar{i} + \lambda(y_2 - y)\bar{j} + \lambda(z_2 - z)\bar{k}$. Отсюда $x - x_1 = \lambda(x_2 - x)$, $y - y_1 = \lambda(y_2 - y)$, $z - z_1 = \lambda(z_2 - z)$ и, наконец,

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}, \quad y = \frac{y_1 + \lambda y_2}{1 + \lambda}, \quad z = \frac{z_1 + \lambda z_2}{1 + \lambda}.$$

§ 2.2. Нелинейные операции над векторами

1. Скалярное произведение двух векторов и его основные свойства.

О п р е д е л е н и е. *Скалярным произведением двух векторов* называется число, равное произведению модулей этих векторов на косинус угла между ними.

Скалярное произведение векторов $\bar{a}$ и $\bar{b}$ обозначается символом $\bar{a}\bar{b}$ или $(\bar{a}, \bar{b})$. Если угол между векторами $\bar{a}$ и $\bar{b}$ равен φ , то

$$\bar{a}\bar{b} = |\bar{a}| |\bar{b}| \cos \varphi. \quad (1)$$

Через $\text{пр}_{\bar{a}} \bar{b}$ обозначим проекцию вектора $\bar{b}$ на ось с направлением вектора $\bar{a}$.

Так как

$$|\bar{b}| \cos \varphi = \text{пр}_{\bar{a}} \bar{b} \quad \text{и} \quad |\bar{a}| \cos \varphi = \text{пр}_{\bar{b}} \bar{a}$$

(см. § 2.1, п. 7), то можно записать:

$$\bar{a}\bar{b} = |\bar{a}| \text{пр}_{\bar{a}} \bar{b} = |\bar{b}| \text{пр}_{\bar{b}} \bar{a}, \quad (2)$$

т. е. *скалярное произведение двух векторов равно модулю одного из них, умноженному на проекцию другого на ось с направлением первого.*

Раскроем физический смысл скалярного произведения. Пусть постоянная сила $\bar{F}$ обеспечивает прямолинейное перемещение $\bar{s} = \overline{MN}$ материальной точки M . Если сила $\bar{F}$ образует угол φ с перемещением $\bar{s}$ (рис. 36), то, как известно из физики, работа A силы $\bar{F}$ при перемещении $\bar{s}$ равна:

$$A = |\bar{F}| |\bar{s}| \cos \varphi,$$

или согласно формуле (1) $A = \bar{F}\bar{s}$. Таким образом, *работа постоянной силы при прямолинейном перемещении ее точки приложения равна скалярному произведению вектора силы на вектор перемещения.*

Скалярное произведение обладает следующими основными свойствами:

1. $\bar{a}\bar{b} = \bar{b}\bar{a}$ (переместительное свойство).

$$2. \bar{a}^2 = \bar{a}\bar{a} = |\bar{a}|^2 \quad (3)$$

($\bar{a}^2$ называется *скалярным квадратом вектора*).

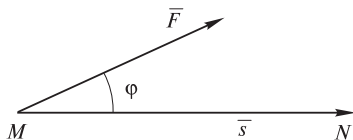


Рис. 36

3. $(\bar{a} + \bar{b})\bar{c} = \bar{a}\bar{c} + \bar{b}\bar{c}$ (распределительное свойство).

4. $(\lambda\bar{a})\bar{b} = \lambda(\bar{a}\bar{b})$ (4)

(сочетательное свойство относительно числового множителя).

Свойства 1 и 2 непосредственно вытекают из определения скалярного произведения.

Докажем свойство 3. На основании формулы (2) и свойства проекций (§ 2.1, (7)) имеем:

$$\begin{aligned}(\bar{a} + \bar{b})\bar{c} &= |\bar{c}| \operatorname{пр}_{\bar{c}}(\bar{a} + \bar{b}) = |\bar{c}|(\operatorname{пр}_{\bar{c}}\bar{a} + \operatorname{пр}_{\bar{c}}\bar{b}) = \\&= |\bar{c}| \operatorname{пр}_{\bar{c}}\bar{a} + |\bar{c}| \operatorname{пр}_{\bar{c}}\bar{b} = \bar{c}\bar{a} + \bar{c}\bar{b} = \bar{a}\bar{c} + \bar{b}\bar{c}.\end{aligned}$$

При доказательстве свойства 4 ограничимся случаем $\lambda > 0$. Замечая, что при $\lambda > 0$ угол φ между векторами $\bar{a}$ и $\bar{b}$ равен углу между векторами $\lambda\bar{a}$ и $\bar{b}$, получим:

$$\lambda(\bar{a}\bar{b}) = \lambda|\bar{a}||\bar{b}|\cos\varphi, \quad (\lambda\bar{a})\bar{b} = |\lambda\bar{a}||\bar{b}|\cos\varphi = \lambda|\bar{a}||\bar{b}|\cos\varphi,$$

откуда следует равенство (4).

П р и м е ч а н и е. Из свойств 1, 3, 4 скалярного умножения и свойств линейных операций над векторами (§ 2.1, п. 2) следует, что векторы можно перемножать скалярно как многочлены.

Из равенства (1) следует, что косинус угла между двумя ненулевыми векторами $\bar{a}$ и $\bar{b}$ равен:

$$\cos\varphi = \frac{\bar{a}\bar{b}}{|\bar{a}||\bar{b}|}. \quad (5)$$

Из формулы (5) получаем, что два вектора $\bar{a}$ и $\bar{b}$ перпендикулярны (ортogonalны), т. е. $\varphi = \frac{\pi}{2}$, тогда и только тогда, когда

$$\bar{a}\bar{b} = 0. \quad (6)$$

Это утверждение справедливо также и в том случае, когда хотя бы один из векторов $\bar{a}$ и $\bar{b}$ нулевой (нулевой вектор имеет неопределенное направление, и его можно считать ортогональным любому вектору).

2. Скалярное произведение векторов в координатной форме. Пусть

$$\bar{a} = a_x\bar{i} + a_y\bar{j} + a_z\bar{k}$$

и

$$\bar{b} = b_x\bar{i} + b_y\bar{j} + b_z\bar{k}.$$

Перемножая эти векторы как многочлены и учитывая вытекающие из равенств (3) и (6) соотношения

$$\bar{i}\bar{j} = \bar{j}\bar{k} = \bar{k}\bar{i} = 0, \quad \bar{i}\bar{i} = \bar{j}\bar{j} = \bar{k}\bar{k} = 1,$$

будем иметь:

$$\bar{a} \bar{b} = a_x b_x + a_y b_y + a_z b_z. \quad (7)$$

Таким образом, *скалярное произведение двух векторов равно сумме парных произведений их одноименных координат.*

П р и м е р. Если $\bar{a}(1; 3; -1)$, $\bar{b}(1; 0; 4)$, то по формуле (7) имеем $\bar{a}\bar{b} = -3$.

Из равенства (7) с учетом формулы (3) имеем:

$$|\bar{a}| = \sqrt{a_x a_x + a_y a_y + a_z a_z} = \sqrt{a_x^2 + a_y^2 + a_z^2}. \quad (8)$$

Отсюда с учетом формул (5) и (7) находим угол между векторами:

$$\cos \varphi = \frac{a_x b_x + a_y b_y + a_z b_z}{\sqrt{a_x^2 + a_y^2 + a_z^2} \sqrt{b_x^2 + b_y^2 + b_z^2}}. \quad (9)$$

З а д а ч а. Найти расстояние между точками $M_1(x_1; y_1; z_1)$ и $M_2(x_2; y_2; z_2)$.
Так как (см. § 2.1, п. 8)

$$\overline{M_1 M_2} (x_2 - x_1; y_2 - y_1; z_2 - z_1),$$

то согласно формуле (8)

$$|\overline{M_1 M_2}| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

В п. 1 было отмечено необходимое и достаточное условие ортогональности векторов в виде равенства (6). Согласно формуле (7) это условие можно представить в виде

$$a_x b_x + a_y b_y + a_z b_z = 0. \quad (10)$$

3. Направляющие косинусы вектора. Пусть дан ненулевой вектор $\bar{a}(a_x; a_y; a_z)$. Обозначим углы наклона этого вектора к осям Ox , Oy и Oz соответственно буквами α , β и γ . Три числа $\cos \alpha$, $\cos \beta$ и $\cos \gamma$ принято называть *направляющими косинусами вектора $\bar{a}$* . Полагая $\bar{b} = \bar{i}(1; 0; 0)$, получим из (9)

$$\cos \alpha = \frac{a_x}{\sqrt{a_x^2 + a_y^2 + a_z^2}}. \quad (11)$$

Аналогично

$$\cos \beta = \frac{a_y}{\sqrt{a_x^2 + a_y^2 + a_z^2}}, \quad (12)$$

$$\cos \gamma = \frac{a_z}{\sqrt{a_x^2 + a_y^2 + a_z^2}}. \quad (13)$$

Из формул (11)—(13) следует:

$$1) \cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1,$$

т. е. сумма квадратов направляющих косинусов любого ненулевого вектора равна единице;

$$2) \frac{\cos \alpha}{a_x} = \frac{\cos \beta}{a_y} = \frac{\cos \gamma}{a_z},$$

т. е. направляющие косинусы этого вектора пропорциональны его соответствующим проекциям.

П р и м е ч а н и е. Из формул (11)—(13) видно, что проекции любого единичного вектора $\vec{a}_0$ на оси координат соответственно совпадают с его направляющими косинусами и, следовательно,

$$\vec{a}_0 = \vec{i} \cos \alpha + \vec{j} \cos \beta + \vec{k} \cos \gamma.$$

П р и м е р. Найти направляющие косинусы вектора $\vec{a}$ (1; 2; 2). По формулам (11)—(13) имеем:

$$\begin{aligned} \cos \alpha &= \frac{1}{\sqrt{1+2^2+2^2}} = \frac{1}{\sqrt{9}} = \frac{1}{3}, \\ \cos \beta &= \frac{2}{3}, \cos \gamma = \frac{2}{3}. \end{aligned}$$

4. Векторное произведение двух векторов и его основные свойства.

О п р е д е л е н и е. Векторным произведением двух векторов $\vec{a}$ и $\vec{b}$ называется новый вектор $\vec{c}$, модуль которого равен площади параллелограмма, построенного на векторах $\vec{a}$ и $\vec{b}$, приведенных к общему началу, и который перпендикулярен перемножаемым векторам (иначе говоря, перпендикулярен плоскости построенного на них параллелограмма) и направлен в такую сторону, чтобы кратчайший поворот от $\vec{a}$ к $\vec{b}$ вокруг полученного вектора $\vec{c}$ представлялся происходящим против часовой стрелки, если смотреть из конца вектора $\vec{c}$ (рис. 37).

Если векторы $\vec{a}$ и $\vec{b}$ коллинеарны, то их векторное произведение считается равным нулевому вектору.

Из этого определения следует, что

$$|\vec{c}| = |\vec{a}| |\vec{b}| \sin \varphi,$$

где φ — угол между векторами $\vec{a}$ и $\vec{b}$ ($0 \leq \varphi \leq \pi$). Векторное произведение векторов $\vec{a}$ и $\vec{b}$ обозначается символом

$$\vec{a} \times \vec{b} \text{ или } [\vec{a}\vec{b}] \text{ или } [\vec{a}, \vec{b}].$$

Выясним физический смысл векторного произведения. В физике момент силы $\vec{F}$ относительно точки O изображают вектором $\vec{OM}$, перпендикулярным плоскости, в которой лежат точка O и вектор $\vec{F}$ (рис. 38). Длину

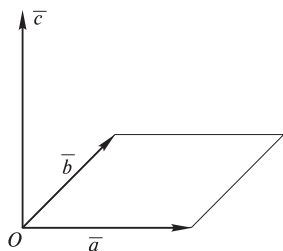


Рис. 37

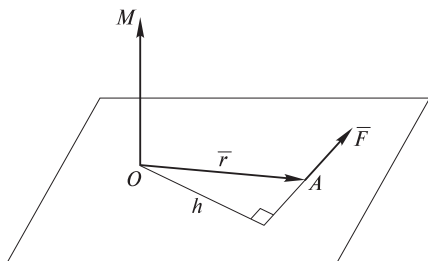


Рис. 38

вектора $\overline{OM}$ определяют как произведение длины вектора $\overline{F}$ на плечо h (h — расстояние от точки O до прямой, на которой изображен вектор силы $\overline{F}$), т. е. $|\overline{OM}| = |\overline{F}|h$, или $|\overline{OM}| = |\overline{F}||\overline{r}| \sin(\overline{F}, \overline{r})$, где $\overline{r} = \overline{OA}$ — радиус-вектор точки приложения силы $\overline{F}$. Таким образом, момент силы $\overline{F}$ относительно некоторой точки O есть $\overline{F} \times \overline{r}$, т. е. *векторное произведение силы $\overline{F}$ на радиус-вектор $\overline{r}$ точки приложения силы $\overline{F}$* .

Свойства векторного произведения

1. При перестановке сомножителей векторное произведение меняет знак, т. е.

$$\overline{a} \times \overline{b} = -(\overline{b} \times \overline{a}). \quad (14)$$

В самом деле, площадь параллелограмма, построенного на векторах $\overline{a}$ и $\overline{b}$, а также и его плоскость не меняются при перестановке $\overline{a}$ и $\overline{b}$. Поэтому векторы $\overline{a} \times \overline{b}$ и $\overline{b} \times \overline{a}$ имеют одинаковые длины и коллинеарны. Направления же этих векторов противоположны; действительно, если смотреть на плоскость векторов $\overline{a}$ и $\overline{b}$ с конца вектора $\overline{a} \times \overline{b}$, то кратчайший поворот от $\overline{b}$ к $\overline{a}$ будет казаться происходящим по часовой стрелке. Следовательно, вектор $\overline{b} \times \overline{a}$ должен быть направлен в противоположную сторону.

Заметим, что в случае коллинеарности векторов $\overline{a}$ и $\overline{b}$ равенство (14) очевидно, так как тогда $\overline{a} \times \overline{b}$ и $\overline{b} \times \overline{a}$ — нулевые векторы.

Таким образом, векторное произведение не обладает переместительным свойством.

$$2. (\lambda \overline{a}) \times \overline{b} = \overline{a} \times (\lambda \overline{b}) = \lambda(\overline{a} \times \overline{b}),$$

где λ — скаляр.

Свойство 2 непосредственно вытекает из смысла произведения вектора на скаляр (см. § 2.1, п. 2) и определения векторного произведения.

3. Векторное произведение подчиняется распределительному закону, т. е.

$$(\overline{a} + \overline{b}) \times \overline{c} = \overline{a} \times \overline{c} + \overline{b} \times \overline{c}.$$

Доказательство этого свойства здесь не приводится (его можно найти, например, в [6]).

4. Если векторное произведение двух векторов равно нулевому вектору, то либо равен нулевому вектору хотя бы один из перемножаемых векторов (тривиальный случай), либо равен нулю синус угла между ними, т. е. векторы коллинеарны.

Обратно, если два не нулевых вектора коллинеарны, то их векторное произведение равно нулевому вектору.

Таким образом, для того чтобы два ненулевых вектора $\vec{a}$ и $\vec{b}$ были коллинеарны, необходимо и достаточно, чтобы их векторное произведение равнялось нулевому вектору.

Отсюда, в частности, следует, что векторное произведение вектора на самого себя равно нулевому вектору:

$$\vec{a} \times \vec{a} = \vec{0}$$

($\vec{a} \times \vec{a}$ еще называют векторным квадратом вектора $\vec{a}$).

5. Смешанное произведение трех векторов и его основные свойства.

Пусть даны три вектора $\vec{a}$, $\vec{b}$ и $\vec{c}$. Представим себе, что вектор $\vec{a}$ умножается векторно на $\vec{b}$ и полученный вектор $\vec{a} \times \vec{b}$ умножается скалярно на вектор $\vec{c}$, тем самым определяется число $(\vec{a} \times \vec{b}) \vec{c}$. Оно называется *векторно-скалярным* или *смешанным произведением* трех векторов $\vec{a}$, $\vec{b}$, $\vec{c}$. Для краткости смешанное произведение $(\vec{a} \times \vec{b}) \vec{c}$ будем обозначать $\vec{a}\vec{b}\vec{c}$ или $(\vec{a}\vec{b}\vec{c})$.

Выясним геометрический смысл смешанного произведения $\vec{a}\vec{b}\vec{c}$. Пусть рассматриваемые векторы $\vec{a}$, $\vec{b}$ и $\vec{c}$ некопланарны. Построим параллелепипед на векторах $\vec{a}$, $\vec{b}$ и $\vec{c}$ как на ребрах. Векторное произведение $\vec{a} \times \vec{b}$ есть вектор $\vec{d}$ ($\vec{d} = \overrightarrow{OE}$), численно равный площади параллелограмма $OADB$ (основание построенного параллелепипеда), построенного на векторах $\vec{a}$ и $\vec{b}$, и направленный перпендикулярно к плоскости параллелограмма (рис. 39).

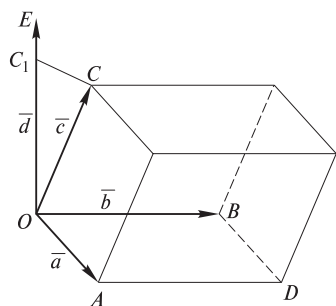


Рис. 39

Скалярное произведение $\vec{a}\vec{b}\vec{c} = \vec{d}\vec{c}$ есть произведение модуля вектора $\vec{d}$ и проекции вектора $\vec{c}$ на $\vec{d}$ (см. п. 1, (2)).

Высота построенного параллелепипеда есть абсолютная величина этой проекции.

Следовательно, произведение $|\vec{a}\vec{b}\vec{c}|$ по абсолютной величине равно произведению площади основания параллелепипеда на его высоту, т. е. объему параллелепипеда, построенного на векторах $\vec{a}$, $\vec{b}$ и $\vec{c}$.

При этом важно отметить, что скалярное произведение $\vec{d}\vec{c}$ дает объем параллелепипеда иногда с положительным, а иногда с отрицательным знаком. Положительный знак получается, если угол между векторами $\vec{d}$ и $\vec{c}$ острый; отрицательный — если тупой. При остром угле между $\vec{d}$ и $\vec{c}$ вектор $\vec{c}$ расположен по ту же сторону плоскости $OADB$, что и вектор $\vec{d}$, и, следовательно, из конца вектора $\vec{c}$ вращение от $\vec{a}$ к $\vec{b}$ будет видно так же, как и из конца вектора $\vec{d}$, т. е. в положительном направлении (против часовой стрелки).

При тупом угле между $\vec{d}$ и $\vec{c}$ вектор $\vec{c}$ расположен по другую сторону плоскости $OADB$, чем вектор $\vec{d}$, и, следовательно, из конца вектора $\vec{c}$ вращение от $\vec{a}$ к $\vec{b}$ будет видно в отрицательном направлении (по часовой стрелке). Иными словами, произведение $\vec{a}\vec{b}\vec{c}$ положительно, если векторы $\vec{a}$, $\vec{b}$ и $\vec{c}$ образуют систему, одноименную с основной $Oxyz$ (взаимно расположены так же, как оси Ox , Oy , Oz), и оно отрицательно, если векторы $\vec{a}$, $\vec{b}$ и $\vec{c}$ образуют систему, разноименную с основной.

Таким образом, *смешанное произведение $\vec{a}\vec{b}\vec{c}$ есть число, абсолютная величина которого выражает объем параллелепипеда, построенного на векторах $\vec{a}$, $\vec{b}$, $\vec{c}$ как на ребрах.*

Знак произведения положителен, если векторы $\vec{a}$, $\vec{b}$, $\vec{c}$ образуют систему, одноименную с основной, и отрицателен в противном случае.

Отсюда следует, что абсолютная величина произведения $\vec{a}\vec{b}\vec{c} = (\vec{a} \times \vec{b})\vec{c}$ останется та же, в каком бы порядке мы ни брали сомножители $\vec{a}$, $\vec{b}$, $\vec{c}$. Что касается знака, то он будет в одних случаях положительным, в других — отрицательным; это зависит от того, образуют ли наши три вектора, взятые в определенном порядке, систему, одноименную с основной, или нет. Заметим, что у нас оси координат расположены так, что они следуют одна за другой против часовой стрелки, если смотреть *во внутреннюю часть трехгранного угла* (рис. 40). Порядок следования не нарушается, если мы начнем обход со второй оси или с третьей, лишь бы он совершался в том же направлении, т. е. против часовой стрелки. При этом множители переставляются в круговом порядке. Таким образом, получаем следующее свойство:

Смешанное произведение не меняется при круговой перестановке его сомножителей. Перестановка двух соседних сомножителей меняет знак произведения:

$$\begin{aligned}\vec{a}\vec{b}\vec{c} &= \vec{b}\vec{c}\vec{a} = \vec{c}\vec{a}\vec{b} = \\ &= -(\vec{b}\vec{a}\vec{c}) = -(\vec{c}\vec{b}\vec{a}) = -(\vec{a}\vec{c}\vec{b}).\end{aligned}$$

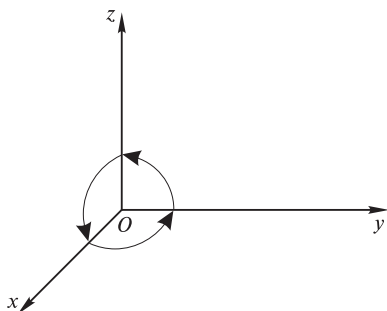


Рис. 40

Наконец, из геометрического смысла смешанного произведения непосредственно следует следующее утверждение:

Необходимым и достаточным условием компланарности векторов $\bar{a}$, $\bar{b}$, $\bar{c}$ является равенство нулю их смешанного произведения:

$$\bar{a} \bar{b} \bar{c} = 0. \quad (15)$$

Глава 3. ЛИНЕЙНАЯ АЛГЕБРА

§ 3.1. Матрицы и действия над ними

1. Понятие о матрице. Таблица чисел a_{ik} вида

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad (1)$$

состоящая из m строк и n столбцов, называется матрицей размера $m \times n$. Числа a_{ik} называются ее *элементами*. Это *прямоугольная* матрица. В частности, когда $m = 1$, $n > 1$, мы имеем однострочечную матрицу $(a_{11} \ a_{12} \dots \ a_{1n})$, которую называют *матрицей-строкой*. Если же $m > 1$, а $n = 1$, то имеем одностолбцовую матрицу

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix},$$

которую называют *матрицей-столбцом*. Матрица, состоящая из одного числа, отождествляется с этим числом.

Если в матрице число строк равно числу столбцов ($m = n$), то такую матрицу называют *квадратной*, причем число ее строк или столбцов

называется *порядком* матрицы. Например, матрица $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ есть квадратная матрица второго порядка, а матрица

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

есть квадратная матрица третьего порядка.

Матрицу для краткости будем обозначать одной буквой, например, буквой A .

Две матрицы A и B называются *равными* ($A = B$), если они одинакового размера (т. е. имеют одинаковое число строк и одинаковое число столбцов) и их соответствующие элементы равны. Так, если

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ и } B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

то $A = B$, если $a_{11} = b_{11}$, $a_{12} = b_{12}$, $a_{21} = b_{21}$, $a_{22} = b_{22}$.

2. Сложение матриц. Матрицы одинакового размера можно складывать.

Суммой двух таких матриц A и B называется матрица C , элементы которой равны сумме соответствующих элементов матриц A и B . Символически будем записывать так: $A + B = C$.

Так, если

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ и } B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

то их суммой является матрица

$$C = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}.$$

Пример.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \end{pmatrix} + \begin{pmatrix} 2 & 4 & 1 \\ 3 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 4 \\ 5 & 4 & 10 \end{pmatrix}.$$

Легко видеть, что сложение матриц подчиняется переместительному и сочетательному законам:

$$\begin{aligned} A + B &= B + A, \\ (A + B) + C &= A + (B + C). \end{aligned}$$

О п р е д е л е н и е. Матрица, все элементы которой равны нулю, называется *нуль-матрицей* и обозначается (0) или просто 0 .

Нуль-матрица при сложении матриц выполняет роль обычного нуля при сложении чисел: $A + 0 = A$.

3. Вычитание матриц. *Разностью* двух матриц A и B одинакового размера называется матрица C , такая, что

$$C + B = A.$$

Из этого определения следует, что элементы матрицы C равны разности соответствующих элементов матриц A и B .

Обозначается разность матриц A и B так: $C = A - B$.

Пример.

$$\begin{pmatrix} 2 & 3 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}.$$

4. Умножение матрицы на число. Произведением матрицы A на число λ называется матрица, элементы которой равны произведению числа λ на соответствующие элементы матрицы A .

Отсюда следует, что при умножении матрицы на нуль получается нуль-матрица.

Пример. Пусть

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 3 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

Найти матрицу $A\lambda + B\mu$.

На основании определения суммы матриц и умножения матрицы на число имеем:

$$A\lambda + B\mu = \begin{pmatrix} \lambda + 2\mu & 0 & 2\lambda + \mu \\ 2\lambda + 2\mu & 3\lambda + \mu & \lambda + \mu \end{pmatrix}.$$

5. Умножение матриц. Рассмотрим правило умножения двух квадратных матриц второго и третьего порядков.

Пусть даны две матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{и} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Произведением матрицы A на матрицу B называется матрица $C = AB$, элементы которой составляются следующим образом:

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

Как видим, элемент матрицы-произведения, находящийся на пересечении i -строки и k -го столбца, представляет собой сумму парных произведений элементов i -й строки первой матрицы на элементы k -го столбца второй матрицы.

Например, элемент, стоящий во второй строке и первом столбце матрицы произведения AB , равен сумме парных произведений элементов второй строки матрицы A на элементы первого столбца матрицы B .

Это правило сохраняется для умножения квадратных матриц третьего и более высокого порядка, а также для умножения прямо-

угольных матриц, в которых число столбцов матрицы-множимого равно числу строк матрицы-множителя.

Пример 1.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}.$$

Пример 2.

$$\begin{pmatrix} 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 + 1 \cdot 2 + 0 \cdot 2 & 2 \cdot 2 + 1 \cdot 1 + 0 \cdot 2 \\ 3 \cdot 1 + 1 \cdot 2 + 1 \cdot 2 & 3 \cdot 2 + 1 \cdot 1 + 1 \cdot 2 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 7 & 9 \end{pmatrix}.$$

Пример 3.

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}.$$

Видим, что в результате перемножения двух матриц получается матрица, содержащая столько строк, сколько их имеет матрица-множимое, и столько столбцов, сколько их имеет матрица-множитель.

Рассмотрим еще пример:

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 \cdot 1 + 6 \cdot 3 & 5 \cdot 2 + 6 \cdot 4 \\ 7 \cdot 1 + 8 \cdot 3 & 7 \cdot 2 + 8 \cdot 4 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}.$$

С другой стороны, как установлено выше:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}.$$

Следовательно, произведение двух матриц, вообще говоря, не подчиняется переместительному закону:

$$AB \neq BA.$$

Можно проверить, что умножение матриц подчиняется *сочетательно-*му закону:

$$A(BC) = (AB)C.$$

При умножении матриц второго порядка особое значение имеет квадратная матрица

$$E = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

При умножении любой квадратной матрицы $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ второго порядка на матрицу E снова получается матрица A .

Действительно,

$$\begin{aligned}
 AE &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \\
 &= \begin{pmatrix} a_{11} \cdot 1 + a_{12} \cdot 0 & a_{11} \cdot 0 + a_{12} \cdot 1 \\ a_{21} \cdot 1 + a_{22} \cdot 0 & a_{21} \cdot 0 + a_{22} \cdot 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.
 \end{aligned}$$

Аналогично

$$EA = A.$$

Матрица E называется *единичной матрицей*. Единичная матрица n -го порядка имеет вид:

$$E = \begin{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}}_{n \text{ столбцов}} \end{pmatrix}.$$

Если в матрице (1), обозначаемой буквой A , сделать все строчки столбцами с тем же номером, то получим матрицу

$$A' = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix},$$

называемую *транспонированной* к матрице A .

§ 3.2. Определители

1. Определители второго порядка. Рассмотрим квадратную матрицу второго порядка

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

О п р е д е л е н и е. *Определителем второго порядка*, соответствующим матрице A , называется число, равное $a_{11}a_{22} - a_{12}a_{21}$.

Определитель обозначают символом:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \text{ (кратко } |A| \text{)}.$$

Таким образом,

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}. \quad (1)$$

Элементы матрицы A называются *элементами определителя* $|A|$, элементы a_{11} , a_{22} образуют *главную диагональ*, а элементы a_{21} , a_{12} — *побочную*.

Из равенства (1) видно, что для вычисления определителя второго порядка нужно из произведения элементов, стоящих на главной диагонали, вычесть произведение элементов, стоящих на побочной диагонали.

Пример 1. Вычислить определитель второго порядка:

$$\begin{vmatrix} 2 & 4 \\ 3 & 7 \end{vmatrix} = 14 - 12 = 2.$$

Пример 2. Имеем:

$$|E| = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1, \text{ т. е. определитель единичной матрицы равен единице.}$$

Легко проверяются следующие свойства определителя (с помощью правила вычисления его по формуле (1)).

Величина определителя $|A|$:

1) не меняется, если заменить его строки соответствующими столбцами;

2) не меняется, если к элементам какой-либо его строки или столбца прибавить соответствующие элементы другой строки или столбца, умноженные на одно и то же число;

3) меняет знак, если поменять местами его строки или столбцы;

4) увеличивается в k раз, если элементы какого-либо его столбца или строки увеличить в k раз, т. е. общий множитель, имеющийся в строке или столбце, можно выносить за знак определителя;

5) равна нулю, если элементы какого-либо его столбца или строки равны нулю,

6) равна нулю, если элементы двух строк или столбцов соответственно равны.

2. Определители третьего порядка. Рассмотрим квадратную матрицу третьего порядка

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

О п р е д е л е н и е 1. *Определителем третьего порядка*, соответствующим матрице A , называется число, равное $a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$ и обозначаемое символом

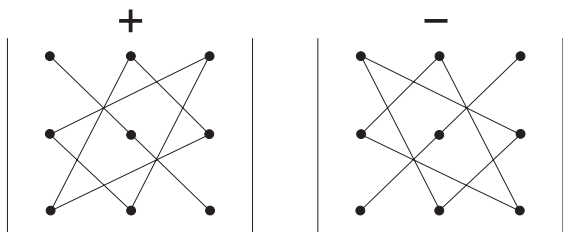
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \text{ (кратко } |A| \text{)}.$$

Итак,

$$|A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \quad (2)$$

Чтобы запомнить, какие произведения в правой части равенства (2) следует брать со знаком «плюс», какие — со знаком «минус», полезно следующее правило, называемое *правилом треугольника*.



П р и м е р 1. По формуле (2) имеем $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{vmatrix} = 15 + 24 + 24 - 27 - 20 - 16 = 0$.

П р и м е р 2. Очевидно что

$$|E| = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Все свойства определителей второго порядка (свойства 1—6) остаются справедливыми и для определителей третьего порядка (проверка их идет по формуле (2)).

О п р е д е л е н и е 2. *Минором* какого либо элемента определителя называется определитель, полученный из данного вычерчиванием той строки и того столбца, которым принадлежит этот элемент.

Например, минором элемента a_{12} определителя $|A|$ является определитель второго порядка

$$\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}. \quad (3)$$

Минор элемента a_{ik} определителя $|A|$ обозначается через M_{ik} .

О п р е д е л е н и е 3. *Алгебраическим дополнением* элемента a_{ik} определителя $|A|$ называется его минор, взятый со знаком $(-1)^{i+k}$.

Например, алгебраическим дополнением элемента a_{12} определителя $|A|$ является определитель (3), взятый со знаком «минус» Алгебраическое дополнение элемента a_{ik} будем обозначать через A_{ik} . Следовательно, $A_{ik} = (-1)^{i+k} M_{ik}$.

Т е о р е м а 1. *Определитель равен сумме произведений элементов какой-либо строки (столбца) на их алгебраические дополнения.*

Д о к а з а т е л ь с т в о. Преобразуем правую часть формулы (2). Так как

$$\begin{aligned} & a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} = \\ & = a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) = \\ & = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}, \end{aligned}$$

то

$$|A| = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}. \quad (4)$$

Формула (4) называется *разложением определителя $|A|$ по элементам первой строки*. Аналогично получается разложение по элементам других строк и столбцов.

Т е о р е м а 2 (теорема аннулирования). *Сумма произведений элементов какой-либо строки (столбца) определителя на алгебраические дополнения соответствующих элементов другой строки (столбца) равна нулю.*

Для определителя $|A|$ покажем, например, что

$$a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23} = 0. \quad (5)$$

Раскладывая определитель

$$|\tilde{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

по элементам второй строки, согласно предыдущей теореме имеем

$$|\tilde{A}| = a_{11}A_{21} + a_{12}A_{22} + a_{13}A_{23}.$$

Так как определитель $|\tilde{A}|$ равен нулю (как содержащий две одинаковые строки), то получаем искомое равенство (5).

3. Понятие определителя n -го порядка. Свойство определителя третьего порядка, выраженное теоремой 1 (п. 2), допускает обобщение, которое может быть принято за определение определителя любого порядка.

В общем случае *определителем n -го порядка*, соответствующим квадратной матрице n -го порядка

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad (6)$$

можно назвать число, равное сумме парных произведений элементов какой-либо строки (столбца) на их алгебраические дополнения (его краткое обозначение $|A|$).

Заметим, что определители любого порядка n обладают всеми полученными выше свойствами (п. 1, 2).

Из свойства 1 (п. 1) определителя следует, что квадратная матрица A и транспонированная к ней матрица A' имеют равные определители, т. е.

$$|A| = |A'|.$$

Пример 1. Вычислить заданный определитель Δ .

$$\Delta = \begin{vmatrix} 3 & 0 & 2 & 0 \\ 2 & 3 & -1 & 4 \\ 0 & 4 & -2 & 3 \\ 5 & 2 & 0 & 1 \end{vmatrix} = 3 \begin{vmatrix} -1 & 4 \\ -2 & 3 \end{vmatrix} + 2 \begin{vmatrix} 2 & 3 \\ 0 & 4 \end{vmatrix} = 3 \cdot 8 - 2 \cdot 39 = -54.$$

Заметим, что если в определителе все элементы какой-либо строки (столбца), кроме одного, равны нулю, то при вычислении определителя выгодно разложить его по элементам этой строки (столбца).

Если же такой строки (столбца) нет, то, используя свойство 2 (п. 1) определителя, его можно преобразовать так, чтобы он имел такую строку (столбец).

Пример 2. Очевидно, что

$$|E| = \begin{matrix} n \text{ строк} \\ \left\{ \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \end{vmatrix} \right. \\ n \text{ столбцов} \end{matrix} = 1.$$

Т е о р е м а. Если A и B — квадратные матрицы одного порядка с определителями $|A|$ и $|B|$, то определитель матрицы $C = AB$ равен произведению определителей перемножаемых матриц, т. е.

$$|C| = |A||B|.$$

Доказательство этого свойства ради краткости проведем для случая матриц второго порядка. Пусть

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Тогда

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix},$$

$$\begin{aligned} |AB| &= (a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{21}b_{11} + a_{22}b_{21})(a_{11}b_{12} + a_{12}b_{22}) = \\ &= a_{11}b_{11}a_{21}b_{12} + a_{12}b_{21}a_{21}b_{12} + a_{11}b_{11}a_{22}b_{22} + a_{12}b_{21}a_{22}b_{22} - a_{21}b_{11}a_{11}b_{12} - \\ &- a_{22}b_{21}a_{11}b_{12} - a_{21}b_{11}a_{12}b_{22} - a_{22}b_{21}a_{12}b_{22} = a_{11}a_{22}b_{11}b_{22} + a_{12}a_{21}b_{12}b_{21} - \\ &- a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22}. \end{aligned} \quad (7)$$

С другой стороны,

$$\begin{aligned} |A||B| &= (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21}) = a_{11}a_{22}b_{11}b_{22} + \\ &+ a_{12}a_{21}b_{12}b_{21} - a_{11}a_{22}b_{12}b_{21} - a_{12}a_{21}b_{11}b_{22}. \end{aligned} \quad (8)$$

Из соотношений (7) и (8) получаем:

$$|AB| = |A||B|.$$

П р и м е р 3. Пусть

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix}.$$

Найти $|AB|$.

Имеем:

$$|A| = 6 - 1 = 5, \quad |B| = 1 - 3 = -2.$$

Согласно только что установленной теореме

$$|AB| = 5 \cdot (-2) = -10.$$

Отметим еще следующий любопытный факт. Как известно, произведение двух отличных от нуля чисел не равно нулю. Для матриц

подобное обстоятельство может и не иметь места, т. е. произведение двух ненулевых матриц может оказаться равным нуль-матрице.

Пример 4. Если

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ и } B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix},$$

то

$$AB = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 1 \cdot (-1) & 1 \cdot 1 + 1 \cdot (-1) \\ 1 \cdot 1 + 1 \cdot (-1) & 1 \cdot 1 + 1 \cdot (-1) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

4. Обратная матрица. Рассмотрим теперь так называемую *обратную* матрицу, понятие которой вводится только для квадратной матрицы.

Если A — квадратная матрица, то обратной для нее матрицей называется матрица, обозначаемая A^{-1} и удовлетворяющая условиям

$$AA^{-1} = E, \quad A^{-1}A = E,$$

где E — единичная матрица.

Примечание. Из этого определения следует, что если матрица A^{-1} является обратной для A , то и A будет обратной для A^{-1} .

Определение. Если определитель $|A|$ матрицы (6), обозначенной через A (п. 3), равен нулю, то матрица A называется *вырожденной*, в противном случае матрица A называется *невырожденной*.

Теорема. Матрица

$$\begin{pmatrix} \frac{A_{11}}{|A|} & \frac{A_{21}}{|A|} & \cdots & \frac{A_{n1}}{|A|} \\ \frac{A_{12}}{|A|} & \frac{A_{22}}{|A|} & \cdots & \frac{A_{n2}}{|A|} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{A_{1n}}{|A|} & \frac{A_{2n}}{|A|} & \cdots & \frac{A_{nn}}{|A|} \end{pmatrix}, \quad (9)$$

где A_{ik} — алгебраическое дополнение элемента a_{ik} невырожденной матрицы A , является обратной для A .

Доказательство. Ради кратности доказательства ограничимся случаем $n = 2$. Умножая матрицу A на матрицу (9), мы получим с использованием известных свойств

$$\begin{pmatrix} a_{11} \frac{A_{11}}{|A|} + a_{12} \frac{A_{21}}{|A|} & a_{11} \frac{A_{21}}{|A|} + a_{12} \frac{A_{22}}{|A|} \\ a_{21} \frac{A_{11}}{|A|} + a_{22} \frac{A_{21}}{|A|} & a_{21} \frac{A_{21}}{|A|} + a_{22} \frac{A_{22}}{|A|} \end{pmatrix} = \begin{pmatrix} \frac{|A|}{|A|} & 0 \\ 0 & \frac{|A|}{|A|} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E.$$

Аналогично проводится доказательство и для того случая, когда матрица (9) является первым множителем, а A — вторым.

Из только что установленной теоремы следует, что для того чтобы построить обратную матрицу для квадратной невырожденной матрицы A , нужно сначала построить транспонированную матрицу A' , а затем каждый элемент A' заменить его алгебраическим дополнением, деленным на $|A|$.

П р и м е р. Найти матрицу, обратную матрице

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Определитель этой матрицы

$$|A| = \begin{vmatrix} 1 & 2 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} = -9.$$

Так как $|A| \neq 0$, то матрица A невырожденная, и, следовательно, существует обратная ей матрица. Вычисляем алгебраические дополнения:

$$A_{11} = (-1)^{1+1} \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3. \text{ Аналогично } A_{12} = -6, A_{13} = 3, A_{21} = -4, A_{22} = 2, A_{23} = -1, A_{31} = 2,$$

$$A_{32} = -1, A_{33} = -4.$$

Составим матрицу

$$C = \begin{pmatrix} -\frac{3}{9} & \frac{6}{9} & -\frac{3}{9} \\ \frac{4}{9} & -\frac{2}{9} & \frac{1}{9} \\ -\frac{2}{9} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}.$$

Сделав в этой матрице ее строки столбцами с тем же номером, получим матрицу

$$A^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{4}{9} & -\frac{2}{9} \\ \frac{2}{3} & -\frac{2}{9} & \frac{1}{9} \\ -\frac{1}{3} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}.$$

5. Преобразование координат, матрица преобразования, ортогональные матрицы. В некоторых случаях приходится одновременно рассматривать две системы координат на плоскости и решать следующую задачу: зная координаты точки в одной системе координат, найти ее

координаты в другой системе. Формулы, выражающие координаты точки в одной системе через ее координаты в другой системе, называются *формулами преобразования координат*.

В главе 1 (§ 1.1, п. 2) были получены формулы преобразования декартовых и полярных координат. Пусть обе системы — декартовы прямоугольные, причем одноименные оси этих систем параллельны и одинаково направлены и на каждой из осей выбрана одна и та же масштабная единица. На рисунке 41 изображены две такие системы xOy и $x'O'y'$. Система $x'O'y'$ может быть получена параллельным переносом осей Ox и Oy . Условимся называть координаты точек в системе xOy *старыми*, а в системе $x'O'y'$ *новыми*. Пусть x_0 и y_0 — координаты нового начала O' в старой системе. Предположим, что точка M на плоскости (рис. 41) имеет старые координаты x и y и новые x' и y' . Из рисунка 41 получаем $x = x' + x_0$.

Аналогично $y = y' + y_0$.

Таким образом, имеем:

$$x = x' + x_0, y = y' + y_0. \quad (10)$$

(Эти формулы верны и при любом другом положении точки M на плоскости.)

Формулы (10) называются *формулами параллельного переноса осей*.

Рассмотрим теперь в плоскости π прямоугольную систему координат x_1Ox_2 с ортами $\bar{e}_1$ и $\bar{e}_2$ (рис. 42). Наряду с системой координат x_1Ox_2 , которую будем называть *старой*, рассмотрим *новую* систему координат $x'_1Ox'_2$ с ортами $\bar{e}'_1$ и $\bar{e}'_2$. Начала координат старой и новой систем совпадают.

Возьмем в плоскости π произвольную точку M . Пусть x_1, x_2 — ее координаты в старой системе и x'_1, x'_2 — в новой. Найдём связь между старыми и новыми координатами. Имеем (см. § 2.1, п. 8):

$$\overline{OM} = x_1\bar{e}_1 + x_2\bar{e}_2, \quad \overline{OM} = x'_1\bar{e}'_1 + x'_2\bar{e}'_2.$$

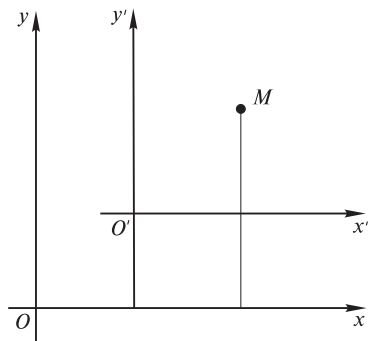


Рис. 41

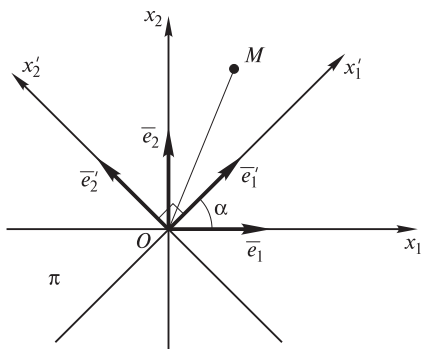


Рис. 42

Таким образом,

$$x_1 \bar{e}_1 + x_2 \bar{e}_2 = x'_1 \bar{e}'_1 + x'_2 \bar{e}'_2. \quad (11)$$

Умножим обе части равенства (11) скалярно на $\bar{e}_1$. Принимая во внимание, что $\bar{e}_1 \bar{e}_1 = 1$ и $\bar{e}_1 \bar{e}_2 = 0$ (§ 2.2, п. 2), получим:

$$x_1 = x'_1 (\bar{e}_1 \bar{e}'_1) + x'_2 (\bar{e}_1 \bar{e}'_2). \quad (12)$$

Умножая обе части равенства (11) скалярно на $\bar{e}_2$, аналогично получим:

$$x_2 = x'_1 (\bar{e}_2 \bar{e}'_1) + x'_2 (\bar{e}_2 \bar{e}'_2). \quad (13)$$

Введем обозначения:

$$\begin{aligned} \alpha_{11} &= \bar{e}_1 \bar{e}'_1 = |\bar{e}_1| |\bar{e}'_1| \cos(\widehat{\bar{e}_1, \bar{e}'_1}) = \cos(\widehat{\bar{e}_1, \bar{e}'_1}), \\ \alpha_{12} &= \bar{e}_1 \bar{e}'_2 = \cos(\widehat{\bar{e}_1, \bar{e}'_2}), \\ \alpha_{21} &= \bar{e}_2 \bar{e}'_1 = \cos(\widehat{\bar{e}_2, \bar{e}'_1}), \\ \alpha_{22} &= \bar{e}_2 \bar{e}'_2 = \cos(\widehat{\bar{e}_2, \bar{e}'_2}). \end{aligned} \quad (14)$$

Тогда равенства (12) и (13) можно записать в виде

$$x_1 = \alpha_{11} x'_1 + \alpha_{12} x'_2, \quad x_2 = \alpha_{21} x'_1 + \alpha_{22} x'_2. \quad (15)$$

Формулы (15) называются *формулами преобразования координат на плоскости*, а матрица

$$L = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} -$$

матрицей преобразования.

Рассмотрим матрицы-столбцы

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{и} \quad X' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}.$$

С их помощью преобразование координат (15) в матричной форме запишем в виде

$$X = LX'.$$

Установим некоторые свойства матрицы L .

Прежде всего найдем разложение векторов $\bar{e}_1$ и $\bar{e}_2$ по новому базису $\bar{e}'_1, \bar{e}'_2$. Так как

$$\text{пр}_{\bar{e}'_1} \bar{e}_1 = \cos(\widehat{\bar{e}_1, \bar{e}'_1}) = \alpha_{11}, \quad \text{пр}_{\bar{e}'_2} \bar{e}_1 = \alpha_{12}, \quad \text{пр}_{\bar{e}'_1} \bar{e}_2 = \alpha_{21}, \quad \text{пр}_{\bar{e}'_2} \bar{e}_2 = \alpha_{22},$$

то

$$\begin{aligned} \bar{e}_1 &= \alpha_{11} \bar{e}'_1 + \alpha_{12} \bar{e}'_2, \\ \bar{e}_2 &= \alpha_{21} \bar{e}'_1 + \alpha_{22} \bar{e}'_2. \end{aligned} \quad (16)$$

Формулы (16) дают разложение векторов $\bar{e}_1, \bar{e}_2$ по базису $\bar{e}'_1, \bar{e}'_2$.

Аналогично получим разложение ортов $\bar{e}'_1$ и $\bar{e}'_2$ по базису $\bar{e}_1, \bar{e}_2$:

$$\begin{aligned} \bar{e}'_1 &= \alpha_{11} \bar{e}_1 + \alpha_{21} \bar{e}_2, \\ \bar{e}'_2 &= \alpha_{12} \bar{e}_1 + \alpha_{22} \bar{e}_2. \end{aligned} \quad (17)$$

Так как $\bar{e}_1 \bar{e}_1 = 1$, $\bar{e}_1 \bar{e}_2 = 0$, $\bar{e}_2 \bar{e}_2 = 1$ (то же и для векторов $\bar{e}'_1, \bar{e}'_2$) (§2.2. п. 2) то, принимая во внимание формулы (16), получим:

$$\bar{e}_1 \bar{e}_1 = \alpha_{11}^2 + \alpha_{12}^2 = 1, \quad \bar{e}_1 \bar{e}_2 = \alpha_{11} \alpha_{21} + \alpha_{12} \alpha_{22} = 0, \quad \bar{e}_2 \bar{e}_2 = \alpha_{21}^2 + \alpha_{22}^2 = 1.$$

Итак,

$$\alpha_{11}^2 + \alpha_{12}^2 = 1, \quad \alpha_{11} \alpha_{21} + \alpha_{12} \alpha_{22} = 0, \quad \alpha_{21}^2 + \alpha_{22}^2 = 1, \quad (18)$$

Аналогично из формул (17) получим

$$\alpha_{11}^2 + \alpha_{21}^2 = 1, \quad \alpha_{11} \alpha_{12} + \alpha_{21} \alpha_{22} = 0, \quad \alpha_{12}^2 + \alpha_{22}^2 = 1. \quad (19)$$

Иными словами, матрица L обладает следующими свойствами:

а) Сумма квадратов элементов строки (или столбца) равна единице.

б) Сумма парных произведений элементов строки (столбца) на соответствующие элементы другой строки (столбца) равна нулю.

Матрица, обладающая этими свойствами, называется *ортogonalной*.

Рассмотрим транспонированную к матрице L матрицу

$$L^* = \begin{pmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{pmatrix}.$$

С учетом равенств (18) имеем:

$$\begin{aligned} LL^* &= \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{pmatrix} = \\ &= \begin{pmatrix} \alpha_{11}^2 + \alpha_{12}^2 & \alpha_{11} \alpha_{21} + \alpha_{12} \alpha_{22} \\ \alpha_{21} \alpha_{11} + \alpha_{22} \alpha_{12} & \alpha_{21}^2 + \alpha_{22}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E. \end{aligned}$$

Аналогично, используя равенства (19), получим:

$$L^* L = E$$

Таким образом, матрица L^* является обратной для матрицы L , т. е.

$$L^{-1} = L^* = \begin{pmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{pmatrix}. \quad (20)$$

Пусть новая система координат получена из старой системы поворотом осей на угол α . В этом случае (рис. 42)

$$\alpha_{11} = \cos(\widehat{\vec{e}'_1, \vec{e}'_1}) = \cos \alpha, \quad \alpha_{12} = \cos(\widehat{\vec{e}'_1, \vec{e}'_2}) = \cos\left(\frac{\pi}{2} + \alpha\right) = -\sin \alpha,$$

$$\alpha_{21} = \cos(\widehat{\vec{e}'_2, \vec{e}'_1}) = \cos\left(\frac{\pi}{2} - \alpha\right) = \sin \alpha, \quad \alpha_{22} = \cos(\widehat{\vec{e}'_2, \vec{e}'_2}) = \cos \alpha,$$

и, следовательно, формулы (15) принимают вид:

$$\begin{aligned} x_1 &= x'_1 \cos \alpha - x'_2 \sin \alpha, \\ x_2 &= x'_1 \sin \alpha + x'_2 \cos \alpha. \end{aligned} \quad (21)$$

Формулы (21) называются *формулами поворота осей*.

6. Ранг матрицы. Рассмотрим матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Выделим в ней какие-либо k строк и k столбцов, где $k \leq m, n$. Элементы, стоящие на пересечении выделенных строк и столбцов, образуют квадратную матрицу, которая порождает определитель k -го порядка. Полученный определитель называется определителем k -го порядка, порожденным матрицей A .

О п р е д е л е н и е. Рангом матрицы A (обозначение $r(A)$) называется наибольшее натуральное число k , для которого существует не равный нулю определитель k -го порядка, порожденный матрицей A .

П р и м е р 1. Матрица

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 3 & 2 \end{pmatrix}$$

имеет ранг 3, так как определитель, порожденный этой матрицей,

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \end{vmatrix} = 1 \neq 0.$$

П р и м е р 2. Ранг матрицы

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

равен 1 (определителями первого порядка служат сами элементы матрицы), а ранг матрицы

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

равен 2.

Очевидно, ранг квадратной матрицы не превосходит ее порядка; равен же ее порядку в том и только в том случае, если эта матрица невырожденная.

При определении ранга матрицы, как правило, приходится вычислять большое число определителей. Чтобы облегчить этот процесс, используют так называемые элементарные преобразования матрицы:

- 1) перемена местами строк (столбцов) матрицы;
- 2) умножение всех элементов какой-либо строки (столбца) матрицы на одно и то же число, отличное от нуля;
- 3) прибавление к элементам какой-либо строки (столбца) матрицы соответствующих элементов другой строки (столбца), умноженных на одно и то же число;
- 4) отбрасывание строк (столбцов) матрицы, все элементы которых равны нулю.

Как можно доказать (см., например, [10]), ранг матрицы при элементарных преобразованиях не меняется.

Не изменится также ранг матрицы, если ее транспонировать.

П р и м е р 3. Найти ранг матрицы $\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$. Умножая первую строку на (-1) и прибавляя ее к третьей строке, получаем $\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 \end{pmatrix}$. Теперь к третьей строке полученной матрицы прибавляем вторую. Будем иметь $\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Отбрасывая здесь строку, состоящую из нулей, получаем матрицу $\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$, ранг которой, очевидно, равен двум. Следовательно, ранг данной матрицы A также равен двум.

§ 3.3. Выражение векторного и смешанного произведений векторов через координаты сомножителей

1. Выражение векторного произведения через координаты перемножаемых векторов. Пусть

$$\vec{a} = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}, \quad \vec{b} = b_x \vec{i} + b_y \vec{j} + b_z \vec{k}.$$

Перемножая векторно эти равенства и используя свойства векторного произведения, получим сумму девяти слагаемых

$$\begin{aligned} \vec{a} \times \vec{b} = & a_x b_x (\vec{i} \times \vec{i}) + a_y b_x (\vec{j} \times \vec{i}) + a_z b_x (\vec{k} \times \vec{i}) + a_x b_y (\vec{i} \times \vec{j}) + \\ & + a_y b_y (\vec{j} \times \vec{j}) + a_z b_y (\vec{k} \times \vec{j}) + a_x b_z (\vec{i} \times \vec{k}) + a_y b_z (\vec{j} \times \vec{k}) + a_z b_z (\vec{k} \times \vec{k}). \end{aligned} \quad (1)$$

Из свойств и определения векторного произведения (§ 2.2, п. 4) следует, что $\vec{i} \times \vec{i} = 0$, $\vec{j} \times \vec{j} = 0$, $\vec{k} \times \vec{k} = 0$ и $\vec{i} \times \vec{j} = -(\vec{j} \times \vec{i}) = \vec{k}$, $\vec{j} \times \vec{k} = -(\vec{k} \times \vec{j}) = \vec{i}$, $\vec{k} \times \vec{i} = -(\vec{i} \times \vec{k}) = \vec{j}$.

Поэтому из равенства (1) получаем:

$$\vec{a} \times \vec{b} = (a_y b_z - a_z b_y) \vec{i} + (a_z b_x - a_x b_z) \vec{j} + (a_x b_y - a_y b_x) \vec{k}$$

или

$$\vec{a} \times \vec{b} = \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} \vec{i} - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} \vec{j} + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \vec{k}. \quad (2)$$

Для удобства запоминания формула (2) записывается в виде определителя третьего порядка (см. § 3.2, (4)):

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}.$$

2. Выражение смешанного произведения через координаты перемножаемых векторов. Пусть

$$\vec{a} = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}, \quad \vec{b} = b_x \vec{i} + b_y \vec{j} + b_z \vec{k}, \quad \vec{c} = c_x \vec{i} + c_y \vec{j} + c_z \vec{k}.$$

Как уже установлено,

$$\vec{a} \times \vec{b} = \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} \vec{i} - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} \vec{j} + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} \vec{k}.$$

Далее по известной формуле (§ 2.2, (7)) для скалярного произведения получаем:

$$(\vec{a} \times \vec{b}) \vec{c} = \begin{vmatrix} a_y & a_z \\ b_y & b_z \end{vmatrix} c_x - \begin{vmatrix} a_x & a_z \\ b_x & b_z \end{vmatrix} c_y + \begin{vmatrix} a_x & a_y \\ b_x & b_y \end{vmatrix} c_z,$$

или

$$(\bar{a} \times \bar{b})\bar{c} = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}.$$

Таким образом, *смешанное произведение равно определителю третьего порядка, в строках которого стоят соответствующие координаты перемножаемых векторов.*

§ 3.4. Системы линейных уравнений

1. Система двух уравнений с двумя неизвестными. Рассмотрим систему двух уравнений первой степени с двумя неизвестными x и y :

$$\begin{cases} a_{11}x + a_{12}y = b_1, \\ a_{21}x + a_{22}y = b_2. \end{cases} \quad (1)$$

В обозначении коэффициента a_{ik} первый индекс обозначает номер уравнения, а второй индекс — номер неизвестного.

Для решения этой системы почленно умножим первое уравнение на a_{22} , а второе на $-a_{12}$ и сложим полученные равенства:

$$(a_{11}a_{22} - a_{21}a_{12})x = b_1a_{22} - b_2a_{12}$$

или

$$\Delta \cdot x = \Delta^{(1)}, \quad (2)$$

$$\text{где } \Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \Delta^{(1)} = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} \quad (\Delta - \text{определитель системы}).$$

Аналогичным образом имеем

$$\Delta \cdot y = \Delta^{(2)}, \quad (3)$$

$$\text{где } \Delta^{(2)} = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

Определитель $\Delta^{(1)}$ ($\Delta^{(2)}$) получается из определителя Δ заменой коэффициентов a_{11} и a_{21} (или a_{12} и a_{22}) при неизвестных x (или y) свободными членами b_1 и b_2 .

Если определитель системы $\Delta \neq 0$, то из формул (2) и (3) получаем единственное решение системы

$$x = \frac{\Delta^{(1)}}{\Delta}, \quad y = \frac{\Delta^{(2)}}{\Delta}. \quad (4)$$

Формулы (4) называются *формулами Крамера* (Г. Крамер (1704—1752) — швейцарский математик).

В случае, когда $\Delta = 0$, система (1) может либо вовсе не иметь решений (быть *несовместной*), как, например, в случае системы

$$\begin{cases} x + y = 0, \\ x + y = 1. \end{cases}$$

(уравнения этой системы противоречат друг другу), либо иметь бесконечное множество различных решений, когда отдельные уравнения системы не противоречат друг другу. Так, система

$$\begin{cases} x + y = 3, \\ 2x + 2y = 6. \end{cases}$$

имеет бесчисленное множество решений. Задавая, например, произвольное значение y , получим соответствующее значение $x = 3 - y$.

2. Система трех линейных уравнений с тремя неизвестными. Рассмотрим систему трех линейных уравнений с тремя неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \end{cases} \quad (5)$$

Справедлива следующая теорема.

Т е о р е м а. Система уравнений (5) имеет единственное решение тогда и только тогда, когда определитель системы

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

отличен от нуля.

В этом случае решение находят по правилу Крамера.

$$x_i = \frac{\Delta^{(i)}}{\Delta} \quad (i = 1, 2, 3), \quad (6)$$

где $\Delta^{(i)}$ — определитель, полученный из определителя Δ заменой его i столбца столбцом свободных членов.

П р и м е р. Система

$$\begin{cases} x - 2y + 3z = 2, \\ 4x - y + 5z = 15, \\ 6x - 8y + 7z = 9. \end{cases}$$

имеет определитель $\Delta = -49 \neq 0$ и поэтому имеет единственное решение, которое можно найти по формулам (6). Поскольку $\Delta^{(1)} = -147$, $\Delta^{(2)} = -98$, $\Delta^{(3)} = -49$, то $x = 3$, $y = 2$, $z = 1$.

Обозначим матрицу-столбец из неизвестных через X и матрицу-столбец из свободных членов через B :

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

$$AX = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}.$$
$$AX=B. \quad (8)$$
$$A^{-1}(AX) = A^{-1}B.$$
$$(A^{-1}A)X = A^{-1}B.$$
$$X=A^{-1}B. \quad (9)$$
$$\begin{cases} x_1 + 2x_2 &= 10, \\ 3x_1 + 2x_2 + x_3 &= 23, \\ x_2 + 2x_3 &= 13. \end{cases}$$
$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad B = \begin{pmatrix} 10 \\ 23 \\ 13 \end{pmatrix}.$$

65

Матрица A называется *матрицей системы*, матрица B — *расширенной матрицей* и отличается от матрицы системы столбцом, состоящим из свободных членов уравнений системы. При решении системы (10) методом Гаусса элементарные преобразования системы заменяются соответствующими элементарными преобразованиями, выполняемыми над ее расширенной матрицей B .

П р и м е р 1. Решить систему уравнений
$$\begin{cases} 2x_1 + x_2 - x_3 = 1, \\ 3x_1 + 2x_2 - 2x_3 = 1, \\ x_1 - x_2 + 2x_3 = 5. \end{cases}$$

Составим расширенную матрицу и выполним над ней элементарные преобразования, указанные в методе Гаусса:

$$\begin{pmatrix} 2 & 1 & -1 & 1 \\ 3 & 2 & -2 & 1 \\ 1 & -1 & 2 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0,5 & -0,5 & 0,5 \\ 3 & 2 & -2 & 1 \\ 1 & -1 & 2 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0,5 & -0,5 & 0,5 \\ 0 & 0,5 & -0,5 & -0,5 \\ 0 & -1,5 & 2,5 & 4,5 \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} 1 & 0,5 & -0,5 & 0,5 \\ 0 & 1 & -1 & -1 \\ 0 & -1,5 & 2,5 & 4,5 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0,5 & -0,5 & 0,5 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

Здесь последовательно произвели следующие преобразования: 1) умножили первую строку на $\frac{1}{2}$; 2) вычли из третьей строки первую и из второй первую, умноженную на 3;

3) умножили третью строку на 2; 4) вычли из третьей строки вторую, умноженную на $-1,5$.

Последней матрице соответствует треугольная система уравнений

$$\begin{cases} x_1 + 0,5x_2 - 0,5x_3 = 0,5, \\ \quad \quad x_2 - x_3 = -1, \\ \quad \quad \quad x_3 = 3. \end{cases} \quad (17)$$

равносильная данной. Из системы (17) последовательно находим:

$$x_3 = 3, x_2 = -1 + x_3 = -1 + 3 = 2, x_1 = 0,5 - 0,5x_2 + 0,5x_3 = 0,5 - 1 + 1,5 = 1.$$

П р и м е р 2. Решить систему уравнений
$$\begin{cases} x_1 + 2x_2 + 4x_3 - x_4 - 3x_5 = 7, \\ 2x_1 \quad \quad + x_3 \quad \quad + x_5 = 4, \\ \quad \quad x_2 \quad \quad + 2x_4 - x_5 = 6. \end{cases}$$

Имеем:

$$\begin{pmatrix} 1 & 2 & 4 & -1 & -3 & 7 \\ 2 & 0 & 1 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 & -1 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 4 & -1 & -3 & 7 \\ 0 & -4 & -7 & 2 & 7 & -10 \\ 0 & 1 & 0 & 2 & -1 & 6 \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 4 & -1 & -3 & 7 \\ 0 & 1 & 0 & 2 & -1 & 6 \\ 0 & -4 & -7 & 2 & 7 & -10 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 4 & -1 & -3 & 7 \\ 0 & 1 & 0 & 2 & -1 & 6 \\ 0 & 0 & -7 & 10 & 3 & 14 \end{pmatrix}$$

(умножили первую строку на 2 и вычли из второй, переставили вторую и третью строки, к третьей строке прибавили вторую, умноженную на 4).

Последней матрице соответствует ступенчатая система уравнений

$$\begin{cases} x_1 + 2x_2 + 4x_3 - x_4 - 3x_5 = 7, \\ x_2 + 2x_4 - x_5 = 6, \\ -7x_3 + 10x_4 + 3x_5 = 14 \end{cases}$$

равносильная данной. Перепишем полученную систему в виде

$$\begin{cases} x_1 + 2x_2 + 4x_3 = 7 - x_4 + 3x_5, \\ x_2 = 6 - 2x_4 + x_5, \\ -7x_3 = 14 - 10x_4 - 3x_5, \end{cases}$$

откуда

$$\begin{cases} x_3 = -2 + \frac{10}{7}x_4 + \frac{3}{7}x_5, \\ x_2 = 6 - 2x_4 + x_5, \\ x_1 = 3 - \frac{5}{7}x_4 - \frac{5}{7}x_5. \end{cases}$$

Придавая x_4 и x_5 произвольные значения, получим соответствующие значения x_1 , x_2 и x_3 .
Неизвестные x_4 и x_5 в этом случае являются свободными неизвестными.

Итак, данная система совместная и неопределенная. Ее решениями, например, будут $x_5 = 0$, $x_4 = 0$, $x_3 = -2$, $x_2 = 6$, $x_1 = 3$ или $x_5 = 7$, $x_4 = 7$, $x_3 = 11$, $x_2 = -1$, $x_1 = -7$ и т. д.

П р и м е р 3. Решить систему уравнений

$$\begin{cases} 2x_1 + x_2 + x_3 - x_4 = 8, \\ 3x_1 + 2x_3 + 5x_4 = 12, \\ x_1 - x_2 + x_3 = 4, \\ 8x_1 + x_2 + 5x_3 + 3x_4 = 10. \end{cases}$$

Имеем

$$\begin{pmatrix} 2 & 1 & 1 & -1 & 8 \\ 3 & 0 & 2 & 5 & 12 \\ 1 & -1 & 1 & 0 & 4 \\ 8 & 1 & 5 & 3 & 10 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} & 4 \\ 3 & 0 & 2 & 5 & 12 \\ 1 & -1 & 1 & 0 & 4 \\ 8 & 1 & 5 & 3 & 10 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} & 4 \\ 0 & -\frac{3}{2} & \frac{1}{2} & \frac{13}{2} & 0 \\ 0 & -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -3 & 1 & 7 & -22 \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} & 4 \\ 0 & 1 & -\frac{1}{3} & -\frac{13}{3} & 0 \\ 0 & -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -3 & 1 & 7 & -22 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} & 4 \\ 0 & 1 & -\frac{1}{3} & -\frac{13}{3} & 0 \\ 0 & 0 & 0 & -6 & 0 \\ 0 & 0 & 0 & -6 & -22 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} & 4 \\ 0 & 1 & -\frac{1}{3} & -\frac{13}{3} & 0 \\ 0 & 0 & 0 & -6 & 0 \\ 0 & 0 & 0 & 0 & -22 \end{pmatrix},$$

откуда

$$\begin{cases} x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 - \frac{1}{2}x_4 = 4, \\ x_2 - \frac{1}{3}x_3 - \frac{13}{3}x_4 = 0, \\ -6x_4 = 0, \\ 0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = -22. \end{cases}$$

Так как последнее уравнение этой системы противоречиво, то она является несовместной. Следовательно, несовместна и равносильная ей исходная система.

7. Критерий совместности системы. При решении системы линейных уравнений методом Гаусса ответ на вопрос, совместна или несовместна данная система, может быть дан лишь в конце вычислений (см. п. 6). Однако часто бывает важно решить вопрос о совместности или несовместности системы вида (10), не находя самих решений. Ответ на этот вопрос дает теорема Кронекера — Капелли. (Л. Кронекер (1823—1891) — немецкий математик; А. Капелли (1855—1910) — итальянский математик.)

Т е о р е м а К р о н е к е р а — К а п е л л и. *Для того чтобы система линейных уравнений (10) была совместной, необходимо и достаточно, чтобы ранг матрицы системы (10) был равен рангу ее расширенной матрицы, т. е. чтобы $r(A) = r(B)$. При этом, если $r(B) = r(A) = n$, то система (10) имеет единственное решение, а если $r(B) = r(A) < n$, то — бесконечное множество решений.*

П р и м е р ы. 1. Как следует из решения системы (п. 4)

$$\begin{cases} x + y = 1, \\ x - y = 2 \end{cases}$$

$r(A) = r(B) = 2$. Следовательно, она имеет единственное решение.

2. В случае системы

$$\begin{cases} x + y = 0, \\ x + y = 1 \end{cases}$$

$r(A) = 1$, $r(B) = 2$, т. е. $r(B) > r(A)$ и, значит, эта система несовместна (см. также п. 1).

3. Для системы

$$\begin{cases} x + y = 3, \\ 2x + 2y = 6 \end{cases}$$

$r(A) = r(B) = 1 < 2$ и, значит, эта система имеет бесчисленное множество решений (см. также п. 1).

Отсюда

$$A^{-1} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -2 & 1 \end{pmatrix}.$$

Пример 2. Найти матрицу A^{-1} , обратную матрице $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

Имеем:

$$\begin{pmatrix} 1 & 2 & 0 & | & 1 & 0 & 0 \\ 3 & 2 & 1 & | & 0 & 1 & 0 \\ 0 & 1 & 2 & | & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & | & 1 & 0 & 0 \\ 0 & -4 & 1 & | & -3 & 1 & 0 \\ 0 & 1 & 2 & | & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & | & 1 & 0 & 0 \\ 0 & -4 & 1 & | & -3 & 1 & 0 \\ 0 & \frac{1}{2} & 1 & | & 0 & 0 & \frac{1}{2} \end{pmatrix} \rightarrow$$

$$\rightarrow \begin{pmatrix} 1 & 2 & 0 & | & 1 & 0 & 0 \\ 0 & -\frac{9}{2} & 0 & | & -3 & 1 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 1 & | & 0 & 0 & \frac{1}{2} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & | & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & | & \frac{2}{3} & -\frac{2}{9} & \frac{1}{9} \\ 0 & \frac{1}{2} & 1 & | & 0 & 0 & \frac{1}{2} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & | & -\frac{1}{3} & \frac{4}{9} & -\frac{2}{9} \\ 0 & \frac{1}{2} & 0 & | & \frac{2}{3} & -\frac{2}{9} & \frac{1}{9} \\ 0 & 0 & 1 & | & -\frac{1}{3} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}$$

Значит,

$$A^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{4}{9} & -\frac{2}{9} \\ \frac{2}{3} & -\frac{2}{9} & \frac{1}{9} \\ -\frac{1}{3} & \frac{1}{9} & \frac{4}{9} \end{pmatrix}$$

что и было найдено другим способом в §3.2 (п. 4).

Глава 4. ПЛОСКОСТЬ И ПРЯМАЯ В ПРОСТРАНСТВЕ

§ 4.1. Плоскость

1. Геометрическое истолкование уравнения между координатами в пространстве. Известно, что уравнение $f(x, y) = 0$, вообще говоря, определяет на плоскости некоторую линию, т. е. множество всех точек плоскости xOy , координаты которых x и y удовлетворяют этому уравнению. Подобно этому уравнение

$$f(x, y, z) = 0, \quad (1)$$

вообще говоря, определяет в пространстве $Oxyz$ некоторую поверхность, т.е. множество всех точек пространства $Oxyz$, координаты

которых x , y и z удовлетворяют уравнению (1). Уравнение (1) называется уравнением этой поверхности, а x , y , z — ее текущими координатами.

Поверхность называется *цилиндрической*, если она может быть образована перемещением прямой параллельно самой себе вдоль некоторой линии L . Эта линия называется *направляющей* цилиндрической поверхности, а всевозможные положения движущейся прямой — ее *образующими*.

2. Уравнение плоскости, проходящей через данную точку перпендикулярно данному вектору. Пусть дана точка $M_0(x_0; y_0; z_0)$ и ненулевой вектор $\vec{n}(A; B; C)$. Требуется составить уравнение плоскости, проходящей через точку M_0 перпендикулярно вектору $\vec{n}$ (его называют *нормальным вектором* этой плоскости).

Рассмотрим произвольную точку $M(x; y; z)$ этой плоскости. Так как вектор $\overrightarrow{M_0M}(x - x_0; y - y_0; z - z_0)$ лежит на плоскости, то он перпендикулярен вектору $\vec{n}$ (рис. 43).

Следовательно, их скалярное произведение равно нулю:

$$\vec{n} \overrightarrow{M_0M} = 0 \quad (2)$$

или в координатной форме:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \quad (3)$$

Уравнение (3) и есть искомое.

3. Общее уравнение плоскости. Введя обозначение $D = -(Ax_0 + By_0 + Cz_0)$, уравнение (3) можно переписать в виде

$$Ax + By + Cz + D = 0. \quad (4)$$

Следовательно, каждая плоскость в пространстве может быть задана уравнением (4), т. е. уравнением первой степени относительно текущих координат.

Обратно: пусть в уравнении (4) по крайней мере один из коэффициентов A , B , C не равен нулю. Предположим для определенности, что $C \neq 0$. Тогда уравнение (4) можно переписать следующим образом:

$$A(x - 0) + B(y - 0) + C\left(z + \frac{D}{C}\right) = 0. \quad (5)$$

Уравнение (5) равносильно уравнению (4). Сравнивая уравнение (5) с уравне-

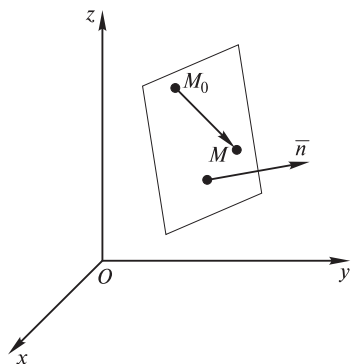


Рис. 43

нием (3), видим, что оно, а следовательно, и равносильное ему уравнение (4), является уравнением плоскости, проходящей через точку $M_0\left(0; 0; -\frac{D}{C}\right)$ и перпендикулярной вектору $\vec{n}(A; B; C)$.

Итак, всякое уравнение первой степени относительно текущих координат, т. е. всякое уравнение вида (4), определяет плоскость.

Уравнение (4) называется *общим уравнением плоскости*.

П р и м е р. Написать уравнение плоскости, проходящей через точку $M_0(1; -5; 6)$ перпендикулярно вектору $\vec{n}(4; 2; -3)$. Согласно равенству (3) имеем:

$$4(x-1) + 2(y+5) - 3(z-6) = 0, \text{ или } 4x + 2y - 3z + 24 = 0.$$

4. Неполные уравнения плоскости.

О п р е д е л е н и е. Общее уравнение (4) называется *полным*, если все его коэффициенты A, B, C, D отличны от нуля. Если хотя бы один из указанных коэффициентов равен нулю, то уравнение (4) называется *неполным*.

Рассмотрим все возможные виды неполных уравнений.

1) $D = 0$; уравнение $Ax + By + Cz = 0$ определяет плоскость, проходящую через начало координат (поскольку координаты начала удовлетворяют этому уравнению).

2) $A = 0$; уравнение $By + Cz + D = 0$ определяет плоскость, параллельную оси Ox (поскольку нормальный вектор этой плоскости $\vec{n}(0; B; C)$ перпендикулярен оси Ox).

Аналогично уравнение $Ax + Cz + D = 0$ ($B = 0$) определяет плоскость, параллельную оси Oy , а уравнение $Ax + By + D = 0$ ($C = 0$) — плоскость, параллельную оси Oz .

3) $A = 0, B = 0$; уравнение $Cz + D = 0$ определяет плоскость, параллельную координатной плоскости xOy (ибо эта плоскость параллельна осям Ox и Oy).

Аналогично уравнение $By + D = 0$ ($A = 0, C = 0$) определяет плоскость, параллельную координатной плоскости xOz , а уравнение $Ax + D = 0$ ($B = 0, C = 0$) — плоскость, параллельную координатной плоскости yOz .

4) $A = 0, B = 0, D = 0$; уравнение $Cz = 0$ определяет координатную плоскость xOy (ибо плоскость $Cz = 0$ параллельна координатной плоскости xOy и проходит через начало координат).

Аналогично уравнение $By = 0$ ($A = 0, C = 0, D = 0$) определяет координатную плоскость xOz , а уравнение $Ax = 0$ ($B = 0, C = 0, D = 0$) — координатную плоскость yOz .

5. Уравнение плоскости в отрезках. Рассмотрим полное уравнение (4). Так как в таком уравнении ни один из коэффициентов A, B, C, D не равен нулю, то его можно переписать в виде

$$\frac{x}{-\frac{D}{A}} + \frac{y}{-\frac{D}{B}} + \frac{z}{-\frac{D}{C}} = 1.$$

Полагая для краткости

$$-\frac{D}{A} = a, \quad -\frac{D}{B} = b, \quad -\frac{D}{C} = c,$$

получаем:

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \quad (6)$$

Уравнение (6) называется *уравнением плоскости в «отрезках»*, так как знаменатели a, b, c есть величины отрезков, отсекаемые плоскостью от осей координат (рис. 44). В самом деле, точка пересечения плоскости с осью Ox определяется из уравнения этой плоскости (6) при дополнительном условии $y = 0, z = 0$, отсюда $x = a$, и, таким образом, величина отрезка, отсекаемая плоскостью (6) на оси Ox , равна a . Аналогично устанавливается, что отрезки, отсекаемые плоскостью (6) на осях Oy и Oz имеют величины, равные соответственно b и c .

6. Расстояние от точки до плоскости. Если в уравнении (2) в качестве нормального вектора плоскости взять единичный вектор

$$\bar{n}_0 = \frac{\bar{n}}{|\bar{n}|},$$

то получим так называемое *нормальное уравнение плоскости*:

$$\bar{n}_0 \overline{M_0 M} = 0, \quad (7)$$

или в координатах:

$$\frac{Ax + By + Cz + D}{\sqrt{A^2 + B^2 + C^2}} = 0, \quad (8)$$

где $D = -(Ax_0 + By_0 + Cz_0)$. Уравнение (8) удобно при нахождении расстояния от точки до плоскости.

З а д а ч а. Найти расстояние d от точки $M_1(x_1; y_1; z_1)$ до плоскости π (рис. 45), заданной уравнением (7).

Из треугольника $M_0 M_1 N_1$ находим:

$$d = \left| \text{пр}_{\bar{n}_0} \overline{M_0 M_1} \right| = \left| \bar{n}_0 \overline{M_0 M_1} \right| = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}. \quad (9)$$

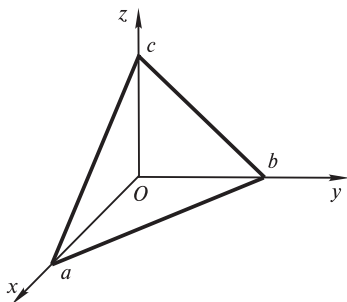


Рис. 44

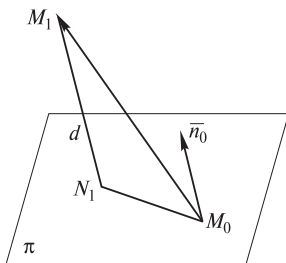


Рис. 45

Следовательно, чтобы найти расстояние от точки до плоскости, нужно в левую часть нормального уравнения плоскости вместо текущих координат подставить координаты данной точки и взять абсолютную величину полученного результата.

П р и м е р. Найти расстояние от точки $M_1(1, 0, -2)$ до плоскости $2x - y + 2z - 4 = 0$. Согласно формуле (9)

$$d = \frac{|2 \cdot 1 - 1 \cdot 0 + 2 \cdot (-2) - 4|}{\sqrt{2^2 + (-1)^2 + 2^2}} = \frac{6}{3} = 2.$$

7. Угол между плоскостями. Условия параллельности и перпендикулярности плоскостей. Пусть уравнения данных плоскостей будут

$$A_1x + B_1y + C_1z + D_1 = 0, \quad A_2x + B_2y + C_2z + D_2 = 0. \quad (10)$$

Углом между плоскостями (10) называется любой из двух смежных двугранных углов, образованных этими плоскостями. (Нам достаточно определить один из этих углов, так как их сумма равна π .) Один из них равен углу φ между нормальными векторами к этим плоскостям $\vec{n}_1(A_1; B_1; C_1)$ и $\vec{n}_2(A_2; B_2; C_2)$. Поэтому

$$\cos \varphi = \frac{A_1 A_2 + B_1 B_2 + C_1 C_2}{\sqrt{A_1^2 + B_1^2 + C_1^2} \sqrt{A_2^2 + B_2^2 + C_2^2}}. \quad (11)$$

П р и м е р 1. Определить угол между плоскостями $x - z = 0$ и $y - z = 0$. Здесь $\vec{n}_1(1; 0; -1)$, $\vec{n}_2(0; 1; -1)$. По формуле (11) имеем:

$$\cos \varphi = \frac{1 \cdot 0 + 0 \cdot 1 + (-1) \cdot (-1)}{\sqrt{1^2 + 0^2 + (-1)^2} \sqrt{0^2 + 1^2 + (-1)^2}} = \frac{1}{\sqrt{2} \sqrt{2}} = \frac{1}{2}.$$

Следовательно, $\varphi = \frac{\pi}{3}$. Итак, один из двух смежных двугранных углов равен $\frac{\pi}{3}$.

Условие параллельности плоскостей (10) совпадает с условием коллинеарности векторов $\bar{n}_1$ и $\bar{n}_2$. Следовательно (§2.1, (10)), оно имеет вид:

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}. \quad (12)$$

Аналогично случаю совпадения прямых на плоскости (§ 1.4, п. 6) условие совпадения плоскостей (10) выражается равенствами

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2} = \frac{D_1}{D_2}. \quad (13)$$

Условие перпендикулярности плоскостей (10) есть вместе с тем условие перпендикулярности нормалей $\bar{n}_1$ и $\bar{n}_2$, следовательно (§ 2.2, (10)),

$$A_1A_2 + B_1B_2 + C_1C_2 = 0. \quad (14)$$

Пример 2. Плоскости $3x + 2y - 7z + 8 = 0$ и $3x + 2y - 7z + 32 = 0$ параллельны, но не совпадают, так как выполнено условие (12), но не выполнено условие (13).

Пример 3. Плоскости $2x + 3y - 2z - 4 = 0$ и $13x - 8y + z + 44 = 0$ перпендикулярны, так как выполнено условие (14).

§ 4.2. Прямая в пространстве

1. Геометрическое истолкование двух уравнений между координатами в пространстве. Пусть имеем два уравнения с тремя переменными

$$f_1(x, y, z) = 0 \quad \text{и} \quad f_2(x, y, z) = 0.$$

Каждое из них, вообще говоря, определяет некоторую поверхность (см. § 4.1, п. 1). Множество точек, общих обеим поверхностям, есть, вообще говоря, некоторая линия.

Пример. Уравнения $x^2 + y^2 = R^2$ и $z = a$ определяют окружность, лежащую в плоскости $z = a$, радиуса R с центром на оси Oz в точке $(0; 0; a)$. Заметим, что эту же линию можно задать параметрически тремя уравнениями $x = R \cos \varphi$, $y = R \sin \varphi$, $z = a$.

В общем случае параметрические уравнения линии имеют вид:

$$x = x(t), \quad y = y(t), \quad z = z(t).$$

2. Общие уравнения прямой. Рассмотрим систему уравнений первой степени

$$\begin{cases} A_1x + B_1y + C_1z + D_1 = 0, \\ A_2x + B_2y + C_2z + D_2 = 0. \end{cases} \quad (1)$$

Каждое из уравнений системы (1) определяет плоскость. Если нормальные векторы $\bar{n}_1(A_1; B_1; C_1)$ и $\bar{n}_2(A_2; B_2; C_2)$ этих плоскостей не

коллинеарны (т. е. плоскости не параллельны и не совпадают), то система (1) определяет некоторую прямую l как линию пересечения двух плоскостей. Уравнения (1) называются *общими уравнениями прямой*.

П р и м е р. Построить прямую, заданную общими уравнениями

$$\begin{cases} x + y + z - 3 = 0, \\ x - 3y - z + 5 = 0. \end{cases}$$

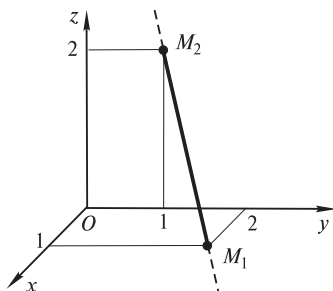


Рис. 46

Для построения прямой достаточно знать две ее точки. Проще выбрать точки пересечения прямой с координатными плоскостями. Точка пересечения прямой с координатными плоскостями называется *следом прямой*. Координаты следа M_1 данной прямой на плоскости xOy получим из уравнений прямой, полагая $z = 0$. Это дает $x = 1, y = 2$. Следовательно, координаты точки M_1 таковы: $x_1 = 1, y_1 = 2, z_1 = 0$. Аналогично, полагая в уравнениях заданной прямой $x = 0$, получаем координаты следа M_2 прямой на плоскости yOz : $x_2 = 0, y_2 = 1, z_2 = 2$. Теперь строим проходящую через точки M_1 и M_2 прямую (рис. 46).

3. Канонические уравнения прямой. Пусть дана точка $M_0(x_0; y_0; z_0)$ и ненулевой вектор $\vec{s}(m; p; q)$. Требуется составить уравнение прямой l , проходящей через точку M_0 и параллельной вектору $\vec{s}$ (этот вектор называют *направляющим вектором прямой*). Для этого заметим, что точка $M(x; y; z)$ лежит на указанной прямой тогда и только тогда, когда векторы $\overline{M_0M}(x - x_0; y - y_0; z - z_0)$ и $\vec{s}(m; p; q)$ коллинеарны, т. е. тогда и только тогда, когда координаты этих векторов пропорциональны (§ 2.1, (10)):

$$\frac{x - x_0}{m} = \frac{y - y_0}{p} = \frac{z - z_0}{q}. \quad (2)$$

Уравнения (2) называются *каноническими уравнениями прямой l* .

П р и м е р. Написать канонические уравнения прямой, проходящей через точку $M_0(-2; -3; -1)$ и имеющей направляющий вектор $\vec{s}(3; 2; 4)$. Согласно равенствам (2) имеем:

$$\frac{x + 2}{3} = \frac{y + 3}{2} = \frac{z + 1}{4}.$$

Рассмотрим теперь вопрос о том, как перейти от общих уравнений прямой к ее каноническим уравнениям. Для этого нужно найти какую-нибудь точку $M_0(x_0; y_0; z_0)$ на прямой и направляющий вектор $\vec{s}$ прямой. Пусть прямая l задана общими уравнениями (1). Координаты

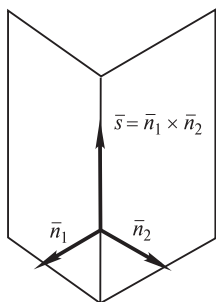


Рис. 47

точки M_0 на прямой l получим из системы уравнений (1), придав одной из координат произвольное значение. Так как прямая l перпендикулярна нормальным векторам $\bar{n}_1 = A_1\bar{i} + B_1\bar{j} + C_1\bar{k}$ и $\bar{n}_2 = A_2\bar{i} + B_2\bar{j} + C_2\bar{k}$ (рис. 47), то за направляющий вектор $\bar{s}$ прямой l можно принять векторное произведение $\bar{n}_1 \times \bar{n}_2$:

$$\bar{s} = \bar{n}_1 \times \bar{n}_2 = \begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{vmatrix}.$$

4. Параметрические уравнения прямой в пространстве. Пусть прямая l задана каноническими уравнениями (2). Примем за параметр t каждое из отношений (2). Так как один из знаменателей в (2) отличен от нуля, а соответствующий числитель может принимать какие угодно значения, то область изменения параметра t является вся числовая ось $-\infty < t < +\infty$. Получим:

$$x - x_0 = mt, \quad y - y_0 = pt, \quad z - z_0 = qt,$$

или

$$x = x_0 + mt, \quad y = y_0 + pt, \quad z = z_0 + qt. \quad (3)$$

Уравнения (3) и есть искомые *параметрические уравнения прямой*.

Пример. Составить параметрические уравнения прямой, проходящей через точку $M_0(2, -3, -7)$ и имеющей направляющий вектор $\bar{s}(4; -6; 5)$. Согласно равенствам (3) имеем:

$$x = 2 + 4t, \quad y = -3 - 6t, \quad z = -7 + 5t.$$

5. Угол между прямыми. Пусть в пространстве даны две прямые:

$$\frac{x - x_1}{m_1} = \frac{y - y_1}{p_1} = \frac{z - z_1}{q_1}, \quad \frac{x - x_2}{m_2} = \frac{y - y_2}{p_2} = \frac{z - z_2}{q_2}. \quad (4)$$

За угол между двумя прямыми принимают один из двух смежных углов, которые образуют прямые, проведенные параллельно данным через какую-нибудь точку пространства. Один из этих смежных углов равен углу между направляющими векторами $\bar{s}_1(m_1; p_1; q_1)$, $\bar{s}_2(m_2; p_2; q_2)$ данных прямых. Поэтому

$$\cos \varphi = \frac{m_1 m_2 + p_1 p_2 + q_1 q_2}{\sqrt{m_1^2 + p_1^2 + q_1^2} \sqrt{m_2^2 + p_2^2 + q_2^2}}. \quad (5)$$

П р и м е р 1. Определить угол между прямыми

$$\frac{x-1}{1} = \frac{y}{-4} = \frac{z+3}{1} \quad \text{и} \quad \frac{x}{2} = \frac{y+2}{-2} = \frac{z}{-1}.$$

Здесь $\vec{s}_1(1; -4; 1)$, $\vec{s}_2(2; -2; -1)$. По формуле (5) получим:

$$\cos \varphi = \frac{1 \cdot 2 + (-4) \cdot (-2) + 1 \cdot (-1)}{\sqrt{1^2 + (-4)^2 + 1^2} \sqrt{2^2 + (-2)^2 + (-1)^2}} = \frac{1}{\sqrt{2}},$$

откуда $\varphi = \frac{\pi}{4}$. Следовательно, один из двух смежных углов равен $\frac{\pi}{4}$.

Условие параллельности прямых (4) совпадает с условием коллинеарности векторов $\vec{s}_1$ и $\vec{s}_2$. Следовательно (§ 2.1, (10)), оно имеет вид:

$$\frac{m_1}{m_2} = \frac{p_1}{p_2} = \frac{q_1}{q_2}. \quad (6)$$

Если при этом точка первой прямой, например $M_1(x_1; y_1; z_1)$, удовлетворяет уравнению второй прямой, т. е. если

$$\frac{x_1 - x_2}{m_2} = \frac{y_1 - y_2}{p_2} = \frac{z_1 - z_2}{q_2}, \quad (7)$$

то эти прямые совпадают.

Условие перпендикулярности прямых (4) есть вместе с тем условие перпендикулярности их направляющих векторов $\vec{s}_1$ и $\vec{s}_2$. Следовательно (§ 2.2, (10)),

$$m_1 m_2 + p_1 p_2 + q_1 q_2 = 0. \quad (8)$$

П р и м е р 2. Прямые $\frac{x+2}{3} = \frac{y-3}{2} = \frac{z+1}{4}$ и $\frac{x-7}{3} = \frac{y-2}{2} = \frac{z}{4}$ параллельны, но не совпадают, так как выполняется условие (6), но не выполняется условие (7).

П р и м е р 3. Прямые $\frac{x+2}{3} = \frac{y-3}{2} = \frac{z+1}{4}$ и $\frac{x+5}{3} = \frac{y-1}{2} = \frac{z+5}{4}$ совпадают, так как наряду с выполнением условия (6) выполняется условие (7).

П р и м е р 4. Прямые $\frac{x+1}{2} = \frac{y+5}{4} = \frac{z-1}{-13}$ и $\frac{x+1}{3} = \frac{y-4}{5} = \frac{z+2}{2}$ перпендикулярны, так как выполняется условие (8).

6. Условия параллельности и перпендикулярности прямой и плоскости.

Пусть даны прямая

$$\frac{x-x_0}{m} = \frac{y-y_0}{p} = \frac{z-z_0}{q} \quad (9)$$

и плоскость

$$Ax + By + Cz + D = 0. \quad (10)$$

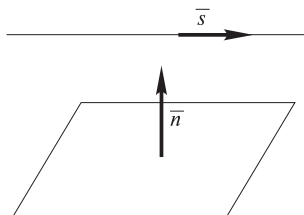


Рис. 48

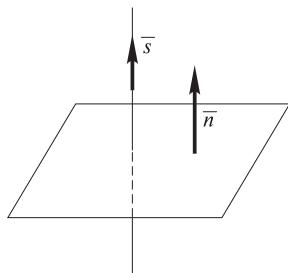


Рис. 49

Прямая (9) параллельна плоскости (10) в том и только в том случае, когда направляющий вектор этой прямой $\vec{s}(m; p; q)$ перпендикулярен к нормальному вектору данной плоскости $\vec{n}(A; B; C)$ (рис. 48).

Отсюда получаем условие параллельности прямой (9) и плоскости (10):

$$Am + Bp + Cq = 0. \quad (11)$$

Прямая (9) перпендикулярна к плоскости (10) в том и только в том случае, когда направляющий вектор этой прямой коллинеарен нормальному вектору данной плоскости (рис. 49). Отсюда получаем условие перпендикулярности прямой и плоскости:

$$\frac{A}{m} = \frac{B}{p} = \frac{C}{q}. \quad (12)$$

Рассмотрим особо случай принадлежности прямой (9) плоскости (10). Эти условия выражаются двумя равенствами:

$$\begin{aligned} Ax_0 + By_0 + Cz_0 + D &= 0, \\ Am + Bp + Cq &= 0, \end{aligned} \quad (13)$$

первое из которых (13) означает, что точка $M_0(x_0; y_0; z_0)$, через которую проходит данная прямая, принадлежит плоскости (10), а второе есть условие параллельности прямой (9) и плоскости (10).

Пример 1. Прямая $\frac{x-2}{1} = \frac{y+1}{1} = \frac{z}{-1}$ параллельна плоскости $3x - y + 2z - 8 = 0$, но ей не принадлежит, так как выполняется условие (11), но не выполняется условие (13).

Пример 2. Прямая $\frac{x-2}{3} = \frac{y+3}{1} = \frac{z-4}{-1}$ перпендикулярна плоскости $6x + 4y - 2z - 9 = 0$, так как выполняется условие (12).

Пр и м е р 3. Прямая $\frac{x-2}{1} = \frac{y+1}{1} = \frac{z}{-1}$ принадлежит плоскости $3x - y + 2z - 7 = 0$, так как выполняются условия (13) и (11).

§ 4.3. Основные задачи на плоскость и прямую в пространстве

1. Составить уравнение произвольной плоскости, проходящей через точку $M_0(x_0; y_0; z_0)$. Пусть

$$Ax + By + Cz + D = 0 \quad (1)$$

уравнение искомой плоскости. Тогда

$$Ax_0 + By_0 + Cz_0 + D = 0. \quad (2)$$

Вычитая из равенства (1) почленно равенство (2), получаем:

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Очевидно, при любых A, B, C этому уравнению удовлетворяют координаты точки M_0 .

2. Составить уравнение произвольной прямой, проходящей через точку $M_0(x_0; y_0; z_0)$. Искомое уравнение

$$\frac{x - x_0}{m} = \frac{y - y_0}{p} = \frac{z - z_0}{q}. \quad (3)$$

Действительно, это уравнение задает прямую, проходящую через точку M_0 , координаты которой, очевидно, удовлетворяют уравнению (3). Давая m, p, q произвольные (не все равные нулю) значения, получаем прямую произвольного направления.

3. Составить уравнение прямой, проходящей через две различные данные точки $M_1(x_1; y_1; z_1)$ и $M_2(x_2; y_2; z_2)$. Уравнение прямой, проходящей через точку $M_1(x_1; y_1; z_1)$, имеет вид:

$$\frac{x - x_1}{m} = \frac{y - y_1}{p} = \frac{z - z_1}{q}. \quad (4)$$

Так как прямая проходит и через точку M_2 , то

$$\frac{x_2 - x_1}{m} = \frac{y_2 - y_1}{p} = \frac{z_2 - z_1}{q}. \quad (5)$$

Исключая из уравнений (4) и (5) m, p, q , получаем искомое уравнение

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

4. Составить уравнение плоскости, проходящей через три точки $M_1(x_1; y_1; z_1)$, $M_2(x_2; y_2; z_2)$ и $M_3(x_3; y_3; z_3)$, не лежащие на прямой. Пусть $M(x; y; z)$ — произвольная точка искомой плоскости. Три вектора $\overline{M_1M}$, $\overline{M_1M_2}$, $\overline{M_1M_3}$ лежат в одной плоскости, т. е. они компланарны. Поэтому (§ 2.2, (15))

$$(\overline{M_1M} \ \overline{M_1M_2} \ \overline{M_1M_3}) = 0,$$

или в координатной форме:

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0.$$

Глава 5. КРИВЫЕ И ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА В КАНОНИЧЕСКОЙ ФОРМЕ

§ 5.1. Кривые второго порядка в канонической форме

1. **Определение и каноническое уравнение эллипса.** Эллипсом называется множество всех точек плоскости, сумма расстояний каждой из которых до двух данных точек F_1 и F_2 (называемых *фокусами эллипса*) есть величина постоянная, равная $2a$.

Выведем уравнение эллипса. Для этого выберем прямоугольную систему координат так, чтобы ось Ox проходила через фокусы F_1 и F_2 (расстояние между фокусами обозначим через $2c$), а начало координат находилось в середине отрезка F_1F_2 (рис. 50). Тогда фокусы будут иметь координаты $F_1(-c; 0)$ и $F_2(c; 0)$.

Пусть $M(x; y)$ — произвольная точка эллипса. Согласно определению эллипса имеем:

$$MF_1 + MF_2 = 2a,$$

или по формуле расстояния между двумя точками запишем:

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a.$$

Это и есть уравнение эллипса. Приведем его к каноническому (т. е. простейшему) виду. Имеем:

$$\sqrt{(x+c)^2 + y^2} = 2a - \sqrt{(x-c)^2 + y^2}.$$

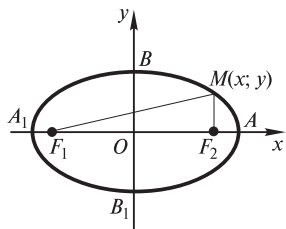


Рис. 50

Возведем обе части последнего равенства в квадрат:

$$x^2 + 2xc + c^2 + y^2 = 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + x^2 - 2xc + c^2 + y^2,$$

откуда

$$a\sqrt{(x-c)^2 + y^2} = a^2 - cx. \quad (1)$$

Возведем теперь в квадрат обе части равенства (1):

$$a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2 = a^4 - 2a^2cx + c^2x^2,$$

откуда

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2).$$

Заметим, что $a^2 - c^2 > 0$, так как $2a > 2c$, или $a > c$ (сумма двух сторон треугольника больше третьей его стороны; случай $2a = 2c$ естественно исключить, так как тогда получаем совокупность всех точек M , для которых $MF_1 + MF_2 = F_1F_2$, т. е. отрезок F_1F_2 .) Поэтому, обозначив $a^2 - c^2$ через b^2 , получаем:

$$b^2x^2 + a^2y^2 = a^2b^2.$$

Деля обе части последнего равенства на a^2b^2 , получаем каноническое уравнение эллипса

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (2)$$

Эллипс, отвечающий уравнению (2), изображен на рисунке 50.

Так как уравнение (2) содержит текущие координаты x и y только в четных степенях, то при замене x на $-x$, а y на $-y$ это уравнение не изменяется, т. е. эллипс симметричен относительно обеих осей координат. Из уравнения (2) при $y = 0$ получаем $x = \pm a$, т. е. эллипс пересекает ось Ox в двух точках: $A(a; 0)$ и $A_1(-a; 0)$; при $x = 0$ получаем $y = \pm b$, т. е. эллипс пересекает ось Oy в двух точках: $B(0; b)$ и $B_1(0, -b)$. Эти четыре точки называются *вершинами* эллипса. Отрезок A_1A называется *большой осью* эллипса, а отрезок B_1B — его *малой осью*. Следовательно, a — длина большой полуоси эллипса, b — длина его малой полуоси.

В частном случае, когда $a = b$, уравнение (2) принимает вид $x^2 + y^2 = a^2$ и определяет окружность с центром в начале координат. В этом случае $c = 0$.

Эксцентриситетом эллипса называется отношение расстояния между фокусами к длине его большой оси, т. е.

$$\varepsilon = \frac{2c}{2a} = \frac{c}{a}. \quad (3)$$

Так как $c < a$, то для любого эллипса $0 \leq \varepsilon < 1$ (случай $\varepsilon = 0$ соответствует окружности). Эксцентриситет характеризует степень сжатия эллипса. Действительно, из (3) и того, что $b^2 = a^2 - c^2$, следует.

$$\varepsilon^2 = \frac{c^2}{a^2} = \frac{a^2 - b^2}{a^2} = 1 - \left(\frac{b}{a}\right)^2,$$

и, значит,

$$\frac{b}{a} = \sqrt{1 - \varepsilon^2}.$$

Отсюда видно, что, чем больше ε , тем меньше отношение $\frac{b}{a}$ и тем больше вытянут эллипс.

Эксцентриситет (ε), длины полуосей (a и b), расстояние между фокусами ($2c$) — параметры, которые полностью определяют эллипс с центром в начале координат.

П р и м е р. Найти параметры эллипса, заданного уравнением $\frac{x^2}{16} + \frac{y^2}{9} = 4$. Для этого приведем данное уравнение к каноническому виду $\frac{x^2}{64} + \frac{y^2}{36} = 1$. Отсюда следует, что $a = 8$, $b = 6$, $c = \sqrt{64 - 36} = \sqrt{28} = 2\sqrt{7}$, $\varepsilon = \frac{2\sqrt{7}}{8} = \frac{\sqrt{7}}{4}$.

2. Определение и каноническое уравнение гиперболы. Гиперболой называется множество всех точек плоскости, разность расстояний каждой из которых до двух данных точек F_1 и F_2 (называемых *фокусами гиперболы*) есть величина постоянная, равная $2a$.

Обозначим через $2c$ расстояние между фокусами F_1 и F_2 (рис. 51).

Пусть $M(x; y)$ — произвольная точка гиперболы. Тогда по определению $MF_1 - MF_2 = 2a$ или $MF_2 - MF_1 = 2a$. Эти условия, определяющие гиперболу, можно записать в виде

$$MF_1 - MF_2 = \pm 2a.$$

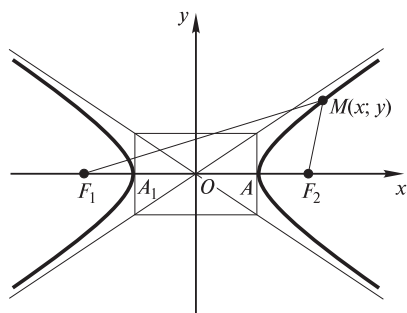


Рис. 51

Заметим, что $a < c$, так как $2a < 2c$ (разность двух сторон треугольника меньше его третьей стороны). Если $a = c$, то мы получаем точки M , для которых или $MF_1 - MF_2 = F_1F_2$, или $MF_2 - MF_1 = F_1F_2$, т. е. совокупность тех точек прямой, проходящей через фокусы, которые лежат вне отрезка F_1F_2 . Поэтому случай $2a = 2c$ естественно исключить.

Далее вывод канонического уравнения гиперболы проводится аналогично выводу канонического уравнения эллипса. Каноническое уравнение гиперболы имеет вид:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad (4)$$

где

$$b^2 = c^2 - a^2.$$

Гипербола, отвечающая уравнению (4), изображена на рисунке 51. Подобно эллипсу, гипербола симметрична относительно обеих осей координат. Она состоит из двух частей, которые называются ее *ветвями*. Из уравнения (4) при $y = 0$ получаем $x = \pm a$, т. е. гипербола пересекает ось Ox в двух точках: $A(a; 0)$ и $A_1(-a; 0)$, называемых *вершинами* гиперболы. Отрезок A_1A называется *действительной осью* гиперболы.

Прямые $y = \pm \frac{b}{a}x$ называются *асимптотами* гиперболы. При увеличении x по абсолютной величине ветви гиперболы все ближе прилегают к своим асимптотам. Для построения асимптот гиперболы целесообразно предварительно построить прямоугольник со сторонами $2a$ и $2b$, параллельными координатным осям и с центром в начале координат (такой прямоугольник называется *основным* прямоугольником гиперболы).

Эксцентриситетом гиперболы называется отношение $\epsilon = \frac{c}{a}$. Так как $a < c$, то для любой гиперболы $\epsilon > 1$. Учитывая, что $b^2 = c^2 - a^2$, имеем:

$$\epsilon^2 = \frac{c^2}{a^2} = \frac{a^2 + b^2}{a^2} = 1 + \left(\frac{b}{a}\right)^2,$$

и, значит,

$$\frac{b}{a} = \sqrt{\epsilon^2 - 1}.$$

Отсюда видно, что, чем меньше эксцентриситет гиперболы, т. е. чем ближе он к единице, тем больше вытянут основной прямоугольник по оси Ox .

Если у гиперболы (4) $a = b$, то она называется *равносторонней* (или *равнобочной*) и ее уравнение имеет вид:

$$x^2 - y^2 = a^2. \quad (5)$$

Асимптотами для равносторонней гиперболы (5) служат взаимно перпендикулярные прямые $y = \pm x$. Поэтому их можно принять за оси прямоугольной системы координат (OX — за ось абсцисс, OY — за ось

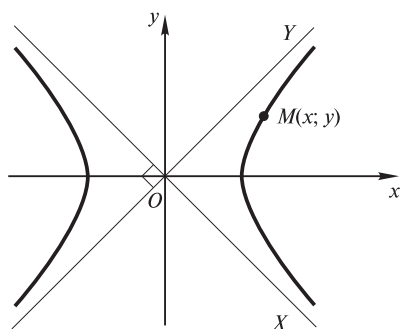


Рис. 52

ординат) и рассматривать эту равностороннюю гиперболу по отношению к этим новым осям. Взяв на указанной гиперболе произвольную точку $M(x, y)$ (рис. 52), выразим новые координаты X и Y точки M через старые x и y (см. § 3.2, п. 5, формулы (21)):

$$X = \frac{\sqrt{2}}{2}(x - y), \quad (6)$$

$$Y = \frac{\sqrt{2}}{2}(x + y). \quad (7)$$

Перемножив равенства (6) и (7) и приняв во внимание равенство (5), получим:

$$XY = \frac{1}{2}(x^2 - y^2) = \frac{1}{2}a^2,$$

или, полагая для краткости $\frac{a^2}{2} = m$, $XY = m$.

Следовательно, уравнению $xy = a$, где $a > 0$, соответствует равносторонняя гипербола, имеющая своими асимптотами оси координат и лежащая в I и III квадрантах (рис. 53). Легко понять, что при $a < 0$ эта гипербола лежит во II и IV квадрантах (рис. 54).

П р и м е р. Найти параметры (a, b, c, ϵ) гиперболы, заданной уравнением $x^2 - 4y^2 = 36$. Для этого приведем данное уравнение к каноническому виду $\frac{x^2}{36} - \frac{y^2}{9} = 1$. Отсюда следует, что $a = 6$, $b = 3$, $c = \sqrt{36 + 9} = \sqrt{45} = 3\sqrt{5}$ и $\epsilon = \frac{3\sqrt{5}}{6} = \frac{\sqrt{5}}{2}$.

Уравнения асимптот гиперболы: $y = \pm \frac{1}{2}x$.

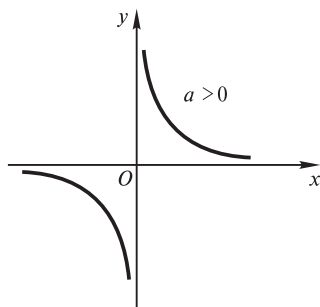


Рис. 53

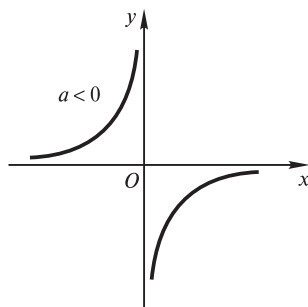


Рис. 54

3. Определение и каноническое уравнение параболы. *Параболой* называется множество всех точек плоскости, равноудаленных от данной точки F (называемой *фокусом параболы*) и от данной прямой l (называемой *директрисой параболы*; предполагается, что F не лежит на l).

Для вывода канонического уравнения параболы проведем ось Ox прямоугольной системы координат через фокус F перпендикулярно директрисе, начало координат O поместим на равных расстояниях от фокуса и директрисы (рис. 55). Расстояние от фокуса до директрисы обозначим через p (оно называется *параметром параболы*). В этом случае фокус будет иметь координаты $F\left(\frac{p}{2}; 0\right)$, а уравнение директрисы будет

$$x = -\frac{p}{2}.$$

Возьмем произвольную точку $M(x; y)$ параболы. Согласно определению параболы имеем:

$$MF = MA$$

(точка A имеет координаты $\left(-\frac{p}{2}; y\right)$), или по формуле расстояния между двумя точками:

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \sqrt{\left(x + \frac{p}{2}\right)^2}.$$

Отсюда

$$\left(x - \frac{p}{2}\right)^2 + y^2 = \left(x + \frac{p}{2}\right)^2,$$

или

$$x^2 - px + \frac{p^2}{4} + y^2 = x^2 + px + \frac{p^2}{4},$$

и окончательно:

$$y^2 = 2px. \quad (8)$$

Формула (8) и есть каноническое уравнение параболы. Парабола, отвечающая уравнению (8), изображена на рисунке 55.

Уравнение (8) имеет смысл только для неотрицательных значений x , т. е. все точки параболы лежат в I и IV квадрантах. Так как уравнение (8) содержит y^2 , то парабола симметрична относительно оси Ox . *Вершиной параболы* называется точка пере-

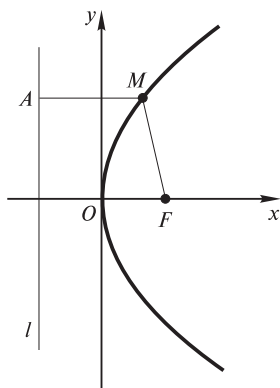


Рис. 55

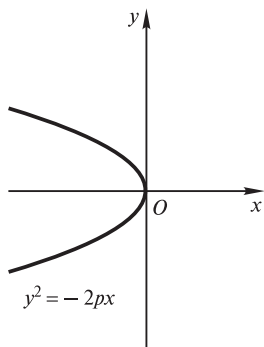


Рис. 56

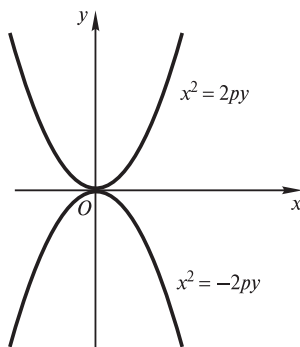


Рис. 57

сечения параболы с ее осью симметрии. При возрастании x значения y возрастают по абсолютной величине. В отличие от гиперболы парабола не имеет асимптот. Ось симметрии параболы называется *осью параболы*. Парабола, определяемая уравнением (8), имеет ось, совпадающую с осью Ox .

П р и м е ч а н и е. Очевидно, что каждому из уравнений $y^2 = -2px$, $x^2 = 2py$, $x^2 = -2py$ ($p > 0$) соответствует парабола, по форме тождественная с параболой (8), но иначе расположенная. На рисунках 56, 57 изображены эти параболы.

К параболом, например, симметричным относительно оси Oy , относятся также кривые, заданные уравнениями $x^2 = 2p(y - c)$, $x^2 = -2p(y + c)$, $c > 0$, $p > 0$ (рис. 58).

П р и м е р. Парабола с вершиной в начале координат проходит через точку $A(9; 3)$ и симметрична относительно оси Ox . Написать ее каноническое уравнение. Подставляя координаты точки A в уравнение (8), найдем, что $p = \frac{1}{2}$. Значит, уравнение искомой параболы $y^2 = x$.

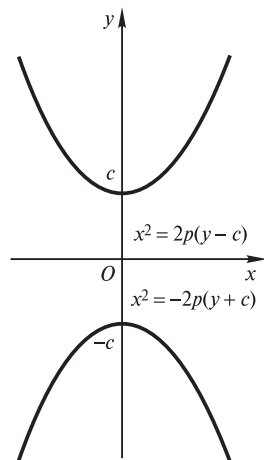


Рис. 58

4. Директрисы эллипса и гиперболы. Рассмотрим эллипс и гиперболу, определяемые соответственно каноническими уравнениями (2) и (4).

Предположим, что этот эллипс вытянут в направлении оси Ox , т. е. что $a > b$.

Две прямые, перпендикулярные к большой оси эллипса и расположенные симметрично относительно центра на расстоянии $\frac{a}{\epsilon}$ от него, называются *директрисами эллипса*. Их уравнения:

$$x = -\frac{a}{\epsilon} \quad \text{и} \quad x = \frac{a}{\epsilon}.$$

Первую из них мы условимся называть *левой*, вторую — *правой*.

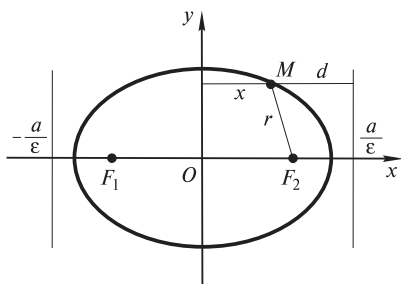


Рис. 59

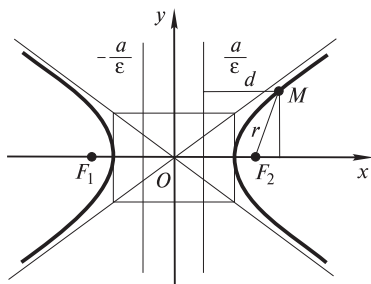


Рис. 60

Так как для эллипса $\epsilon < 1$, то $\frac{a}{\epsilon} > a$. Отсюда следует, что правая директриса расположена правее правой вершины эллипса; аналогично левая директриса расположена левее его левой вершины (рис. 59).

Две прямые, перпендикулярные к той оси гиперболы, которая ее пересекает, и расположенные симметрично относительно центра на расстоянии $\frac{a}{\epsilon}$ от него, называются *директрисами гиперболы*.

Уравнения директрис гиперболы (4) имеют вид:

$$x = -\frac{a}{\epsilon} \quad \text{и} \quad x = \frac{a}{\epsilon}.$$

Первую из них условимся называть левой, вторую — правой.

Так как для гиперболы $\epsilon > 1$, то $\frac{a}{\epsilon} < a$. Отсюда следует, что правая директриса расположена между центром и правой вершиной гиперболы; аналогично левая директриса расположена между центром и левой вершиной (рис. 60).

Имеет место следующая теорема:

Т е о р е м а. Если r — расстояние от произвольной точки эллипса (гиперболы) до какого-нибудь фокуса, d — расстояние от той же точки до соответствующей этому фокусу директрисы, то

$$\frac{r}{d} = \epsilon, \tag{9}$$

ϵ — эксцентриситет эллипса (гиперболы).

Докажем теорему для эллипса, так как для гиперболы схема доказательства аналогична. Предположим для определенности, что речь идет о правом фокусе и правой директрисе. Пусть $M(x, y)$ — произвольная точка эллипса (см. рис. 59). Расстояние от M до правой директрисы выражается равенством

$$d = \frac{a}{\epsilon} - x, \tag{10}$$

которое легко усматривается из рисунка 59. Расстояние r от точки M до правого фокуса, как следует из формулы (1), выражается равенством

$$r = a - \frac{c}{a}x,$$

или

$$r = a - \varepsilon x. \quad (11)$$

Из соотношений (10) и (11) имеем:

$$\frac{r}{d} = \frac{a - \varepsilon x}{\frac{a}{\varepsilon} - x} = \frac{(a - \varepsilon x)\varepsilon}{a - \varepsilon x} = \varepsilon.$$

П р и м е ч а н и е. Так как для параболы $\frac{r}{d} = 1$, то говорят, что парабола имеет эксцентриситет $\varepsilon = 1$.

Установленная теорема позволяет сформулировать *общее определение* эллипса, гиперболы и параболы.

Множество точек плоскости, для каждой из которых расстояние r до некоторой фиксированной точки (фокуса) и расстояние d до некоторой фиксированной прямой (директрисы) находятся в постоянном отношении $\frac{r}{d} = \varepsilon$, есть эллипс, если $\varepsilon < 1$; гипербола, если $\varepsilon > 1$; парабола, если $\varepsilon = 1$.

5. Уравнение кривых второго порядка в полярных координатах. Пусть дана какая-нибудь из линий: эллипс, гипербола или парабола (если данная линия — гипербола, то будем рассматривать какую-нибудь ее ветвь); обозначим ее буквой L . Пусть F — фокус линии, g — соответствующая этому фокусу директриса (в случае гиперболы в качестве F и g возьмем фокус и директрису, ближайшие к рассматриваемой ветви). Введем полярную систему координат с полюсом в точке F и полярной осью, которая перпендикулярна директрисе g и направлена от g (рис. 61). Обозначим через ρ , φ полярные координаты произвольной точки M линии L . Для точек линии L выполнено условие (9) (см. теорему и примечание из п. 4). Здесь, как видно из рисунка, $r = \rho$,

$$d = DF + FN = DF + \rho \cos \varphi. \quad (12)$$

Пусть точка P имеет полярные координаты p и $\frac{\pi}{2}$, где $p = FP$ — заданное число, называемое *фокальным параметром* линии L . Для этой точки согласно равенству (9) имеем:

$$\frac{p}{SP} = \varepsilon, \quad \text{или} \quad \frac{p}{DF} = \varepsilon,$$

откуда

$$DF = \frac{p}{\varepsilon}.$$

Из последнего равенства и равенства (12) имеем:

$$d = \frac{p}{\varepsilon} + \rho \cos \varphi.$$

Подставляя теперь в левую часть равенства (9) выражения для r и d , найдем:

$$\frac{\rho}{\frac{p}{\varepsilon} + \rho \cos \varphi} = \varepsilon,$$

откуда

$$\rho = \frac{p}{1 - \varepsilon \cos \varphi}.$$

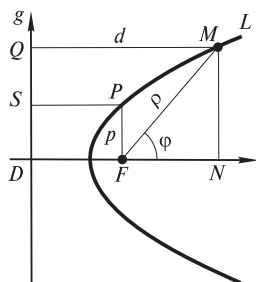


Рис. 61

Это и есть полярное уравнение эллипса, гиперболы (одной ее ветви) и параболы. Здесь p — фокальный параметр, ϵ — эксцентриситет кривой L . Это уравнение используется к механике.

§ 5.2. Изучение поверхностей второго порядка по их каноническим уравнениям

1. Эллипсоид и гиперboloиды. *Эллипсоидом* (рис. 62) называется поверхность, определяемая в прямоугольной системе координат $Oxyz$ уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \quad (1)$$

Уравнение (1) называется *каноническим уравнением эллипсоида*. Величины a , b , c называются *полуосями эллипсоида*.

Из уравнения (1) видно, что координатные плоскости являются плоскостями симметрии эллипсоида, а начало координат — центром симметрии. Точки пересечения осей координат с эллипсоидом называются *вершинами эллипсоида*.

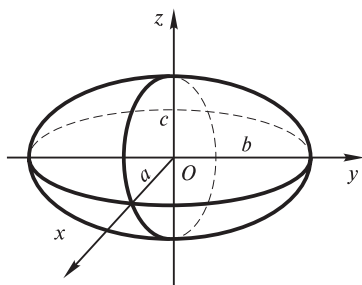


Рис. 62

Рассмотрим сечение эллипсоида плоскостью, параллельной плоскости xOy ; пусть это будет плоскость $z = h$ и пусть при этом $|h| < c$. Тогда линия, которая получается в сечении, определяется двумя уравнениями:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 - \frac{h^2}{c^2}, \quad z = h. \quad (2)$$

Обозначив через k^2 положительное число $1 - \frac{h^2}{c^2}$, уравнения (2) перепишем в виде

$$\frac{x^2}{(ak)^2} + \frac{y^2}{(bk)^2} = 1, \quad z = h.$$

Мы видим, что сечение эллипсоида (1) плоскостью $z = h$ ($|h| < c$) представляет собой эллипс с полуосями ak и bk , уменьшающимися с увеличением $|h|$; при $|h| = c$ этот эллипс стягивается в точку — вершину эллипсоида. Совершенно аналогичная картина выявляется при рассмотрении сечений эллипсоида плоскостями, параллельными координатным плоскостям xOz и yOz . Отметим только, что сама плоскость

xOz пересекает эллипсоид по эллипсу, который определяется уравнениями

$$\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1, \quad y = 0;$$

плоскость yOz — по эллипсу, который определяется уравнениями

$$\frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad x = 0.$$

Если две полуоси эллипсоида равны, например $a = b$, то получаем уравнение

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{c^2} = 1. \quad (3)$$

Пересекая этот эллипсоид плоскостью $z = h$, параллельной плоскости xOy , получим окружность

$$x^2 + y^2 = \left(1 - \frac{h^2}{c^2}\right)a^2, \quad z = h$$

с центром на оси Oz . Поэтому такой эллипсоид может быть получен вращением эллипса

$$\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1,$$

расположенного в плоскости xOz , вокруг оси Oz . Эллипсоид (3) называется *эллипсоидом вращения*.

Если же все три полуоси эллипсоида (1) равны: $a = b = c$, то получаем:

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

т. е. *сферу*, которая, таким образом, оказывается частным случаем эллипсоида.

Однополостным гиперboloидом (рис. 63) называется поверхность, определяемая в прямоугольной системе координат $Oxyz$ уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \quad (4)$$

а *двуполостным гиперboloидом* (рис. 64) — поверхность, определяемая уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1. \quad (5)$$

Уравнения (4) и (5) называются *каноническими уравнениями гиперboloидов*. Величины a, b, c называются *полуосями* однополостного (двуполостного) гиперboloида. Оба гиперboloида имеют координатные плоскости симметрии, а начало координат — центр симметрии.

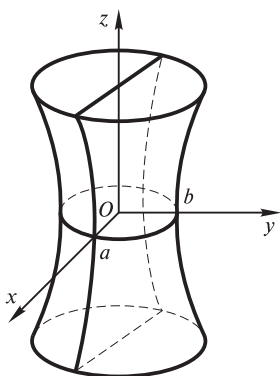


Рис. 63

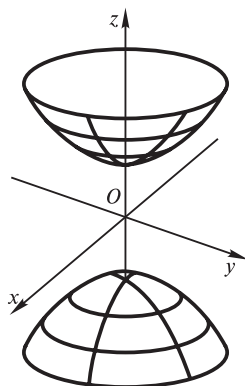


Рис. 64

Отметим некоторые сечения гиперboloидов плоскостями, параллельными координатным плоскостями. Например, однополостный гиперboloид (4) плоскость $z = h$ пересекает по эллипсу

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + \frac{h^2}{c^2}, \quad z = h,$$

плоскость $y = h$ ($|h| \neq b$) — по гиперболе

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{h^2}{b^2}, \quad y = h,$$

плоскость $y = b$ — по двум прямым

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 0, \quad y = b.$$

По аналогии с эллипсоидом, если полуоси a и b гиперboloида (однополостного или двуполостного) равны, то он называется *гиперboloидом вращения* и получается вращением около оси Oz гиперболы

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1, \quad y = 0$$

в случае однополостного гиперboloида и гиперболы

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = -1, \quad y = 0$$

в случае двуполостного гиперboloида.

2. Параболоиды. *Эллиптическим параболоидом* (рис. 65) называется поверхность, которая в прямоугольной системе координат $Oxyz$ определяется уравнением

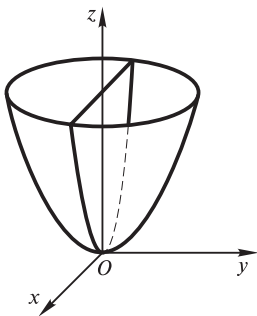


Рис. 65

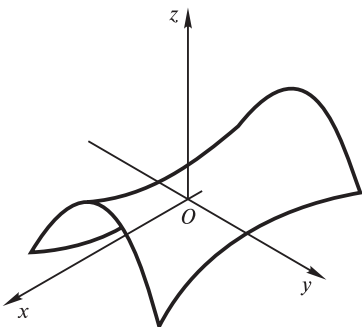


Рис. 66

$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2}, \quad (6)$$

а *гиперболическим параболоидом* (рис. 66) — поверхность, определяемая уравнением

$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2}. \quad (7)$$

Уравнения (6) и (7) называются *каноническими уравнениями параболоидов*.

Плоскости xOz и yOz являются плоскостями симметрии параболоидов. Пересечение этих плоскостей (ось Oz) называется *осью параболоида*, а пересечение оси Oz с поверхностью параболоида — *вершиной*.

Оба параболоида (эллиптический и гиперболический) плоскостями, параллельными координатным плоскостям xOz и yOz , пересекаются по параболам. Так, плоскость $x = h$ пересекает эллиптический параболоид по параболе

$$z - \frac{h^2}{a^2} = \frac{y^2}{b^2}, \quad x = h.$$

Из уравнения (6) следует, что плоскость $z = h$ ($h > 0$), параллельная плоскости xOy , пересекает эллиптический параболоид по эллипсу, а из уравнения (7) следует, что плоскость $z = h$ ($h \neq 0$) пересекает гиперболический параболоид по гиперболе. Плоскость xOy пересекает гиперболический параболоид по двум прямым.

При $a = b$ эллиптический параболоид называется *параболоидом вращения*. Он получается при вращении параболы $z = \frac{x^2}{a^2}, y = 0$ около оси Oz .

3. Цилиндры второго порядка. Цилиндры второго порядка или, кратко, цилиндры определяются в прямоугольной системе координат $Oxyz$ уравнениями:

$$а) \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (8)$$

(эллиптический цилиндр, в частности при $a = b$ круговой);

$$б) \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \text{ (гиперболический цилиндр);} \quad (9)$$

$$в) y^2 = 2px \text{ (параболический цилиндр).} \quad (10)$$

Уравнения (8)—(10) называются *каноническими уравнениями цилиндров*. Уравнения (8)—(10) не содержат переменной z . На плоскости xOy уравнение (8) определяет эллипс с полуосями a и b . Если точка $(x; y)$ лежит на этом эллипсе, то при любом z точка $(x; y; z)$ лежит на поверхности (8). Совокупность таких точек есть поверхность, описанная прямой, параллельной оси Oz и пересекающей эллипс

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

в плоскости xOy .

Этот эллипс называют *направляющей линией* данной поверхности, а все возможные положения движущейся прямой — *образующими*.

Вообще поверхность, описываемая прямой, остающейся параллельной некоторому заданному направлению и пересекающей данную линию L , называется *цилиндрической*. Поверхность (8) изображена на рисунке 67.

В случае гиперболического и параболического цилиндров ((9) и (10)) направляющими линиями поверхностей являются гипербола и парабола, а образующими — прямые, параллельные оси Oz и проходящие через гиперболу и параболу в плоскости xOy . Поверхности (9) и (10) изображены на рисунках 68 и 69.

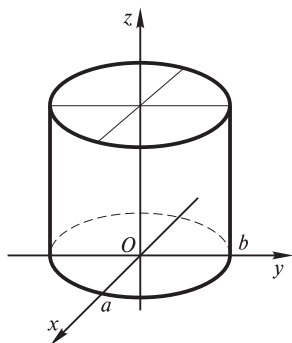


Рис. 67

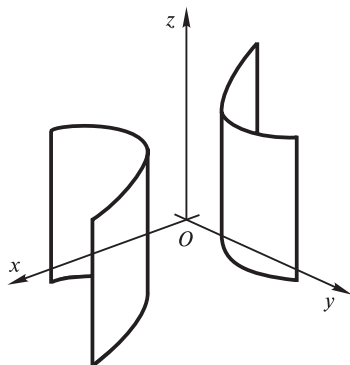


Рис. 68

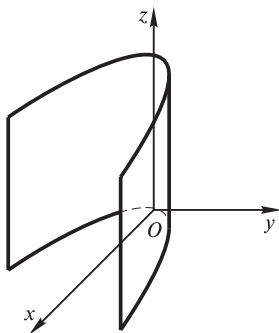


Рис. 69

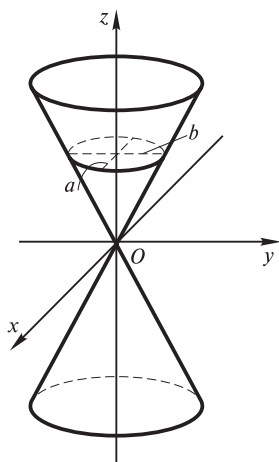


Рис. 70

4. Конус второго порядка. Конусом второго порядка или, кратко, конусом (рис. 70) называется поверхность, определяемая в прямоугольной системе координат $Oxyz$ уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0. \quad (11)$$

Уравнение (11) называется *каноническим уравнением конуса*. Эта поверхность симметрична относительно координатных плоскостей. Начало координат, являющееся центром симметрии, принадлежит этой поверхности и называется *вершиной* конуса. Сечениями конуса плоскостями $x = 0$ и $y = 0$ являются прямые $z = \pm \frac{c}{b}y$ и $z = \pm \frac{c}{a}x$.

В плоскости $z = h$ ($h \neq 0$) имеем эллипс $\frac{x^2}{a_*^2} + \frac{y^2}{b_*^2} = 1$ с полуосями $a_* = \frac{a|h|}{c}$, $b_* = \frac{b|h|}{c}$. Если $a = b$, то конус называется *конусом вращения*. Для конуса вращения в плоскости $z = h$ ($h \neq 0$) имеем окружность $x^2 + y^2 = a_*^2$.

5. Прямолинейные образующие поверхностей второго порядка. Поверхность, образованная движением прямой, называется *линейчатой*, а лежащие на ней прямые — *прямолинейными образующими*.

Примерами таких поверхностей являются рассмотренные выше цилиндры и конус. Среди поверхностей второго порядка прямолинейными образующими обладают (кроме

цилиндров и конуса) однополостный гиперболоид и гиперболический параболоид. Проведем доказательство для однополостного гиперболоида.

Представим каноническое уравнение однополостного гиперболоида

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \text{ в виде } \frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{y^2}{b^2},$$

$$\text{или } \left(\frac{x}{a} + \frac{z}{c}\right)\left(\frac{x}{a} - \frac{z}{c}\right) = \left(1 + \frac{y}{b}\right)\left(1 - \frac{y}{b}\right). \quad (12)$$

Составим систему уравнений первой степени

$$\begin{cases} \frac{x}{a} + \frac{z}{c} = k \left(1 + \frac{y}{b}\right), \\ \frac{x}{a} - \frac{z}{c} = \frac{1}{k} \left(1 - \frac{y}{b}\right), \end{cases} \quad (13)$$

где k — произвольное число ($k \neq 0$).

При определенном значении k эти уравнения определяют прямую линию. Меняя параметр k , мы получим совокупность прямых (семейство прямых). Если мы перемножим уравнения (13) почленно, то получим уравнение (12). Поэтому любая точка $(x; y; z)$, координаты которой удовлетворяют системе (13), лежит на поверхности (12). Таким образом, каждая из прямых семейства (13) целиком лежит на поверхности однополостного гиперболоида.

Аналогично система

$$\begin{cases} \frac{x}{a} + \frac{z}{c} = l \left(1 - \frac{y}{b}\right), \\ \frac{x}{a} - \frac{z}{c} = \frac{1}{l} \left(1 + \frac{y}{b}\right), \end{cases} \quad (14)$$

где l — произвольный параметр ($l \neq 0$) определяет семейство прямых, отличное от семейства (13), принадлежащее поверхности (12).

Кроме того, можно доказать (см., например, [1]), что через каждую точку однополостного гиперболоида проходит по одной прямой из каждого из этих семейств (одно из них изображено на рисунке 71). Например, через точку $\left(\sqrt{\frac{3}{2}}a; \frac{\sqrt{2}}{2}b; c\right)$ поверхности (12)

проходит прямая (13) при $k = \frac{2 + \sqrt{6}}{2 + \sqrt{2}}$ и прямая (14) при $l = \frac{2 + \sqrt{6}}{2 - \sqrt{2}}$.

На поверхности гиперболического параболоида

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z$$

также располагаются два семейства прямолинейных образующих (одно из них изображено на рисунке 72). Их уравнения:

$$\begin{aligned} \frac{x}{a} + \frac{y}{b} &= kz, & \frac{x}{a} - \frac{y}{b} &= \frac{1}{k}, \\ \frac{x}{a} + \frac{y}{b} &= \frac{1}{l}, & \frac{x}{a} - \frac{y}{b} &= lz. \end{aligned}$$

Через каждую точку этой поверхности проходит по одной прямой каждого семейства.

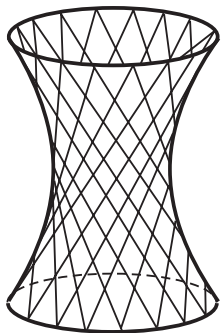


Рис. 71

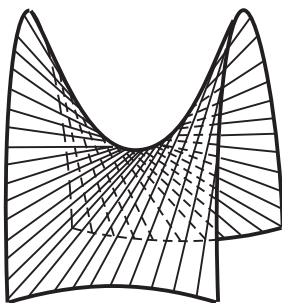


Рис. 72

Отметим, что однополостные гиперboloиды нашли применение в практике строительства. Сооружение различных высотных башен с использованием прямолинейных образующих однополостного гиперboloида сочетает в себе прочность конструкции с простотой ее исполнения. Идея использования однополостного гиперboloида в строительстве принадлежит знаменитому русскому инженеру Владимиру Григорьевичу Шухову (1853—1939). По проекту Шухова построена телевизионная башня в Москве, она состоит из секций однополостных гиперboloидов вращения.

Глава 6. ОБЩЕЕ УРАВНЕНИЕ КРИВОЙ ВТОРОГО ПОРЯДКА

§ 6.1. Приведение матрицы квадратичной формы к диагональному виду

1. Определение квадратичной формы и ее матрицы. *Квадратичной формой от двух переменных x_1 и x_2 называется однородный многочлен второй степени относительно этих двух переменных:*

$$F = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2. \quad (1)$$

Здесь a_{11} , a_{12} , a_{22} — числа, задание которых определяет форму; их называют *коэффициентами формы*. (Двойка перед a_{12} поставлена для упрощения получающихся формул.)

Покажем, как квадратичную форму (1) можно записать в матричной форме. Прежде всего, полагая $a_{12} = a_{21}$, запишем ее в виде

$$F = (a_{11}x_1 + a_{12}x_2)x_1 + (a_{21}x_1 + a_{22}x_2)x_2.$$

Матрица $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ называется *матрицей квадратичной формы* (1).

Введя матрицу-столбец $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ и матрицу-строку $X^* = (x_1 \ x_2)$, можно убедиться, что

$$F = X^*AX. \quad (2)$$

Действительно, по правилу умножения матриц (§ 3.1, п. 5) последовательно находим:

$$AX = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix},$$

$$X^*AX = (x_1 \ x_2) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix} = x_1(a_{11}x_1 + a_{12}x_2) + x_2(a_{21}x_1 + a_{22}x_2) = F.$$

Будем трактовать переменные x_1 и x_2 как координаты точек в прямоугольной системе координат x_1Ox_2 . Рассмотрим новую прямоугольную систему координат $x_1'Ox_2'$. Пусть координаты точек в старой и новой системах связаны между собой формулами преобразования (15) из § 3.2 (п. 5)

$$x_1 = \alpha_{11}x_1' + \alpha_{12}x_2', \quad x_2 = \alpha_{21}x_1' + \alpha_{22}x_2' \quad (3)$$

с ортогональной матрицей преобразования

$$L = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}.$$

Как установлено в § 3.2 (п. 5), формулы преобразования (3) можно записать в следующей матричной форме:

$$X = LX'. \quad (4)$$

Здесь

$$X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad X' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}.$$

Если вместо x_1 и x_2 в квадратичную форму (1) подставить их выражения через x_1' и x_2' согласно (3), то получим квадратичную форму от переменных x_1' и x_2' .

2. Приведение квадратичной формы к каноническому виду. Поставим задачу выбрать новую систему координат $x_1'Ox_2'$ так, чтобы в квадратичной форме от переменных x_1' , x_2' отсутствовал член с произведением координат, т. е. чтобы она приняла вид

$$F = \lambda_1 x_1'^2 + \lambda_2 x_2'^2, \quad (5)$$

который называется *каноническим*.

Для сокращения записи преобразования будем проводить в матричной форме. Рассмотрим матрицу-строку

$$X^{**} = (x_1' \ x_2').$$

Имеет место равенство

$$X^* = X^{**}L^{-1}. \quad (6)$$

Действительно, в § 3.2 (п. 5, (20)) установлено:

$$L^{-1} = L^* = \begin{pmatrix} \alpha_{11} & \alpha_{21} \\ \alpha_{12} & \alpha_{22} \end{pmatrix}.$$

Поэтому

$$X^{**}L^{-1} = (x_1' \ x_2') \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = (\alpha_{11}x_1' + \alpha_{12}x_2' \ \alpha_{21}x_1' + \alpha_{22}x_2'),$$

откуда в силу равенств (3) $X^*L^{-1} = (x_1 \ x_2) = X^*$.

Подставляя в правую часть равенства (2) выражения для X^* и X из равенств (6) и (4), будем иметь:

$$F = X^*AX = (X^*L^{-1})A(LX) = X^*(L^{-1}AL)X.$$

Отсюда следует, что в новой системе координат матрица квадратичной формы $F = X^*(L^{-1}AL)X$ имеет вид $A' = L^{-1}AL$.

Выберем теперь новую систему координат x'_1, x'_2 так, чтобы матрица A' приняла следующую форму:

$$A' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

В этом случае говорят, что матрица приведена к *диагональному виду*. При этом квадратичная форма F в переменных x'_1 и x'_2 запишется в виде (5).

Итак, новую систему координат надо выбрать так, чтобы матрица L преобразования удовлетворяла соотношению

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = L^{-1}AL.$$

Умножим обе части этого равенства слева на матрицу L :

$$L \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = LL^{-1}AL = EAL = AL.$$

Отсюда

$$\begin{pmatrix} \alpha_{11}\lambda_1 & \alpha_{12}\lambda_2 \\ \alpha_{21}\lambda_1 & \alpha_{22}\lambda_2 \end{pmatrix} = \begin{pmatrix} a_{11}\alpha_{11} + a_{12}\alpha_{21} & a_{11}\alpha_{12} + a_{12}\alpha_{22} \\ a_{21}\alpha_{11} + a_{22}\alpha_{21} & a_{21}\alpha_{12} + a_{22}\alpha_{22} \end{pmatrix}$$

и, значит,

$$\begin{cases} (a_{11} - \lambda_1)\alpha_{11} + a_{12}\alpha_{21} = 0, \\ a_{21}\alpha_{11} + (a_{22} - \lambda_1)\alpha_{21} = 0 \end{cases} \quad (7)$$

и

$$\begin{cases} (a_{11} - \lambda_2)\alpha_{12} + a_{12}\alpha_{22} = 0, \\ a_{21}\alpha_{12} + (a_{22} - \lambda_2)\alpha_{22} = 0. \end{cases} \quad (8)$$

Таким образом, неизвестные коэффициенты преобразования α_{11} , α_{12} , α_{21} , α_{22} находятся из систем уравнений (7) и (8). Каждая из этих систем является однородной. Для того чтобы они имели отличные от нуля решения, необходимо и достаточно (см. § 3.4, п. 3), чтобы определитель каждой из этих систем был равен нулю.

Таким образом, числа λ_1 и λ_2 являются корнями уравнения

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0, \quad (9)$$

или

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0. \quad (9')$$

Дискриминант D этого квадратного уравнения всегда неотрицателен. Действительно,

$$\begin{aligned} D &= (a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21}) = a_{11}^2 + a_{22}^2 - 2a_{11}a_{22} + 4a_{12}a_{21} = \\ &= (a_{11} - a_{22})^2 + 4a_{12}^2, \end{aligned}$$

так как $a_{12} = a_{21}$ (п. 1) Итак, уравнение (9) (или (9')) всегда имеет действительные корни. Оно называется *характеристическим* уравнением матрицы A . Корни этого уравнения называются *собственными* значениями матрицы A . Подставляя найденные из уравнения (9) значения λ_1 и λ_2 в системы (7), (8) и решая их, найдем коэффициенты преобразования координат $\alpha_{12}, \alpha_{12}, \alpha_{21}, \alpha_{22}$.

П р и м е р. Привести к каноническому виду квадратичную форму

$$F = 5x_1^2 + 4x_1x_2 + 2x_2^2. \quad (10)$$

Здесь $a_{11} = 5, a_{12} = a_{21} = 2, a_{22} = 2$. Матрица данной квадратичной формы имеет вид:

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}.$$

Составим характеристическое уравнение $\lambda^2 - 7\lambda + 6 = 0$, откуда $\lambda_1 = 1, \lambda_2 = 6$. Поэтому заданная квадратичная форма приводится к каноническому виду

$$F = x_1'^2 + 6x_2'^2.$$

Если числа λ_1 и λ_2 одного знака, то будем говорить, что квадратичная форма (1) принадлежит *эллиптическому типу*; если λ_1 и λ_2 разных знаков, то *гиперболическому типу*; если же одно из чисел λ_1 или λ_2 равно нулю, то *параболическому типу*.

Из (9') видно, что $\lambda_1\lambda_2 = a_{11}a_{22} - a_{12}^2$. Поэтому квадратичная форма (1) будет эллиптической, гиперболической или параболической, если выражение $a_{11}a_{22} - a_{12}^2$ соответственно больше, меньше или равно нулю.

Квадратичная форма (10) будет эллиптического типа.

§ 6.2. Общее уравнение кривой второго порядка, его приведение к каноническому виду

1. Упрощение общего уравнения кривой второго порядка. Уравнение

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0, \quad (1)$$

где коэффициенты $a_{11}, 2a_{12}, a_{22}, 2a_{13}, 2a_{23}, a_{33}$ — вещественные постоянные, причем $a_{11}^2 + a_{12}^2 + a_{22}^2 > 0$, называется *общим уравнением кривой второго порядка*. Первые три члена уравнения (1), т. е. члены второй степени, называются ее *старшими членами*. Коэффициенты при xy, x и y обозначены соответственно через $2a_{12}, 2a_{13}$ и $2a_{23}$ для удобства преобразований уравнения (1) в дальнейшем.

Важнейшей задачей аналитической геометрии является исследование общего уравнения кривой второго порядка и приведение его к простейшим (каноническим) формам.

Применим преобразование квадратичной формы, описанное в предыдущем пункте, к упрощению общего уравнения кривой второго порядка.

Переход к новой декартовой системе координат с тем же началом, как показано в § 3.2 (п. 5), сводится к замене переменных

$$x = \alpha_{11}x' + \alpha_{12}y', \quad y = \alpha_{21}x' + \alpha_{22}y' \quad (2)$$

с ортогональной матрицей перехода L . При подстановке этих выражений в уравнение (1) группа членов второй степени и группа членов первой степени преобразуются

независимо друг от друга. Если следить сначала за группой членов второй степени (квадратичной формой), то на основании предыдущего пункта всегда можно выбрать систему координат $x'Oy'$ так, что эта группа членов приобретает канонический вид. Поэтому все уравнение (1) после преобразования будет иметь вид:

$$\lambda_1 x'^2 + \lambda_2 y'^2 + 2\mu_1 x' + 2\mu_2 y' + a_{33} = 0. \quad (3)$$

Здесь λ_1 и λ_2 — корни уравнения

$$\lambda^2 + (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0,$$

а μ_1 и μ_2 — некоторые новые коэффициенты при членах первой степени, которые сами получаются после подстановки выражений (2).

2. Упрощение уравнения $\lambda_1 x'^2 + \lambda_2 y'^2 + 2\mu_1 x' + 2\mu_2 y' + a_{33} = 0$. Упрощение уравнения (3) производится в зависимости от типа квадратичной формы старших членов.

Введем обозначение: $\delta = a_{11}a_{22} - a_{12}^2$.

1. Предположим, что $\delta \neq 0$. Так как $\delta = \lambda_1 \lambda_2$ (§ 6.1, п. 2), то $\lambda_1 \neq 0$ и $\lambda_2 \neq 0$, и поэтому можно уравнение (3) переписать так:

$$\lambda_1 \left(x' + \frac{\mu_1}{\lambda_1} \right)^2 + \lambda_2 \left(y' + \frac{\mu_2}{\lambda_2} \right)^2 = -a_{33} + \frac{\mu_1^2}{\lambda_1} + \frac{\mu_2^2}{\lambda_2}. \quad (4)$$

Выполним параллельный перенос системы повернутых осей на величину $-\frac{\mu_1}{\lambda_1}$ в направлении оси Ox' и на величину $-\frac{\mu_2}{\lambda_2}$ в направлении оси Oy' ; переход к новым координатам дают формулы

$$x' = x'' - \frac{\mu_1}{\lambda_1}, \quad y' = y'' - \frac{\mu_2}{\lambda_2}.$$

Если мы обозначим правую часть (4) буквой H , то уравнение данной линии в последних координатах примет канонический вид

$$\lambda_1 x''^2 + \lambda_2 y''^2 = H. \quad (5)$$

Возможны два случая:

1) λ_1 и λ_2 одного знака; будем считать, что $\lambda_1 > 0$, $\lambda_2 > 0$ (в противном случае изменим знаки всех членов уравнения). Если $H \neq 0$, то уравнение (5) приводится к виду

$$\frac{x''^2}{\frac{H}{\lambda_1}} + \frac{y''^2}{\frac{H}{\lambda_2}} = 1.$$

Отсюда видно, что при $H > 0$ уравнение (5) определяет эллипс с полуосями $a = \sqrt{\frac{H}{\lambda_1}}$, $b = \sqrt{\frac{H}{\lambda_2}}$.

Если $H = 0$, то уравнение (5) определяет единственную вещественную точку: $x'' = 0$, $y'' = 0$. Однако в этом случае принято говорить, что уравнение (5) есть уравнение *вырожденного эллипса*.

Если $H < 0$ (при $\lambda_1 > 0$, $\lambda_2 > 0$), то уравнение (5) никакого вещественного образа не определяет. В этом случае говорят также, что уравнение (5) есть уравнение *мнимого эллипса*.

2) Числа λ_1, λ_2 разных знаков; будем считать, что $\lambda_1 > 0, \lambda_2 < 0$. Если при этом $H > 0$, то уравнение (5) определяет гиперболу, которая пересекает ось абсцисс и имеет полуоси $a = \sqrt{\frac{H}{\lambda_1}}, b = \sqrt{-\frac{H}{\lambda_2}}$. Если $H < 0$, то также получается гипербола, но пересекающая ось ординат. Если $H = 0$, то уравнение (5) определяет пару прямых, проходящих через начало координат, так как оно приводится к виду

$$(\sqrt{\lambda_1}x'' + \sqrt{-\lambda_2}y'')(\sqrt{\lambda_1}x'' - \sqrt{-\lambda_2}y'') = 0.$$

В этом случае говорят, что уравнение (5) определяет *вырожденную гиперболу*.

Возвращаясь теперь к исходному уравнению (1), напомним, что $\delta = a_{11}a_{22} - a_{12}^2 = \lambda_1\lambda_2$. Поэтому если $\delta > 0$, то λ_1, λ_2 — числа одного знака; если $\delta < 0$, то λ_1, λ_2 — числа разных знаков. Отсюда заключаем: *если данное уравнение (1) имеет $\delta = a_{11}a_{22} - a_{12}^2 > 0$, то оно определяет эллипс (действительный, мнимый или вырожденный); если $\delta < 0$, то уравнение определяет гиперболу (настоящую или вырожденную).*

2. Рассмотрим теперь уравнение (1) при условии $\delta = 0$. Так как $\delta = \lambda_1\lambda_2$, то на этот раз одно из чисел λ_1, λ_2 равно нулю. Будем считать, что $\lambda_1 \neq 0, \lambda_2 = 0$. Уравнение (3) можно переписать в виде

$$\lambda_1 \left(x' + \frac{\mu_1}{\lambda_1} \right)^2 + 2\mu_2 y' + \left(a_{33} - \frac{\mu_1^2}{\lambda_1} \right) = 0.$$

Обозначая величину, стоящую в последней скобке, одной буквой k , получим:

$$\lambda_1 \left(x' + \frac{\mu_1}{\lambda_1} \right)^2 + 2\mu_2 y' + k = 0. \quad (6)$$

Дальнейшее преобразование уравнения (6) зависит от μ_2 .

1) Если $\mu_2 \neq 0$, то уравнение (6) можно переписать так:

$$\lambda_1 \left(x' + \frac{\mu_1}{\lambda_1} \right)^2 + 2\mu_2 \left(y' + \frac{k}{2\mu_2} \right) = 0.$$

Выполним параллельный перенос системы осей Ox', Oy' :

$$x' = x'' - \frac{\mu_1}{\lambda_1}, \quad y' = y'' - \frac{k}{2\mu_2};$$

получим $\lambda_1 x''^2 + 2\mu_2 y'' = 0$, или

$$x''^2 = 2py'', \quad (7)$$

где $p = -\frac{\mu_2}{\lambda_1}$.

Уравнение (7) есть каноническое уравнение параболы.

2) Если $\mu_2 = 0$, то, выполняя параллельный перенос $x' = x'' - \frac{\mu_1}{\lambda_1}, y' = y''$, уравнение (6) примет вид:

$$\lambda_1 x''^2 + k = 0. \quad (8)$$

Предположим, что $\lambda_1 > 0$. Тогда если $k < 0$, то уравнение (8) определяет пару параллельных прямых:

$$\sqrt{\lambda_1}x'' + \sqrt{-k} = 0, \quad \sqrt{\lambda_1}x'' - \sqrt{-k} = 0 \quad (9)$$

(например, уравнение $4x''^2 - 9 = 0$ определяет пару параллельных прямых $2x'' + 3 = 0, 2x'' - 3 = 0$). Если $k = 0$, то прямые (9) сливаются. Если $k > 0$, то уравнение (8) никакого

вещественного образа не определяет. Однако в этом случае говорят, что уравнение (8) определяет *пару мнимых параллельных прямых* (при $k > 0$ в уравнениях (9) число $\sqrt{-k}$ будет мнимым, см. § 7.7). Вообще уравнение (8) называют уравнением *вырожденной параболы*.

Возвращаясь к исходному уравнению (1), приходим к следующему заключению: *если $\delta = a_{11}a_{22} - a_{12}^2 = 0$, то уравнение (1) является уравнением параболы* (обыкновенной или вырожденной).

3. Подводя итог всему исследованию, приходим к следующему общему выводу: *каждое уравнение второй степени $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0$ определяет либо эллипс (если $\delta = a_{11}a_{22} - a_{12}^2 > 0$), либо гиперболу (если $\delta < 0$), либо параболу (если $\delta = 0$)*. При этом следует учитывать мнимый и вырожденный эллипсы, а также вырожденные гиперболы и параболы.

П р и м е р. Привести к каноническому виду общее уравнение кривой второго порядка $5x^2 + 8xy + 5y^2 - 18x - 18y + 9 = 0$.

Р е ш е н и е. Квадратичная форма, составленная из старших членов данного уравнения, имеет вид: $F = 5x^2 + 8xy + 5y^2$.

Здесь $a_{11} = 5$, $a_{12} = a_{21} = 4$, $a_{22} = 5$, матрица

$$A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}.$$

Так как $\delta = a_{11}a_{22} - a_{12}^2 = 5 \cdot 5 - 4^2 = 9 > 0$, то данная кривая определяет эллипс.

Составим характеристическое уравнение (см. § 6.1, п. 2, (9')):

$$\lambda^2 - 10\lambda + 9 = 0.$$

Корни этого уравнения $\lambda_1 = 9$, $\lambda_2 = 1$. Значит (§ 6.1, п. 2), квадратичная форма F в новой системе координат $x'Oy'$ запишется в виде

$$F = \lambda_1 x'^2 + \lambda_2 y'^2 = 9x'^2 + y'^2.$$

Найдем матрицу $L = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$ перехода от старой системы координат xOy к новой

$x'Oy'$. Для этого составим системы уравнений:

$$\begin{cases} (5-9)\alpha_{11} + 4\alpha_{21} = 0, \\ 4\alpha_{11} + (5-9)\alpha_{21} = 0, \\ (5-1)\alpha_{12} + 4\alpha_{22} = 0, \\ 4\alpha_{12} + (5-1)\alpha_{22} = 0. \end{cases}$$

Каждая из этих систем сводится к одному уравнению: первая система сводится к уравнению $\alpha_{11} = \alpha_{21}$, а вторая — к уравнению $\alpha_{22} = -\alpha_{12}$. Матрица L является ортогональной (см. § 3.2, п. 5). Поэтому должны иметь место равенства

$$\alpha_{11}^2 + \alpha_{21}^2 = 1 \text{ и } \alpha_{12}^2 + \alpha_{22}^2 = 1.$$

Следовательно,

$$\alpha_{11} = \alpha_{21} = \frac{1}{\pm\sqrt{2}} \text{ и } \alpha_{12} = -\frac{1}{\pm\sqrt{2}}, \quad \alpha_{22} = \frac{1}{\pm\sqrt{2}}.$$

Выбирая для определенности перед корнем знак «+», получим:

$$\alpha_{11} = \frac{1}{\sqrt{2}}, \quad \alpha_{21} = \frac{1}{\sqrt{2}}, \quad \alpha_{12} = -\frac{1}{\sqrt{2}}, \quad \alpha_{22} = \frac{1}{\sqrt{2}}.$$

Итак, формулы преобразования координат в данном случае принимают вид:

$$x = \frac{1}{\sqrt{2}}(x' - y'), \quad y = \frac{1}{\sqrt{2}}(x' + y')$$

и потому

$$\begin{aligned} -18x - 18y + 9 &= -\frac{18}{\sqrt{2}}(x' - y') - \\ &- \frac{18}{\sqrt{2}}(x' + y') + 9 = -\frac{36}{\sqrt{2}}x' + 9 \end{aligned}$$

Таким образом, в новой системе координат $x'Oy'$ уравнение данной кривой запишем в виде

$$9x'^2 + y'^2 - \frac{36x'}{\sqrt{2}} + 9 = 0, \quad \text{или в виде} \quad \frac{(x' - \sqrt{2})^2}{1} + \frac{y'^2}{9} = 1.$$

Таким образом, данная линия является эллипсом, центр которого в новой системе координат находится в точке $O_1(\sqrt{2}; 0)$. Для того чтобы установить расположение эллипса относительно старой системы координат, надо определить положение новых осей относительно старой системы. Для этого достаточно установить углы между ортами $\vec{e}_1$ и $\vec{e}_2$ старой системы и $\vec{e}'_1$ и $\vec{e}'_2$ новой системы координат. По формулам (14) § 3.2 (п. 5) находим:

$$\cos(\widehat{\vec{e}_1, \vec{e}'_1}) = \alpha_{11} = \frac{1}{\sqrt{2}}, \quad \cos(\widehat{\vec{e}_2, \vec{e}'_1}) = \alpha_{21} = \frac{1}{\sqrt{2}},$$

$$\cos(\widehat{\vec{e}_1, \vec{e}'_2}) = \alpha_{12} = -\frac{1}{\sqrt{2}}, \quad \cos(\widehat{\vec{e}_2, \vec{e}'_2}) = \alpha_{22} = \frac{1}{\sqrt{2}}.$$

Следовательно, углы, которые оси новой системы образуют с осями старой, таковы:

$$(\widehat{\vec{e}_1, \vec{e}'_1}) = \frac{\pi}{4}, \quad (\widehat{\vec{e}_1, \vec{e}'_2}) = \frac{3\pi}{4}, \quad (\widehat{\vec{e}_2, \vec{e}'_1}) = \frac{\pi}{4}, \quad (\widehat{\vec{e}_2, \vec{e}'_2}) = \frac{\pi}{4}.$$

Расположение осей и эллипса приведено на рисунке 73.

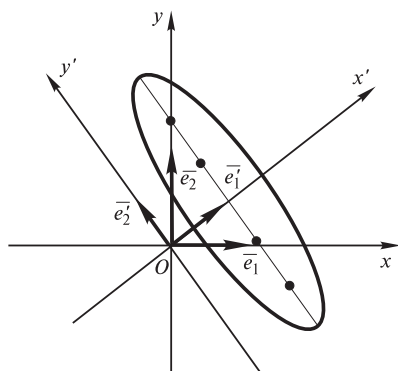


Рис. 73

§ 6.3. Инварианты кривых второго порядка

1. Преобразование коэффициентов уравнения линии второго порядка при параллельном переносе. Пусть декартова прямоугольная система координат $x'O'y'$ получена параллельным переносом системы xOy . Как известно (см. 3.2, п. 5), старые и новые координаты точки M связаны соотношениями

$$x = x' + x_0, \quad y = y' + y_0,$$

где x_0, y_0 — координаты начала O' в системе xOy . Подставляя эти выражения для x и y в левую часть уравнения

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0, \quad (1)$$

получим:

$$a_{11}x'^2 + 2a_{12}x'y' + a_{22}y'^2 + 2a'_{13}x' + 2a'_{23}y' + a'_{33} = 0, \quad (1')$$

где

$$\begin{aligned} a'_{13} &= a_{11}x_0 + a_{12}y_0 + a_{13}, & a'_{23} &= a_{12}x_0 + a_{22}y_0 + a_{23}, \\ a'_{33} &= a_{11}x_0^2 + 2a_{12}x_0y_0 + a_{22}y_0^2 + 2a_{13}x_0 + 2a_{23}y_0 + a_{33} \end{aligned} \quad (2)$$

Итак, при параллельном переносе системы координат коэффициенты группы старших членов не изменяются, а остальные коэффициенты преобразуются по формулам (2).

П р и м е ч а н и е. Используя первую и вторую из формул (2), можно, очевидно, выражению для a'_{33} придать следующий вид:

$$a'_{33} = (a'_{13} + a_{13})x_0 + (a'_{23} + a_{23})y_0 + a_{33}. \quad (3)$$

2. Преобразование коэффициентов при повороте. Пусть декартова прямоугольная система координат $x'Oy'$ получена поворотом системы xOy на угол φ (при этом не исключается поворот на угол φ , равный нулю). Как известно (§ 3.2, п. 5), старые и новые координаты точки связаны соотношениями

$$x = x' \cos \varphi - y' \sin \varphi, \quad y = x' \sin \varphi + y' \cos \varphi.$$

Подставляя эти выражения для x и y в левую часть уравнения (1) и группируя коэффициенты при различных степенях x' и y' , получим:

$$a'_{11}x'^2 + 2a'_{12}x'y' + a'_{22}y'^2 + 2a'_{13}x' + 2a'_{23}y' + a_{33} = 0,$$

т. е. при повороте системы координат свободный член не изменяется (остальные коэффициенты преобразуются по соответствующим формулам; см., например, [6]).

3. Инварианты уравнения линий второго порядка.

Т е о р е м а. Величины

$$I_1 = a_{11} + a_{22}, \quad I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}, \quad I_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} \quad (4)$$

не меняются при любом преобразовании уравнения (1) к новым прямоугольным координатам (поэтому их называют инвариантами уравнения (1) относительно преобразования прямоугольных координат).

Д о к а з а т е л ь с т в о. Очевидно, инвариантность величин I_1, I_2, I_3 потребует доказать отдельно для параллельного переноса системы координат и для поворота.

Ради краткости рассмотрим лишь параллельный перенос. При этом преобразовании координат коэффициенты группы старших членов не изменяются (см. п. 1). Поэтому не изменяются и величины I_1 и I_2 . Займемся вычислением I_3 . В новой системе координат $x'Oy'$ величина I_3 равна

$$\begin{vmatrix} a_{11} & a_{12} & a'_{13} \\ a_{12} & a_{22} & a'_{23} \\ a'_{13} & a'_{23} & a'_{33} \end{vmatrix}.$$

Вычитая из последней строки этого определителя первую строку, умноженную на x_0 , и вторую, умноженную на y_0 (x_0 и y_0 — координаты нового начала O'), и используя при этом

выражения для a'_{13} и a'_{23} из формул (2) и выражение для a'_{33} из формулы (3), найдем, что этот определитель равен

$$\begin{vmatrix} a_{11} & a_{12} & a'_{13} \\ a_{12} & a_{22} & a'_{23} \\ a_{13} & a_{23} & a_{13}x_0 + a_{23}y_0 + a_{33} \end{vmatrix}.$$

Напомним, что при этих преобразованиях значение определителя не меняется (см. § 3.2, п. 2).

Если теперь вычесть из последнего столбца полученного определителя первый столбец, умноженный на x_0 , и второй, умноженный на y_0 , и использовать при этом выражения для a'_{13} и a'_{23} из формул (2), то в результате получится определитель, стоящий в правой части выражения для I_3 в формулах (4).

§ 6.4. Уравнение центра. Вырождение кривых второго порядка

1. Уравнение центра. Некоторую точку S называют *центром* данной линии второго порядка, если после переноса начала координат в точку S уравнение этой линии не будет содержать членов первой степени. Левая часть такого уравнения не меняется при одновременном изменении знаков текущих координат. Это означает, что точки линии расположены парами симметрично относительно точки S , т. е. что точка S является центром симметрии линии.

Пусть дана линия второго порядка

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0. \quad (1)$$

Требуется найти ее центр (если он есть).

Предполагая, что центр имеется, обозначим через x_0, y_0 искомые координаты центра S (в данной координатной системе). Перенесем начало координат в точку S . При этом координаты произвольной точки изменятся по формулам

$$x = x' + x_0, \quad y = y' + y_0,$$

где x', y' — новые координаты той же точки. Перейдем в уравнении (1) к новым координатам. Получим (см. § 6.3, п. 1):

$$a_{11}x'^2 + 2a_{12}x'y' + a_{22}y'^2 + 2a'_{13}x' + 2a'_{23}y' + a'_{33} = 0, \quad (2)$$

где a'_{13}, a'_{23} и a'_{33} определяются по формулам (2) из § 6.3. Точка S будет центром, если $a'_{13} = 0$ и $a'_{23} = 0$. Отсюда и из (2) (§ 6.3) получаем *уравнения центра*:

$$\begin{cases} a_{11}x_0 + a_{12}y_0 + a_{13} = 0, \\ a_{12}x_0 + a_{22}y_0 + a_{23} = 0. \end{cases} \quad (3)$$

Решая их совместно, найдем центр $S(x_0, y_0)$. Система (3) может оказаться несовместной, тогда центра у данной линии нет.

Определитель системы (3) есть

$$I_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}. \quad (4)$$

Если $I_2 \neq 0$, то система (3) совместна и имеет единственное решение. Следовательно, если $I_2 \neq 0$, то данная линия имеет единственный центр. Такая линия второго порядка называется *центральной*. Координаты центра выражаются формулами

$$x_0 = \frac{\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}}{I_2}, \quad y_0 = \frac{\begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{12} \end{vmatrix}}{I_2}. \quad (5)$$

Отсюда и из (2) § 6.3 найдем a'_{33} . Вычисление a'_{33} можно значительно упростить, если выразить a'_{33} со следующей группировкой членов:

$$a'_{33} = (a_{11}x_0 + a_{12}y_0 + a_{13})x_0 + (a_{12}x_0 + a_{22}y_0 + a_{23})y_0 + a_{13}x_0 + a_{23}y_0 + a_{33}.$$

Тогда в силу системы (3) имеем: $a'_{33} = a_{13}x_0 + a_{23}y_0 + a_{33}$.

Отсюда, используя формулы (5) и определитель (4), найдем:

$$a'_{33} = \frac{a_{13} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + a_{23} \begin{vmatrix} a_{13} & a_{11} \\ a_{23} & a_{12} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}}{I_2} = \frac{I_3}{I_2}.$$

I_3 называется *дискриминантом* левой части общего уравнения (1).

Итак, если линия, заданная уравнением (1), является центральной ($I_2 \neq 0$), то после переноса начала координат в ее центр данное уравнение (1) (как следует из (2)) приводится к виду

$$a_{11}x'^2 + 2a_{12}x'y' + a_{22}y'^2 + \frac{I_3}{I_2} = 0 \quad (6)$$

(коэффициенты a_{11} , a_{12} , a_{22} прежние). Совершая теперь надлежащий поворот осей, можно привести уравнение (6) к каноническому виду:

$$\lambda_1 x''^2 + \lambda_2 y''^2 + \frac{I_3}{I_2} = 0.$$

Таким образом, свободный член H в уравнении (5) § 6.2 может быть подсчитан по данному уравнению (1) сразу $\left(H = -\frac{I_3}{I_2} \right)$ без преобразования координат. Тем самым может быть написано и все уравнение (5) § 6.2, поскольку числа λ_1 , λ_2 также непосредственно находятся по коэффициентам старших членов (см. § 6.2).

2. Вырождение кривых второго порядка. В § 6.2 установлено, что уравнение (5) из § 6.2 определяет вырожденную линию при $H=0$. Отсюда заключаем: *центральная линия второго порядка является вырожденной тогда и только тогда, когда $I_3 = 0$.*

Такое же условие характеризует и вырожденную параболу (см. [7]).

Таким образом, *уравнение (1) определяет вырожденную линию второго порядка тогда и только тогда, когда дискриминант его левой части равен нулю, т. е. $I_3 = 0$.*

Раздел II

МАТЕМАТИЧЕСКИЙ АНАЛИЗ

Глава 7. ВВЕДЕНИЕ В АНАЛИЗ

§ 7.1. Определение и способы задания функции

1. Действительные числа. Будем считать, что нам известны основные свойства *целых* чисел ($0, \pm 1, \pm 2, \dots$).

Число x называется *рациональным*, если его можно представить как частное двух целых чисел m и n ($n \neq 0$): $x = \frac{m}{n}$. Любое рациональное число x представимо в виде конечной или бесконечной периодической десятичной дроби.

Число x называется *иррациональным*, если оно представимо в виде бесконечной непериодической десятичной дроби $x = a_0, a_1 a_2 \dots a_n \dots$ (например, $\sqrt{2}, \sqrt{3}, \pi$). Каждое иррациональное число можно с любой заданной степенью точности приблизить рациональными числами; для этого достаточно брать в десятичном разложении этого числа конечное число знаков после запятой. Поэтому на практике при различных измерениях оперируют рациональными числами. Но в общих математических законах и формулах нельзя обойтись без иррациональных чисел (например, формула длины окружности $l = 2\pi R$ включает иррациональное число π).

Множество (совокупность) всех рациональных и иррациональных чисел называют *множеством действительных чисел*. Действительные числа изображаются на числовой оси Ox точками (рис. 74). При этом каждому действительному числу соответствует определенная точка числовой оси и каждой точке оси соответствует определенное действительное число. Поэтому вместо слов «действительное число» можно говорить «точка».

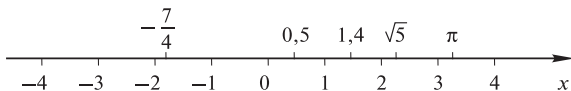


Рис. 74

Абсолютной величиной (или *модулем*) действительного числа x называется неотрицательное число $|x|$, определяемое соотношением

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

Непосредственно из определения абсолютной величины числа вытекают свойства 1 и 2:

1. $|-x| = |x|$.

2. $-|x| \leq x \leq |x|$.

3. *Неравенства $|x| \leq a$ и $-a \leq x \leq a$ равносильны.*

Докажем свойство 3. Из $|x| \leq a$ и свойства 2 имеем $x \leq a$. В то же время $|x| \leq a$ равносильно $-a \leq -|x|$, откуда с учетом свойства 2 следует $-a \leq x$. Таким образом, получаем $-a \leq x \leq a$. Обратно: из неравенства $-a \leq x \leq a$ вытекает, что одновременно $-x \leq a$ и $x \leq a$, т. е. по определению абсолютной величины $|x| \leq a$.

4. *Модуль суммы двух действительных чисел меньше или равен сумме модулей этих чисел: $|x+y| \leq |x| + |y|$.*

Действительно, если $x+y \geq 0$, то по определению абсолютной величины и свойству 2 $|x+y| = x+y \leq |x| + |y|$. Если $x+y < 0$, то $|x+y| = -(x+y) = -x + (-y) \leq |x| + |y|$.

Примечание 1. Свойство 4 справедливо для любого конечного числа слагаемых.

5. *Модуль разности двух действительных чисел больше или равен разности модулей этих чисел: $|x-y| \geq |x| - |y|$.*

По свойству 4 имеем:

$$|x| = |y + (x-y)| \leq |y| + |x-y|,$$

откуда

$$|x-y| \geq |x| - |y|.$$

6. *Модуль произведения двух действительных чисел равен произведению модулей этих чисел: $|xy| = |x||y|$.*

Примечание 2. Свойство 6 справедливо для любого конечного числа множителей.

7. *Модуль частного двух действительных чисел (если делитель отличен от нуля) равен частному модулей этих чисел:*

$$\left| \frac{x}{y} \right| = \frac{|x|}{|y|}.$$

Свойства 6, 7 непосредственно следуют из определения абсолютной величины числа.

Множество действительных чисел x , удовлетворяющих неравенствам $a \leq x \leq b$ ($a < x < b$), называется *сегментом* или *отрезком* (*интервалом*) и обозначается $[a; b]$ ($(a; b)$).

Полусегментом $[a; b)$ ($(a; b]$) называют множество действительных чисел x , удовлетворяющих неравенствам $a \leq x < b$ ($a < x \leq b$). Множества действительных чисел x , удовлетворяющих условию $x < a$ ($x \leq a$) или $x > b$ ($x \geq b$), обозначаются соответственно $(-\infty; a)$ ($(-\infty; a]$) или $(b; +\infty)$ ($[b; +\infty)$). Множество всех действительных чисел обозначается символом $(-\infty; +\infty)$ или $|x| < +\infty$. Все указанные множества называют *промежутками*. *Окрестностью* точки x_0 называют любой интервал, содержащий эту точку; ε — *окрестностью* ($\varepsilon > 0$) точки x_0 называется интервал $(x_0 - \varepsilon, x_0 + \varepsilon)$, т. е. множество чисел, удовлетворяющих условию $|x - x_0| < \varepsilon$.

2. Погрешности вычисления. Пусть некоторая величина имеет точное значение a . В результате измерения этой величины получено ее *приближенное* значение x . *Абсолютной погрешностью* Δ_0 приближенного значения x называется модуль разности между числом x и точным значением a : $\Delta_0 = |x - a|$.

Если число a неизвестно (что бывает в большинстве измерений), то абсолютную погрешность вычислить нельзя. В этом случае используется *предельная абсолютная погрешность* — положительное число Δ , такое, что $\Delta_0 \leq \Delta$. Очевидно, что $x - \Delta \leq a \leq x + \Delta$. Кратко последнее неравенство записывают так: $a = x \pm \Delta$.

П р и м е р 1. Если x_1 и x_2 — приближенные значения точного значения числа a , причем известно, что $x_1 \leq a \leq x_2$, то в этом случае можно положить $a = x \pm \Delta$, где

$$x = \frac{1}{2}(x_1 + x_2); \Delta = \frac{1}{2}(x_2 - x_1).$$

Точность измерения характеризуется с помощью *относительной погрешности*. *Относительной погрешностью* δ_0 приближенного значения x называется отношение абсолютной погрешности этого значения к модулю точного значения a : $\delta_0 = \frac{\Delta_0}{a}$.

Если точное значение a неизвестно, то используют предельную относительную погрешность — положительное число δ такое, что $\delta_0 \leq \delta$.

Для вычисления относительных погрешностей часто используются приближенные формулы

$$\delta_0 \approx \frac{\Delta_0}{|x|} \text{ и } \delta \approx \frac{\Delta}{|x|}.$$

Эти формулы тем точнее, чем ближе значение x к точному значению a , т. е. чем меньше погрешность Δ_0 или Δ .

Пример 2. Каковы предельные абсолютная и относительная погрешности числа 1,41 — приближенного значения числа $\sqrt{2}$? Так как $1,410 < \sqrt{2} < 1,415$, то $\Delta_0 = \sqrt{2} - 1,410 < 0,005$. Следовательно, можно положить $\Delta = 0,005$. Далее,

$$\delta_0 = \frac{\Delta_0}{\sqrt{2}} < \frac{0,005}{1,41} < 0,0036,$$

откуда $\delta = 0,0036$ или $\delta = 0,36\%$.

Говорят, что приближенное значение x (записанное в виде десятичной дроби) имеет n *верных знаков*, если абсолютная погрешность этого числа меньше или равна половине единицы его n -го разряда. Например, если 9,263 имеет 3 верных знака 9,2 и 6, то абсолютная погрешность этого числа $\Delta_0 \leq 0,005$.

3. Понятие функции. При изучении природных и технических процессов исследователи сталкиваются с величинами, одни из которых сохраняют одно и то же числовое значение — они называются *постоянными*, а другие принимают различные числовые значения и называются *переменными*. Примерами постоянных величин могут служить: температура кипения воды при нормальном давлении, скорость тела, движущегося равномерно и прямолинейно. Скорость камня, брошенного вверх, есть переменная величина: сначала она уменьшается, и, когда камень достигает наивысшей точки полета, скорость его становится равной нулю. Затем начинается свободное падение под действием силы тяжести и скорость камня увеличивается.

В практических задачах изменение переменной величины обычно связано с изменением одной или нескольких других переменных величин. Например, путь, пройденный телом с постоянной скоростью, прямо пропорционален времени движения: $s = vt$. Этой формулой выражена зависимость переменной s — пути, пройденного телом, от переменной t — времени движения. Видно, что переменные s и t не могут принимать произвольные значения независимо друг от друга. Придав определенное значение переменной t , мы тем самым единственным образом определим значение переменной s .

О п р е д е л е н и е. Если каждому значению, которое может принять переменная x , по некоторому правилу или закону ставится в соответствие одно определенное значение переменной y , то говорят, что y есть однозначная *функция* от x ; обозначают $y = f(x)$ (читается: «Игрек равно эф от икс»).

Используются и другие обозначения функции: $y = \varphi(x)$, $y = \psi(x)$, $y = \chi(x)$ и т. п.

Переменная x называется *независимой переменной* или *аргументом*.

Совокупность всех значений аргумента x , для которых функция $y = f(x)$ определена, называется *областью определения* этой функции.

Совокупность всех значений, принимаемых переменной y , называют *областью значений* функции $y = f(x)$.

П р и м е р. Найти область определения функции $y = \sqrt{4 - x^2}$. Эта функция имеет смысл, если $4 - x^2 \geq 0$. Отсюда $x^2 \leq 4$, или $|x| \leq 2$. Следовательно, область определения данной функции есть сегмент $[-2; 2]$. Множество значений этой функции есть сегмент $[0; 2]$.

4. Способы задания функции. *Аналитический* способ — это способ задания функции при помощи формул. Например, $y = 2x$, $y = x + 1$, $y = \lg x$, $y = \sin x$, $y = x^2$. Если уравнение, с помощью которого задается функция, не разрешено относительно y , то функция называется *неявной*. Когда такое решение возможно, неявная функция может быть приведена к явной форме, т. е. к виду $y = f(x)$. Например, уравнение $2x + 3y - 5 = 0$ можно рассматривать как неявно задающую функцию. Решив его относительно y , мы получаем ту же функцию, но уже в явном виде: $y = \frac{5 - 2x}{3}$.

Отметим, что при аналитическом способе задания функции встречаются случаи, когда функция задана не одной, а несколькими формулами, например:

$$y = \begin{cases} x^2, & \text{если } x \leq 0, \\ -x, & \text{если } x > 0. \end{cases}$$

Табличный способ — это способ задания функции при помощи таблицы. Примерами такого задания являются таблицы тригонометрических функций, логарифмов и т. п. Табличный способ задания функции широко используется в различного рода экспериментах и наблюдениях. Таблицы просты в обращении, для нахождения значения функции не надо производить вычисления. Недостатком табличного способа является то, что функция задается не для всех значений аргумента.

Графический способ — это способ задания функции при помощи графика. *Графиком* функции $y = f(x)$ называется множество точек $(x; y)$ плоскости xOy , координаты которых связаны соотношением $y = f(x)$. Само равенство $y = f(x)$ называется *уравнением* этого *графика*.

С построением графиков мы уже встречались в главе 1. Например, графиком функции $y = x$ является прямая (см. рис. 7).

Заметим, что если имеется график некоторой функции $y = f(x)$, то для нахождения значения $y = f(x_0)$, отвечающего какому-нибудь заданному значению x_0 , надо отложить это значение x_0 по оси абсцисс и из полученной точки восставить перпендикуляр до пересечения с графиком. Длина этого перпендикуляра, взятая с соответствующим

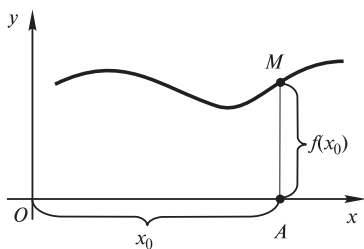


Рис. 75

знаком, и равна $f(x_0)$. Например, на рисунке 75 имеем $OA = x_0$, $AM = f(x_0)$.

Преимуществом графического способа задания функции по сравнению с аналитическим и табличным является его наглядность.

Графический способ задания функции обычно используется в практике физических измерений,

когда соответствие между переменными x и y задается посредством графика. Во многих случаях графики чертятся с помощью самопишущих приборов. Так, например, для измерения атмосферного давления на различных высотах пользуются специальным самопишущим прибором — барографом, который на движущейся ленте записывает в виде кривой линии изменения давления в зависимости от высоты.

§ 7.2. Обзор элементарных функций и их графиков

1. Целая рациональная функция. Многочлен вида

$$y = a_0 + a_1x + a_2x^2 + \dots + a_mx^m$$

($a_0, a_1, a_2, \dots, a_m$ — постоянные числа, называемые *коэффициентами* многочлена, m — натуральное число, называемое *степенью* многочлена) — целая рациональная функция. Эта функция определена при всех значениях x .

Пример. $y = kx + b$ — *линейная* функция. Ее график — прямая линия (см § 1. 4, п. 1). При $b = 0$ линейная функция $y = kx$ выражает прямо пропорциональную зависимость y от x . В этом случае ее график проходит через начало координат.

2. Дробно-рациональная функция. Эта функция определяется как отношение двух многочленов:

$$y = \frac{a_0 + a_1x + \dots + a_mx^m}{b_0 + b_1x + \dots + b_nx^n}.$$

Она определена при всех значениях x , кроме тех, при которых знаменатель обращается в нуль. Дробно-рациональной функцией является, например, функция $y = \frac{k}{x}$, выражающая обратно пропорциональную зависимость между x и y . Ее график есть равносторонняя гипербола (см. § 5. 1, п. 2).

3. Степенная функция. *Степенная функция* — это функция вида

$$y = x^{\alpha}, \quad (1)$$

где α — действительное число. Она определена при всех значениях x , если α — натуральное число, при всех x , не равных нулю, если α — целое отрицательное число, и при всех x , больших нуля, если α — произвольное действительное число.

Если $\alpha = \frac{1}{q}$, где q — натуральное число, то функция (1) примет вид:

$$y = \sqrt[q]{x}. \quad (2)$$

(Символ $\sqrt[q]{}$ называют *корнем степени q* или *радикалом*.)

Функция (2) определена при всех неотрицательных x , если q четное, и при всех x , если q нечетное.

Пр и м е р. $y = \sqrt{x}$. График этой функции — верхняя ветвь параболы $y^2 = x$ (см. § 5.1, п. 3).

4. Показательная функция. Функция вида $y = a^x$, где $a > 0$ и $a \neq 1$, называется *показательной*. Она определена при всех x (рис. 76).

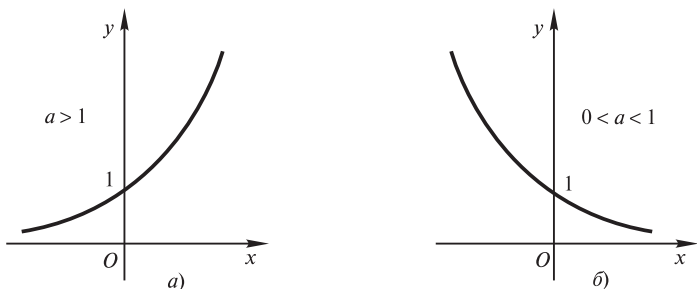


Рис. 76

5. Логарифмическая функция. Функция вида $y = \log_a x$, где $a > 0$ и $a \neq 1$, называется *логарифмической*. Она определена при $x > 0$ (рис. 77).

6. Понятие обратной функции. Между степенной функцией и радикалом, а также между показательной и логарифмической функциями существует связь, выражаемая через понятие обратной функции.

Пусть

$$y = f(x) \quad (3)$$

есть функция независимой переменной x . Это значит, что, задавая значения x , мы вполне определяем значения зависимой переменной y . Поступим наоборот, а именно: будем считать независимой переменную y , а зависимой переменную x . Тогда x будет являться функцией переменной y , которая называется функцией, *обратной* к данной.

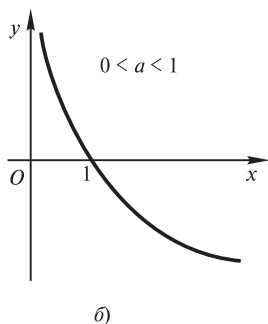
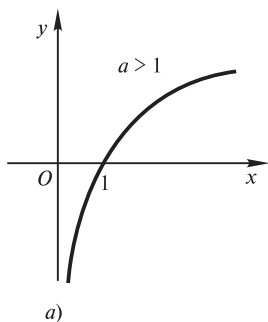


Рис. 77

Предполагая, что уравнение (3) разрешимо относительно x , получим явное выражение обратной функции:

$$x = \varphi(y). \quad (4)$$

Обратная функция однозначной функции может быть многозначной, т. е. данному значению y может соответствовать несколько значений переменной x . Иногда удается сделать обратную функцию однозначной, вводя дополнительные ограничения на ее значения.

П р и м е р. Двухзначная функция $x = \pm\sqrt{y}$ является обратной по отношению к функции $y = x^2$. Если условиться для корня брать лишь его арифметическое значение, то обратная функция будет однозначной.

Очевидно, что если функция, заданная формулой (4), есть функция, обратная к (3), то и функция (3) будет обратной по отношению к функции (4), т. е. эти функции являются взаимно обратными.

Иногда придерживаются стандартных обозначений: под x понимают независимую переменную, а под y — функцию, т. е. зависимую переменную. В таком случае обратную функцию следует писать в виде $y = \varphi(x)$. Например, можно говорить, что функции $y = 2^x$ и $y = \log_2 x$ являются взаимно обратными.

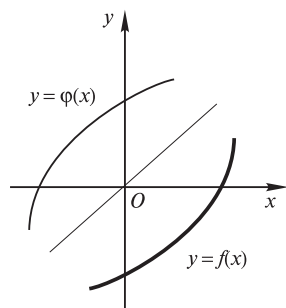


Рис. 78

Чтобы из графика данной функции $y = f(x)$ получить график обратной ей функции $y = \varphi(x)$, очевидно, достаточно график первой функции симметрично отобразить относительно биссектрисы I и III координатных углов (рис. 78).

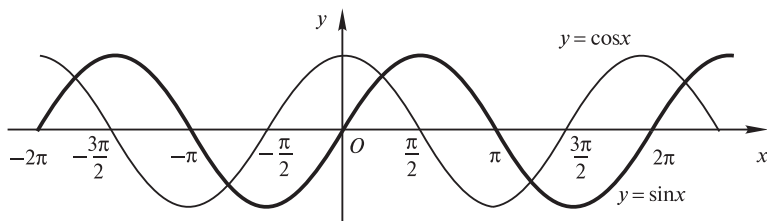


Рис. 79

7. Тригонометрические функции. Функции $y = \sin x$, $y = \cos x$ определены для всех значений x . Они являются периодическими с периодом 2π , т. е. при изменении аргумента на число, кратное 2π , значение функции остается прежним. Кроме того, функция $\sin x$ нечетная ($\sin(-x) = -\sin x$), $\cos x$ четная ($\cos(-x) = \cos x$). Графики этих функций — синусоида и косинусоида (рис. 79).

Функция $y = \operatorname{tg} x$ не определена только в точках, где $\cos x = 0$, т. е. в точках $x = \frac{2k+1}{2}\pi$ ($k = 0, \pm 1, \pm 2, \dots$), а функция $y = \operatorname{ctg} x$ не определена только в точках, где $\sin x = 0$, т. е. в точках $x = k\pi$ ($k = 0, \pm 1, \pm 2, \dots$). При этом $\operatorname{tg} x$ и $\operatorname{ctg} x$ — нечетные «функции». Графики функций $y = \operatorname{tg} x$ и $y = \operatorname{ctg} x$, имеющие период π , изображены на рисунке 80.

Отметим, что в тригонометрических функциях переменная x обычно выражается в радианах.

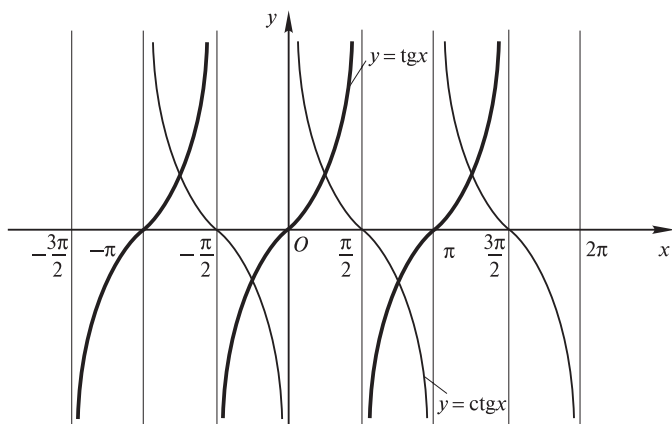


Рис. 80

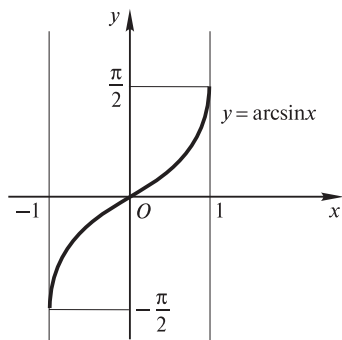


Рис. 81

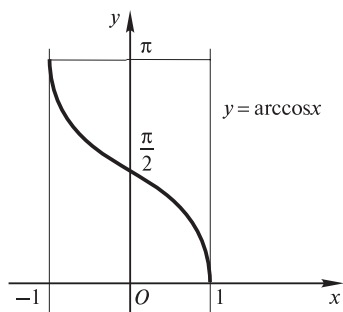


Рис. 82

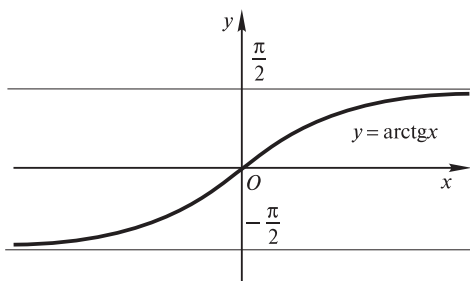


Рис. 83

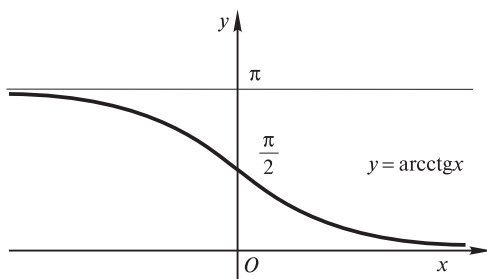


Рис. 84

8. Обратные тригонометрические функции. Функция $y = \arcsin x$.

Здесь y — переменная из сегмента $-\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$, синус которой равен x ,

т. е. $x = \sin y$. Область определения этой функции — сегмент $-1 \leq x \leq 1$, а ее график изображен на рисунке 81.

Функция $y = \arccos x$ означает, что $x = \cos y$, причем $|x| \leq 1$ и $0 \leq y \leq \pi$. График $y = \arccos x$ изображен на рисунке 82.

Функция $y = \operatorname{arctg} x$ есть переменная, тангенс которой равен x , т. е. $x = \operatorname{tg} y$, причем x любое и $|y| < \frac{\pi}{2}$ (рис. 83), а функция $y = \operatorname{arcctg} x$ есть переменная, для которой $x = \operatorname{ctg} y$, где x — любое и $0 < y < \pi$ (рис. 84).

9. Сложная функция. Элементарные функции. Пусть переменная y зависит от переменной u , которая, в свою очередь, зависит от переменной x , т. е. $y = f(u)$, $u = \varphi(x)$. Тогда при изменении x будет меняться u , а потому будет меняться и y . Значит, y является функцией x : $y = f(\varphi(x))$. Эта функция называется *сложной функцией* (или *функцией от функции*), переменная u называется *промежуточной* переменной. Указанную сложную функцию называют также *суперпозицией функций* f и φ .

П р и м е р. Если $y = \sin u$, а $u = x^2$, то $y = \sin x^2$ есть сложная функция независимой переменной x .

Функции: степенная, показательная, логарифмическая, тригонометрические и обратные тригонометрические функции, постоянная (константа) $y = C$ — называются *основными элементарными функциями*.

Всякая функция, которая получается из основных элементарных функций путем конечного числа суперпозиций и четырех арифметических действий (сложения, вычитания, умножения и деления), называется *элементарной функцией*.

Например, элементарными функциями будут рассмотренные выше целая рациональная и дробно-рациональная функции.

Важное значение элементарных функций состоит в том, что в математическом анализе, применяемом в основных задачах физики и техники, употребляются чаще всего элементарные функции.

10. Гармонические колебания. В природе и технике часто происходят явления и процессы, повторяющиеся периодически, например колебания маятника, переменный ток, электромагнитные колебания и др.

Рассмотрим простейший вид колебаний, так называемое *гармоническое колебание*:

$$y = A \sin \omega t, \quad (5)$$

где A и ω — положительные постоянные.

График функции, заданной формулой (5), изображен на рисунке 85.

Коэффициент A , представляющий наибольшую величину, которую может иметь y , называется *амплитудой* колебания, а ω — *частотой* колебания. Функция (5) является периодической с периодом $\frac{2\pi}{\omega}$:

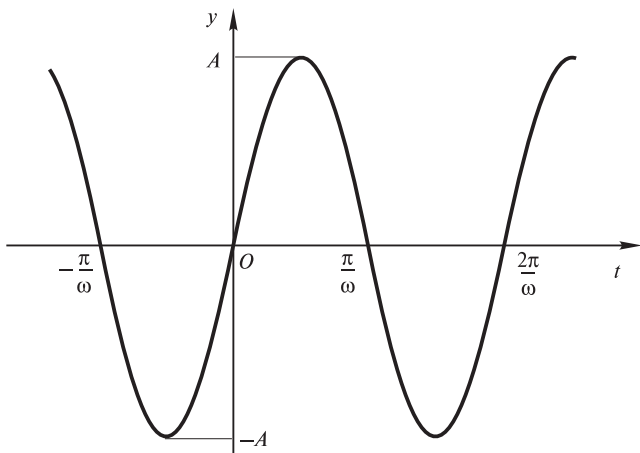


Рис. 85

значения y в точках $t + k \frac{2\pi}{\omega}$ ($k = 0, \pm 1, \pm 2, \dots$) одни и те же. Если считать, что t — время, то период $T = \frac{2\pi}{\omega}$ показывает время, в течение которого совершается одно колебание. Поэтому $\omega = \frac{2\pi}{T}$ — число колебаний за время 2π . График гармонического колебания (рис. 85) называется *простой гармоникой*.

Однако далеко не всегда периодическое явление описывается простой гармоникой. Многие из таких явлений есть результат сложения нескольких простых гармоник, который называется *сложным гармоническим колебанием*, а его график — *сложной гармоникой*.

На рисунке 86 изображена сложная гармоника $y = \sin t + \sin 2t$ — результат сложения двух простых гармоник $y = \sin t$ и $y = \sin 2t$.

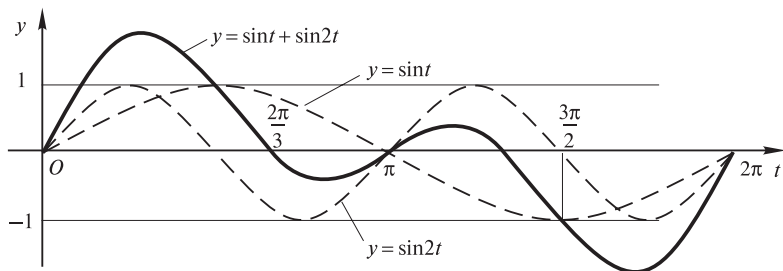


Рис. 86

§ 7.3. Предел функции

1. Предел числовой последовательности. *Бесконечной числовой последовательностью* (или просто числовой последовательностью) называется функция $a_n = f(n)$, определенная на множестве всех натуральных чисел $1, 2, \dots, n, \dots$. Значения последовательности $a_1, a_2, \dots, a_n, \dots$ называются ее *членами*.

Последовательность $a_n = f(n)$ иногда обозначают так: $\{a_n\}$. Это означает, что задана последовательность с *общим* членом a_n . По данному общему члену всегда можно найти любой член последовательности a_k , подставив в a_n вместо n число k . Ниже приведены примеры последовательностей, причем сначала приведена форма записи $\{a_n\}$, а затем записаны первые члены:

1) $\{(-1)^n n\}; -1, 2, -3, \dots;$

5) $\left\{\frac{n}{n+1}\right\}; \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots;$

2) $\{3n + 1\}; 4, 7, 10, \dots;$

6) $\left\{\frac{n+1}{2n}\right\}; 1, \frac{3}{4}, \frac{2}{3}, \dots;$

3) $\{2 - n\}; 1, 0, -1, \dots;$

4) $\left\{\frac{1}{n}\right\}; 1, \frac{1}{2}, \frac{1}{3}, \dots;$

7) $\left\{(-1)^n \frac{1}{n}\right\}; -1, \frac{1}{2}, -\frac{1}{3}, \dots$

Для числовой последовательности, как и для любой функции, можно построить график. Он не является линией, а состоит из отдельных точек, расположенных справа от оси Oy . На рисунках 87, 88 и 89 построены графики последовательностей 1, 2 и 5.

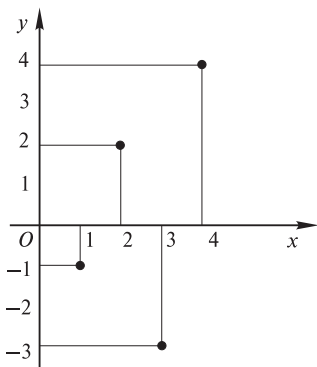


Рис. 87

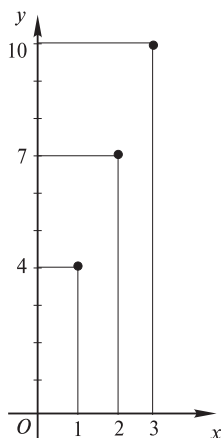


Рис. 88

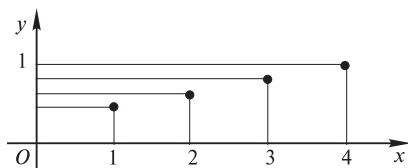


Рис. 89

Числовая последовательность $\{a_n\}$ называется *невозрастающей* (неубывающей), если для любого номера n справедливо неравенство $a_n \geq a_{n+1}$ ($a_n \leq a_{n+1}$). Если $a_n > a_{n+1}$ ($a_n < a_{n+1}$), то последовательность $\{a_n\}$ — *убывающая* (*возрастающая*). Например, последовательность 3 убывающая, последовательность 2 возрастающая. Невозрастающие и неубывающие последовательности называются *монотонными*.

Последовательность $\{a_n\}$ называется *ограниченной сверху* (снизу), если существует такое число M , что для любого номера n выполняется неравенство $a_n \leq M$ ($a_n \geq M$). Последовательность 3 ограничена сверху, например, числом 1. Последовательности, одновременно ограниченные сверху и снизу, называются *ограниченными*. Последовательность 4 ограничена.

По графику последовательности 5 (рис. 89) видно, что ординаты точек с увеличением номера n приближаются к единице. Члены последовательности 4 с возрастанием номера становятся близкими к нулю.

О п р е д е л е н и е. Число a называется *пределом* числовой последовательности $\{a_n\}$, если для любого числа $\varepsilon > 0$ существует такой номер $N = N(\varepsilon)$, зависящий от ε , что для всех $n > N$ выполняется неравенство $|a_n - a| < \varepsilon$. Это обозначают так:

$$\lim_{n \rightarrow \infty} a_n = a \text{ или } a_n \rightarrow a \text{ при } n \rightarrow \infty.$$

П р и м е р ы. Доказать, что

$$\text{а) } \lim_{n \rightarrow \infty} \frac{1}{n} = 0;$$

$$\text{в) } \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1;$$

$$\text{б) } \lim_{n \rightarrow \infty} (-1)^n \frac{1}{n} = 0;$$

$$\text{г) } \lim_{n \rightarrow \infty} \frac{n+1}{2n} = \frac{1}{2}.$$

Ограничимся доказательством первого из этих четырех равенств, так как доказательства трех других проводятся аналогично.

Пусть $\varepsilon > 0$ — произвольное число. Тогда

$$\left| \frac{1}{n} - 0 \right| = \frac{1}{n} < \varepsilon, \text{ если } n > \frac{1}{\varepsilon}.$$

Из последнего неравенства следует, что в качестве номера N можно взять целую часть числа $\frac{1}{\varepsilon}$, т. е. $N = \left[\frac{1}{\varepsilon} \right]$. Итак, $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$.

Характер стремления последовательности к своему пределу различен. Последовательности 4 и 6 стремятся к своим пределам убывая; последовательность 5 стремится к единице возрастая; последовательность 7 стремится к нулю так, что ее члены становятся поочередно то больше, то меньше нуля.

Отметим следующие важные свойства пределов последовательностей:

1. *Последовательность, имеющая предел, ограничена.*

Действительно, из определения предела следует: если последовательность $\{a_n\}$ имеет своим пределом число a , то, например, для $\varepsilon = 1$ найдется натуральное число N такое, что при $n > N$ $|a_n - a| < 1$, или, что то же, $a - 1 < a_n < a + 1$. Обозначим через m и M соответственно наименьшее и наибольшее из чисел $a_1, a_2, \dots, a_N, a - 1, a + 1$. Тогда для всех $n = 1, 2, \dots$ $m \leq a_n \leq M$, т. е. последовательность $\{a_n\}$ ограничена.

2. *Последовательность может иметь только один предел.*

Это следует из более общего свойства (см. § 7. 5, п. 1, следствие 1).

3. *Любая неубывающая (невозрастающая) и ограниченная сверху (снизу) числовая последовательность имеет предел.*

Строгое доказательство этого предложения дано в [11]. Здесь приведем разъяснение его справедливости. В целях простоты и наглядности такого разъяснения ограничимся рассмотрением возрастающей последовательности $y_1 < y_2 < \dots < y_n < \dots$ (ее члены рассматриваем как точки оси Oy), ограниченной сверху: $y_n < M$ ($n = 1, 2, \dots$). Тогда так как точка y_n , двигаясь в одном направлении — вверх по оси Oy , не переходит точку $y = M$, то точка y_n должна неограниченно приближаться к некоторой точке $y = a$, при этом ясно, что $a \leq M$.

2. Число e . Рассмотрим числовую последовательность

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

Для доказательства существования предела этой последовательности воспользуемся свойством 3 предыдущего пункта. Для этого покажем сначала, что последовательность (1) возрастающая. Разложим общий член последовательности по формуле бинома Ньютона (об этой формуле см., например, в главе 8: § 8. 9, п. 2; Исаак Ньютон (1643—1727) — английский математик и физик):

$$a_n = \left(1 + \frac{1}{n}\right)^n = 1 + n \cdot \frac{1}{n} + \frac{n(n-1)}{1 \cdot 2} \cdot \frac{1}{n^2} + \\ + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} \cdot \frac{1}{n^3} + \dots + \frac{n(n-1) \dots (n-(n-1))}{1 \cdot 2 \cdot 3 \cdot \dots \cdot n} \cdot \frac{1}{n^n},$$

или

$$a_n = 2 + \frac{1}{2} \left(1 - \frac{1}{n}\right) + \frac{1}{2 \cdot 3} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) + \dots + \\ + \frac{1}{2 \cdot 3 \cdot \dots \cdot n} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{n-1}{n}\right). \quad (2)$$

Из равенства (2) видно, что с увеличением номера n каждое слагаемое, кроме первого, увеличивается. Возрастает также число таких слагаемых. Следовательно, $a_n < a_{n+1}$ для всех n , и поэтому последовательность возрастающая.

Покажем теперь, что последовательность (1) ограничена сверху. Заменим во всех членах разложения (2) выражения в круглых скобках единицами. Тогда

$$a_n < 2 + \frac{1}{2} + \frac{1}{2 \cdot 3} + \dots + \frac{1}{2 \cdot 3 \cdot \dots \cdot n}.$$

Подставляя вместо множителей 3, 4, ..., n в знаменателях число 2, мы еще больше увеличим правую часть:

$$a_n < 2 + \frac{1}{2} + \dots + \frac{1}{2^{n-1}}.$$

Но по формуле суммы членов геометрической прогрессии

$$\frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}} = \frac{\frac{1}{2} - \frac{1}{2^n}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^{n-1}} < 1.$$

Поэтому $a_n < 3$ при любом n .

По свойству 3 предыдущего пункта последовательность (1), как возрастающая ограниченная сверху, имеет предел. Этот предел принято обозначать буквой e (*второй замечательный предел*):

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n. \quad (3)$$

Число e является иррациональным и приблизительно равно 2,71828.

3. Натуральные логарифмы. Число e принято за основание системы логарифмов, называемых *натуральными* логарифмами. Оказалось, что

с помощью натуральных логарифмов некоторые формулы записываются проще. Для обозначения натурального логарифма числа N пользуются символом $\ln N$.

Для отыскания приближенных значений натуральных логарифмов по таблицам десятичных логарифмов найдем связь между натуральными и десятичными логарифмами.

Если $\ln N = a$, то $N = e^a$ и логарифмирование обеих частей последнего равенства по основанию 10 дает $\lg N = a \lg e$ ($\lg e \approx 0,43429$), или $\lg N = \ln N \lg e$, откуда

$$\ln N = \frac{\lg N}{\lg e} \left(\frac{1}{\lg e} \approx 2,30258 \right).$$

4. Гиперболические функции. Наряду с показательными функциями в математике и ее приложениях находят применение различные комбинации показательных функций, среди которых особое значение имеют некоторые линейные и дробно-линейные функции e^x и e^{-x} — так называемые *гиперболические* функции. Для них введены следующие специальные наименования и обозначения:

$$\begin{aligned} \operatorname{sh} x &= \frac{e^x - e^{-x}}{2} \text{ (гиперболический синус),} \\ \operatorname{ch} x &= \frac{e^x + e^{-x}}{2} \text{ (гиперболический косинус),} \\ \operatorname{th} x &= \frac{e^x - e^{-x}}{e^x + e^{-x}} \text{ (гиперболический тангенс),} \\ \operatorname{cth} x &= \frac{e^x + e^{-x}}{e^x - e^{-x}} \text{ (гиперболический котангенс).} \end{aligned}$$

Значит,

$$\begin{aligned} \operatorname{th} x &= \frac{\operatorname{sh} x}{\operatorname{ch} x}, \\ \operatorname{cth} x &= \frac{\operatorname{ch} x}{\operatorname{sh} x}. \end{aligned}$$

Из определения $\operatorname{sh} x$ имеем: $\operatorname{sh} 0 = 0$; областью определения и областью значений функции $\operatorname{sh} x$ является интервал $(-\infty; +\infty)$.

Из определения $\operatorname{ch} x$ имеем: $\operatorname{ch} 0 = 1$; областью определения функции $\operatorname{ch} x$ является интервал $(-\infty; +\infty)$, а областью значений — полусегмент $[1; +\infty)$.

Различные тригонометрические функции одного и того же аргумента связаны между собой рядом известных соотношений. Анало-

гичные соотношения имеют место и для гиперболических функций. При этом основному тригонометрическому тождеству $\sin^2 x + \cos^2 x = 1$ соответствует тождество, связывающее гиперболический синус и косинус:

$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1,$$

в справедливости которого легко убедиться простой проверкой:

$$\begin{aligned} \operatorname{ch}^2 x - \operatorname{sh}^2 x &= \left(\frac{e^x + e^{-x}}{2} \right)^2 - \left(\frac{e^x - e^{-x}}{2} \right)^2 = \\ &= \left(\frac{e^x + e^{-x} + e^x - e^{-x}}{2} \right) \left(\frac{e^x + e^{-x} - e^x + e^{-x}}{2} \right) = \\ &= e^x \cdot e^{-x} = e^{x-x} = e^0 = 1. \end{aligned}$$

5. Предел функции. Выше рассмотрено понятие предела для частного вида функций — числовых последовательностей. Обобщим его на произвольные функции.

Пусть функция $f(x)$ определена в некоторой окрестности точки $x = a$, кроме, может быть, самой точки a .

О п р е д е л е н и е. Число A называется *пределом* функции $f(x)$ при стремлении x к a (или в точке a), если для любого числа $\varepsilon > 0$ существует такое число $\delta = \delta(\varepsilon) > 0$, что для всех x , удовлетворяющих условию $0 < |x - a| < \delta$, имеет место неравенство $|f(x) - A| < \varepsilon$. Обозначают это так: $\lim_{x \rightarrow a} f(x) = A$ или $f(x) \rightarrow A$ при $x \rightarrow a$.

Другими словами, число A есть предел функции $f(x)$ в точке $x = a$, если для всех x , достаточно близких к числу a и отличных от него, соответствующие им значения функции $f(x)$ оказываются сколь угодно близкими к числу A (естественно, в тех точках x , в которых функция $f(x)$ определена).

П р и м е р 1. Показать, что $\lim_{x \rightarrow 1} x = 1$. Пусть ε — произвольное положительное число.

Выбрав $\delta = \varepsilon$, получим, что $|x - 1| < \varepsilon$, как только $0 < |x - 1| < \delta$. Следовательно, согласно определению предела функции $\lim_{x \rightarrow 1} x = 1$.

П р и м е р 2. Показать, что $\lim_{x \rightarrow 1} x^2 = 1$. Пусть ε — произвольное положительное число.

Найдем такое число $\delta > 0$, что для всех x , удовлетворяющих неравенству $0 < |x - 1| < \delta$, выполняется неравенство $|x^2 - 1| < \varepsilon$.

Если $|x - 1| < \delta$, то $|x + 1| = |(x - 1) + 2| \leq |x - 1| + 2 < \delta + 2$. Следовательно, $|x^2 - 1| = |x - 1| \times |x + 1| < \delta \cdot (\delta + 2)$. Для выполнения неравенства $|x^2 - 1| < \varepsilon$ достаточно потребовать, чтобы $\delta(\delta + 2) = \varepsilon$, т. е. чтобы $\delta^2 + 2\delta - \varepsilon = 0$. Отсюда $\delta = -1 + \sqrt{1 + \varepsilon}$ (второй корень $\delta = -1 - \sqrt{1 + \varepsilon}$ отбрасываем, так как $\delta > 0$). По определению предела функции $\lim_{x \rightarrow 1} x^2 = 1$.

П р и м е ч а н и е. Если в формуле (3) положить $\frac{1}{n} = z$, то она примет вид:

$$e = \lim_{z \rightarrow 0} (1 + z)^{\frac{1}{z}}. \quad (4)$$

Оказывается, что формула (4) верна не только, когда переменная z пробегает последовательность значений $z_n = \frac{1}{n}$, но и при любом другом законе стремления z к нулю.

При изучении свойств функций приходится рассматривать и предел функции при стремлении аргумента x к бесконечности.

О п р е д е л е н и е. Число A называется *пределом* функции $f(x)$ при стремлении x к бесконечности (или в бесконечности), если для любого числа $\varepsilon > 0$ существует такое положительное число $N = N(\varepsilon)$, что для всех x , удовлетворяющих условию $|x| > N$, имеет место неравенство $|f(x) - A| < \varepsilon$. При этом пишут: $\lim_{x \rightarrow \infty} f(x) = A$. Рассматривают также как $\lim_{x \rightarrow +\infty} f(x)$, так и $\lim_{x \rightarrow -\infty} f(x)$.

Предел функции $f(x)$ при $x \rightarrow +\infty$ ($x \rightarrow -\infty$) определяется аналогично $\lim_{x \rightarrow \infty} f(x)$, только в самой формулировке определения $\lim_{x \rightarrow \infty} f(x)$ условие $|x| > N$ следует заменить на $x > N$ ($x < -N$).

§ 7.4. Бесконечно малые и бесконечно большие величины

1. Бесконечно малые и их свойства. При изучении свойств пределов функций особую роль играют функции, предел которых при стремлении аргумента к какой-либо точке равен нулю.

Числовая последовательность $\{a_n\}$ называется *бесконечно малой*, если ее предел равен нулю: $\lim_{n \rightarrow \infty} a_n = 0$. Последовательности $\left\{\frac{1}{n}\right\}$, $\left\{(-1)^n \frac{1}{n}\right\}$ являются бесконечно малыми: их пределами является нуль (см. § 7. 3, п. 1). Понятие бесконечно малой последовательности можно перенести на функции.

О п р е д е л е н и е. Функция $\alpha(x)$ называется *бесконечно малой* при $x \rightarrow a$, если $\lim_{x \rightarrow a} \alpha(x) = 0$, т. е. для любого числа $\varepsilon > 0$ существует такое число $\delta > 0$, что для всех x , удовлетворяющих неравенству $0 < |x - a| < \delta$, выполняется неравенство $|\alpha(x)| < \varepsilon$.

Бесконечно малую функцию $\alpha(x)$ называют также бесконечно малой величиной или просто бесконечно малой.

П р и м е р. Показать, что функция $y = x^2 - 1$ является бесконечно малой при $x \rightarrow 1$. Пусть ε — произвольное положительное число. Найдем такое число $\delta > 0$, что для всех x , удовлетворяющих неравенству $0 < |x - 1| < \delta$, выполняется неравенство $|x^2 - 1| < \varepsilon$. Как

показано ранее (см. § 7. 3, п. 4, пример 2), таким δ является $\delta = -1 + \sqrt{1 + \varepsilon}$. Следовательно, функция $y = x^2 - 1$ является бесконечно малой при $x \rightarrow 1$.

В дальнейшем в этом параграфе при рассмотрении бесконечно малых будем иметь в виду, что они являются бесконечно малыми при $x \rightarrow a$.

Остановимся на основных свойствах бесконечно малых функций. Эти свойства будут верны также и для бесконечно малых последовательностей.

1. Если функции $\alpha_1(x)$ и $\alpha_2(x)$ являются бесконечно малыми, то функция $\alpha_1(x) + \alpha_2(x)$ также есть бесконечно малая.

Д о к а з а т е л ь с т в о. Пусть ε — произвольное положительное число. Так как функции $\alpha_1(x)$ и $\alpha_2(x)$ бесконечно малые, то найдутся такие числа δ_1 и δ_2 , что при $0 < |x - a| < \delta_1$ и $0 < |x - a| < \delta_2$ имеют место соответственно неравенства

$$|\alpha_1(x)| < \frac{\varepsilon}{2} \text{ и } |\alpha_2(x)| < \frac{\varepsilon}{2}. \quad (1)$$

Обозначим через δ наименьшее из двух чисел δ_1 и δ_2 . Тогда при $0 < |x - a| < \delta$ будут верны неравенства (1), и, следовательно,

$$|\alpha_1(x) + \alpha_2(x)| \leq |\alpha_1(x)| + |\alpha_2(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Итак, для любого $\varepsilon > 0$ существует такое $\delta > 0$, что при $0 < |x - a| < \delta$ выполняется неравенство $|\alpha_1(x) + \alpha_2(x)| < \varepsilon$, а это и означает, что $\alpha_1(x) + \alpha_2(x)$ есть функция бесконечно малая.

П р и м е ч а н и е 1. Свойство 1 распространяется на случай алгебраической суммы любого конечного числа бесконечно малых.

Функция $f(x)$ называется *ограниченной* при $x \rightarrow a$, если существуют положительные числа M и δ , такие, что при условии $0 < |x - a| < \delta$ выполняется неравенство $|f(x)| \leq M$. Например, любая бесконечно малая $\alpha(x)$ является ограниченной функцией при $x \rightarrow a$.

Температура воздуха T в данной местности — ограниченная функция времени t . Изменяясь днем и ночью, зимой и летом, она никогда не достигнет $+100^\circ\text{C}$ и -100°C . Таким образом, $|T| < 100$.

2. Произведение ограниченной при $x \rightarrow a$ функции на бесконечно малую есть функция бесконечно малая.

Д о к а з а т е л ь с т в о. Пусть $f(x)$ — ограниченная при $x \rightarrow a$ функция и $\alpha(x)$ — бесконечно малая. Тогда существует такое число $M > 0$, что $|f(x)| \leq M$ для всех x , достаточно близких к a . Возьмем любое $\varepsilon > 0$.

Для ε существует такое $\delta > 0$, что при условии $0 < |x - a| < \delta$ одновременно выполняются неравенства $|f(x)| \leq M$, $|\alpha(x)| < \frac{\varepsilon}{M}$. Поэтому $|f(x) \cdot \alpha(x)| =$
 $= |f(x)| \cdot |\alpha(x)| < M \cdot \frac{\varepsilon}{M} = \varepsilon$.

Непосредственно из свойства 2 следуют свойства 3 и 4.

3. Произведение постоянной на бесконечно малую есть бесконечно малая.

4. Произведение двух бесконечно малых есть бесконечно малая.

Примечание 2. Свойство 4 распространяется на любое конечное число бесконечно малых.

2. Бесконечно большие. Числовая последовательность $\{a_n\}$ называется *бесконечно большой*, если для любого положительного числа M найдется такое натуральное число N , что для любого $n > N$ выполняется неравенство $|a_n| > M$. В этом случае пишут: $\lim_{n \rightarrow \infty} a_n = \infty$.

Последовательности $\{n\}$, $\{(-1)^n n\}$ являются бесконечно большими.

Понятие бесконечно большой последовательности можно перенести на функции.

Функция $f(x)$ называется *бесконечно большой* при $x \rightarrow a$, если для любого числа $M > 0$ существует такое число $\delta > 0$, что $|f(x)| > M$ для всех x , удовлетворяющих неравенству $0 < |x - a| < \delta$. Обозначается это так: $\lim_{x \rightarrow a} f(x) = \infty$.

Если при этом $f(x)$ положительна (отрицательна) в $0 < |x - a| < \delta$, то пишут: $\lim_{x \rightarrow a} f(x) = +\infty$ ($\lim_{x \rightarrow a} f(x) = -\infty$).

Примечания: 1. Бесконечность (∞) не число, а символ, который употребляется, например, для того, чтобы указать, что соответствующая функция есть бесконечно большая.

2. Бесконечно большая функция $f(x)$ при $x \rightarrow a$ не имеет предела, так как предел переменной (если он существует) — некоторое число. То же в случае бесконечно большой числовой последовательности.

В дальнейшем всегда под пределом последовательности (функции) будем понимать конечный предел, т. е. число, если не оговорено противное.

Ниже рассматриваются бесконечно большие функции при $x \rightarrow a$.

Как видно из следующих свойств, которые верны и для последовательностей, бесконечно большие и бесконечно малые функции тесно связаны между собой.

1. Если функция $f(x)$ бесконечно большая, то $\frac{1}{f(x)}$ бесконечно малая.

Доказательство. Возьмем любое $\varepsilon > 0$ и обозначим $\frac{1}{\varepsilon} = M$. Так как $f(x)$ бесконечно большая, то числу M соответствует $\delta > 0$, такое, что при $0 < |x - a| < \delta$ выполняется неравенство $|f(x)| > M = \frac{1}{\varepsilon}$, откуда $\frac{1}{|f(x)|} < \varepsilon$.

2. Если функция $\alpha(x)$ бесконечно малая и не обращается в нуль, то $\frac{1}{\alpha(x)}$ бесконечно большая.

Доказательство. Возьмем любое $M > 0$ и обозначим $\frac{1}{M} = \varepsilon$. Так как $\alpha(x)$ бесконечно малая, то числу $\varepsilon > 0$ соответствует $\delta > 0$, такое, что при $0 < |x - a| < \delta$ выполняется неравенство $|\alpha(x)| < \varepsilon = \frac{1}{M}$, откуда $\frac{1}{|\alpha(x)|} > M$.

Пример. Функция $\frac{1}{x^2 - 1}$ бесконечно большая при $x \rightarrow 1$, так как функция $y = x^2 - 1$ является бесконечно малой при $x \rightarrow 1$ (п. 1, пример).

Замечание. В данном параграфе были рассмотрены функции аргумента x для случая, когда $x \rightarrow a$. Однако все предложения, установленные здесь, остаются в силе и для случая, когда $x \rightarrow \infty$. При этом все доказательства аналогичны.

§ 7.5. Основные теоремы о пределах и их применение

1. Основные теоремы о пределах. Прежде чем излагать материал, сделаем следующее замечание. Ниже рассматриваются функции аргумента x , при этом $x \rightarrow a$ или $x \rightarrow \infty$. Все устанавливаемые в этом пункте предложения о пределах имеют место в обоих случаях, они верны также и для последовательностей. Здесь приводятся доказательства для одного из этих случаев ($x \rightarrow a$), так как для другого доказательства аналогичны. Это замечание относится и к п. 4 этого параграфа.

Теорема 1. Для того чтобы число A было пределом функции $f(x)$ при $x \rightarrow a$, необходимо и достаточно, чтобы эта функция была представима в виде $f(x) = A + \alpha(x)$, где $\alpha(x)$ — бесконечно малая.

Доказательство. 1) Пусть $\lim_{x \rightarrow a} f(x) = A$. Это значит, что для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для всех x , удовлетворяющих условию $0 < |x - a| < \delta$, имеет место неравенство $|f(x) - A| < \varepsilon$, т. е. функция $\alpha(x) = f(x) - A$ есть бесконечно малая и $f(x) = A + \alpha(x)$.

2) Пусть $f(x) = A + \alpha(x)$, где $\alpha(x)$ — бесконечно малая. Тогда для любого $\varepsilon > 0$ существует такое $\delta > 0$, что для x из $0 < |x - a| < \delta$ будет $|\alpha(x)| = |f(x) - A| < \varepsilon$, т. е. A — предел функции $f(x)$ при $x \rightarrow a$.

С л е д с т в и е 1. *Функция не может в одной точке иметь два различных предела.*

Д о к а з а т е л ь с т в о. Если $\lim_{x \rightarrow a} f(x) = A, \lim_{x \rightarrow a} f(x) = B (A \neq B)$, то по теореме 1 $f(x) = A + \alpha(x)$ и $B = \beta(x)$ ($\alpha(x)$ и $\beta(x)$ — бесконечно малые). Отсюда $A + \alpha(x) = B + \beta(x)$ и $A - B = \beta(x) - \alpha(x)$, где $\beta(x) - \alpha(x)$ — бесконечно малая, а $A - B$ — постоянная. Этой постоянной может быть только нуль. Поэтому $A = B$.

Т е о р е м а 2. *Предел постоянной величины равен самой постоянной.*

Это предложение непосредственно вытекает из определения предела.

Т е о р е м а 3. *Если функция $f(x) \geq 0$ ($f(x) \leq 0$) для всех x в некоторой окрестности точки a , кроме, быть может, самой точки a , и в точке a имеет предел, то $\lim_{x \rightarrow a} f(x) \geq 0$ ($\lim_{x \rightarrow a} f(x) \leq 0$).*

Д о к а з а т е л ь с т в о. Пусть, например, $f(x) \geq 0$ и $\lim_{x \rightarrow a} f(x) = A$.

Если бы было $A < 0$, то для $\varepsilon = \frac{|A|}{2}$ неравенство $|f(x) - A| < \varepsilon$ при $0 < |x - a| < \delta$ было бы невозможно ни при каком $\delta > 0$, так как влекло бы за собой отрицательность $f(x)$.

П р и м е ч а н и е 1. Заметим, что при условии, что $\lim_{x \rightarrow a} f(x) = A$ существует, из $f(x) > 0$ ($f(x) < 0$), вообще говоря, не вытекает $\lim_{x \rightarrow a} f(x) > 0$ ($\lim_{x \rightarrow a} f(x) < 0$), а вытекает только $\lim_{x \rightarrow a} f(x) \geq 0$ ($\lim_{x \rightarrow a} f(x) \leq 0$). Так, $|x| > 0$ для всех $x \neq 0$, но $\lim_{x \rightarrow 0} |x| = 0$.

Т е о р е м а 4. *Если функции $f_1(x)$ и $f_2(x)$ имеют пределы при $x \rightarrow a$, то при $x \rightarrow a$ имеют пределы также их сумма $f_1(x) + f_2(x)$, произведение $f_1(x) \cdot f_2(x)$ и при условии $\lim_{x \rightarrow a} f_2(x) \neq 0$ частное $\frac{f_1(x)}{f_2(x)}$, причем*

$$\lim_{x \rightarrow a} (f_1(x) + f_2(x)) = \lim_{x \rightarrow a} f_1(x) + \lim_{x \rightarrow a} f_2(x), \quad (1)$$

$$\lim_{x \rightarrow a} (f_1(x) \cdot f_2(x)) = \lim_{x \rightarrow a} f_1(x) \cdot \lim_{x \rightarrow a} f_2(x), \quad (2)$$

$$\lim_{x \rightarrow a} \frac{f_1(x)}{f_2(x)} = \frac{\lim_{x \rightarrow a} f_1(x)}{\lim_{x \rightarrow a} f_2(x)}. \quad (3)$$

Д о к а з а т е л ь с т в о. Ограничимся рассмотрением случая суммы. Все остальные утверждения доказываются аналогично. Пусть $\lim_{x \rightarrow a} f_1(x) = A_1$, $\lim_{x \rightarrow a} f_2(x) = A_2$. Тогда согласно теореме 1

$$f_1(x) = A_1 + \alpha_1(x), f_2(x) = A_2 + \alpha_2(x),$$

где $\alpha_1(x)$, $\alpha_2(x)$ — бесконечно малые. Отсюда

$$f_1(x) + f_2(x) = (A_1 + A_2) + (\alpha_1(x) + \alpha_2(x)).$$

По свойству 1 бесконечно малых (§ 7. 4, п. 1) сумма $\alpha_1(x) + \alpha_2(x)$ бесконечно мала. Следовательно, по теореме 1

$$\lim_{x \rightarrow a} (f_1(x) + f_2(x)) = A_1 + A_2.$$

П р и м е ч а н и е 2. Формула (1) распространяется на случай алгебраической суммы любого конечного числа слагаемых, а формула (2) — на случай любого конечного числа сомножителей.

С л е д с т в и е 2. Если функция $f(x)$ имеет предел при $x \rightarrow a$, то

$$\lim_{x \rightarrow a} (f(x))^n = [\lim_{x \rightarrow a} f(x)]^n,$$

где n — натуральное число.

С л е д с т в и е 3. Постоянный множитель можно выносить за знак предела:

$$\lim_{x \rightarrow a} C f(x) = C \cdot \lim_{x \rightarrow a} f(x), \quad C — \text{const.}$$

Т е о р е м а 5. Если для функций $f(x)$, $f_1(x)$ и $f_2(x)$ в некоторой окрестности точки a выполняется неравенство

$$f_1(x) \leq f(x) \leq f_2(x) \tag{4}$$

и $\lim_{x \rightarrow a} f_1(x) = \lim_{x \rightarrow a} f_2(x) = A$, то $\lim_{x \rightarrow a} f(x) = A$.

Д о к а з а т е л ь с т в о. Из определения предела вытекает, что существует окрестность точки a (при $x \neq a$), в которой одновременно выполняются следующие неравенства:

$$|f_1(x) - A| < \varepsilon, |f_2(x) - A| < \varepsilon,$$

где ε — произвольное положительное число. Запишем эти неравенства, освободившись от знака абсолютной величины:

$$A - \varepsilon < f_1(x) < A + \varepsilon, \tag{5}$$

$$A - \varepsilon < f_2(x) < A + \varepsilon. \tag{6}$$

Из неравенств (4) и (5) имеем: $A - \varepsilon < f_1(x) \leq f(x)$, откуда

$$A - \varepsilon < f(x). \tag{7}$$

Аналогично из неравенств (4) и (6) получим:

$$f(x) < A + \varepsilon. \quad (8)$$

Из неравенств (7) и (8) следует, что $A - \varepsilon < f(x) < A + \varepsilon$, или

$$|f(x) - A| < \varepsilon, \text{ т. е. } \lim_{x \rightarrow a} f(x) = A.$$

2. Примеры нахождения пределов.

Пример 1. Найти $\lim_{x \rightarrow 1} (2x^2 + 1)$.

Используя теоремы 4 и 2 и следствия 3 и 2, последовательно получаем:

$$\lim_{x \rightarrow 1} (2x^2 + 1) = \lim_{x \rightarrow 1} 2x^2 + \lim_{x \rightarrow 1} 1 = 2[\lim_{x \rightarrow 1} x]^2 + 1.$$

Но, как уже установлено (§ 7. 3, п. 4, пример 1), $\lim_{x \rightarrow 1} x = 1$. Поэтому $\lim_{x \rightarrow 1} (2x^2 + 1) = 3$.

Пример 2. Найти $\lim_{x \rightarrow 1} \frac{x^2 - 5x + 1}{x^3 + 1}$. Как и в примере 1, находим, что $\lim_{x \rightarrow 1} (x^3 + 1) = \lim_{x \rightarrow 1} x^3 + \lim_{x \rightarrow 1} 1 = (\lim_{x \rightarrow 1} x)^3 + 1 = 2$. Теперь по теореме 4

$$\lim_{x \rightarrow 1} \frac{x^2 - 5x + 1}{x^3 + 1} = \frac{\lim_{x \rightarrow 1} (x^2 - 5x + 1)}{\lim_{x \rightarrow 1} (x^3 + 1)}.$$

Но $\lim_{x \rightarrow 1} (x^2 - 5x + 1) = \lim_{x \rightarrow 1} x^2 - \lim_{x \rightarrow 1} 5x + \lim_{x \rightarrow 1} 1 = [\lim_{x \rightarrow 1} x]^2 - 5 \lim_{x \rightarrow 1} x + 1 = -3$. Следовательно,

$$\lim_{x \rightarrow 1} \frac{x^2 - 5x + 1}{x^3 + 1} = -\frac{3}{2}.$$

Как показывают решения приведенных примеров, в простейших случаях нахождение предела сводится к подстановке в данное выражение предельного значения аргумента. Однако не всегда можно вычислить предел с помощью формул (1), (2), (3). Так, формулы (1) и (2) утрачивают смысл, если хотя бы одна из функций $f_1(x)$ или $f_2(x)$ не имеет предела. Формула (3) неверна, если знаменатель дроби стремится к нулю. Здесь могут представиться два случая.

а) Предел числителя не равен нулю.

Пример 3. Найти $\lim_{x \rightarrow 1} \frac{x^2}{1 - x^2}$.

Имеем $\lim_{x \rightarrow 1} (1 - x^2) = 0$. Поэтому формулу (3) в этом примере использовать нельзя. Так как

$$\lim_{x \rightarrow 1} \frac{1 - x^2}{x^2} = \frac{\lim_{x \rightarrow 1} (1 - x^2)}{\lim_{x \rightarrow 1} x^2} = \frac{0}{1} = 0,$$

то функция $\frac{1-x^2}{x^2}$ бесконечно малая при $x \rightarrow 1$. Тогда функция $\frac{x^2}{1-x^2}$ бесконечно большая при $x \rightarrow 1$, т. е. $\lim_{x \rightarrow 1} \frac{x^2}{1-x^2} = \infty$.

Можно отметить, что, когда x приближается к 1 слева, т. е. оставаясь все время меньше 1 (что записывают: $x \rightarrow 1 - 0$), функция $\frac{x^2}{1-x^2}$ остается все время положительной.

В этом случае записывают:

$$\lim_{x \rightarrow 1-0} \frac{x^2}{1-x^2} = +\infty.$$

Если же x приближается к 1 справа, т. е. оставаясь все время больше 1 (что записывают: $x \rightarrow 1 + 0$), то эта функция остается все время отрицательной. В этом случае записывают: $\lim_{x \rightarrow 1+0} \frac{x^2}{1-x^2} = -\infty$.

б) Предел числителя равен нулю.

Пример 4. Найти $\lim_{x \rightarrow 0} \frac{x^2 + 3x}{x^2 + x}$.

Здесь $\lim_{x \rightarrow 0} (x^2 + 3x) = 0 + 3 \cdot 0 = 0$, $\lim_{x \rightarrow 0} (x^2 + x) = 0 + 0 = 0$.

Говорят, что в этом случае имеем *неопределенность вида* $\frac{0}{0}$. Однако предел $\lim_{x \rightarrow 0} \frac{x^2 + 3x}{x^2 + x}$ существует и его можно найти. Для его нахождения, т. е. раскрытия *неопределенности вида* $\frac{0}{0}$, нужно предварительно преобразовать дробь $\frac{x^2 + 3x}{x^2 + x}$, разделив числитель и знаменатель почленно на x , что возможно, так как до перехода к предельному значению $x \neq 0$. Следовательно, получим:

$$\lim_{x \rightarrow 0} \frac{x^2 + 3x}{x^2 + x} = \lim_{x \rightarrow 0} \frac{x + 3}{x + 1}.$$

Но $\lim_{x \rightarrow 0} \frac{x + 3}{x + 1} = 3$. В результате имеем: $\lim_{x \rightarrow 0} \frac{x^2 + 3x}{x^2 + x} = 3$.

Пример 5. Найти $\lim_{x \rightarrow 0} \frac{\sqrt{4-x} - \sqrt{4+x}}{x}$.

Так как $\lim_{x \rightarrow 0} (\sqrt{4-x} - \sqrt{4+x}) = 0$, $\lim_{x \rightarrow 0} x = 0$, то здесь также имеем неопределенность вида $\frac{0}{0}$. Для того чтобы раскрыть эту неопределенность, преобразуем дробь, стоящую под знаком предела, умножив числитель и знаменатель этой дроби на $\sqrt{4-x} + \sqrt{4+x}$ и сделав после этого необходимые упрощения:

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\sqrt{4-x} - \sqrt{4+x}}{x} &= \lim_{x \rightarrow 0} \frac{(\sqrt{4-x} - \sqrt{4+x})(\sqrt{4-x} + \sqrt{4+x})}{x(\sqrt{4-x} + \sqrt{4+x})} = \\ &= \lim_{x \rightarrow 0} \frac{-2x}{x(\sqrt{4-x} + \sqrt{4+x})} = -2 \cdot \lim_{x \rightarrow 0} \frac{1}{\sqrt{4-x} + \sqrt{4+x}} = -2 \cdot \frac{1}{2+2} = -\frac{1}{2}. \end{aligned}$$

Рассмотрим теперь примеры на вычисление пределов функций при $x \rightarrow \infty$.

Пример 6. Найти $\lim_{x \rightarrow \infty} \frac{4}{3x+2}$.

Очевидно, $\lim_{x \rightarrow \infty} (3x+2) = \infty$. Поэтому (§ 7.4) функция $\frac{1}{3x+2}$, значит, и функция $\frac{4}{3x+2}$ бесконечно малые при $x \rightarrow \infty$, т. е. $\lim_{x \rightarrow \infty} \frac{4}{3x+2} = 0$.

Пример 7. Найти $\lim_{x \rightarrow \infty} \frac{3x+5}{4x+1}$.

Здесь $\lim_{x \rightarrow \infty} (3x+5) = \infty$, $\lim_{x \rightarrow \infty} (4x+1) = \infty$. Говорят, что в этом случае имеем *неопределенность вида* $\frac{\infty}{\infty}$. Для ее раскрытия предварительно числитель и знаменатель дроби $\frac{3x+5}{4x+1}$ почленно разделим на x . Следовательно, получим:

$$\lim_{x \rightarrow \infty} \frac{3x+5}{4x+1} = \lim_{x \rightarrow \infty} \frac{3 + \frac{5}{x}}{4 + \frac{1}{x}}.$$

Но $\lim_{x \rightarrow \infty} \frac{1}{x} = 0$ и $\lim_{x \rightarrow \infty} \frac{5}{x} = 0$. В результате имеем, что $\lim_{x \rightarrow \infty} \frac{3x+5}{4x+1} = \frac{3}{4}$.

Аналогично устанавливается, что при $x \rightarrow \infty$ дробно-рациональная функция стремится либо к нулю, либо к бесконечности, либо к конечному числу, отличному от нуля, в зависимости от того, будет ли степень числителя меньше степени знаменателя, больше ее или равна ей.

3. Первый замечательный предел. Справедливо равенство

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1, \quad (9)$$

называемое *первым замечательным пределом*. С его помощью можно вычислять пределы различных функций, содержащих тригонометрические функции и степени x .

Перейдем к доказательству равенства (9). Возьмем круг единичного радиуса и предположим, что угол x , выраженный в радианах, заключен в интервале $\left(0; \frac{\pi}{2}\right)$ (рис. 90). Обозна-

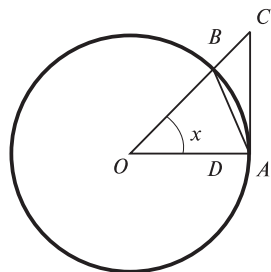


Рис. 90

чим площади треугольников OAB и OAC соот-

ветственно через S_1 и S_2 , а площадь сектора OAB через S . Из рисунка 90 видно, что

$$S_1 < S < S_2. \quad (10)$$

Замечая, что $AC = OA \cdot \operatorname{tg} x = \operatorname{tg} x$, имеем $S_1 = \frac{1}{2} \cdot OA \cdot OB \cdot \sin x = \frac{1}{2} \sin x$, $S = \frac{1}{2} \cdot OA^2 \cdot x = \frac{1}{2} x$, $S_2 = \frac{1}{2} \cdot OA \cdot AC = \frac{1}{2} \operatorname{tg} x$. Поэтому с учетом неравенств (10) получаем:

$$\frac{1}{2} \sin x < \frac{1}{2} x < \frac{1}{2} \operatorname{tg} x,$$

откуда после деления на $\sin x$ и сокращения на $\frac{1}{2}$ находим:

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

или

$$\cos x < \frac{\sin x}{x} < 1 \quad (11)$$

Неравенства (11) получены для $0 < x < \frac{\pi}{2}$. Однако $\cos x$ и $\frac{\sin x}{x}$ — четные функции: $\cos(-x) = \cos x$, $\frac{\sin(-x)}{-x} = \frac{-\sin x}{-x} = \frac{\sin x}{x}$. Тем самым неравенства (11) справедливы и в интервале $-\frac{\pi}{2} < x < 0$. Так как $\lim_{x \rightarrow 0} \cos x = 1$ (это следует из геометрического определения косинуса), то из неравенств (11) на основании теоремы 5 заключаем, что действительно имеет место равенство (9).

П р и м е р. Найти $\lim_{x \rightarrow 0} \frac{\sin 2x}{\sin 5x}$.

Имеем:

$$\lim_{x \rightarrow 0} \frac{\sin 2x}{\sin 5x} = \lim_{x \rightarrow 0} \frac{2 \sin 2x}{5 \sin 5x} = \frac{2}{5} \lim_{x \rightarrow 0} \frac{\sin 2x}{2x} : \lim_{x \rightarrow 0} \frac{\sin 5x}{5x} = \frac{2}{5}.$$

4. Сравнение бесконечно малых. Рассмотрим отношение двух бесконечно малых $\alpha(x)$ и $\beta(x)$ при $x \rightarrow a$ (для компактности записи будем обозначать их просто α и β). Выделим три случая.

1. $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = 0$. В этом случае говорят, что α — бесконечно малая более

высокого порядка, чем β .

Пример 1. При $x \rightarrow 2$ функция $(x-2)^3$ бесконечно малая более высокого порядка, чем $x-2$, так как $\lim_{x \rightarrow 2} \frac{(x-2)^3}{x-2} = 0$.

2. $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = A \neq 0$. В этом случае функции α и β называются бесконечно малыми *одного* порядка.

Пример 2. При $x \rightarrow 0$ функции $5x^2$ и x^2 являются бесконечно малыми одного порядка, так как $\lim_{x \rightarrow 0} \frac{5x^2}{x^2} = 5$.

3. $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = \infty$. В этом случае говорят, что α — бесконечно малая более *низкого* порядка, чем β . Можно сказать также, что β — бесконечно малая более высокого порядка, чем α .

Пример 3. При $x \rightarrow -1$ функция $x+1$ бесконечно малая более низкого порядка, чем $(x-1)(x+1)^2$, так как

$$\lim_{x \rightarrow -1} \frac{x+1}{(x-1)(x+1)^2} = \infty.$$

Определение. Если функции α и β бесконечно малые одного порядка, причем $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = 1$, то они называются *эквивалентными* бесконечно малыми. Символически это записывается так: $\alpha \sim \beta$.

Из определения, в частности, следует, что если $\lim_{x \rightarrow a} \frac{\alpha}{\beta} = A \neq 0$, т. е. если α и β — бесконечно малые одного порядка, то α и $A\beta$ будут являться эквивалентными бесконечно малыми: $\alpha \sim A\beta$.

Пример 4. Как установлено в п. 3, $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$, т. е. $\sin x$ и x при $x \rightarrow 0$ являются эквивалентными бесконечно малыми.

Теорема. Если существует предел отношения двух бесконечно малых α и β , то он равен пределу отношения соответствующих им эквивалентных бесконечно малых.

Доказательство. Действительно, если $\alpha \sim \alpha_1$, $\beta \sim \beta_1$ и существует $\lim_{x \rightarrow a} \frac{\alpha}{\beta}$, то, перейдя к пределу в равенстве $\frac{\alpha}{\beta} = \frac{\alpha}{\alpha_1} \cdot \frac{\alpha_1}{\beta_1} \cdot \frac{\beta_1}{\beta}$, получим:

$$\begin{aligned} \lim_{x \rightarrow a} \frac{\alpha}{\beta} &= \lim_{x \rightarrow a} \left(\frac{\alpha}{\alpha_1} \cdot \frac{\alpha_1}{\beta_1} \cdot \frac{\beta_1}{\beta} \right) = \lim_{x \rightarrow a} \frac{\alpha}{\alpha_1} \cdot \lim_{x \rightarrow a} \frac{\alpha_1}{\beta_1} \cdot \lim_{x \rightarrow a} \frac{\beta_1}{\beta} = \\ &= 1 \cdot \lim_{x \rightarrow a} \frac{\alpha_1}{\beta_1} \cdot 1 = \lim_{x \rightarrow a} \frac{\alpha_1}{\beta_1}. \end{aligned}$$

Доказанная теорема позволяет во многих случаях упрощать отыскание предела.

Пример 5. $\lim_{x \rightarrow 0} \frac{\sin 5x}{3x} = \lim_{x \rightarrow 0} \frac{5x}{3x} = \frac{5}{3}$, так как $\sin 5x \sim 5x$ при $x \rightarrow 0$ (см. пример 4).

Замечание. При сравнении бесконечно малых часто используют символ o («о малое»). Если α — бесконечно малая более высокого порядка, чем бесконечно малая β , то это условно записывают так:

$$\alpha = o(\beta).$$

§ 7.6. Непрерывность функции

1. Понятие непрерывности. Мы видели, что графиками последовательностей являются множества точек. Эти точки всегда находятся на некотором расстоянии друг от друга. Графиком же, например, степенной функции является кривая, которая похожа на сплошную непрерывную линию. Оказывается, эту разницу характеризует точное математическое понятие непрерывности, к введению которого и перейдем.

Пусть функция $y = f(x)$ определена в некотором интервале, x_0 и x — два произвольных значения аргумента из этого интервала. Обозначим $x - x_0 = \Delta x$, откуда $x = x_0 + \Delta x$. Говорят, что для перехода от значения аргумента x_0 к значению x первоначальному значению придано *приращение* Δx .

Приращением Δy функции $y = f(x)$, соответствующим приращению Δx аргумента x в точке x_0 , называется разность

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

Например, приращением функции $y = x^3$, которое соответствует приращению Δx аргумента x в точке x_0 , будет величина

$$\Delta y = (x_0 + \Delta x)^3 - x_0^3 = 3x_0^2 \Delta x + 3x_0 (\Delta x)^2 + (\Delta x)^3.$$

О п р е д е л е н и е. Функция $y = f(x)$ называется *непрерывной в точке* x_0 , если бесконечно малому приращению Δx аргумента x в точке x_0 соответствует бесконечно малое приращение функции Δy , т. е. $\lim_{\Delta x \rightarrow 0} \Delta y = \lim_{\Delta x \rightarrow 0} (f(x_0 + \Delta x) - f(x_0)) = 0$.

Другими словами, функция $y = f(x)$ непрерывна в точке x_0 , если $\lim_{x \rightarrow x_0} f(x) = f(x_0)$, т. е. предел функции в точке x_0 равен значению функции в этой точке.

Пример 1. Функция $y = x$ непрерывна при любом значении $x = x_0$. В самом деле, $\Delta y = x_0 + \Delta x - x_0 = \Delta x$, и, значит, $\lim_{\Delta x \rightarrow 0} \Delta y = \lim_{\Delta x \rightarrow 0} \Delta x = 0$.

Пример 2. Функция $y = \sin x$ непрерывна при любом значении $x = x_0$. В самом деле, $\Delta y = \sin(x_0 + \Delta x) - \sin x_0 = 2 \cos\left(x_0 + \frac{\Delta x}{2}\right) \sin \frac{\Delta x}{2}$. Отсюда $\lim_{x \rightarrow 0} \Delta y = 0$, так как

$$\left| \cos\left(x_0 + \frac{\Delta x}{2}\right) \right| \leq 1, \text{ а } \lim_{\Delta x \rightarrow 0} \sin \frac{\Delta x}{2} = \lim_{\Delta x \rightarrow 0} \frac{\sin \frac{\Delta x}{2}}{\frac{\Delta x}{2}} \cdot \frac{\Delta x}{2} = 1 \cdot \frac{1}{2} \lim_{\Delta x \rightarrow 0} \Delta x = 0.$$

Аналогично доказывается непрерывность функции $\cos x$.

Функция, непрерывная в каждой точке интервала, называется непрерывной на этом интервале.

Теорема 1. Если функции $f_1(x)$ и $f_2(x)$ непрерывны в точке x_0 , то непрерывны в этой точке также их алгебраическая сумма $f_1(x) \pm f_2(x)$, произведение $f_1(x)f_2(x)$ и при условии $f_2(x_0) \neq 0$ частное $\frac{f_1(x)}{f_2(x)}$.

Эта теорема вытекает из аналогичной теоремы о пределах.

Примечание. Для алгебраической суммы и произведения теорема 1 распространяется на любое конечное число функций.

Теорема 2. Если функция $u = \varphi(x)$ непрерывна в точке x_0 , а функция $y = f(u)$ непрерывна в точке $u_0 = \varphi(x_0)$, то сложная функция $y = f(\varphi(x))$ непрерывна в точке x_0 .

Доказательство. Согласно непрерывности функции $u = \varphi(x)$ имеем $\lim_{x \rightarrow x_0} \varphi(x) = \varphi(x_0) = u_0$, т. е. при $x \rightarrow x_0$ также и $u \rightarrow u_0$. Поэтому в силу непрерывности функции $f(u)$ $\lim_{x \rightarrow x_0} f(\varphi(x)) = \lim_{u \rightarrow u_0} f(u) = f(u_0) = f(\varphi(x_0))$, что и доказывает теорему 2.

Таким образом, сложная функция $y = f(\varphi(x))$, образованная из двух непрерывных функций $f(u)$ и $\varphi(x)$, является непрерывной функцией.

Имеет место и следующая теорема.

Теорема 3. Если $f(x)$ — непрерывная функция, имеющая однозначную обратную функцию, то обратная функция тоже непрерывна.

Вместо доказательства ограничимся следующим наглядным соображением: если график функции $y = f(x)$ — непрерывная кривая, то график обратной к ней функции тоже непрерывная кривая.

Теорема 4. Все основные элементарные функции непрерывны там, где они определены.

Доказательство. Постоянная функция $y = C$ непрерывна при любом значении $x = x_0$, так как $\Delta y = C - C = 0$, и, следовательно, $\lim_{\Delta x \rightarrow 0} \Delta y = 0$. Так как функция $y = x$ непрерывна при любом x (см. пример 1), то согласно теореме 1 степенная функция $y = x^n$, где n — натуральное число, также непрерывна при любом x .

Непрерывность тригонометрических функций $\sin x$ и $\cos x$ имеет место всюду (см. пример 2); $\operatorname{tg} x$ и $\operatorname{ctg} x$ непрерывны всюду, где они определены как отношения двух непрерывных функций $\sin x$ и $\cos x$.

Можно доказать непрерывность x^α (α — действительное) и других основных элементарных функций там, где они определены.

Из теорем 1, 2 и 4 вытекает:

С л е д с т в и е. *Всякая элементарная функция непрерывна во всех точках, принадлежащих ее области определения.*

Имеет место (см., например, [11]) и следующее предложение:

Т е о р е м а 5. *Функция $f(x)$, непрерывная в точке x_0 , не равная нулю в этой точке, сохраняет знак $f(x_0)$ в некоторой окрестности точки x_0 .*

2. Точки разрыва функции. Если функция $f(x)$ не является непрерывной в точке x_0 , то говорят, что в точке x_0 функция $f(x)$ *разрывна*, а точка x_0 называется *точкой разрыва* функции $f(x)$.

В качестве конкретного примера функции, имеющей точку разрыва, рассмотрим скорость тела, падающего на землю. Эта скорость вообще является непрерывной функцией времени, но для момента удара можно условно считать, что она мгновенно (скачком) падает до нуля, т. е. скорость терпит разрыв.

Пределом функции $f(x)$ в точке x_0 *слева (справа)* называется предел, вычисляемый в предположении, что x стремится к x_0 , оставаясь все время меньше (больше) значения x_0 . Пределы слева и справа, называемые *односторонними* пределами, соответственно обозначают:

$$\lim_{x \rightarrow x_0 - 0} f(x) \text{ и } \lim_{x \rightarrow x_0 + 0} f(x).$$

Точка x_0 разрыва функции $f(x)$ называется точкой разрыва *первого рода*, если существуют односторонние пределы функции:

$$\lim_{x \rightarrow x_0 - 0} f(x), \lim_{x \rightarrow x_0 + 0} f(x)$$

(при этом функция $f(x)$ необязательно должна быть определена в точке x_0).

Все прочие точки разрыва функции $f(x)$ называются ее точками разрыва *второго рода*.

П р и м е р 1. Рассмотрим функцию $y=f(x)$, определенную на сегменте $[0; 4]$ следующим образом:

$$f(x) = \begin{cases} x - 1, & \text{если } 0 \leq x < 3, \\ 3 - x, & \text{если } 3 \leq x \leq 4. \end{cases}$$

Эта функция (рис. 91) определена во всех точках сегмента $[0; 4]$, и ее значение при $x=3$ равно 0. Однако в точке $x=3$ функция претерпевает разрыв первого рода, так как при $x \rightarrow 3$

$$\lim_{x \rightarrow 3 - 0} f(x) = 2, \lim_{x \rightarrow 3 + 0} f(x) = 0.$$

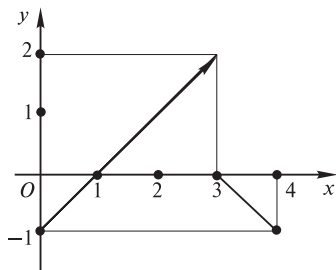


Рис. 91

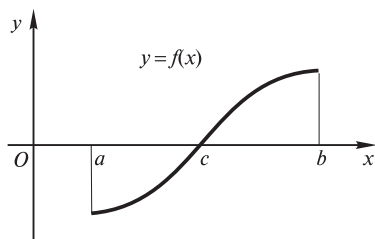


Рис. 92

Точка x_0 разрыва первого рода, в которой

$$\lim_{x \rightarrow x_0+0} f(x) = \lim_{x \rightarrow x_0-0} f(x),$$

называется точкой *устранимого* разрыва.

П р и м е р 2. Функция $\frac{\sin x}{x}$ в точке $x = 0$ имеет устранимый разрыв, так как

$$\lim_{x \rightarrow 0+0} \frac{\sin x}{x} = \lim_{x \rightarrow 0-0} \frac{\sin x}{x} = 1.$$

Пусть x_0 — точка разрыва первого рода.

Скачком функции $f(x)$ в точке x_0 называют разность

$$\lim_{x \rightarrow x_0+0} f(x) - \lim_{x \rightarrow x_0-0} f(x).$$

Так, функция, рассмотренная в примере 1, имеет в точке $x = 3$ скачок, равный $0 - 2 = -2$.

П р и м е р 3. Функция $y = \frac{1}{x}$ (см. рис. 53) в точке $x = 0$ имеет разрыв второго рода, так как эта функция при $x \rightarrow 0$ не имеет предела ни слева, ни справа.

Функция $f(x)$ называется непрерывной в точке x_0 *слева (справа)*, если $\lim_{x \rightarrow x_0-0} f(x) = f(x_0)$ ($\lim_{x \rightarrow x_0+0} f(x) = f(x_0)$).

3. Свойства функций, непрерывных на сегменте. Функция $f(x)$ называется *непрерывной на сегменте* $[a; b]$, если она непрерывна в интервале $(a; b)$ и, кроме того, в точке a непрерывна справа, а в точке b — слева.

Ниже при рассмотрении свойств функций, непрерывных на сегменте, ограничимся формулировками и пояснениями, не проводя доказательств (доказательства см., например, в [9], т. 1).

Т е о р е м а 1 (теорема Вейерштрасса о достижении функций своего наибольшего и наименьшего значений; Карл Вейерштрасс

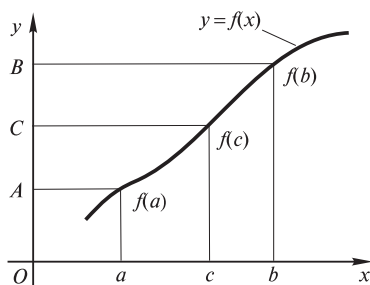


Рис. 93

(1815—1897) — немецкий математик). Функция $f(x)$, непрерывная на сегменте $[a; b]$, достигает в этом сегменте своего наибольшего и своего наименьшего значений, т. е. существуют такие точки x_1 и x_2 отрезка $[a; b]$, что для всех x из $[a; b]$ выполняются неравенства $f(x_1) \geq f(x)$ и $f(x) \geq f(x_2)$.

С л е д с т в и е. Если функция $f(x)$ непрерывна на сегменте $[a; b]$, то она ограничена на нем, т. е. существу-

ет такое положительное число M , что $|f(x)| \leq M$ при $a \leq x \leq b$.

Д о к а з а т е л ь с т в о. Обозначим через m и $\tilde{m}$ соответственно наибольшее и наименьшее значения функции $f(x)$ на сегменте $[a; b]$. Тогда для любого x , принадлежащего сегменту $[a; b]$, имеют место неравенства $\tilde{m} \leq f(x) \leq m$. Пусть M — наибольшее из чисел $|\tilde{m}|, |m|$. Тогда $|f(x)| \leq M$ при $a \leq x \leq b$.

Т е о р е м а 2. Если функция $f(x)$ непрерывна на сегменте $[a; b]$ и на концах его принимает значения разных знаков, то между точками a и b найдется точка c , такая, что $f(c) = 0$.

Эта теорема имеет простой геометрический смысл (рис. 92): если непрерывная кривая переходит с одной стороны оси Ox на другую, то она пересекает ось Ox .

Т е о р е м а 3 (теорема Коши о промежуточных значениях; Огюстен Коши (1789—1857) — французский математик). Пусть функция $f(x)$ непрерывна на сегменте $[a; b]$ и $f(a) = A, f(b) = B$. Тогда для любого числа C , заключенного между A и B , найдется внутри этого сегмента такая точка c , что $f(c) = C$.

Эта теорема геометрически очевидна. Рассмотрим график функции $y = f(x)$ (рис. 93). Так как $f(a) = A, f(b) = B$, то прямая $y = C$, где C любое число, заключенное между A и B , пересечет его по крайней мере в одной точке.

Таким образом, непрерывная функция, переходя от одного значения к другому, обязательно проходит через все промежуточные значения.

§ 7.7. Комплексные числа

1. Определение комплексных чисел и основные операции над ними. К комплексным числам обычно приходят, рассматривая уравнение $x^2 + 1 = 0$. Очевидно, не существует действительных чисел, удовлетворяющих этому уравнению. Корнями его (как и целого ряда других уравнений) оказываются комплексные числа.

Под *комплексным числом* понимается выражение

$$z = x + iy, \quad (1)$$

где x и y — действительные числа, а i — *мнимая единица*.

Числа $x + i0 = x$ отождествляются с действительными числами; в частности, $0 + i0 = 0$.

Числа $0 + iy = iy$ называются *чисто мнимыми*.

Действительные числа x и y называются соответственно *действительной* и *мнимой частями* числа z и обозначаются следующим образом:

$$x = \operatorname{Re} z, \quad y = \operatorname{Im} z.$$

Под *модулем* комплексного числа z понимается неотрицательное число $|z| = \sqrt{x^2 + y^2}$.

Сопряженным числом $\bar{z}$ к числу (1) называется комплексное число $\bar{z} = x - iy$. Два комплексных числа $z_1 = x_1 + iy_1$ и $z_2 = x_2 + iy_2$ равны тогда и только тогда, когда $x_1 = x_2$ и $y_1 = y_2$.

Сложение, вычитание, умножение и деление комплексных чисел определяются следующим образом:

$$\text{I. } z_1 \pm z_2 = (x_1 \pm x_2) + i(y_1 \pm y_2).$$

$$\text{II. } z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

Отсюда, в частности,

$$i^2 = (0 + i1)(0 + i1) = (0 - 1) + i(0 + 0) = -1,$$

$$\bar{z}z = x^2 + y^2 = |z|^2.$$

$$\text{III. } \frac{z_1}{z_2} = \frac{\bar{z}_2 \cdot z_1}{z_2 \cdot \bar{z}_2} = \frac{(x_1 x_2 + y_1 y_2) + i(x_2 y_1 - x_1 y_2)}{x_2^2 + y_2^2} \quad (z_2 \neq 0).$$

2. Геометрическое изображение комплексных чисел. Рассмотрим плоскость с прямоугольной системой координат xOy (рис. 94). Так как комплексное число $z = x + iy$ является парой $(x; y)$ действительных чисел, а каждой паре $(x; y)$ действительных чисел соответствует одна точка плоскости и наоборот (см. § 1.1, п. 1), то каждую точку $(x; y)$ плоскости можно принять за изображение комплексного числа $z = x + iy$. В этом случае эта плоскость называется *комплексной плоскостью*, а z — точкой этой плоскости.

На оси Ox расположены действительные числа: $z = x + i0 = x$; поэтому она называется *действительной осью*. На оси Oy расположены чисто мнимые числа: $z = 0 + iy = iy$, она называется *мнимой осью*.

Заметим, что $r = |z|$ представляет собой расстояние от точки z до начала координат (см. § 1.2, п. 1).

Удобной является интерпретация комплексного числа $z = x + iy$ как радиуса-вектора $\overrightarrow{OM}$ (рис. 94). Очевидно, каждому радиусу-вектору в плоскости с концом в точке $M(x; y)$ соответствует комплексное число $z = x + iy$ и наоборот. Нулевому вектору соот-

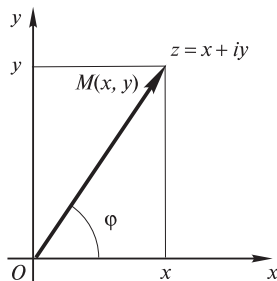


Рис. 94

ветствует комплексное число $0 + i0$. Сумме и разности комплексных чисел будет соответствовать сумма и разность радиусов-векторов, изображающих эти числа, при этом

$$|z_1 + z_2| \leq |z_1| + |z_2|,$$

$$|z_1 - z_2| \geq ||z_1| - |z_2||.$$

Положение точки z на плоскости, кроме ее прямоугольных координат x , y , может быть определено также и полярными координатами r , φ , при этом (см. § 1.1, п. 2)

$$x = r \cos \varphi, y = r \sin \varphi. \quad (2)$$

Число φ будем называть *аргументом* комплексного числа z . Аргумент считается положительным или отрицательным в зависимости от того, ведется ли его отсчет от положительного направления действительной оси против или по часовой стрелке соответственно.

По заданной точке z ее модуль определяется единственным образом, а аргумент — с точностью до слагаемого $2k\pi$, $k = 0, \pm 1, \pm 2, \dots$. Значение аргумента φ , удовлетворяющее условию $-\pi < \varphi \leq \pi$, называется *главным* и обозначается $\arg z$.

Точка $z = 0$ является единственной точкой комплексной плоскости, для которой аргумент не определен.

3. Умножение и деление комплексных чисел, записанных в тригонометрической форме. Из формул (2) получается тригонометрическая форма записи комплексного числа:

$$z = r(\cos \varphi + i \sin \varphi). \quad (3)$$

Пользуясь записью (3) для комплексных чисел

$$z_1 = r_1(\cos \varphi_1 + i \sin \varphi_1), z_2 = r_2(\cos \varphi_2 + i \sin \varphi_2),$$

имеем:

$$\begin{aligned} z_1 z_2 &= r_1 r_2 ((\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2) + i(\sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2)) = \\ &= r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)), \\ \frac{z_1}{z_2} &= \frac{r_1}{r_2} \cdot \frac{\cos \varphi_1 + i \sin \varphi_1}{\cos \varphi_2 + i \sin \varphi_2} = \frac{r_1}{r_2} \cdot \frac{(\cos \varphi_1 + i \sin \varphi_1)(\cos(-\varphi_2) + i \sin(-\varphi_2))}{(\cos \varphi_2 + i \sin \varphi_2)(\cos \varphi_2 - i \sin \varphi_2)} = \\ &= \frac{r_1}{r_2} (\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)) \quad (r_2 \neq 0). \end{aligned} \quad (4)$$

Следовательно, при умножении комплексных чисел их модули перемножаются, а аргументы складываются. При делении комплексных чисел их модули делятся, а аргументы вычитаются.

4. Возведение в степень и извлечение корня. Следствием формулы (4) является формула

$$z^n = r^n (\cos n\varphi + i \sin n\varphi), \quad (5)$$

где n — натуральное число.

Пусть

$$\sqrt[n]{z} = \rho(\cos \psi + i \sin \psi),$$

где $z = r(\cos \varphi + i \sin \varphi)$. Тогда на основании формулы (5) имеем:

$$z = (\rho (\cos \psi + i \sin \psi))^n = \rho^n (\cos n\psi + i \sin n\psi),$$

Отсюда

$$\rho^n = r, \quad n\psi = \varphi + 2k\pi \quad (k = 0, \pm 1, \pm 2, \dots)$$

и, следовательно,

$$\rho = \sqrt[n]{r} \text{ (под } \sqrt[n]{r} \text{ понимается арифметическое значение корня), } \psi = \frac{\varphi + 2k\pi}{n}.$$

Здесь в качестве k достаточно брать лишь значения $k = 0, 1, 2, \dots, n-1$, так как при прочих значениях k получаются повторения уже найденных значений корня. Таким образом, окончательно

$$\sqrt[n]{z} = \sqrt[n]{r}(\cos \varphi + i \sin \varphi) = \sqrt[n]{r} \left(\cos \frac{\varphi + 2k\pi}{n} + i \sin \frac{\varphi + 2k\pi}{n} \right), \quad (6)$$

$$k = 0, 1, 2, \dots, n-1.$$

П р и м е р. Найти $\omega = \sqrt{-1}$.

Так как $-1 = \cos \pi + i \sin \pi$, то на основании формулы (6) имеем:

$$\sqrt{-1} = \cos \frac{\pi + 2k\pi}{2} + i \sin \frac{\pi + 2k\pi}{2}, \quad k = 0, 1.$$

Отсюда

$$\omega_0 = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = i, \quad \omega_1 = \cos \frac{3\pi}{2} + i \sin \frac{3\pi}{2} = -i.$$

5. Формула Муавра. Формула Эйлера. Формула (5) может быть переписана в виде

$$r^n (\cos \varphi + i \sin \varphi)^n = r^n (\cos n\varphi + i \sin n\varphi).$$

Полагая здесь $r = 1$, получим формулу

$$(\cos \varphi + i \sin \varphi)^n = \cos n\varphi + i \sin n\varphi,$$

называемую *формулой Муавра* (Абрахам Муавр (1667—1754)— английский математик).

Справедлива и следующая формула:

$$\cos \varphi + i \sin \varphi = e^{i\varphi},$$

называемая *формулой Эйлера* (Леонард Эйлер (1707—1783)— великий математик, большую часть своей жизни провел в России, по происхождению швейцарец). Ее справедливость будет установлена позднее (см. § 15.5, п. 1).

6. Разложение многочленов на множители. Пусть

$$P_n(x) = A_0 + A_1x + A_2x^2 + \dots + A_{n-1}x^{n-1} + A_nx^n \quad (7)$$

многочлен степени n (т. е. $A_n \neq 0$) с вещественными коэффициентами $A_j, j = 0, 1, 2, \dots, n$. Число a (действительное или комплексное), такое, что $P_n(a) = 0$, называется *корнем многочлена* (7).

Из курса алгебры известно, что число a тогда и только тогда будет корнем многочлена $P_n(x)$, если $P_n(x)$ делится на $x - a$.

Если $P_n(x)$ делится на $(x - a)^k$ (k — натуральное) и не делится на $(x - a)^{k+1}$, то число k называется *кратностью корня* a .

Известно также, что всякий многочлен (7) степени n можно представить, и притом единственным образом, в следующем виде:

$$P_n(x) = A_n(x - a_1)^{k_1}(x - a_2)^{k_2} \dots (x - a_m)^{k_m},$$

где $a_1, a_2, \dots, a_m$ — различные корни многочлена (7), а числа $k_1, k_2, \dots, k_m$ — кратности соответственно корней $a_1, a_2, \dots, a_m$, причем $k_1 + k_2 + \dots + k_m = n$.

Заметим, что если многочлен с действительными коэффициентами имеет корнем комплексное число $a = b + ic$ кратности k , то (см. [10]) сопряженное комплексное число $\bar{a} = b - ic$ также является корнем многочлена той же кратности.

Имеем:

$$\begin{aligned} (x - a)(x - \bar{a}) &= (x - (b + ic))(x - (b - ic)) = ((x - b) - ci)((x - b) + ci) = \\ &= x^2 - 2bx + b^2 + c^2 = x^2 + px + q, \end{aligned}$$

где $p = -2b, q = b^2 + c^2$, причем $\frac{p^2}{4} - q < 0$.

Следовательно, произведение линейных множителей, соответствующих сопряженным корням, можно заменить квадратным трехчленом с действительными коэффициентами.

Таким образом, многочлен (7) можно представить в следующей форме:

$$\begin{aligned} P_n(x) &= A_n(x - \alpha_1)^{\lambda_1}(x - \alpha_2)^{\lambda_2} \dots \\ &\dots (x - \alpha_r)^{\lambda_r} \cdot (x^2 + p_1x + q_1)^{l_1}(x^2 + p_2x + q_2)^{l_2} \dots (x^2 + p_sx + q_s)^{l_s}. \end{aligned} \quad (8)$$

Здесь линейные двучлены соответствуют действительным корням ($\lambda_1, \lambda_2, \dots, \lambda_r$ — их кратности), а квадратные трехчлены — комплексным корням ($l_1, l_2, \dots, l_s$ — их кратности) многочлена. При этом $\lambda_1 + \lambda_2 + \dots + \lambda_r + 2l_1 + 2l_2 + \dots + 2l_s = n$,

$$\frac{p_j^2}{4} - q_j < 0, j = 1, 2, \dots, s.$$

7. Применение комплексных чисел в расчете физических величин. Так как комплексные числа геометрически представляются векторами на плоскости, то все векторные физические величины могут быть охарактеризованы при помощи комплексных чисел.

Представление векторных физических величин комплексными числами облегчает выполнение расчетов этих величин. При этом действия над векторами, которые выполняются графическим путем, заменяются соответствующими действиями над комплексными числами, что значительно проще.

Особенно широкое применение комплексные числа получили в электротехнике при расчете электрических цепей.

Отметим, что в электротехнике мнимая единица i обозначается буквой j , так как буквой i традиционно обозначается сила тока в цепи.

Рассмотрим один пример. На рисунке 95 дана векторная диаграмма неразветвленной цепи переменного тока. Пусть вектор $\overline{OM}$ представляет вектор напряжения $\vec{u}_1$, модуль которого $|\vec{u}_1| = 220$ (В); вектор $\overline{ON}$ представляет вектор напряжения $\vec{u}_2$, модуль которого $|\vec{u}_2| = 127$ (В). Тогда $\vec{u}_1 = 220 (\cos 60^\circ + j \sin 60^\circ) = 110 + 190,5j$. Также

$$\vec{u}_2 = 127 (\cos(-90^\circ) + j \sin(-90^\circ)) = -127j.$$

Если электрическая цепь составлена из двух последовательно включенных участков с напряжениями $\vec{u}_1$ и $\vec{u}_2$, то на зажимах будем иметь напряжение $\vec{u} = \vec{u}_1 + \vec{u}_2 = 110 + 190,5j - 127j = 110 + 63,5j$.

Аналогично применяются комплексные числа при выражении других характеристик электрических цепей.

Заметим, что в электротехнике модуль $|\vec{u}|$ вектора напряжения называется просто напряжением и обозначается через u , а соответствующее этому вектору комплексное число называется *комплексом* напряжения и обозначается $\dot{u}$ (ставится точка над символом).

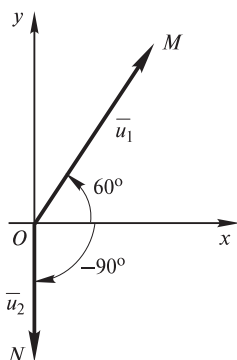


Рис. 95

Глава 8. ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ

§ 8.1. Понятие производной и ее механический и геометрический смысл

1. Задачи, приводящие к понятию производной.

Задача о скорости движущейся точки. Пусть $s = s(t)$ представляет закон прямолинейного движения материальной

точки. Это уравнение выражает путь s , пройденный точкой, как функцию времени t . Обозначим через Δs путь, пройденный за промежуток времени Δt от момента t до $t + \Delta t$, т. е. $\Delta s = s(t + \Delta t) - s(t)$. Отношение $\frac{\Delta s}{\Delta t}$ называется *средней скоростью* точки за время от t до $t + \Delta t$. Чем

меньше Δt , т. е. чем короче промежуток времени от t до $t + \Delta t$, тем лучше средняя скорость характеризует движение точки в момент времени t . Поэтому естественно ввести понятие скорости v в данный момент t , определив ее как предел средней скорости за промежуток от t до $t + \Delta t$, когда $\Delta t \rightarrow 0$: $v = \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t}$. Величина v называется *мгновенной скоростью* точки в данный момент t .

Задача о касательной к данной кривой. Пусть на плоскости xOy дана кривая уравнением $y = f(x)$. Требуется провести касательную к данной кривой в данной точке $M_0(x_0; f(x_0))$. Так как точка касания M_0 дана, то для решения задачи потребуются найти только угловой коэффициент искомой касательной, т. е. $\operatorname{tg} \varphi$ — тангенс угла наклона касательной к положительному направлению оси Ox (рис. 96).

Через точки $M_0(x_0; f(x_0))$ и $M'(x_0 + \Delta x; f(x_0 + \Delta x))$ проведем секущую M_0M' . Из рисунка 96 видно, что угловой коэффициент $\operatorname{tg} \alpha$ секущей M_0M' равен отношению:

$$\operatorname{tg} \alpha = \frac{\Delta y}{\Delta x},$$

где

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

Угловой коэффициент касательной M_0T к данной кривой в точке M_0

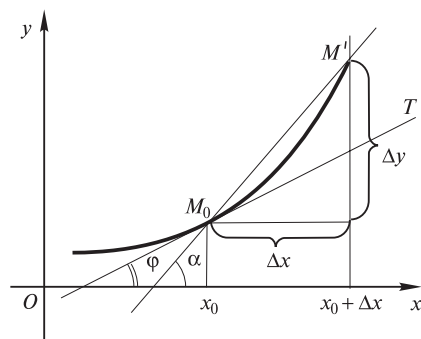


Рис. 96

может быть найден на основании следующего определения: *касательной к кривой в точке M_0* называется прямая M_0T , угловой коэффициент которой равен пределу углового коэффициента секущей M_0M' , когда $\Delta x \rightarrow 0$. Отсюда следует, что $\operatorname{tg} \varphi = \lim_{\Delta x \rightarrow 0} \operatorname{tg} \alpha = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$.

2. Определение производной.

Математическая операция, тре-

буемая для решения рассмотренных выше задач, одна и та же. Выясним аналитическую сущность этой операции, отвлекаясь от вызвавших ее конкретных вопросов.

Пусть функция $y = f(x)$ определена в интервале $(a; b)$. Возьмем какое-нибудь значение x из $(a; b)$. Затем возьмем новое значение аргумента $x + \Delta x$ из этого промежутка, придав первоначальному значению x приращение Δx (положительное или отрицательное). Этому новому значению аргумента соответствует и новое значение функции $y + \Delta y = f(x + \Delta x)$, где

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

Теперь составим отношение:

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

Оно является функцией от Δx .

Если существует предел отношения $\frac{\Delta y}{\Delta x}$ приращения функции Δy к вызвавшему его приращению аргумента Δx , когда Δx стремится к нулю, то этот предел называется *производной* функции $y = f(x)$ в данной точке x и обозначается через y' или $f'(x)$ (читается: «Игрек штрих» или «Эф штрих от икс»):

$$y' = f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \quad (1)$$

Для обозначения производной принят также и следующий символ: $\frac{dy}{dx}$ (читается: «Дэ игрек по дэ икс»). Этот символ надо рассматривать пока как целый символ, а не как частное.

Если существует предел справа $\lim_{\Delta x \rightarrow 0+0} \frac{\Delta y}{\Delta x}$ (или предел слева $\lim_{\Delta x \rightarrow 0-0} \frac{\Delta y}{\Delta x}$), то он называется *правой* (или *левой*) производной функции $f(x)$ в точке x .

Действие нахождения производной функции называется ее *дифференцированием*, а функцию, имеющую производную в точке x , называют *дифференцируемой* в этой точке. Функция, дифференцируемая в каждой точке промежутка, называется дифференцируемой в этом промежутке. При этом если промежуток от a до b есть отрезок $[a; b]$, то в точке a речь идет о правой производной, а в точке b — о левой производной.

Пример 1. Найти производную функции $y = C$, где C — постоянная.

Имеем $y + \Delta y = C$, $\Delta y = 0$, $\frac{\Delta y}{\Delta x} = 0$, $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = 0$, т. е. $y' = 0$. Следовательно, производная

постоянной равна нулю: $(C)' = 0$.

Пример 2. Найти производную функции $y = x$.

Имеем $y + \Delta y = x + \Delta x$, $\Delta y = \Delta x$, $\frac{\Delta y}{\Delta x} = 1$, $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = 1$, т. е. $y' = 1$. Итак, $(x)' = 1$.

Пример 3. Найти производную функции $y = \sin x$.

$$\begin{aligned} \text{Имеем } y + \Delta y &= \sin(x + \Delta x), \Delta y = \sin(x + \Delta x) - \sin x = 2 \sin \frac{\Delta x}{2} \cos \left(x + \frac{\Delta x}{2} \right), \frac{\Delta y}{\Delta x} = \\ &= \frac{\sin \frac{\Delta x}{2}}{\frac{\Delta x}{2}} \cos \left(x + \frac{\Delta x}{2} \right), \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\sin \frac{\Delta x}{2}}{\frac{\Delta x}{2}} \lim_{\Delta x \rightarrow 0} \cos \left(x + \frac{\Delta x}{2} \right) = \cos x \quad (\text{здесь используется} \end{aligned}$$

формула (9) из § 7.5 и непрерывность функции $\cos x$). Следовательно, $(\sin x)' = \cos x$.

Из рассмотренных выше задач, приводящих к понятию производной, следует:

1. *Скорость v прямолинейного движения есть производная пути s по времени t : $v = \frac{ds}{dt}$.* В этом состоит механический смысл производной. По

анalogии с этим производную любой функции часто называют скоростью изменения этой функции.

2. *Угловым коэффициентом касательной к кривой $y = f(x)$ в точке с абсциссой x_0 есть производная $f'(x_0)$.* В этом состоит геометрический смысл производной.

Поэтому с учетом формулы (9) из § 1.4 уравнение касательной к кривой $y = f(x)$ в точке $M_0(x_0; y_0)$ примет вид:

$$y - y_0 = f'(x_0)(x - x_0).$$

Нормалью к кривой в некоторой ее точке называется перпендикуляр к касательной к этой кривой в той же точке. Следовательно, если $f'(x_0) \neq 0$, то, имея в виду формулу (11) из § 1.4, уравнение нормали к кривой $y = f(x)$ в точке $M_0(x_0; y_0)$ можно записать так:

$$y - y_0 = -\frac{1}{f'(x_0)}(x - x_0).$$

Т е о р е м а. Если функция $y = f(x)$ дифференцируема в некоторой точке x , то она непрерывна в этой точке.

Д о к а з а т е л ь с т в о. Согласно формуле (1) и теореме 1 (§ 7.5, п. 1) имеем:

$$\frac{\Delta y}{\Delta x} = y' + \alpha(\Delta x),$$

где $\alpha(\Delta x)$ — бесконечно малая. Отсюда $\Delta y = y' \Delta x + \alpha(\Delta x) \Delta x$ и $\lim_{\Delta x \rightarrow 0} \Delta y = 0$, т. е. функция $y = f(x)$ непрерывна в данной точке x .

П р и м е ч а н и е. Обратное утверждение уже не имеет места, что видно из следующего примера:

П р и м е р 4. Функция $y = |x|$, график которой приведен на рисунке 97, непрерывна в точке $x = 0$, но ясно, что в этой точке в соответствии с геометрическим смыслом производной функция $y = |x|$ не дифференцируема, так как в ней нет определенной касательной.

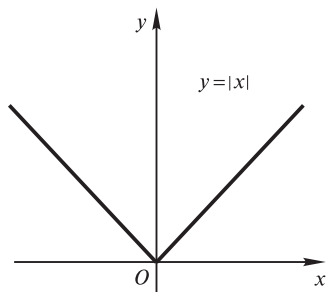


Рис. 97

§ 8.2. Правила дифференцирования функций и производные элементарных функций

1. Вывод общих правил дифференцирования. Пусть u и v — две функции аргумента x , имеющие производные u' и v' .

П р о и з в о д н а я с у м м ы. Пусть $y = u + v$. Тогда имеем:

$$y + \Delta y = (u + \Delta u) + (v + \Delta v), \Delta y = \Delta u + \Delta v,$$

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}, \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} + \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x},$$

т. е. $y' = u' + v'$, или $(u + v)' = u' + v'$.

П р и м е ч а н и е. Правило дифференцирования суммы двух слагаемых распространяется на случай алгебраической суммы любого конечного числа слагаемых, что доказывается аналогично.

П р о и з в о д н а я п р о и з в е д е н и я. Пусть $y = uv$. Тогда имеем:

$$y + \Delta y = (u + \Delta u)(v + \Delta v) = uv + v\Delta u + u\Delta v + \Delta v \cdot \Delta u,$$

$$\Delta y = v\Delta u + u\Delta v + \Delta u \cdot \Delta v, \frac{\Delta y}{\Delta x} = v \frac{\Delta u}{\Delta x} + u \frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta x} \Delta v,$$

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = v \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} + u \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} + \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} \cdot \lim_{\Delta x \rightarrow 0} \Delta v = vu' + uv' + u' \cdot 0,$$

т. е. $y' = u'v + uv'$, или

$$(uv)' = u'v + uv'. \quad (1)$$

При выводе формулы воспользовались тем, что в силу непрерывности функции v (ее непрерывность следует из ее дифференцируемости) $\lim_{\Delta x \rightarrow 0} \Delta v = 0$.

Вынесение постоянного множителя за знак производной. Так как $(C)' = 0$ (см. § 8.1, п. 2, пример 1), то из формулы (1) непосредственно получаем:

$$(Cu)' = Cu'.$$

Производная частного. Пусть $y = \frac{u}{v}$, где $v \neq 0$. Тогда имеем:

$$\Delta y + y = \frac{u + \Delta u}{v + \Delta v}, \Delta y = \frac{u + \Delta u}{v + \Delta v} - \frac{u}{v} = \frac{(u + \Delta u)v - (v + \Delta v)u}{(v + \Delta v)v} = \frac{v\Delta u - u\Delta v}{(v + \Delta v)v},$$

$$\frac{\Delta y}{\Delta x} = \frac{\frac{\Delta u}{\Delta x}v - u\frac{\Delta v}{\Delta x}}{(v + \Delta v)v}, \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{v \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} - u \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x}}{(v + \lim_{\Delta x \rightarrow 0} \Delta v)v} = \frac{vu' - uv'}{(v + 0)v},$$

т. е. $y' = \frac{u'v - uv'}{v^2}$, или

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

Производная сложной функции. Пусть $y = f(u)$, где $u = \varphi(x)$, причем $f(u)$ имеет производную по u , а $\varphi(x)$ — по x . Требуется найти производную y по x . Имеем $\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \cdot \frac{\Delta u}{\Delta x}$ (предполагается, что Δu при достаточно малых значениях Δx не обращается в нуль), откуда

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} \cdot \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x},$$

или

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}. \quad (2)$$

Пусть теперь $y = f(u)$, где $u = \varphi(v)$, а $v = \Psi(x)$, причем $f(u)$ имеет производную по u , $\varphi(v)$ — по v , а $\Psi(x)$ — по x . Здесь имеем цепочку трех зависимостей:

$$y = f(u), u = \varphi(v), v = \Psi(x).$$

Аналогично формуле (2) имеем:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dv} \cdot \frac{dv}{dx}.$$

Производная обратной функции. Пусть $y = f(x)$ и $x = \varphi(y)$ — взаимно обратные функции. Тогда если функция $y = f(x)$ имеет не равную нулю производную $f'(x)$, то обратная функция имеет производную $\varphi'(y)$ и

$$\varphi'(y) = \frac{1}{f'(x)},$$

или, короче,

$$x'_y = \frac{1}{y'_x}. \quad (3)$$

Действительно, так как $y = f(x)$ и $x = \varphi(y)$ — взаимно обратные функции, то $x = \varphi(f(x))$. Отсюда, используя формулу (2) дифференцирования сложной функции, получаем:

$$1 = \varphi'(y)f'(x),$$

откуда и следует искомая формула (3).

2. Производные элементарных функций. Пусть u , как и выше, — функция аргумента x , имеющая производную u' .

Производные тригонометрических функций. Как установлено ранее (§ 8.1, п. 2, пример 3),

$$(\sin x)' = \cos x.$$

Отсюда с учетом формулы (2)

$$(\sin u)' = u' \cos u. \quad (4)$$

На основании формулы (4) имеем:

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

Отсюда с учетом формулы (2) получаем:

$$(\cos u)' = -u' \sin u.$$

Далее имеем:

$$(\operatorname{tg} x)' = \left(\frac{\sin x}{\cos x} \right)' = \frac{(\sin x)' \cos x - \sin x (\cos x)'}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

Отсюда с учетом формулы (2)

$$(\operatorname{tg} u)' = \frac{u'}{\cos^2 u}. \quad (5)$$

Используя формулу (5), находим:

$$(\operatorname{ctg} x)' = \left(\operatorname{tg} \left(\frac{\pi}{2} - x \right) \right)' = \left(\frac{\pi}{2} - x \right)' \frac{1}{\cos^2 \left(\frac{\pi}{2} - x \right)} = -\frac{1}{\sin^2 x}.$$

Отсюда с учетом формулы (2)

$$(\operatorname{ctg} u)' = -\frac{u'}{\sin^2 u}.$$

П р о и з в о д н а я л о г а р и ф м а. Пусть $y = \ln x$. Тогда имеем:

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

$$\frac{\Delta y}{\Delta x} = \frac{\ln \left(1 + \frac{\Delta x}{x} \right)}{\Delta x} = \frac{1}{x} \cdot \frac{x}{\Delta x} \ln \left(1 + \frac{\Delta x}{x} \right) = \frac{1}{x} \ln \left(1 + \frac{\Delta x}{x} \right)^{\frac{x}{\Delta x}}.$$

Пользуясь известным пределом $e = \lim_{z \rightarrow 0} (1 + z)^{\frac{1}{z}}$ (§ 7.3, примечание),

будем иметь $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{1}{x}$, т. е. $y' = \frac{1}{x}$, или

$$(\ln x)' = \frac{1}{x}. \quad (6)$$

Пусть теперь $y = \log_a x$ ($a > 0, a \neq 1$). Тогда $a^y = x$. Отсюда $y \ln a = \ln x$, или $y = \frac{\ln x}{\ln a}$, откуда согласно формуле (6)

$$y' = \frac{1}{x \ln a},$$

или

$$(\log_a x)' = \frac{1}{x \ln a}. \quad (7)$$

Из формул (6), (7) с учетом формулы (2) получаем:

$$(\ln u)' = \frac{u'}{u}, \quad (8)$$

$$(\log_a u)' = \frac{u'}{u \ln a}.$$

В частности,

$$(\lg u)' = (\log_{10} u)' = \frac{u'}{u \ln 10}.$$

П р и м е р 1. $y = \ln \cos x, y' = \frac{(\cos x)'}{\cos x} = -\frac{\sin x}{\cos x} = -\operatorname{tg} x.$

Производная степенной функции. Пусть $y = x^\alpha$ (α — действительное число и $x > 0$). Тогда $\ln y = \alpha \ln x$, и согласно формуле (8) $\frac{y'}{y} = \alpha \frac{1}{x}$. Отсюда $y' = \alpha y \frac{1}{x} = \alpha x^\alpha \frac{1}{x} = \alpha x^{\alpha-1}$, т. е.

$$(x^\alpha)' = \alpha x^{\alpha-1}. \quad (9)$$

Формула (9) верна и в случае, когда функция $y = x^\alpha$ определена на всей числовой оси (например, когда α — натуральное число).

Из формулы (9) с учетом формулы (2) получаем:

$$(u^\alpha)' = \alpha u^{\alpha-1} u'. \quad (10)$$

Пример 2. Если $y = \sqrt{\sin x}$, то $y' = \frac{\cos x}{2\sqrt{\sin x}}$.

Производная показательной функции. Пусть $y = a^u$ ($0 < a \neq 1$). Тогда $\ln y = u \ln a$, и согласно формулам (8), (10) $\frac{y'}{y} = u' \ln a$. Отсюда

$$y' = y u' \ln a = a^u u' \ln a,$$

т. е.

$$(a^u)' = a^u u' \ln a.$$

В частности,

$$(e^u)' = e^u u'.$$

Пример 3. $y = 2^{x^2}$, $y' = 2^{x^2} \cdot 2x \ln 2 = 2^{x^2+1} x \ln 2$.

Производные обратных тригонометрических функций. Функция $y = \arcsin x$ является обратной по отношению к функции $x = \sin y$. Поэтому по правилу дифференцирования обратной функции получаем:

$$(\arcsin x)' = \frac{1}{(\sin y)'_y} = \frac{1}{\cos y} = \frac{1}{+\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}} \left(-\frac{\pi}{2} < y < \frac{\pi}{2} \right).$$

Таким же приемом получаем:

$$(\arccos x)' = \frac{1}{(\cos y)'_y} = -\frac{1}{\sin y} = -\frac{1}{+\sqrt{1-\cos^2 y}} = -\frac{1}{\sqrt{1-x^2}} \quad (0 < y < \pi),$$

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

$$(\operatorname{arctg} x)' = \frac{1}{(\operatorname{ctg} y)'_y} = -\sin^2 y = -\frac{1}{1+\operatorname{ctg}^2 y} = -\frac{1}{1+x^2}.$$

Отсюда с учетом формулы (2) получаем:

$$\begin{aligned}(\arcsin u)' &= \frac{u'}{\sqrt{1-u^2}}, \quad (\arctg u)' = \frac{u'}{1+u^2}, \\(\arccos u)' &= -\frac{u'}{\sqrt{1-u^2}}, \quad (\operatorname{arccctg} u)' = -\frac{u'}{1+u^2}.\end{aligned}$$

Пр и м е р 4. $(\arctg x^2)' = \frac{(x^2)'}{1+(x^2)^2} = \frac{2x}{1+x^4}.$

Для удобства нахождения производных различных функций сведем все правила и формулы дифференцирования в одну таблицу.

Правила дифференцирования и производные основных элементарных функций

- | | |
|---|--|
| 1. $(C)' = 0.$ | 10. $(\ln u)' = \frac{u'}{u}.$ |
| 2. $(u+v)' = u' + v'.$ | 11. $(\sin u)' = u' \cos u.$ |
| 3. $(uv)' = u'v + uv'.$ | 12. $(\cos u)' = -u' \sin u.$ |
| 4. $(Cu)' = Cu'.$ | 13. $(\operatorname{tg} u)' = \frac{u'}{\cos^2 u}.$ |
| 5. $\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}.$ | 14. $(\operatorname{ctg} u)' = -\frac{u'}{\sin^2 u}.$ |
| 6. $(u^a)' = \alpha u^{a-1} u'.$ | 15. $(\arcsin u)' = \frac{u'}{\sqrt{1-u^2}}.$ |
| 7. $(a^u)' = a^u u' \ln a.$ | 16. $(\arccos u)' = -\frac{u'}{\sqrt{1-u^2}}.$ |
| 8. $(e^u)' = e^u u'.$ | 17. $(\arctg u)' = \frac{u'}{1+u^2}.$ |
| 9. $(\log_a u)' = \frac{u'}{u \ln a}.$ | 18. $(\operatorname{arccctg} u)' = -\frac{u'}{1+u^2}.$ |

Здесь $u = u(x)$ и $v = v(x)$ — дифференцируемые функции.

§ 8.3. Дифференциал функции

1. Понятие дифференциала. Из определения производной

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = y'$$

с учетом теоремы 1 (§ 7.5, п. 1) получаем:

$$\frac{\Delta y}{\Delta x} = y' + \alpha, \quad (1)$$

где $\alpha = \alpha(\Delta x)$ — бесконечно малая при $\Delta x \rightarrow 0$.

Умножим обе части равенства (1) на Δx :

$$\Delta y = y' \Delta x + \alpha \Delta x.$$

Пусть $y' \neq 0$. Тогда первое слагаемое $y' \Delta x$ линейно по Δx , поскольку y' не зависит от Δx . При $\Delta x \rightarrow 0$ это слагаемое бесконечно мало, но порядок его малости ниже порядка малости второго слагаемого, так как для всех значений $y' \neq 0$

$$\lim_{\Delta x \rightarrow 0} \frac{\alpha \Delta x}{y' \Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\alpha}{y'} = 0.$$

Поэтому слагаемое $y' \Delta x$ является главной частью приращения функции. Это слагаемое называют *дифференциалом* функции $y = f(x)$ и обозначают символом dy или $df(x)$. Итак, $dy = y' \Delta x$.

2. Геометрический смысл дифференциала. Для выяснения геометрического смысла дифференциала к графику функции $y = f(x)$ в точке $M(x; y)$ проведем касательную MT , обозначив через φ ее угол наклона к положительному направлению оси Ox (рис. 98).

Так как $\operatorname{tg} \varphi = f'(x)$, то $dy = \operatorname{tg} \varphi \cdot \Delta x$. Поэтому из треугольника MLN следует, что дифференциал dy есть приращение ординаты касательной, соответствующее приращению аргумента Δx .

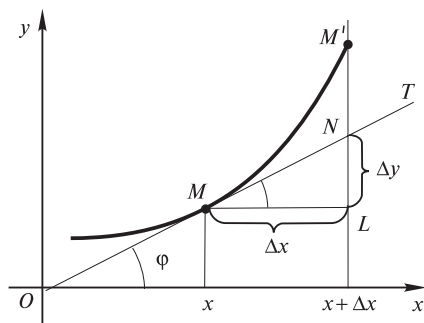


Рис. 98

Замечая, что $dx = x' \Delta x = \Delta x$, т. е. что дифференциал независимой переменной равен ее приращению, получаем:

$$dy = y' dx. \quad (2)$$

Таким образом, дифференциал функции равен произведению ее производной на дифференциал (или приращение) независимой переменной.

Из формулы (2) имеем:

$$y' = \frac{dy}{dx},$$

т. е. производная функции равна отношению дифференциала этой функции к дифференциалу аргумента.

3. Дифференциал сложной функции. Найдем выражение для дифференциала сложной функции. Пусть $y = f(u)$, где $u = \varphi(x)$, причем $f(u)$ имеет производную по u , $\varphi(x)$ — по x . Тогда по правилу дифференцирования сложной функции

$$\frac{dy}{dx} = f'_u(u)\varphi'(x).$$

Следовательно,

$$dy = f'_u(u)\varphi'(x)dx.$$

Но

$$\varphi'(x)dx = du.$$

Поэтому

$$dy = f'(u)du.$$

Таким образом, дифференциал сложной функции имеет тот же вид, какой он имел бы в том случае, если бы промежуточный аргумент ее был независимой переменной. Иначе говоря, форма записи дифференциала не зависит от того, является аргумент функции независимой переменной или функцией другого аргумента. Это свойство дифференциала называется *инвариантностью формы дифференциала*.

П р и м е р. Найти du сложной функции

$$y = \sin u, u = \sqrt{x}.$$

Имеем:

$$dy = \cos u du = \frac{1}{2\sqrt{x}} \cos u dx.$$

4. Таблица формул для дифференциалов. Согласно формуле (2) для получения дифференциала нужно умножить производную на dx (дифференциал независимой переменной). Это позволяет нам из таблицы формул для производных сразу получить соответствующую таблицу формул для дифференциалов. Например, из формулы $(u + v)' = u' + v'$, умножив обе части на dx , получим:

$$(u + v)' dx = u' dx + v' dx, \text{ или } d(u + v) = du + dv.$$

§ 8.4. Производные и дифференциалы высших порядков

1. Производные высших порядков. Производная $y' = f'(x)$ данной дифференцируемой функции $y = f(x)$, называемая *производной первого порядка*, представляет собой некоторую новую функцию. Возможно, что эта функция сама имеет производную. Тогда производная от

производной первого порядка называется *производной второго порядка* или *второй производной* и обозначается так: $y'' = (y')'$ или $f''(x)$. Аналогично, если существует производная от производной второго порядка, то она называется *производной третьего порядка* или *третьей производной* и обозначается так: $y''' = (y'')'$ или $f'''(x)$ и т. д.

Вообще производная от производной порядка $n - 1$ называется *производной порядка n* и обозначается $y^{(n)} = (y^{(n-1)})'$.

Функция $y = f(x)$ называется непрерывно дифференцируемой n раз, если существуют ее производные до порядка n включительно и эти производные непрерывны.

Производные четвертого, пятого и более высоких порядков обозначаются также с помощью римских цифр: $y^{\text{IV}}, y^{\text{V}}, y^{\text{VI}}$ и т. д. В таком случае порядок производной можно писать без скобок.

П р и м е р ы. 1) $y = x^k, y' = kx^{k-1}, y'' = k(k-1)x^{k-2}, \dots, y^{(n)} = k(k-1)(k-2)\dots(k-n+1)x^{k-n}$; если k — натуральное число, то

$$y^{(k)} = k! \text{ и } y^{(k+1)} = y^{(k+2)} = \dots = 0.$$

$$2) y = e^{kx}, y' = k e^{kx}, y'' = k^2 e^{kx}, \dots, y^{(n)} = k^n e^{kx}.$$

$$3) y = a^x, y' = a^x \ln a, y'' = a^x \ln^2 a, \dots, y^{(n)} = a^x \ln^n a.$$

$$4) y = \ln x, y' = \frac{1}{x}, y'' = -\frac{1}{x^2}, \dots, y^{(n)} = \frac{(-1) \cdot (-2) \cdot \dots \cdot (-n+1)}{x^n} = (-1)^{n-1} \frac{(n-1)!}{x^n}$$

$$5) y = \sin x, y' = \cos x = \sin\left(x + \frac{\pi}{2}\right), y'' = -\sin x = \sin(x + \pi) = \sin\left(x + 2 \cdot \frac{\pi}{2}\right), \dots, y^{(n)} = \sin\left(x + n \frac{\pi}{2}\right).$$

Совершенно аналогично устанавливается формула

$$(\cos x)^{(n)} = \cos\left(x + n \frac{\pi}{2}\right).$$

На случай производных любого порядка легко обобщаются правила дифференцирования суммы и вынесения постоянного множителя за знак производной (§ 8.2, п. 1):

$$(u + v)^{(n)} = u^{(n)} + v^{(n)}, (Cu)^{(n)} = Cu^{(n)}.$$

Выведем теперь формулу, дающую возможность вычислить производную порядка n от произведения двух функций uv . Применяя правила дифференцирования произведения и суммы, получим:

$$y' = u'v + uv',$$

$$y'' = u''v + u'v' + u'v' + uv'' = u''v + 2u'v' + uv'',$$

$$y''' = u'''v + u''v' + 2u''v' + 2u'v'' + u'v'' + uv''' = u'''v + 3u''v' + 3u'v'' + uv''''.$$

Продолжая процесс дифференцирования, придем к следующей формуле:

$$(uv)^{(n)} = u^{(n)}v + nu^{(n-1)}v' + \frac{n(n-1)}{1 \cdot 2}u^{(n-2)}v'' + \dots \\ \dots + \frac{n(n-1) \dots (n-k+1)}{1 \cdot 2 \dots k} \cdot u^{(n-k)}v^{(k)} + \dots + uv^{(n)},$$

называемой *формулой Лейбница* (Готфрид Лейбниц (1646—1716) — немецкий философ и математик).

П р и м е р. Если $u = e^x$, $v = x^2$, то по формуле Лейбница получим:

$$(e^x x^2)^{(20)} = e^x x^2 + 20e^x \cdot 2x + 190e^x \cdot 2 = e^x(x^2 + 40x + 380).$$

2. Дифференциалы высших порядков. Пусть имеем функцию $y = f(x)$, где x — независимая переменная. Ее дифференциал

$$dy = f'(x)dx$$

есть некоторая функция от x , но от x может зависеть только первый сомножитель $f'(x)$, второй же сомножитель является приращением независимой переменной x и от значения этой переменной не зависит. Так как dy есть функция от x , то мы имеем право говорить о дифференциале этой функции.

Дифференциал от дифференциала функции называется *дифференциалом второго порядка* или *вторым дифференциалом* этой функции и обозначается через d^2y :

$$d(dy) = d^2y.$$

Найдем выражение второго дифференциала. В силу определения дифференциала имеем:

$$d^2y = (f'(x)dx)'dx.$$

Так как dx от x не зависит, то dx при дифференцировании выносится за знак производной и мы получаем:

$$d^2y = f''(x)(dx)^2.$$

Принято, записывая степень дифференциала, опускать скобки; так, например, вместо $(dx)^2$ принято писать dx^2 , подразумевая под этим квадрат выражения dx ; вместо $(dx)^3$ пишут dx^3 и т. д.

Дифференциалом третьего порядка или *третьим дифференциалом* функции называется дифференциал от ее второго дифференциала:

$$d^3y = d(d^2y) = (f''(x) dx^2)'dx = f'''(x)dx^3.$$

Вообще *дифференциалом n -го порядка* называется первый дифференциал от дифференциала $(n - 1)$ -го порядка:

$$d^n y = d(d^{n-1} y) = (f^{(n-1)}(x) dx^{n-1})' dx = f^{(n)}(x) dx^n. \quad (1)$$

Отсюда получается другая запись для n -й производной:

$$f^{(n)}(x) = \frac{d^n y}{dx^n}. \quad (2)$$

П р и м е ч а н и е. Равенства (1) и (2) (при $n > 1$) верны, вообще говоря, лишь для того случая, когда x является независимой переменной.

Действительно, пусть имеем сложную функцию

$$f = f(u), \quad u = \varphi(x).$$

Тогда (§ 8.3, п. 3)

$$dy = f'(u) du. \quad (3)$$

Далее, используя формулу (3), получаем:

$$d^2 y = d(f'_u(u) du).$$

Но здесь $du = f'(x) dx$, вообще говоря, зависит от x , и потому получаем в силу формулы для дифференциала произведения (§ 8.3, п. 4):

$$d^2 y = d(f'_u(u)) du + f''_u(u) d(du),$$

или

$$d^2 y = f''_{u^2}(u) (du)^2 + f''_u(u) d^2 u,$$

где

$$d^2 u = u''(x) dx^2.$$

Аналогично находятся $d^3 y$ и т. д.

Таким образом, второй и последующие дифференциалы не обладают, вообще говоря, свойством инвариантности формы (§ 8.3, п. 3).

3. Физический смысл второй производной. Пусть $s = s(t)$ — уравнение прямолинейного движения материальной точки. Как установлено ранее (см. § 8.1, п. 1, 2), мгновенная скорость v этого движения есть производная пути s по времени t , т. е. $v = \frac{ds}{dt}$. Если теперь эту скорость рассматривать как функцию времени, то так же, как и в п. 1 (§ 8.1), установим, что $\frac{dv}{dt}$ есть ускорение a в момент t . Таким образом, получаем, что $a = \frac{d^2 s}{dt^2}$, т. е. вторая производная пути s по времени t есть ускорение a движущейся точки в момент t . В этом и заключается физический смысл второй производной.

З а д а ч а. Точка движется по прямой по закону $s = t^3$, где s — путь (см), а t — время (с).
Найти скорость и ускорение движения точки в момент $t = 2$ с.

Р е ш е н и е. Имеем $v = s' = 3t^2$, $a = s'' = 6t$.

В частности, при $t = 2$ с, $v = 12$ см/с и $a = 12$ см/с².

§ 8.5. Параметрическое задание функции и ее дифференцирование

Будем говорить, что переменная y как функция аргумента x задана *параметрически*, если обе переменные x и y заданы как функции некоторой третьей переменной t : $x = \varphi(t)$, $y = \psi(t)$. При этом указанную переменную t обычно называют *параметром*.

Примеры такого задания функций уже встречались (см. § 1.6, п. 2) при рассмотрении параметрических уравнений линий (парабола, окружность, циклоида), дающих выражения текущих координат x и y в виде функций от параметра t .

Будем предполагать, что функции $\varphi(t)$ и $\psi(t)$ имеют нужное число производных по параметру t в рассматриваемом промежутке изменения этого параметра. Кроме того, будем считать также, что функция $x = \varphi(t)$ имеет обратную функцию $t = \Phi(x)$, что позволяет рассматривать переменную y как функцию переменной x .

Рассмотрим вопрос о вычислении производных функции $y = y(x)$ по аргументу x . Имеем (в силу инвариантности формы первого дифференциала; § 8.3, п. 3):

$$y'(x) = \frac{dy}{dx}, \quad dy = \psi'(t)dt, \quad dx = \varphi'(t)dt. \quad (1)$$

Отсюда следует, что

$$y'(x) = \frac{\psi'(t)}{\varphi'(t)}. \quad (2)$$

Для вычисления второй производной $y''(x)$ представим ее (в силу инвариантности формы первого дифференциала) в виде

$$y''(x) = \frac{d(y'(x))}{dx}. \quad (3)$$

Теперь, используя в правой части (3) формулу (2), третью из формул (1) и правило дифференцирования частного, получим:

$$y''(x) = \frac{\left(\frac{\psi'(t)}{\varphi'(t)} \right)' dt}{\varphi'(t)dt} = \frac{\psi''(t)\varphi'(t) - \varphi''(t)\psi'(t)}{(\varphi'(t))^3}. \quad (4)$$

По такому же принципу вычисляются производные третьего и последующих порядков.

П р и м е р. Вычислить первую и вторую производные функции $y(x)$, заданной параметрически:

$$x = R(t - \sin t), y = R(1 - \cos t) \quad (-\infty < t < +\infty).$$

Как известно (§ 1.6, п. 2), кривая, определяемая этими уравнениями, называется циклоидой. Пользуясь формулами (2) и (4), получим:

$$y'(x) = \frac{R \sin t}{R(1 - \cos t)} = \operatorname{ctg} \frac{t}{2};$$

$$y''(x) = \frac{\left(\operatorname{ctg} \frac{t}{2}\right)'}{R(1 - \cos t)} = -\frac{1}{4R \sin^4 \frac{t}{2}}$$

($t \neq 2\pi k$, где k — целое число).

§ 8.6. Свойства дифференцируемых функций

1. Теорема Ферма (Пьер Ферма (1601—1665) — французский математик). Если функция $y = f(x)$, определенная в интервале $(a; b)$, достигает в некоторой точке c этого интервала наибольшего (или наименьшего) значения и существует производная $f'(c)$, то $f'(c) = 0$.

Д о к а з а т е л ь с т в о. Допустим, что в точке c функция $f(x)$ достигает наибольшего значения. Придадим значению c достаточно малое приращение Δx . Тогда $f(c + \Delta x) < f(c)$. Отсюда при $\Delta x < 0$

$$\frac{\Delta y}{\Delta x} = \frac{f(c + \Delta x) - f(c)}{\Delta x} > 0, \text{ и, следовательно,}$$

$$\lim_{\Delta x \rightarrow 0-0} \frac{\Delta y}{\Delta x} = f'(c) \geq 0. \quad (1)$$

При $\Delta x > 0$ $\frac{\Delta y}{\Delta x} < 0$, и, следовательно,

$$\lim_{\Delta x \rightarrow 0+0} \frac{\Delta y}{\Delta x} = f'(c) \leq 0. \quad (2)$$

Из неравенств (1) и (2) следует, что $f'(c) = 0$.

Геометрический смысл заключения теоремы состоит в том, что касательная к графику функции $y = f(x)$ в точке c абсциссой c параллельна оси абсцисс (рис. 99).

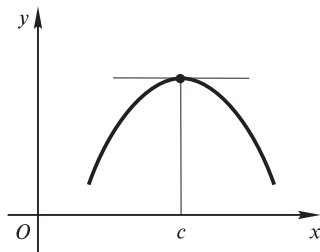


Рис. 99

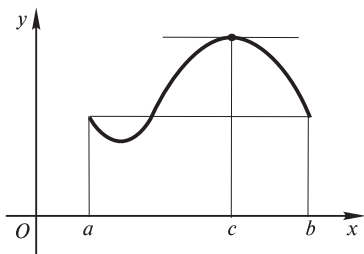


Рис. 100

цируемая в интервале $(a; b)$, принимает на концах этого сегмента равные значения $f(a) = f(b)$, то в интервале $(a; b)$ существует точка c , такая, что $f'(c) = 0$.

Доказательство. Так как функция $f(x)$ непрерывна на сегменте $[a; b]$, то, как известно (см. § 7.6, п. 3, теорема 1), она принимает на этом сегменте как свое наибольшее значение M , так и свое наименьшее значение m . Возможны только два случая.

1. $M = m$. Тогда $f(x)$ постоянна на $[a; b]$: в самом деле, неравенство $m \leq f(x) \leq M$ в этом случае дает $f(x) = M$ для всех x из $[a; b]$. Поэтому в любой точке интервала $(a; b)$ $f'(x) = 0$.

2. $M > m$. Так как $f(a) = f(b)$, то хоть одно из значений M и m достигается в некоторой точке c ($a < c < b$). Следовательно, согласно теореме Ферма $f'(c) = 0$. Теорема доказана.

Геометрически теорема Ролля означает следующее: если крайние ординаты кривой $y = f(x)$ равны, то на кривой найдется точка, где касательная параллельна оси абсцисс (рис. 100).

Примечание 2. Условия теоремы Ролля являются существенными. Так, например, для функции $f(x) = |x|$, $-1 \leq x \leq 1$ (см. рис. 97), выполнены все условия теоремы Ролля, кроме существования производной в точке $x = 0$. Одновременно с этим замечаем, что в интервале $(-1; 1)$ нет такой точки, где производная равна нулю: $f'(x) = -1$, если $-1 < x < 0$, $f'(x) = 1$, если $0 < x < 1$, а при $x = 0$ производная $f'(x)$, как уже отмечалось, не существует.

3. Теорема Лагранжа (Жозеф Луи Лагранж (1736—1813) — французский математик и механик). Если функция $y = f(x)$ непрерывна на сегменте $[a; b]$ и дифференцируема в интервале $(a; b)$, то в интервале $(a; b)$ найдется такая точка c , что

$$\frac{f(b) - f(a)}{b - a} = f'(c). \quad (3)$$

Примечание 1. Все условия теоремы Ферма существенны. Например, функция $y = \sin x$ в промежутке $0 \leq x \leq \frac{\pi}{2}$ достигает в точке $x_0 = 0$ наименьшего значения, но ее производная в этой точке равна единице.

2. Теорема Ролля (Мишель Ролль (1652—1719) — французский математик). Если функция $y = f(x)$, непрерывная на сегменте $[a; b]$ и дифференцируемая в интервале $(a; b)$, принимает на концах этого сегмента равные значения $f(a) = f(b)$, то в интервале $(a; b)$ существует точка c , такая, что $f'(c) = 0$.

цируемая в интервале $(a; b)$, принимает на концах этого сегмента равные значения $f(a) = f(b)$, то в интервале $(a; b)$ существует точка c , такая, что $f'(c) = 0$.

Доказательство. Так как функция $f(x)$ непрерывна на сегменте $[a; b]$, то, как известно (см. § 7.6, п. 3, теорема 1), она принимает на этом сегменте как свое наибольшее значение M , так и свое наименьшее значение m . Возможны только два случая.

1. $M = m$. Тогда $f(x)$ постоянна на $[a; b]$: в самом деле, неравенство $m \leq f(x) \leq M$ в этом случае дает $f(x) = M$ для всех x из $[a; b]$. Поэтому в любой точке интервала $(a; b)$ $f'(x) = 0$.

2. $M > m$. Так как $f(a) = f(b)$, то хоть одно из значений M и m достигается в некоторой точке c ($a < c < b$). Следовательно, согласно теореме Ферма $f'(c) = 0$. Теорема доказана.

Геометрически теорема Ролля означает следующее: если крайние ординаты кривой $y = f(x)$ равны, то на кривой найдется точка, где касательная параллельна оси абсцисс (рис. 100).

Примечание 2. Условия теоремы Ролля являются существенными. Так, например, для функции $f(x) = |x|$, $-1 \leq x \leq 1$ (см. рис. 97), выполнены все условия теоремы Ролля, кроме существования производной в точке $x = 0$. Одновременно с этим замечаем, что в интервале $(-1; 1)$ нет такой точки, где производная равна нулю: $f'(x) = -1$, если $-1 < x < 0$, $f'(x) = 1$, если $0 < x < 1$, а при $x = 0$ производная $f'(x)$, как уже отмечалось, не существует.

3. Теорема Лагранжа (Жозеф Луи Лагранж (1736—1813) — французский математик и механик). Если функция $y = f(x)$ непрерывна на сегменте $[a; b]$ и дифференцируема в интервале $(a; b)$, то в интервале $(a; b)$ найдется такая точка c , что

$$\frac{f(b) - f(a)}{b - a} = f'(c). \quad (3)$$

Доказательство. Положим

$$\frac{f(b) - f(a)}{b - a} = \lambda \quad (4)$$

и рассмотрим вспомогательную функцию $\varphi(x) = f(x) - f(a) - \lambda(x - a)$. Эта функция удовлетворяет первым двум условиям теоремы Ролля как алгебраическая сумма трех непрерывных и дифференцируемых функций. При этом $\varphi(a) = \varphi(b) = 0$. Следовательно, к функции $\varphi(x)$ применима теорема Ролля, т. е. существует точка c , $a < c < b$, такая, что $\varphi'(c) = 0$. Но $\varphi'(x) = f'(x) - \lambda$. Поэтому $f'(c) - \lambda = 0$, или $\lambda = f'(c)$. Отсюда с учетом формулы (4) получаем искомое равенство (3).

Теорема Лагранжа имеет простой геометрический смысл (рис. 101): на графике функции $y = f(x)$ между точками A и B есть внутренняя точка C , такая, что касательная к нему в точке C параллельна хорде AB . В самом деле, левая часть равенства (3) — угловой коэффициент хорды AB , а правая — угловой коэффициент касательной к графику в точке C .

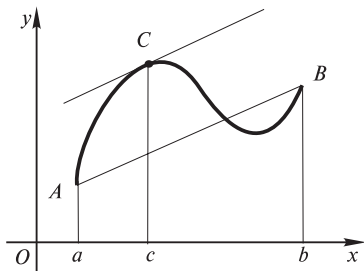


Рис. 101

Примечание 3. Теорема Лагранжа является обобщением теоремы Ролля, так как если $f(a) = f(b)$, то из равенства (3) следует $f'(c) = 0$.

Формула (3) называется **формулой Лагранжа** или **формулой конечных приращений**. Из нее получаем $f(b) - f(a) = f'(c)(b - a)$. Наконец, взяв вместо a и b соответственно x_0 и x и обозначив $\Delta x = x - x_0$, $\Delta y = f(x) - f(x_0)$, формулу Лагранжа запишем так:

$$\Delta y = f'(c)\Delta x.$$

Из теоремы Лагранжа вытекает:

Следствие. Если $f'(x) = 0$ в интервале $(a; b)$, то в этом интервале функция $f(x)$ постоянна.

Доказательство. Для любых значений x_1 и x_2 ($x_1 < x_2$) из рассматриваемого интервала справедлива теорема Лагранжа, т. е. $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$, где $x_1 < c < x_2$. Но $f'(c) = 0$, а потому и $f(x_2) - f(x_1) = 0$, т. е. $f(x_1) = f(x_2)$, а это значит, что $f(x) = \text{const}$ в интервале $(a; b)$.

4. Теорема Коши. Если функции $f(x)$ и $g(x)$: 1) непрерывны на отрезке $[a; b]$; 2) дифференцируемы в интервале $(a; b)$; 3) $g'(x) \neq 0$ в этом интервале —

ле, то в интервале $(a; b)$ существует точка c , такая, что имеет место равенство

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}. \quad (5)$$

Д о к а з а т е л ь с т в о. Отметим, что $g(b) - g(a) \neq 0$, так как в противном случае имели бы, что $g(a) = g(b)$, и тогда по теореме Ролля $g'(c_1) = 0$, где c_1 — некоторая точка из интервала $(a; b)$, что противоречит условию 3.

Рассмотрим вспомогательную функцию

$$\varphi(x) = f(x) - f(a) - \lambda(g(x) - g(a)),$$

где

$$\lambda = \frac{f(b) - f(a)}{g(b) - g(a)}. \quad (6)$$

Имеем:

$$\varphi(a) = 0, \varphi(b) = 0.$$

т. е.

$$\varphi(a) = \varphi(b).$$

Функция $\varphi(x)$ удовлетворяет и остальным условиям теоремы Ролля. В самом деле, $\varphi(x)$ непрерывна на $[a; b]$, так как непрерывны на $[a; b]$ $f(x)$ и $g(x)$; производная $\varphi'(x)$ существует в $(a; b)$, она равна:

$$\varphi'(x) = f'(x) - \lambda g'(x).$$

Следовательно, в интервале $(a; b)$ существует такая точка c , что $\varphi'(c) = 0$ или $f'(c) - \lambda g'(c) = 0$, откуда

$$\lambda = \frac{f'(c)}{g'(c)}.$$

Подставляя в равенство (6) значение λ , получаем искомое равенство (5).

П р и м е ч а н и е. Заметим, что теорема Лагранжа является частным случаем теоремы Коши, соответствующим $g(x) \equiv x$.

5. Правило Лопиталья. В главе 7 (см. § 7.5, п. 2) мы познакомились с некоторыми приемами нахождения пределов отношения двух бесконечно малых или бесконечно больших, т. е. раскрытия неопределенности соответственно вида $\frac{0}{0}$ или $\frac{\infty}{\infty}$. Здесь будет рассмотрен новый простой прием для раскрытия этих неопределенностей, называемый

правилом Лопиталья (Гийом Лопиталь (1661—1704) — французский математик).

Рассмотрим отношение

$$\frac{f(x)}{\varphi(x)},$$

где функции $f(x)$ и $\varphi(x)$ определены и дифференцируемы в некоторой окрестности точки a , исключая, быть может, саму точку a . Пусть далее эти функции одновременно являются бесконечно малыми или бесконечно большими.

Т е о р е м а. *Предел отношения двух бесконечно малых или бесконечно больших функций равен пределу отношения их производных, если последний существует, т. е.*

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{\varphi'(x)}.$$

Д о к а з а т е л ь с т в о. Доказательство проведем для неопределенности $\frac{0}{0}$, причем для простоты будем предполагать, что функции $f(x)$ и $\varphi(x)$ на некотором отрезке $[a; b]$ удовлетворяют условиям теоремы Коши, при этом $f(a) = \varphi(a) = 0$. В случае неопределенности $\frac{\infty}{\infty}$ доказательство несколько сложнее (см., например, [11]).

Возьмем на отрезке $[a; b]$ какую-нибудь точку $x \neq a$. Применяя формулу Коши, будем иметь:

$$\frac{f(x) - f(a)}{\varphi(x) - \varphi(a)} = \frac{f'(c)}{\varphi'(c)},$$

где $a < c < x$. Но по условию $f(a) = \varphi(a) = 0$, и, значит,

$$\frac{f(x)}{\varphi(x)} = \frac{f'(c)}{\varphi'(c)}.$$

Если x стремится к a , то и c стремится к a , так как c заключено между a и x . Поэтому можем записать

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow a} \frac{f'(c)}{\varphi'(c)} = \lim_{c \rightarrow a} \frac{f'(c)}{\varphi'(c)} = \lim_{x \rightarrow a} \frac{f'(x)}{\varphi'(x)}$$

и окончательно

$$\lim_{x \rightarrow a} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{\varphi'(x)}.$$

П р и м е р 1. $\lim_{x \rightarrow 0} \frac{2^x - 1}{\sin x} = \lim_{x \rightarrow 0} \frac{(2^x - 1)'}{(\sin x)'} = \lim_{x \rightarrow 0} \frac{2^x \ln 2}{\cos x} = \ln 2.$

Пр и м е р 2. $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = \lim_{x \rightarrow 0} \frac{e^x}{1} = 1.$

Иногда правило Лопиталья приходится применять несколько раз.

Пр и м е р 3. $\lim_{x \rightarrow \frac{\pi}{2}} \frac{1 - \sin x}{\left(\frac{\pi}{2} - x\right)^2} = \lim_{x \rightarrow \frac{\pi}{2}} \frac{\cos x}{\pi - 2x} = \lim_{x \rightarrow \frac{\pi}{2}} \frac{\sin x}{2} = \frac{1}{2}.$

Пр и м е ч а н и е. Правило Лопиталья верно и в том случае, когда a — символ ∞ .

Пр и м е р 4. $\lim_{x \rightarrow +\infty} \frac{x}{e^x} = \lim_{x \rightarrow +\infty} \frac{1}{e^x} = 0.$

Случаи других неопределенностей

$$0 \cdot \infty, \infty^0, 1^\infty, \infty - \infty$$

с помощью тождественных преобразований сводятся к основным типам неопределенностей $\frac{0}{0}$ или $\frac{\infty}{\infty}$.

Пр и м е р 5. $\lim_{x \rightarrow 0+0} x \ln x.$

Здесь мы имеем неопределенность $0 \cdot \infty$. Переписывая данное выражение в виде

$$\lim_{x \rightarrow 0+0} x \ln x = \lim_{x \rightarrow 0+0} \frac{\ln x}{\frac{1}{x}},$$

получаем неопределенность вида $\frac{\infty}{\infty}$. Отсюда, применяя правило Лопиталья, находим:

$$\lim_{x \rightarrow 0+0} x \ln x = \lim_{x \rightarrow 0+0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = - \lim_{x \rightarrow 0+0} x = 0.$$

Пр и м е р 6. $\lim_{x \rightarrow 0} \left(\operatorname{ctg} x - \frac{1}{x} \right).$

Данное выражение представляет собой неопределенность вида $\infty - \infty$. Пользуясь тем, что $\operatorname{ctg} x = \frac{\cos x}{\sin x}$, имеем:

$$\lim_{x \rightarrow 0} \left(\operatorname{ctg} x - \frac{1}{x} \right) = \lim_{x \rightarrow 0} \left(\frac{\cos x}{\sin x} - \frac{1}{x} \right) = \lim_{x \rightarrow 0} \frac{x \cos x - \sin x}{x \sin x}.$$

Так как получилась неопределенность $\frac{0}{0}$, то применяем правило Лопиталья:

$$\begin{aligned} \lim_{x \rightarrow 0} \left(\operatorname{ctg} x - \frac{1}{x} \right) &= \lim_{x \rightarrow 0} \frac{\cos x - x \sin x - \cos x}{\sin x + x \cos x} = - \lim_{x \rightarrow 0} \frac{x \sin x}{\sin x + x \cos x} = \\ &= - \lim_{x \rightarrow 0} \frac{\sin x + x \cos x}{\cos x + \cos x - x \sin x} = - \frac{0}{2} = 0. \end{aligned}$$

Пр и м е р 7. $\lim_{x \rightarrow 0+0} x^x$. Здесь неопределенность вида 0^0 . Положив $y = x^x$, имеем $\ln y = x \ln x$, откуда в силу непрерывности логарифмической функции

$$\ln \lim_{x \rightarrow 0+0} y = \lim_{x \rightarrow 0+0} \ln y = \lim_{x \rightarrow 0+0} x \ln x = 0 \text{ (пример 5).}$$

Следовательно, $\ln \lim_{x \rightarrow 0+0} y = 0$, откуда

$$\lim_{x \rightarrow 0+0} y = e^0 = 1, \text{ т.е. } \lim_{x \rightarrow 0+0} x^x = 1.$$

Использовавшийся в последнем примере прием логарифмирования применяется также и в случае неопределенностей ∞^0 , 1^∞ .

§ 8.7. Возрастание и убывание функций. Максимум и минимум

1. Возрастание и убывание функций. Функция $f(x)$ называется *возрастающей* (*убывающей*) в интервале $(a; b)$, если, каковы бы ни были значения x_1 и x_2 из этого интервала, из неравенства $x_2 > x_1$, вытекает неравенство $f(x_2) > f(x_1)$ (соответственно $f(x_2) < f(x_1)$). Если же для таких x_1 и x_2 из неравенства $x_2 > x_1$ следует неравенство $f(x_2) \geq f(x_1)$ ($f(x_2) \leq f(x_1)$), то функция $f(x)$ называется *неубывающей* (*невозрастающей*) в интервале $(a; b)$.

Функции всех этих типов носят общее название *монотонные*.

Монотонные функции часто встречаются в различных исследованиях. Например, освещенность, меняющаяся по мере удаления от источника света, — монотонно убывающая функция расстояния.

Т е о р е м а 1. Если функция $y = f(x)$, дифференцируемая в интервале $(a; b)$, неубывающая (невозрастающая) на нем, то ее производная в этом интервале не отрицательна (не положительна), т. е.

$$f'(x) \geq 0 \text{ (} f'(x) \leq 0 \text{)}.$$

Д о к а з а т е л ь с т в о. Пусть x — произвольное значение из интервала $(a; b)$. Придадим этому значению x приращение Δx , такое, чтобы точка $x + \Delta x$ принадлежала интервалу $(a; b)$. Если $f(x)$ — неубывающая функция, то $\Delta y \geq 0$ при $\Delta x > 0$ и $\Delta y \leq 0$ при $\Delta x < 0$. В обоих случаях $\frac{\Delta y}{\Delta x} \geq 0$, и, следовательно, $f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} \geq 0$. Если же $f(x)$ — невозрастающая функция, то $\frac{\Delta y}{\Delta x} \leq 0$ и $f'(x) \leq 0$.

Т е о р е м а 2. Если функция $f(x)$, дифференцируемая в интервале $(a; b)$, удовлетворяет в нем условию $f'(x) > 0$ ($f'(x) < 0$), то эта функция *возрастает* (*убывает*) в интервале $(a; b)$.

Д о к а з а т е л ь с т в о. Согласно формуле Лагранжа для произвольных x_1 и x_2 ($x_1 < x_2$) из $(a; b)$ имеем: $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$, где $x_1 < c < x_2$. Следовательно, если $f'(x) > 0$ в $(a; b)$, то $f(x_2) - f(x_1) > 0$, или $f(x_2) > f(x_1)$, и заданная функция возрастает в $(a; b)$. Если же $f'(x) < 0$ в $(a; b)$, то $f(x_2) - f(x_1) < 0$, или $f(x_2) < f(x_1)$, и данная функция убывает.

Приведем несколько примеров исследования функций на возрастание и убывание.

Пример 1. Функция $y = e^x$ всюду возрастает, так как $y' = e^x > 0$ для всех x .

Пример 2. Функция $y = x^2$ убывает в промежутке $(-\infty; 0)$, так как в этом промежутке $y' = 2x < 0$. Эта же функция в промежутке $(0; +\infty)$ возрастает, так как в последнем промежутке $y' = 2x > 0$.

2. Максимумы и минимумы функций.

Определение 1. Говорят, что функция $f(x)$ имеет в точке x_0 *максимум (минимум)*, если существует такая окрестность точки $x_0(x_0 - \delta, x_0 + \delta)$, что для всех x из этой окрестности, отличных от x_0 , выполняется неравенство $f(x) < f(x_0)$ ($f(x) > f(x_0)$).

Иначе говоря, функция $f(x)$ имеет в точке x_0 максимум (минимум), если для достаточно малого приращения Δx (любого знака) выполняется неравенство

$$f(x_0 + \Delta x) < f(x_0) \quad (f(x_0 + \Delta x) > f(x_0)).$$

Максимум или минимум функции называется *экстремумом* функции.

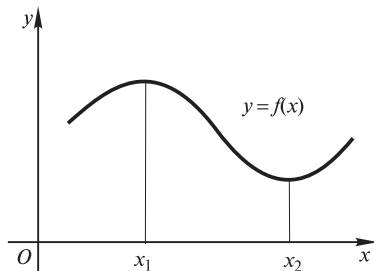


Рис. 102

По определению максимумов и минимумов функции они могут достигаться лишь внутри области определения, концы сегментов области определения не могут служить точками, в которых функция принимает экстремум.

На рисунке 102 изображен график функции, которая принимает в точке x_1 максимум, а в точке x_2 минимум.

Если исследуемая на экстремум функция дифференцируема, то изучение свойств ее производной дает возможность находить точки, в которых функция принимает экстремум.

Теорема 1 (необходимое условие существования экстремума). Если функция $f(x)$, дифференцируемая в интервале $(a; b)$, имеет в точке x_0 , $a < x_0 < b$, экстремум, то ее производная в этой точке равна нулю:

$$f'(x_0) = 0. \quad (1)$$

Эта теорема есть непосредственное следствие теоремы Ферма.

Примечание. Условие (1), будучи необходимым условием экстремума, не является достаточным условием экстремума, что показывает следующий пример.

Пример 1. Функция $f(x) = x^3$ не имеет экстремума в точке $x_0 = 0$ (разность $f(x) - f(0)$ меняет знак при изменении знака аргумента x), хотя ее производная $y' = 3x^2$ обращается в этой точке в нуль.

Теорема 2 (достаточное условие существования экстремума). Если производная функции $f(x)$ обращается в точке x_0 в нуль (такие точки называются стационарными) и при переходе через эту точку в направлении возрастания x меняет знак «плюс» («минус») на «минус» («плюс»), то в точке x_0 эта функция имеет максимум (минимум). Если же при переходе через точку x_0 производная функции $f(x)$ не меняет знака, то в этой точке функция $f(x)$ экстремума не имеет.

Доказательство. Допустим, что $f'(x)$ меняет знак «плюс» на «минус». Тогда согласно теореме 2 (п. 1) в достаточно малой окрестности точки x_0 слева от x_0 функция $f(x)$ возрастает и $f(x) < f(x_0)$, а справа от нее функция $f(x)$ убывает и снова $f(x) < f(x_0)$. Следовательно, для всех x из достаточно малой окрестности точки x_0 (кроме самой этой точки) будет выполняться неравенство $f(x) < f(x_0)$, т. е. в точке x_0 функция $f(x)$ имеет максимум.

Аналогичное доказательство и в случае обратной смены знака.

Предположим теперь, что при переходе через точку x_0 производная функция $f(x)$ не меняет знака. Тогда по теореме 2 п. 1 как слева, так и справа от x_0 функция $f(x)$ либо возрастает, либо убывает и, следовательно, не может иметь экстремума в точке x_0 .

Отсюда следует такое *правило исследования функции на экстремум* с помощью первой производной. Пусть в интервале $(a; b)$ дана дифференцируемая функция $f(x)$:

- 1) находим ее производную $f'(x)$;
- 2) находим корни уравнения $f'(x) = 0$;
- 3) выясняем знак $f'(x)$ слева и справа от каждого из этих корней и согласно теореме 2 выносим заключение об экстремуме;
- 4) вычисляем значения функции в точках экстремума.

Пример 2. Исследовать на экстремум функцию $f(x) = x^3 - 3x + 2$. Для этого вычисляем производную $f'(x) = 3x^2 - 3 = 3(x^2 - 1) = 3(x + 1)(x - 1)$ и находим корни уравнения $f'(x) = 0$. Имеем $x_1 = -1$, $x_2 = 1$. Затем согласно полученному правилу последовательно заполняем строки таблицы:

x	$x < -1$	$x_1 = -1$	$-1 < x < 1$	$x_2 = 1$	$x > 1$
$f'(x)$	+	0	-	0	+
$f(x)$		Максимум $f(x_1) = 4$		Минимум $f(x_2) = 0$	

3. Исследование функций на экстремум с помощью второй производной. Следующая теорема является вторым достаточным условием существования экстремума.

Т е о р е м а. Пусть функция $f(x)$ имеет в точке x_0 и ее окрестности непрерывные первую и вторую производные, причем $f'(x_0) = 0, f''(x_0) \neq 0$. Тогда функция $f(x)$ имеет в точке x_0 минимум (максимум), если $f''(x_0) > 0$ ($f''(x_0) < 0$).

Д о к а з а т е л ь с т в о. Пусть $f''(x_0) > 0$. Так как $f''(x)$ непрерывна в точке x_0 , то (см. § 7.6, п. 1, теорема 5) $f''(x) > 0$ и в некоторой окрестности точки x_0 . В этой окрестности точки x_0 функция $z = f'(x)$ возрастает, так как $z' = f''(x) > 0$. Но $f'(x_0) = 0$. Следовательно, при переходе через точку x_0 в направлении возрастания x $f'(x)$ меняет знак «минус» на «плюс», и потому согласно теореме 2 п. 2 настоящего параграфа $f(x)$ имеет в точке x_0 минимум. Доказательство в случае $f''(x_0) < 0$ аналогично.

Эта теорема позволяет сформулировать второе правило отыскания экстремума функции, в котором меняется лишь пункт 3. Этот пункт заменяется на следующий:

находим вторую производную $f''(x)$, вычисляем ее значения для каждого из найденных корней уравнения $f'(x) = 0$ и согласно теореме выносим заключение об экстремуме.

Заметим, что пользоваться вторым правилом обычно проще, чем первым. Но если вторая производная при значении корня первой производной обращается в нуль, то используют первое правило отыскания экстремума.

П р и м е р. Исследовать на экстремум функцию $f(x) = x^3 - 3x + 2$. Имеем $f'(x) = 3x^2 - 3$, $f''(x) = 6x$. Как уже отмечалось (см. пример 2 предыдущего пункта), корни уравнения $f'(x) = 0$: $x_1 = -1, x_2 = 1$. В точке $x_1 = -1$ функция $f(x)$ имеет максимум, так как $f''(-1) = -6 < 0$, а в точке $x^2 = 1$ — минимум, так как $f''(1) = 6 > 0$.

4. Наибольшее и наименьшее значения функции на отрезке. Пусть функция $f(x)$ непрерывна на отрезке $[a; b]$. Тогда (см. § 7.6, п. 3, теорема 1) на этом отрезке функция $f(x)$ достигает наибольшего и наименьшего значений. Остановимся для определенности на наибольшем значении. Если эта функция достигает наибольшего значения в интервале $(a; b)$, то это, очевидно, будет максимум функции $f(x)$. Но функция может достигать своего наибольшего значения также на одном из концов отрезка $[a; b]$ (см. рис. 92). Таким образом, чтобы найти наибольшее значение функции $f(x)$ на отрезке $[a; b]$, надо найти на интервале $(a; b)$ все максимумы этой функции, затем вычислить значения функции $f(x)$

на концах отрезка $[a; b]$, т. е. $f(a)$ и $f(b)$. Наибольшее из всех этих чисел и будет наибольшим значением функции $f(x)$ на отрезке $[a; b]$.

З а м е ч а н и е. Очевидно, что если дифференцируемая в интервале $(a; b)$ и непрерывная на отрезке $[a; b]$ функция имеет в интервале $(a; b)$ только одну стационарную точку и экстремум в ней, то в этой точке она имеет наибольшее значение в случае максимума и наименьшее в случае минимума.

	x	x	
x			x
		$a - 2x$	

Рис. 103

З а д а ч а. Из квадратного листа жести со стороной a , вырезая по углам равные квадраты и сгибая края (рис. 103), необходимо сделать прямоугольную коробку наибольшего объема.

Р е ш е н и е. Обозначим сторону вырезаемого квадрата через x . Тогда объем коробки выразится равенством $V = x(a - 2x)^2$, где x изменяется в интервале $\left(0; \frac{a}{2}\right)$. Производная $V' = (a - 2x)(a - 6x)$ между 0 и $\frac{a}{2}$ обращается в нуль в единственной точке $x = \frac{a}{6}$. При этом значении x объем V имеет максимум, значит, согласно замечанию и наибольшее значение, так как при $x = \frac{a}{6}$ $V'' = -4a < 0$. Таким образом, объем коробки будет наибольшим при стороне вырезаемого квадрата $x = \frac{a}{6}$.

5. Выпуклость и вогнутость графика функции. Точки перегиба. На рисунке 104 построен график дифференцируемой функции $y = f(x)$. В точках A , B и C построим касательные к графику. Видно, что все точки графика, достаточно близкие к точке A и лежащие по обе стороны от нее, расположены ниже касательной. В этом случае график функции $y = f(x)$ называется *выпуклым в точке A* . Все точки графика, достаточно близкие к точке C и лежащие по обе стороны от нее, расположены выше касательной. В таком случае график функции $y = f(x)$ называется *вогнутым в точке C* .

График дифференцируемой функции $y = f(x)$ называется *выпуклым (вогнутым) в интервале $(a; b)$* , если он является выпуклым (вогнутым) в каждой своей точке с первой координатой из $(a; b)$. На рисунке 104 график между точками A и B является выпуклым, а между точками B и C — вогнутым.

Касательная к графику в точке B (рис. 104) пересекает график. При этом во всех точках графика, близких к точке B и лежащих слева от нее, график является выпук-

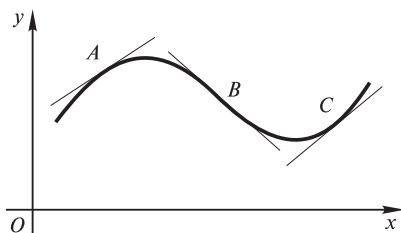


Рис. 104

лым, а во всех точках графика, лежащих справа от точки B и близких к ней, график является вогнутым. Точка графика дифференцируемой функции $y = f(x)$, при переходе через которую график меняет выпуклость на вогнутость и наоборот, называется *точкой перегиба*. В частности, на рисунке 104 точка B — точка перегиба.

Сформулируем пока без доказательства теоремы (их доказательства будут даны в § 8.9, п. 4), позволяющие находить интервалы выпуклости и вогнутости, а также точки перегиба.

Т е о р е м а 1. Если вторая производная $f''(x)$ функции $y = f(x)$ положительна (отрицательна) в интервале $(a; b)$, то график этой функции является вогнутым (выпуклым) в этом интервале.

Т е о р е м а 2. Если вторая производная $f''(x)$ функции $y = f(x)$ обращается в точке x_0 в нуль и при переходе через эту точку меняет знак, то точка $(x_0; f(x_0))$ графика данной функции является точкой перегиба.

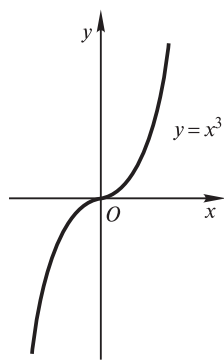


Рис. 105

П р и м е р. Кривая $y = x^3$ выпукла в промежутке $(-\infty; 0)$, так как в этом промежутке $y'' = 6x < 0$, и вогнута в промежутке $(0; +\infty)$, так как в нем $y'' = 6x > 0$; при $x = 0$ $y'' = 0$, следовательно, точка $(0; 0)$ — точка перегиба (рис. 105).

6. Асимптоты. В § 5.1 (п. 2) рассматривались асимптоты гиперболы. Многие другие линии также имеют асимптоты, т. е. прямые, к которым неограниченно приближается данная линия, когда ее точка неограниченно удаляется от начала координат.

Различают асимптоты *вертикальные* (т. е. параллельные оси ординат) и *наклонные* (т. е. непараллельные оси ординат).

Прямая $x = a$ называется вертикальной асимптотой графика функции $y = f(x)$, если хотя бы одно из предельных значений

$$\lim_{x \rightarrow a+0} f(x), \lim_{x \rightarrow a-0} f(x)$$

является бесконечным, т. е. равно $+\infty$ или $-\infty$.

П р и м е р 1. Прямая $x = 0$ является вертикальной асимптотой графика функции $y = \frac{1}{x}$, так как $\lim_{x \rightarrow 0+0} \frac{1}{x} = +\infty$, $\lim_{x \rightarrow 0-0} \frac{1}{x} = -\infty$.

Предположим, что функция $y = f(x)$ определена при сколь угодно больших (по модулю) значениях аргумента.

Прямая

$$y = kx + b$$

называется *наклонной* (при $k = 0$ *горизонтальной*) *асимптотой* графика функции $y = f(x)$ при $x \rightarrow +\infty$, если эта функция представима в виде

$$f(x) = kx + b + \alpha(x),$$

где

$$\lim_{x \rightarrow +\infty} \alpha(x) = 0.$$

Примечание. Аналогично определяется наклонная асимптота для случая $x \rightarrow -\infty$.

Пример 2. Прямая $y = 0$ является горизонтальной асимптотой графика функции $y = \frac{1}{x}$ при $x \rightarrow +\infty$ и при $x \rightarrow -\infty$.

§ 8.8. Построение графиков функций

Наиболее наглядное представление о ходе изменения функции дает ее график. Поэтому построение графика является заключительным этапом исследования функции, в котором используются все результаты ее исследования.

Пример 1. Построить график функции $y = \frac{x^2 + 1}{x - 1}$.

Эта функция определена и непрерывна для всех значений x , кроме $x = 1$. Функция не является ни четной, ни нечетной. Ее график не имеет точек пересечения с осью Ox , так как $x^2 + 1 > 0$ для всех вещественных x . При $x = 0$ $y = -1$. Далее,

$$\lim_{x \rightarrow 1-0} \frac{x^2 + 1}{x - 1} = -\infty, \quad \lim_{x \rightarrow 1+0} \frac{x^2 + 1}{x - 1} = +\infty,$$

т. е. прямая $x = 1$ является вертикальной асимптотой. При $x \rightarrow +\infty$ $y \rightarrow +\infty$, а при $x \rightarrow -\infty$

$y \rightarrow -\infty$. Производная данной функции $y' = \frac{x^2 - 2x - 1}{(x - 1)^2}$ обращается в нуль в точках

$x_1 = 1 - \sqrt{2}$, $x_2 = 1 + \sqrt{2}$. Эти точки разбивают всю числовую ось на три промежутка $(-\infty; 1 - \sqrt{2})$, $(1 - \sqrt{2}; 1 + \sqrt{2})$, $(1 + \sqrt{2}; +\infty)$, внутри каждого из которых производная y' сохраняет постоянный знак. Очевидно, что в первом и третьем промежутках $y' > 0$ и, следовательно, здесь функция y возрастает, во втором промежутке $y' < 0$ и, следовательно, в этом промежутке данная функция убывает. Ее вторая производная $y'' = \frac{4}{(x - 1)^3}$ всюду отлична от нуля (значит, точек перегиба график рассматриваемой

функции не имеет), в промежутке $(-\infty; 1)$ $y'' < 0$, и, следовательно, здесь график данной функции является выпуклым и в точке x_1 эта функция имеет максимум, в промежутке $(1; +\infty)$ $y'' > 0$, и, следовательно, в последнем промежутке этот график является вогнутым и в точке x_2 данная функция имеет минимум. Наконец, поскольку

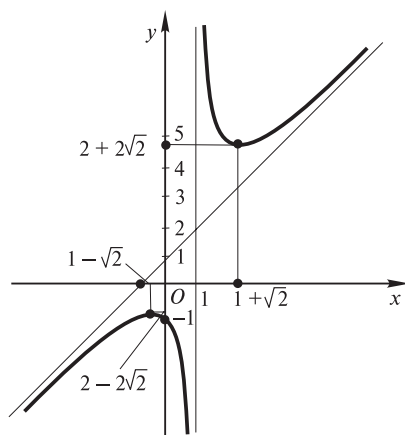


Рис. 106

$\frac{x^2 + 1}{x - 1} = x + 1 + \frac{2}{x - 1}$ и $\lim_{x \rightarrow \pm\infty} \frac{2}{x - 1} = 0$, то график данной функции имеет наклонную асимптоту $y = x + 1$ и при $x \rightarrow +\infty$, и при $x \rightarrow -\infty$. График функции $y = \frac{x^2 + 1}{x - 1}$ изображен на рисунке 106.

Пример 2. Построить график функции $y = e^{-x^2}$.

Эта функция определена, непрерывна, положительна на всей числовой оси и является четной. Поэтому достаточно построить ее график в первом квадранте. При $x = 0$ $y = 1$. При $x \rightarrow +\infty$ $y \rightarrow 0$, значит, прямая $y = 0$ является горизонтальной асимптотой и при $x \rightarrow +\infty$, и при $x \rightarrow -\infty$. Производная $y' = -2xe^{-x^2}$ обращается в нуль только в точке $x_0 = 0$; при $x > 0$ $y' < 0$, т. е. при $x > 0$ данная функция убывает. Ее вторая производная $y'' = 2(2x^2 - 1)e^{-x^2}$ в точке $x_1 = \frac{1}{\sqrt{2}}$ обращается в нуль, в промежутке $(\frac{1}{\sqrt{2}}; +\infty)$ $y'' > 0$ и, следовательно, здесь график данной функции является вогну-

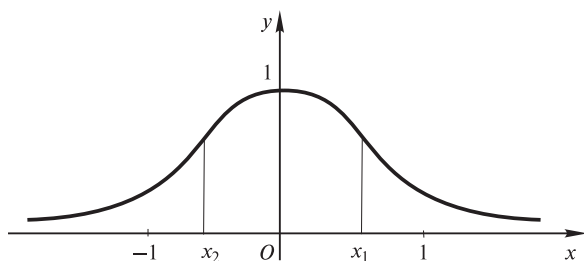


Рис. 107

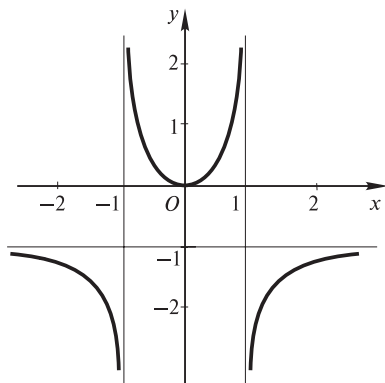


Рис. 108

тым, а в промежутке $[x_0; x_1)$ $y'' < 0$, и, следовательно, в нем этот же график является выпуклым, x_1 — абсцисса его точки перегиба и, наконец, в точке x_0 данная функция имеет максимум. График данной функции изображен на рисунке 107. Это так называемая *кривая Гаусса* (Карл Гаусс (1777—1855) — немецкий математик).

Пример 3. Построить график функции $y = \frac{x^2}{1 - x^2}$.

График этой функции изображен на рисунке 108.

§ 8.9. Формула Тейлора

1. Формула Тейлора с остаточным членом в форме Лагранжа.

Теорема. Пусть в интервале $(\alpha; \beta)$ функция $f(x)$ имеет производные до $(n + 1)$ -го порядка включительно. Тогда для всякого x из этого интервала и фиксированного a этого интервала имеет место формула

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}. \quad (1)$$

Здесь c находится между x и a , т. е. $c = a + \Theta(x-a)$, где Θ — число, заключенное между 0 и 1, т. е. $0 < \Theta < 1$.

Формула (1) называется *формулой Тейлора* (Брук Тейлор (1685—1731) — английский математик), а последнее слагаемое в правой части этой формулы — *остаточным членом формулы Тейлора в форме Лагранжа*.

Доказательство. Доказательство проведем в предположении, что взятое из интервала $(\alpha; \beta)$ значение $x > a$ (случай $x < a$ доказывается аналогично). Составим разность

$$r_n(x) = f(x) - (f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n) \quad (2)$$

и рассмотрим вспомогательные функции

$$F(t) = f(x) - (f(t) + \frac{f'(t)}{1!}(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \dots + \frac{f^{(n)}(t)}{n!}(x-t)^n), \quad (3)$$

$$\Phi(t) = (x-t)^{n+1}. \quad (4)$$

Очевидно, эти функции, рассматриваемые как функции t (при постоянном x), дифференцируемы на отрезке $[a; x]$ (дифференцируемость функции $F(t)$ вытекает из наличия для функции $f(t)$ производной $(n+1)$ -го порядка), следовательно, и непрерывны на нем, причем $\Phi'(t) \neq 0$ в $(a; x)$. Поэтому согласно теореме Коши (§ 8.6, п. 4) существует такое c из интервала $(a; x)$, что

$$\frac{F(x) - F(a)}{\Phi(x) - \Phi(a)} = \frac{F'(c)}{\Phi'(c)}. \quad (5)$$

Но из формул (3) и (4) имеем:

$$\begin{aligned} F(x) &= 0, \quad F(a) = r_n(x), \quad \Phi(x) = 0, \\ \Phi(a) &= (x-a)^{n+1}, \\ F'(t) &= 0 - (f'(t) - f'(t) + \frac{f''(t)}{1!}(x-t) - \frac{f''(t)}{1!}(x-t) + \dots \\ &\quad + \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} - \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} + \\ &\quad + \frac{f^{(n+1)}(t)}{n!}(x-t)^n) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n, \\ \Phi'(t) &= -(n+1)(x-t)^n, \end{aligned}$$

и, значит,

$$F'(c) = -\frac{f^{(n+1)}(c)}{n!}(x-c)^n, \quad \Phi'(c) = -(n+1)(x-c)^n.$$

Поэтому соотношение (5) принимает вид:

$$\frac{0 - r_n(x)}{0 - (x - a)^{n+1}} = \frac{f^{(n+1)}(c)(x - c)^n}{(n + 1)!(x - c)^n},$$

откуда

$$r_n(x) = \frac{f^{(n+1)}(c)}{(n + 1)!} (x - a)^{n+1}. \quad (6)$$

Наконец, равенства (2) и (6) приводят к искомой формуле (1).

З а м е ч а н и е. Формулу (1) можно записать в виде

$$\begin{aligned} f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \dots \\ \dots + \frac{f^{(n)}(a)}{n!}(x - a)^n + o((x - a)^n). \end{aligned} \quad (1')$$

Эта форма остаточного члена была указана Пеано (Джузеппе Пеано (1858—1932) — итальянский математик); формула (1') полезная, например, при вычислении пределов.

Так, при нахождении предела $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2}$ используем разложение $\cos x = 1 - \frac{x^2}{2!} + o(x^2)$.

Имеем:

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \lim_{x \rightarrow 0} \frac{\frac{x^2}{2!} - o(x^2)}{x^2} = \frac{1}{2},$$

поскольку $\lim_{x \rightarrow 0} \frac{o(x^2)}{x^2} = 0$.

2. Применение формулы Тейлора к элементарным функциям. Приближенные формулы. Частным случаем формулы Тейлора, соответствующим случаю $a = 0$, является формула

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(c)}{(n + 1)!}x^{n+1}, \quad (7)$$

где $c = Ox$. Эту формулу часто называют *формулой Маклорена* (Колин Маклорен (1698—1746) — шотландский математик). Здесь остаточный член

$$r_n(x) = \frac{f^{(n+1)}(c)}{(n + 1)!} x^{n+1}.$$

Разложим по этой формуле некоторые элементарные функции.

1. Пусть $f(x) = e^x$, тогда $f^{(k)}(x) = e^x$ при любом натуральном k и при любом x . В частности, при $x = 0$ имеем $f(0) = 1$ и $f^{(k)}(0) = 1$. По формуле (7) получим разложение функции e^x :

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n + 1)!} e^c. \quad (8)$$

Из равенства (8) получаем приближенную формулу

$$e^x \approx 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}.$$

Так как остаточный член здесь

$$r_n(x) = \frac{x^{n+1}}{(n+1)!} e^c,$$

то, например, при $x > 0$ погрешность $r_n(x)$ оценивается так:

$$0 < r_n(x) < \frac{x^{n+1}}{(n+1)!} e^x. \quad (9)$$

В частности, при $x = 1$ получаем приближенное значение числа e :

$$e \approx 1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!}.$$

При этом из (9) следует, что $0 < r_n(1) < \frac{3}{(n+1)!}$, так как $e < 3$. Если требуется вычислить значение e с точностью до 0,001, то число n определяется из неравенства $\frac{3}{(n+1)!} < 0,001$, или $(n+1)! > 3000$. Следовательно, если взять $n = 6$, то требуемое неравенство верно.

Таким образом, используя формулу Маклорена, можно вычислить число e с любой точностью, при этом алгоритм вычисления числа e легко реализуется на ЭВМ.

2. Пусть $f(x) = \sin x$. Тогда $f'(x) = \cos x$, $f''(x) = -\sin x$, $f'''(x) = -\cos x$, $f^{IV}(x) = \sin x$, $f^V(x) = \cos x$, Следовательно, $f(0) = 0$, $f'(0) = 1$, $f''(0) = 0$, $f'''(0) = -1$, $f^{IV}(0) = 0$, $f^V(0) = 1$, и далее такое же чередование значений.

По формуле (7) (если взять $n = 2m$) имеем:

$$\begin{aligned} \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^{m-1} \frac{x^{2m-1}}{(2m-1)!} + \\ + (-1)^m \frac{x^{2m+1}}{(2m+1)!} \cos c. \end{aligned} \quad (10)$$

Из разложения (10) получим приближенную формулу

$$\sin x \approx x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^{m-1} \frac{x^{2m-1}}{(2m-1)!}$$

с погрешностью $r_n(x)$, оцениваемой неравенством

$$|r_n(x)| \leq \frac{|x|^{2m+1}}{(2m+1)!}$$

(так как $|\cos c| \leq 1$ при любом c).

В частности, имеем приближенную формулу

$$\sin x \approx x$$

с погрешностью, по модулю не превосходящей $\frac{|x|^3}{3!}$.

3. Аналогично при $f(x) = \cos x$ имеем:

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^m \frac{x^{2m}}{(2m)!} + (-1)^{m+1} \frac{x^{2m+2}}{(2m+2)!} \cos c, \quad (11)$$

откуда получаем приближенную формулу

$$\cos x \approx 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^m \frac{x^{2m}}{(2m)!}$$

с погрешностью

$$|r_n(x)| \leq \frac{x^{2m+2}}{(2m+2)!}.$$

4. Пусть $f(x) = (1+x)^n$, где n — натуральное число. Последовательно дифференцируя $f(x)$, получим:

$f'(x) = n(1+x)^{(n-1)}$, $f''(x) = n(n-1)(1+x)^{(n-2)}$, ..., $f^{(k)}(x) = n(n-1) \dots (n-k+1)(1+x)^{(n-k)}$, ..., $f^{(n)}(x) = n!$ (производные порядка выше n все равны нулю).

Отсюда $f(0) = 1$, $f'(0) = n$, $f''(0) = n(n-1)$, ..., $f^{(k)}(0) = n(n-1) \dots (n-k+1)$, ..., $f^{(n)}(0) = n!$

По формуле (7) имеем равенство

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \dots + \frac{n(n-1) \dots (n-k+1)}{k!}x^k + \dots + x^n, \quad (12)$$

называемое *биномом Ньютона*.

Очевидно, равенства (8), (10), (11), (12) верны при всех значениях x .

3. Исследование функции на экстремум с помощью формулы Тейлора. Оказывается, если в стационарной точке функции $f(x)$ вторая производная $f''(x)$ равна нулю, то вопрос о существовании максимума или минимума функции $f(x)$ в этой точке может быть решен при помощи производных более высокого порядка, если, конечно, они существуют в этой точке.

Пусть в некоторой окрестности точки x_0 функция $f(x)$ имеет производные до $(n+1)$ -го порядка включительно, причем производная $f^{(n+1)}(x)$ непрерывна в точке x_0 . Пусть далее $f'(x_0) = f''(x_0) = \dots = f^{(n)}(x_0) = 0$, $f^{(n+1)}(x_0) \neq 0$.

Так как $f^{(n+1)}(x)$ непрерывна в точке x_0 , то согласно известной теореме (§ 7.6, п. 1, теорема 5) найдется такая окрестность точки x_0 (эту окрестность обозначим через $(\alpha; \beta)$), в

которой $f^{(n+1)}(x)$ сохраняет свой знак. Пусть x — произвольная точка из $(\alpha; \beta)$. Формула Тейлора (1) примет вид:

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

или

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

Последняя формула дает возможность решить вопрос об экстремуме функции $f(x)$ в точке x_0 . Действительно, если $(n+1)$ — четное число и $f^{(n+1)}(x_0) > 0$, то $f(x) - f(x_0) > 0$ для всех x рассматриваемого промежутка, кроме x_0 , а потому x_0 — точка минимума; если $(n+1)$ — четное число и $f^{(n+1)}(x_0) < 0$, то $f(x) - f(x_0) < 0$ для всех x рассматриваемого промежутка, кроме x_0 , а потому x_0 — точка максимума. Наконец, если $(n+1)$ — нечетное число, то $f(x) - f(x_0)$ меняет знак (при переходе x через x_0) вместе с величиной $(x - x_0)^{(n+1)}$, и, следовательно, в точке x_0 нет экстремума.

П р и м е р. Если $y = x^4$, то $y' = 4x^3$, $y'' = 12x^2$, $y''' = 24x$, $y^{IV} = 24$. В точке $x_0 = 0$ данная функция имеет минимум, так как $y^{IV} > 0$.

4. Достаточные условия выпуклости, вогнутости и точек перегиба.

Применим формулу Тейлора для доказательства сформулированных ранее (§ 8.7, п. 5) теорем 1 и 2.

Т е о р е м а 1. Если вторая производная $f''(x)$ функции $y = f(x)$ положительна (отрицательна) в интервале $(a; b)$, то график этой функции является вогнутым (выпуклым) в этом интервале.

Д о к а з а т е л ь с т в о. Фиксируем точку x_0 из интервала $(a; b)$ и докажем, что в точке $(x_0; f(x_0))$ график данной функции является вогнутым (выпуклым). По формуле Тейлора (1) для $n = 1$

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + f''(c) \frac{(x - x_0)^2}{2!}, \quad (13)$$

где c лежит между x_0 и x .

Далее напишем уравнение касательной к кривой $y = f(x)$ в точке $(x_0; f(x_0))$, обозначив текущую ординату касательной через Y :

$$Y - f(x_0) = f'(x_0)(x - x_0). \quad (14)$$

Вычитая из равенства (13) равенство (14), получаем:

$$f(x) - Y = f''(c) \frac{(x - x_0)^2}{2!}.$$

Так как $f''(c) > 0$ (< 0), то и $f(x) - Y > 0$ (< 0).

С учетом теоремы 1 непосредственно получаем теорему 2.

Глава 9. ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ

§ 9.1. Первообразная функция и неопределенный интеграл

1. Понятие первообразной функции и неопределенного интеграла.

В главе 8 было введено новое действие — дифференцирование: нахождение по заданной функции ее производной. Оказывается, что для дифференцирования существует обратное действие — интегрирование: отыскание функции по заданной ее производной. К этому приводят многочисленные задачи из физики и других областей науки и техники.

Ранее (см. § 8.1) было установлено, что если известен закон $s = s(t)$ прямолинейного движения материальной точки, выражающий зависимость пути s от времени движения t , то скорость точки выражается производной пути по времени: $v = s'(t)$. Обратная задача: известна скорость прямолинейного движения точки $v = v(t)$ как функция времени. Найти закон движения. Ясно, что искомой функцией $s = s(t)$ будет такая, для которой $s'(t) = v(t)$.

О п р е д е л е н и е 1. Функция $F(x)$ называется *первообразной функцией* для данной функции $f(x)$ (или, короче, *первообразной* данной функции $f(x)$) на данном промежутке, если на этом промежутке $F'(x) = f(x)$.

П р и м е р. Функция $F(x) = x^3$ является первообразной функции $f(x) = 3x^2$ на всей числовой оси, так как при любом x $(x^3)' = 3x^2$. Отметим при этом, что вместе с функцией $F(x) = x^3$ первообразной для $f(x) = 3x^2$ является любая функция $\Phi(x) = x^3 + C$, где C — произвольное постоянное число (это следует из того, что производная постоянной равна нулю). Это свойство имеет место и в общем случае.

Т е о р е м а 1. Если $F_1(x)$ и $F_2(x)$ — две первообразные для функции $f(x)$ в некотором промежутке, то разность между ними в этом промежутке равна постоянному числу.

Д о к а з а т е л ь с т в о. Пусть, например, указанный промежуток — интервал $(a; b)$. Из определения первообразной имеем $F_1'(x) = f(x)$ и $F_2'(x) = f(x)$ для любого x из $(a; b)$.

Пусть $\alpha(x) = F_2(x) - F_1(x)$. Тогда для любого x из $(a; b)$

$$\alpha'(x) = F_2'(x) - F_1'(x) = f(x) - f(x) = 0,$$

следовательно (см. следствие из § 8.6, п. 3), $\alpha(x) = C$.

Из теоремы 1 следует, что если известна какая-нибудь первообразная $F(x)$ данной функции $f(x)$, то все множество первообразных для $f(x)$ исчерпывается функциями $F(x) + C$.

Подчеркнем важный факт: если производная для функции одна, т. е. операция дифференцирования однозначна, то нахождение первообразной для функции возможно лишь с точностью до некоторого постоянного слагаемого.

О п р е д е л е н и е 2. Выражение $F(x) + C$, где $F(x)$ — первообразная функции $f(x)$ и C — произвольная постоянная, называется *неопределенным интегралом* от функции $f(x)$ и обозначается символом $\int f(x)dx$, причем $f(x)$ называется *подынтегральной функцией*, $f(x)dx$ — *подынтегральным выражением*, x — *переменной интегрирования*; $\int$ — *знак неопределенного интеграла*. Таким образом, по определению

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

если $F'(x) = f(x)$.

Возникает вопрос: для всякой ли функции $f(x)$ существует первообразная, а значит, и неопределенный интеграл? В связи с этим вопросом приведем без доказательства следующую теорему (см. [9], т. I):

Т е о р е м а 2. Если функция $f(x)$ непрерывна на сегменте $[a; b]$, то на этом сегменте у функции $f(x)$ существует первообразная.

Ниже мы будем говорить о первообразных лишь для непрерывных функций. Поэтому рассматриваемые нами далее в этом параграфе интегралы существуют.

2. Свойства неопределенного интеграла. Из определения неопределенного интеграла непосредственно вытекают следующие два свойства:

$$1. \frac{d}{dx} \int f(x)dx = f(x),$$

и, значит, $d \int f(x)dx = f(x)dx$.

$$2. \int F'(x)dx = F(x) + C,$$

что может быть записано так: $\int dF(x) = F(x) + C$.

3. Постоянный множитель можно вынести за знак интеграла:

$$\int kf(x)dx = k \int f(x)dx. \quad (1)$$

Действительно, имеем:

$$\left[k \int f(x)dx \right]' = k \left[\int f(x)dx \right]' = kf(x).$$

Аналогично доказывается следующее свойство:

4. Интеграл от суммы двух функций равен сумме интегралов от этих функций:

$$\int [f_1(x) + f_2(x)]dx = \int f_1(x)dx + \int f_2(x)dx. \quad (2)$$

П р и м е ч а н и е 1. Равенства (1) и (2) следует понимать с точностью до постоянного слагаемого.

П р и м е ч а н и е 2. Свойство 4 распространяется на случай алгебраической суммы любого конечного числа функций.

3. Таблица основных интегралов. Таблица содержит формулы, легко проверяемые непосредственным дифференцированием:

$$\text{I. } \int x^{\mu} dx = \frac{x^{\mu+1}}{\mu+1} + C \quad (\mu \neq -1). \quad \text{V. } \int \sin ax dx = -\frac{1}{a} \cos ax + C.$$

$$\text{II. } \int \frac{dx}{x} = \ln|x| + C. \quad \text{VI. } \int \cos ax dx = \frac{1}{a} \sin ax + C.$$

$$\text{III. } \int a^x dx = \frac{a^x}{\ln a} + C. \quad \text{VII. } \int \frac{dx}{\cos^2 ax} = \frac{1}{a} \operatorname{tg} ax + C.$$

$$\text{IV. } \int e^{kx} dx = \frac{e^{kx}}{k} + C. \quad \text{VIII. } \int \frac{dx}{\sin^2 ax} = -\frac{1}{a} \operatorname{ctg} ax + C.$$

$$\text{IX. } \int \frac{x dx}{x^2 + a} = \frac{1}{2} \ln|x^2 + a| + C.$$

$$\text{X. } \int \frac{dx}{x^2 + a^2} = \frac{1}{a} \operatorname{arctg} \frac{x}{a} + C.$$

$$\text{XI. } \int \frac{dx}{\sqrt{a^2 - x^2}} = \arcsin \frac{x}{a} + C.$$

$$\text{XII. } \int \frac{dx}{\sqrt{x^2 + a}} = \ln|x + \sqrt{x^2 + a}| + C.$$

$$\text{XIII. } \int \frac{dx}{x^2 - a^2} = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| + C.$$

Проверим, например, формулу II. Если $x > 0$, то $|x| = x$ и $(\ln|x|)' = (\ln x)' = \frac{1}{x}$. Если $x < 0$, то $|x| = -x$ и $(\ln|x|)' = (\ln(-x))' = \frac{1}{x}$. Следовательно, формула II справедлива как при $x > 0$, так и при $x < 0$.

4. Примеры непосредственного интегрирования. Метод непосредственного интегрирования связан с приведением подынтегрального выражения к табличной форме путем преобразований и применения свойств неопределенного интеграла.

$$\text{П р и м е р 1. } \int (5x^4 - 3x^2 + 1) dx = 5 \int x^4 dx - 3 \int x^2 dx + \int dx = x^5 - x^3 + x + C.$$

$$\text{П р и м е р 2. } \int \frac{x^6 - x^5 + 1}{x^2} dx = \int (x^4 - x^3 + x^{-2}) dx = \frac{x^5}{5} - \frac{x^4}{4} - \frac{1}{x} + C.$$

$$\text{П р и м е р 3. } \int (\sqrt{x} - \sqrt[3]{x})^2 dx = \int (x - 2x^{\frac{5}{6}} + x^{\frac{2}{3}}) dx = \frac{x^2}{2} - \frac{12}{11} x^{\frac{11}{6}} + \frac{3}{5} x^{\frac{5}{3}} + C = \frac{x^2}{2} - \frac{12}{11} x^{\frac{11}{6}} + \frac{3}{5} x^{\frac{5}{3}} + C.$$

§ 9.2. Основные методы интегрирования

1. Замена переменной интегрирования. Этот способ часто бывает полезным в тех случаях, когда интеграл $\int f(x)dx$ ($f(x)$ непрерывна) не может быть непосредственно преобразован к виду табличного. Сделаем подстановку $x = \varphi(t)$, где $\varphi(t)$ — функция, имеющая непрерывную производную. Тогда

$$f(x) = f(\varphi(t)), \quad dx = \varphi'(t)dt$$

и

$$\int f(x)dx = \int f(\varphi(t))\varphi'(t)dt. \quad (1)$$

Формула (1) называется *формулой замены переменной в неопределенном интеграле*.

Пример 1. $\int \frac{e^{\sqrt{x}}}{\sqrt{x}} dx$ найдем подстановкой $x = t^2$. Тогда $dx = 2tdt$ и $\int \frac{e^{\sqrt{x}}}{\sqrt{x}} dx = 2 \int e^t dt = 2e^t + C = 2e^{\sqrt{x}} + C$.

Иногда вместо подстановки $x = \varphi(t)$ лучше выполнить замену переменной вида $t = \psi(x)$.

Пример 2. $\int \frac{dx}{e^x + e^{-x}}$. Полагая $t = e^x$, получаем: $dt = e^x dx = tdx$, $dx = \frac{dt}{t}$ и $\int \frac{dx}{e^x + e^{-x}} = \int \frac{dt}{t(t + t^{-1})} = \int \frac{dt}{t^2 + 1} = \arctg t + C = \arctg e^x + C$.

Во многих случаях нет необходимости записывать, какое выражение мы принимаем за новую переменную.

Пример 3. $\int \sqrt{x+3} dx = \int (x+3)^{\frac{1}{2}} d(x+3) = \frac{2}{3}(x+3)^{\frac{3}{2}} + C = \frac{2}{3}(x+3)\sqrt{x+3} + C$.

Пример 4. $\int xe^{x^2} dx = \frac{1}{2} \int e^{x^2} d(x^2) = \frac{1}{2} e^{x^2} + C$.

Пример 5. $\int \frac{\ln^3 x}{x} dx = \int \ln^3 x d(\ln x) = \frac{\ln^4 x}{4} + C$.

2. Интегрирование по частям. Пусть $u = u(x)$ и $v = v(x)$ — непрерывно дифференцируемые функции. Как известно (см. § 8.3, п. 4), $d(uv) = vdu + u dv$, откуда $u dv = d(uv) - vdu$. Интегрируя последнее соотношение, получим:

$$\int u dv = \int d(uv) - \int v du,$$

или

$$\int u dv = uv - \int v du \quad (2)$$

(произвольная постоянная интегрирования C здесь включена в слагаемое $\int v du$). Это и есть *формула интегрирования по частям*. Применение способа интегрирования по частям целесообразно в том случае, когда интеграл в правой части (2) окажется более простым для вычисления, чем исходный интеграл.

Пример 1. $\int \underbrace{\ln x}_u \underbrace{dx}_{dv} = x \ln x - \int dx = x \ln x - x + C.$

Пример 2. $\int x \underbrace{e^x dx}_{dv} = xe^x - \int e^x dx = xe^x - e^x + C.$

Пример 3. $\int \underbrace{\arctg x}_u \underbrace{dx}_{dv} = x \arctg x - \int \frac{xdx}{1+x^2} = x \arctg x - \frac{1}{2} \ln(1+x^2) + C.$

Иногда полезно повторное интегрирование по частям.

Пример 4. $\int \underbrace{x^2 \cos x dx}_{dv} = x^2 \sin x - 2 \int \underbrace{x \sin x dx}_{dv} = x^2 \sin x + 2x \cos x - 2 \int \cos x dx = x^2 \sin x + 2x \cos x - 2 \sin x + C.$

§ 9.3. Интегрирование дробно-рациональных функций

1. Выделение правильной рациональной дроби. Как известно (см. § 7.2, п. 2), дробно-рациональной функцией (или рациональной дробью) называется отношение двух многочленов.

Рациональная дробь называется *правильной* (*неправильной*), если степень многочлена, стоящего в числителе, меньше (больше или равна) степени многочлена, стоящего в знаменателе.

Например, дроби

$$\frac{3x+2}{x^2-4x+12}, \frac{x+2}{x^2-1}, \frac{x^5}{x^6-1} \text{ правильные,}$$

а дроби

$$\frac{x^5}{x^2+1}, \frac{x^3-1}{x+1}, \frac{x+1}{3x-1} \text{ неправильные.}$$

Неправильную рациональную дробь всегда можно свести к правильной, разделив числитель на знаменатель «столбиком» и выделив из дроби целую часть, т. е. многочлен.

Пример. $\frac{x^4-x^3+1}{x^2+x+2} = (x^2-2x) + \frac{4x+1}{x^2+x+2}.$

Поэтому будем рассматривать задачу интегрирования правильной рациональной дроби, так как интегрирование многочлена не представляет труда.

Как будет видно ниже (см. п. 3), всякую правильную рациональную дробь можно представить в виде суммы конечного числа так называемых *простейших* дробей следующих четырех типов:

$$\begin{array}{ll} \text{I.} & \frac{A}{x-a}. \\ \text{II.} & \frac{A}{(x-a)^k} \quad (k=2, 3, \dots). \\ \text{III.} & \frac{Mx+N}{x^2+px+q}. \\ \text{IV.} & \frac{Mx+N}{(x^2+px+q)^l} \quad (l=2, 3, \dots), \end{array}$$

где A, a, p, q, M, N — постоянные действительные числа, а трехчлен $x^2 + px + q$ не имеет действительных корней, т. е. $\frac{p^2}{4} - q < 0$.

2. Интегрирование простейших рациональных дробей. Интегрирование простейших дробей I и II типов не представляет труда. В самом деле,

$$\begin{aligned} \int \frac{A dx}{x-a} &= A \int \frac{d(x-a)}{x-a} = A \ln|x-a| + C, \\ \int \frac{A dx}{(x-a)^k} &= A \int (x-a)^{-k} d(x-a) = \frac{A}{(1-k)(x-a)^{k-1}} + C. \end{aligned}$$

Перейдем к интегрированию простейшей дроби III типа. Подстановка $t = \frac{1}{2}(x^2 + px + q)$, т. е. $t = x + \frac{p}{2}$, и обозначение $q - \frac{p^2}{4} = \alpha^2$ дают $x^2 + px + q = \left(x + \frac{p}{2}\right)^2 + q - \frac{p^2}{4} = t^2 + \alpha^2$, $Mx + N = Mt + D$, где $D = N - \frac{1}{2}Mp$. Поэтому

$$\begin{aligned} \int \frac{Mx+N}{x^2+px+q} dx &= \int \frac{Mt+D}{t^2+\alpha^2} dt = \frac{M}{2} \int \frac{2t dt}{t^2+\alpha^2} + \\ &+ D \int \frac{dt}{t^2+\alpha^2} = \frac{M}{2} \ln(t^2+\alpha^2) + \frac{D}{\alpha} \operatorname{arctg} \frac{t}{\alpha} + C, \end{aligned}$$

где $t = x + \frac{p}{2}$.

Рассмотрим, наконец, интегрирование рациональных дробей IV типа. С помощью той же подстановки $t = x + \frac{p}{2}$ при $l > 1$ и $\frac{p^2}{4} - q < 0$ получаем:

$$\begin{aligned} \int \frac{Mx+N}{(x^2+px+q)^l} dx &= \int \frac{Mt+D}{(t^2+\alpha^2)^l} dt = M \int \frac{t dt}{(t^2+\alpha^2)^l} + D \int \frac{dt}{(t^2+\alpha^2)^l} = \\ &= \frac{M}{2} \int (t^2+\alpha^2)^{-l} d(t^2+\alpha^2) + D \int \frac{dt}{(t^2+\alpha^2)^l}. \end{aligned}$$

Интеграл $\int (t^2 + \alpha^2)^{-l} d(t^2 + \alpha^2)$ табличный.

Для интеграла $I_l = \int \frac{dt}{(t^2 + \alpha^2)^l}$ получим *рекуррентную* формулу, т. е.

формулу, сводящую вычисление I_l к вычислению I_{l-1} . С этой целью рассмотрим интеграл

$$I_{l-1} = \int \frac{dt}{(t^2 + \alpha^2)^{l-1}}$$

и применим к нему формулу интегрирования по частям, положив

$$u = \frac{1}{(t^2 + \alpha^2)^{l-1}}, \quad dv = dt,$$

откуда

$$du = -\frac{(l-1)2tdt}{(t^2 + \alpha^2)^l}, \quad v = t.$$

Тогда

$$\begin{aligned} I_{l-1} &= \frac{t}{(t^2 + \alpha^2)^{l-1}} + 2(l-1) \int \frac{t^2 dt}{(t^2 + \alpha^2)^l} = \\ &= \frac{t}{(t^2 + \alpha^2)^{l-1}} + 2(l-1) \int \frac{(t^2 + \alpha^2) - \alpha^2}{(t^2 + \alpha^2)^l} dt = \\ &= \frac{t}{(t^2 + \alpha^2)^{l-1}} + 2(l-1)I_{l-1} - 2(l-1)\alpha^2 I_l. \end{aligned}$$

Отсюда

$$2(l-1)\alpha^2 I_l = \frac{t}{(t^2 + \alpha^2)^{l-1}} + (2l-3)I_{l-1},$$

и, следовательно,

$$I_l = \frac{t}{\alpha^2(2l-2)(t^2 + \alpha^2)^{l-1}} + \frac{1}{\alpha^2} \frac{2l-3}{2l-2} I_{l-1}. \quad (1)$$

Формула (1) и есть рекуррентная формула. Переходя от l к $l-1$, от $l-1$ к $l-2$ и т. д., дойдем до табличного интеграла

$$I_1 = \int \frac{dt}{t^2 + \alpha^2}.$$

В результате будет известен и I_l .

3. Разложение правильной рациональной дроби на простейшие. Пусть дана правильная рациональная дробь

$$\frac{P(x)}{Q(x)}. \quad (2)$$

Без ограничения общности можно считать, что коэффициент у старшего члена многочлена $Q(x)$ равен единице, так как в случае, когда он равен какому-то другому числу (отличному от нуля), можно разделить числитель и знаменатель дроби $\frac{P(x)}{Q(x)}$ на это число, после чего у по-

лучившегося в знаменателе многочлена коэффициент у старшего члена окажется равным единице.

Будем предполагать, что коэффициенты входящих в нее многочленов — действительные числа. Как установлено в § 7.7 (п. 6), для многочлена с действительными коэффициентами имеет место соответствующее разложение на множители (см. § 7.7, (8)). Пусть для определенности знаменатель $Q(x)$ разлагается на множители следующим образом:

$$Q(x) = (x - a)^k (x^2 + px + q)^l, \quad (3)$$

где $\frac{p^2}{4} - q < 0$. Имеет место следующая теорема (см., например, [9], т. 1):

Т е о р е м а. Для дроби (2), знаменатель которой имеет вид (3), справедливо следующее разложение на сумму простейших дробей:

$$\begin{aligned} \frac{P(x)}{Q(x)} = & \frac{A_1}{x - a} + \frac{A_2}{(x - a)^2} + \dots + \frac{A_k}{(x - a)^k} + \frac{M_1 x + N_1}{x^2 + px + q} + \frac{M_2 x + N_2}{(x^2 + px + q)^2} + \dots \\ & \dots + \frac{M_l x + N_l}{(x^2 + px + q)^l}, \end{aligned} \quad (4)$$

где $A_1, A_2, \dots, A_k, M_1, N_1, M_2, N_2, \dots, M_l, N_l$ — постоянные действительные числа.

Из формулы (4) видно, что линейному множителю $x - a$ знаменателя $Q(x)$ соответствуют в разложении (4) простейшие дроби I и II типов, а квадратичному множителю $x^2 + px + q$ — простейшие дроби III и IV типов. При этом число простейших дробей, соответствующих данному множителю (линейному или квадратичному), равно показателю степени, с которым этот множитель входит в разложение знаменателя дроби на множители. Правило разложения правильной рациональной дроби остается справедливым ([9], т. I) и при

любом конечном числе линейных и квадратичных множителей, входящих в разложение знаменателя $Q(x)$.

4. Метод неопределенных коэффициентов. Одним из наиболее простых методов определения коэффициентов в разложении правильной дроби на простейшие является метод неопределенных коэффициентов. Поясним применение этого метода на примере.

П р и м е р. Разложить на простейшие дроби правильную дробь

$$\frac{2x+3}{(x+1)(x^2+x+1)}.$$

Убедившись, что квадратный трехчлен x^2+x+1 имеет комплексные корни, ищем согласно теореме из п. 3 разложение в виде

$$\frac{2x+3}{(x+1)(x^2+x+1)} = \frac{A}{x+1} + \frac{Mx+N}{x^2+x+1}, \quad (5)$$

где коэффициенты A , M , N необходимо определить. Для этого приводим к общему знаменателю дроби, стоящие в правой части соотношения (5). Приравниваем числители исходной дроби и той, которая будет получена:

$$2x+3 = A(x^2+x+1) + (Mx+N)(x+1). \quad (6)$$

Коэффициенты при одинаковых степенях в правой и левой частях тождества (6) должны быть равны. Записывая это условие последовательно для коэффициентов при x^2 , x и x^0 , получим:

$$\begin{cases} A + M = 0, \\ A + M + N = 2, \\ A + N = 3. \end{cases}$$

Решая эту систему, найдем $A=1$, $M=-1$, $N=2$. Наконец, подставляя в равенство (5) найденные значения A , M , N , окончательно получим:

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

5. Интегрирование правильных рациональных дробей. Для вычисления интеграла от правильной рациональной дроби надо эту дробь разложить на простейшие согласно теореме из п. 3 и найти интеграл от полученной суммы простейших дробей. Следовательно, интеграл от любой правильной рациональной дроби есть функция элементарная, потому что представляет сумму элементарных функций, которые получаются в результате интегрирования простейших дробей.

П р и м е р. Найти

$$\int \frac{5x^2 - 14x + 11}{x^3 - 3x^2 + 3x - 1} dx.$$

Так как $x^3 - 3x^2 + 3x - 1 = (x - 1)^3$, то, получив разложение дроби $\frac{5x^2 - 14x + 11}{(x - 1)^3}$ на простейшие и проинтегрировав их, получим:

$$\begin{aligned}\int \frac{5x^2 - 14x + 11}{x^3 - 3x^2 + 3x - 1} dx &= \int \left(\frac{5}{x - 1} - \frac{4}{(x - 1)^2} + \frac{2}{(x - 1)^3} \right) dx = \\ &= 5 \ln |x - 1| + \frac{4}{x - 1} - \frac{1}{(x - 1)^2} + C.\end{aligned}$$

§ 9.4. Интегрирование тригонометрических выражений

1. Интегралы вида $\int \sin ax \sin bxdx$, $\int \cos ax \cos bxdx$, $\int \sin ax \cos bxdx$. Эти интегралы с помощью известных тригонометрических формул приводятся к интегралам

$$\begin{aligned}\int \cos \gamma x dx &= \frac{1}{\gamma} \sin \gamma x + C, \\ \int \sin \gamma x dx &= -\frac{1}{\gamma} \cos \gamma x + C.\end{aligned}$$

Пример. Найти $\int \cos 8x \cos 6x dx$.

Так как $\cos 8x \cos 6x = \frac{1}{2}(\cos 2x + \cos 14x)$, то

$$\int \cos 8x \cos 6x dx = \frac{1}{2} \int (\cos 2x + \cos 14x) dx = \frac{\sin 2x}{4} + \frac{\sin 14x}{28} + C.$$

2. Интегралы вида $I_{n, m} = \int \sin^n x \cos^m x dx$, где n и m — натуральные числа. Если n и m четные, то интегралы $I_{n, m}$ находятся с помощью тригонометрических формул

$$\sin^2 x = \frac{1}{2}(1 - \cos 2x), \quad \cos^2 x = \frac{1}{2}(1 + \cos 2x), \quad \sin x \cos x = \frac{1}{2} \sin 2x.$$

Если хотя бы одно из чисел n и m нечетное, то от нечетной степени отделяется множитель первой степени и вводится новая переменная.

Пример. Найти $\int \sin^8 x \cos^5 x dx$. Имеем:

$$\int \sin^8 x \cos^5 x dx = \int \sin^8 x (1 - \sin^2 x)^2 d \sin x = \frac{\sin^9 x}{9} - \frac{2 \sin^{11} x}{11} + \frac{\sin^{13} x}{13} + C.$$

3. Интегралы вида $\int R(\sin x, \cos x) dx$, где $R(u, v)$ — рациональная функция двух аргументов u и v . Заметим, что рациональная функция двух аргументов — это отношение двух многочленов двух переменных, а многочленом двух переменных u и v называется сумма произведений вида $a_{mn} u^m v^n$ (m, n — целые неотрицательные числа, a_{mn} — постоянные числа). Пока-

жем, что интеграл $\int R(\sin x, \cos x) dx$ может быть сведен к интегралу от рациональной функции аргумента t подстановкой $t = \operatorname{tg} \frac{x}{2}$. Действительно,

$$\sin x = \frac{2 \sin \frac{x}{2} \cos \frac{x}{2}}{\sin^2 \frac{x}{2} + \cos^2 \frac{x}{2}} = \frac{2 \operatorname{tg} \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}} = \frac{2t}{1+t^2},$$

$$\cos x = \frac{\cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}}{\sin^2 \frac{x}{2} + \cos^2 \frac{x}{2}} = \frac{1 - \operatorname{tg}^2 \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}} = \frac{1-t^2}{1+t^2}.$$

Из подстановки $t = \operatorname{tg} \frac{x}{2}$ следует, что $x = 2 \arctg t$, $dx = \frac{2dt}{1+t^2}$. Таким образом,

$$\int R(\sin x, \cos x) dx = \int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right) \frac{2dt}{1+t^2} = \int R_1(t) dt,$$

где $R_1(t)$ — рациональная функция t .

П р и м е р. $\int \frac{dx}{\sin x} = \int \frac{1+t^2}{2t} \cdot \frac{2dt}{1+t^2} = \int \frac{dt}{t} = \ln|t| + C = \ln \left| \operatorname{tg} \frac{x}{2} \right| + C.$

§ 9.5. Интегрирование простейших иррациональностей

1. Интегралы с линейной иррациональностью. Если подынтегральное выражение содержит лишь линейную иррациональность $\sqrt[n]{ax+b}$ ($a \neq 0$), то полезна подстановка

$$t = \sqrt[n]{ax+b}.$$

П р и м е р. Найти $\int \frac{dx}{\sqrt[3]{x+1}}.$

Полагаем, что $t = \sqrt[3]{x+1}$, откуда

$$x = t^3 - 1, \quad dx = 3t^2 dt.$$

Имеем:

$$\int \frac{dx}{\sqrt[3]{x+1}} = \int \frac{3t^2 dt}{t} = 3 \int t dt = \frac{3}{2} t^2 + C = \frac{3}{2} (x+1)^{\frac{2}{3}} + C.$$

2. Интегралы с квадратичной иррациональностью.

1. Интеграл

$$\int \frac{dx}{\sqrt{Ax^2 + Bx + C}} \quad (1)$$

с помощью дополнения квадратного трехчлена $Ax^2 + Bx + C$ до полного квадрата приводится в зависимости от знака A к одному из двух табличных интегралов XI, XII (см. таблицу основных интегралов).

2. Интеграл

$$\int \frac{Mx + N}{\sqrt{Ax^2 + Bx + C}} dx$$

приводится к интегралу (1) следующим путем:

$$\begin{aligned} \int \frac{(Mx + N)dx}{\sqrt{Ax^2 + Bx + C}} &= \int \frac{\left(Mx + \frac{MB}{2A}\right) - \frac{MB}{2A} + N}{\sqrt{Ax^2 + Bx + C}} dx = \frac{M}{2A} \int \frac{(2Ax + B)dx}{\sqrt{Ax^2 + Bx + C}} + \\ &+ \left(N - \frac{MB}{2A}\right) \int \frac{dx}{\sqrt{Ax^2 + Bx + C}} = \frac{M}{A} \sqrt{Ax^2 + Bx + C} + \\ &+ \left(N - \frac{MB}{2A}\right) \int \frac{dx}{\sqrt{Ax^2 + Bx + C}}. \end{aligned}$$

3. Найдем интегралы

$$\int \sqrt{x^2 + b} dx, \quad (2)$$

$$\int \sqrt{a^2 - x^2} dx. \quad (3)$$

Имеем:

$$\int \sqrt{x^2 + b} dx = \int \frac{x^2 + b}{\sqrt{x^2 + b}} dx = \int \frac{x^2 dx}{\sqrt{x^2 + b}} + b \int \frac{dx}{\sqrt{x^2 + b}}. \quad (4)$$

Для вычисления интеграла $\int \frac{x^2 dx}{\sqrt{x^2 + b}}$ применим метод интегрирования

по частям, полагая $u = x$, $dv = \frac{xdx}{\sqrt{x^2 + b}}$, тогда $du = dx$, $v = \sqrt{x^2 + b}$. Следова-

тельно,

$$\int \frac{x^2 dx}{\sqrt{x^2 + b}} = x\sqrt{x^2 + b} - \int \sqrt{x^2 + b} dx.$$

Подставляя это выражение в равенство (4), получим:

$$\int \sqrt{x^2 + b} dx = x\sqrt{x^2 + b} - \int \sqrt{x^2 + b} dx + b \int \frac{dx}{\sqrt{x^2 + b}},$$

или

$$2 \int \sqrt{x^2 + b} dx = x\sqrt{x^2 + b} + b \int \frac{dx}{\sqrt{x^2 + b}},$$

откуда

$$\int \sqrt{x^2 + b} dx = \frac{1}{2} (x\sqrt{x^2 + b} + b \ln |x + \sqrt{x^2 + b}|) + C.$$

Аналогично можно показать, что

$$\int \sqrt{a^2 - x^2} dx = \frac{1}{2} \left(x\sqrt{a^2 - x^2} + a^2 \arcsin \frac{x}{a} \right) + C.$$

4. Интеграл $\int \sqrt{Ax^2 + Bx + C} dx$ в зависимости от знака A приводится к одному из интегралов вида (2), (3).

П р и м е ч а н и е. Заканчивая методы интегрирования, заметим, что хотя для всякой непрерывной функции существует первообразная (§ 9.1, п. 1, теорема 2), но эта первообразная не для всякой функции является элементарной функцией. Например, для функции e^{-x^2} первообразная не выражается в элементарных функциях. В этом случае говорят, что интеграл $\int e^{-x^2} dx$ не берется в элементарных функциях.

§ 9.6. Понятие определенного интеграла

1. Задачи, приводящие к понятию определенного интеграла.

Задача о пройденном пути. Требуется найти путь, пройденный движущейся по прямой точкой за отрезок времени $[t_0; T]$, если известен закон изменения мгновенной скорости $v = v(t)$. Разобьем отрезок времени $[t_0; T]$ моментами времени (точками) $t_0 < t_1 < t_2 < \dots < t_n = T$ на n отрезков времени (частичных отрезков) и положим $\Delta t_k = t_k - t_{k-1}$, $k = 1, 2, \dots, n$. Наибольшую из этих разностей обозначим через λ : $\lambda = \max \Delta t_k$. Если эти отрезки достаточно малы, то без большой ошибки на каждом из них движение можно считать равномерным, что дает приближенное выражение для пути

$$s \approx v(\tau_1)\Delta t_1 + v(\tau_2)\Delta t_2 + \dots + v(\tau_n)\Delta t_n,$$

где τ_n — одна из точек сегмента $[t_{k-1}; t_k]$. Эта сумма (ее кратко будем обозначать $\sum_{k=1}^n v(\tau_k)\Delta t_k$) будет тем точнее выражать искомый путь s , чем меньше будет каждый из временных отрезков $[t_{k-1}; t_k]$ $k = 1, 2, \dots, n$. Поэтому за путь s , пройденный точкой за время $T - t_0$ со скоростью $v = v(t)$, естественно принять:

$$s = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n v(\tau_k)\Delta t_k. \quad (1)$$

Работа переменной силы. Пусть материальная точка под действием постоянной силы F перемещается по направлению этой

силы. Если пройденный путь равен s , то, как известно из курса физики, работа A этой силы F вычисляется по формуле

$$A = Fs.$$

Пусть теперь материальная точка движется по оси Ox от точки $A(a)$ до точки $B(b)$ ($b > a$) под действием переменной силы, направленной по оси Ox и являющейся функцией от x :

$$F = f(x).$$

Для нахождения работы A в этом случае разобьем отрезок $[a; b]$ точками $a = x_0 < x_1 < \dots < x_n = b$ на n частичных отрезков и положим $\Delta x_k = x_k - x_{k-1}$, $k = 1, 2, \dots, n$. Наибольшую из этих разностей обозначим через $\lambda = \max \Delta x_k$. Если эти отрезки достаточно малы, то без большой ошибки на каждом из них силу F можно считать постоянной (равной $f(\tau_k)$), что дает приближенное выражение для работы

$$A \approx \sum_{k=1}^n f(\tau_k) \Delta x_k,$$

где τ_k — одна из точек сегмента $[x_{k-1}; x_k]$. Отсюда

$$A = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\tau_k) \Delta x_k. \quad (2)$$

Задача о площади криволинейной трапеции. Пусть требуется найти площадь плоской фигуры $aABb$ (рис. 109), ограниченной графиком функции $y = f(x)$, непрерывной и неотрицательной на сегменте $[a; b]$, и отрезками прямых $y = 0$, $x = a$, $x = b$. Эта фигура называется *криволинейной трапецией*. Разобьем отрезок $[a; b]$ точками $a = x_0 < x_1 < \dots < x_n = b$ на n частичных отрезков и положим $\Delta x_k = x_k - x_{k-1}$, $k = 1, 2, \dots, n$. Наибольшую из этих разностей

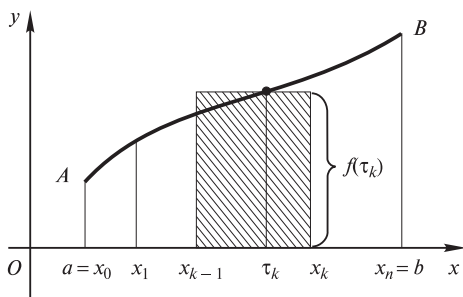


Рис. 109

обозначим через λ : $\lambda = \max \Delta x_k$. На каждом частичном сегменте $[x_{k-1}; x_k]$, $k = 1, 2, \dots, n$, выберем произвольную точку τ_k ($x_{k-1} \leq \tau_k \leq x_k$). Произведение $f(\tau_k)\Delta x_k$ даст площадь прямоугольника, имеющего основание Δx_k и высоту $f(\tau_k)$, а сумма $\sum_{k=1}^n f(\tau_k)\Delta x_k$ — приближенно площадь S криволинейной трапеции $aAbb$. Отсюда, как и в предыдущих задачах,

$$S = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\tau_k) \Delta x_k. \quad (3)$$

2. Понятие определенного интеграла. Из решения приведенных трех задач мы видим, что хотя они имеют различный смысл, но математический аппарат для их решения один и тот же. Во всех этих задачах мы получаем выражение одного и того же вида

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\tau_k) \Delta x_k. \quad (4)$$

Если существует предел (4), не зависящий от способа разбиения отрезка $[a; b]$ и выбора точек τ_k , то этот предел будем называть *определенным интегралом* функции $f(x)$ на отрезке $[a; b]$ и обозначать символом

$$\int_a^b f(x) dx = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\tau_k) \Delta x_k.$$

Функция $f(x)$ в этом случае называется *интегрируемой* на отрезке $[a; b]$. При этом $f(x)$ называется *подынтегральной функцией*, $f(x)dx$ — *подынтегральным выражением*, числа a и b — *пределами интегрирования* (a — *нижний предел*, b — *верхний предел*), а сумма $\sum_{k=1}^n f(\tau_k) \Delta x_k$ — *интегральной суммой*.

Очевидно, если функция $f(x)$ интегрируема на отрезке $[a; b]$, то она на нем ограничена.

Справедлива следующая теорема (она доказывается в более полных курсах математического анализа — см., например, [11]):

Т е о р е м а. *Если функция $f(x)$ непрерывна на отрезке $[a; b]$ то она интегрируема на этом отрезке.*

В условиях рассмотренных выше задач, приведших к понятию определенного интеграла, выражения вида (1) — (3) (пределы сумм) являются определенными интегралами. Рассмотрим это подробнее.

1. Путь s , пройденный точкой по прямой за время $T - t_0$ со скоростью $v = v(t)$ ($v(t)$ непрерывна на $[t_0; T]$), есть $\int_{t_0}^T v(t) dt$.

2. Если переменная сила $F = f(x)$ действует в направлении оси Ox ($f(x)$ непрерывна на $[a; b]$), то работа этой силы на отрезке $[a; b]$ оси Ox равна

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

При решении задач на вычисление работы силы часто используется закон Гука

$$F = kx,$$

где F — сила (Н), x — величина растяжения или сжатия (м), вызванного силой F , а k — коэффициент пропорциональности.

3. Если функция $f(x)$ непрерывна и неотрицательна на отрезке $[a; b]$, то $\int_a^b f(x)dx$ геометрически представляет собой *площадь криволинейной трапеции, ограниченной сверху графиком функции $y = f(x)$, снизу — отрезком $[a; b]$ оси Ox , с боков — отрезками прямых $x = a$, $x = b$.*

§ 9.7. Основные свойства определенного интеграла

1. Свойства определенного интеграла. Ниже рассматриваем функции, непрерывные на отрезке $[a; b]$. По определению полагают, что определенный интеграл от функции с равными верхним и нижним пределами интегрирования равен нулю:

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

1. *Постоянный множитель можно выносить за знак определенного интеграла:*

$$\int_a^b Af(x)dx = A \int_a^b f(x)dx.$$

Действительно, по определению определенного интеграла как предела интегральной суммы имеем:

$$\begin{aligned} \int_a^b Af(x)dx &= \lim_{\lambda \rightarrow 0} \sum_{k=1}^n Af(\tau_k) \Delta x_k = \lim_{\lambda \rightarrow 0} A \sum_{k=1}^n f(\tau_k) \Delta x_k = \\ &= A \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\tau_k) \Delta x_k = A \int_a^b f(x)dx. \end{aligned}$$

Аналогично устанавливается свойство:

2. *Определенный интеграл от суммы двух функций равен сумме определенных интегралов от этих функций:*

$$\int_a^b (f_1(x) + f_2(x))dx = \int_a^b f_1(x)dx + \int_a^b f_2(x)dx.$$

П р и м е ч а н и е. Свойство 2 распространяется на случай алгебраической суммы любого конечного числа функций.

3. При перестановке пределов интегрирования определенный интеграл меняет знак на противоположный:

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

Действительно, здесь соответствующие интегральные суммы различаются по знаку, ибо в одной из них все $\Delta x_k = x_k - x_{k-1}$ положительны, а в другой аналогичные разности все отрицательны.

4. Интеграл по отрезку равен сумме интегралов по его частям:

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx, \text{ где } a < c < b.$$

Это свойство вытекает из определения определенного интеграла.

5. Если функция $f(x) > 0$ на отрезке $[a; b]$, то

$$\int_a^b f(x)dx > 0.$$

Последнее свойство непосредственно следует из геометрического смысла определенного интеграла.

2. Формула Ньютона — Лейбница.

Т е о р е м а. Если функция $f(x)$ непрерывна на сегменте $[a; b]$ и $F(x)$ — первообразная функции $f(x)$ на этом отрезке, то

$$\int_a^b f(x)dx = F(b) - F(a). \quad (1)$$

Формула (1) называется *формулой Ньютона — Лейбница*. (Ньютон и Лейбниц — создатели дифференциального и интегрального исчислений.) Эта формула дает удобное правило вычисления определенного интеграла. Кроме того, она устанавливает связь между определенным интегралом и неопределенным интегралом.

Д о к а з а т е л ь с т в о. Разобьем сегмент $[a; b]$ на n частичных отрезков точками $a = x_0 < x_1 < x_2 < \dots < x_n = b$. В силу формулы Лагранжа (§ 8.6, п. 3) и формулы $F'(x) = f(x)$ имеем:

$$\begin{aligned} F(x_1) - F(a) &= F'(c_1)(x_1 - a) = f(c_1)\Delta x_1, \quad a < c_1 < x_1, \\ F(x_2) - F(x_1) &= F'(c_2)(x_2 - x_1) = f(c_2)\Delta x_2, \quad x_1 < c_2 < x_2, \\ &\dots\dots\dots \\ F(b) - F(x_{n-1}) &= F'(c_n)(b - x_{n-1}) = f(c_n)\Delta x_n, \quad x_{n-1} < c_n < b. \end{aligned}$$

Суммируя эти равенства, получим:

$$F(b) - F(a) = \sum_{k=1}^n f(c_k) \Delta x_k. \quad (2)$$

Так как функция $f(x)$ непрерывна на сегменте $[a; b]$, то

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(c_k) \Delta x_k = \int_a^b f(x) dx.$$

Поэтому, переходя в равенстве (2) к пределу при $\lambda \rightarrow 0$, получим искомую формулу (1).

П р и м е ч а н и е. Для краткости записи употребляется обозначение $F(x)|_a^b = F(b) - F(a)$ или $[F(x)]_a^b = F(b) - F(a)$. Поэтому формула Ньютона — Лейбница принимает вид:

$$\int_a^b f(x) dx = F(x)|_a^b \quad \text{или} \quad \int_a^b f(x) dx = [F(x)]_a^b.$$

Заметим, что в формуле (1) можно взять любую из первообразных для $f(x)$, так как $[F(x) + C]_a^b = F(b) + C - F(a) - C = F(b) - F(a)$.

П р и м е р 1. $\int_0^1 2x dx = x^2|_0^1 = 1.$

П р и м е р 2. $\int_0^{\frac{\pi}{2}} \cos x dx = \sin x|_0^{\frac{\pi}{2}} = \sin \frac{\pi}{2} - \sin 0 = 1.$

П р и м е р 3. $\int_1^2 (3x^4 + 2x^2 - 5) dx = 3 \int_1^2 x^4 dx + 2 \int_1^2 x^2 dx - 5 \int_1^2 dx = \frac{3}{5} x^5|_1^2 + \frac{2}{3} x^3|_1^2 - 5x|_1^2 =$
 $= \frac{3}{5} (2^5 - 1) + \frac{2}{3} (2^3 - 1) - 5(2 - 1) = \frac{274}{15}.$

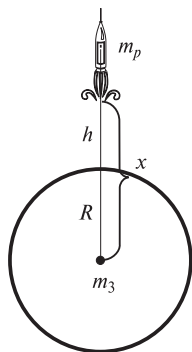


Рис. 110

З а д а ч а. Определить работу A , необходимую для запуска ракеты массой m_p с поверхности Земли вертикально вверх на высоту h (рис. 110).

Р е ш е н и е. Обозначим через F величину силы притяжения ракеты Землей. Пусть m_3 — масса Земли. Согласно закону Ньютона

$$F = k \frac{m_p m_3}{x^2},$$

где x — расстояние от ракеты до центра Земли. Полагая $km_p m_3 = \gamma$, получим $F(x) = \frac{\gamma}{x^2}$, $R \leq x \leq h + R$, R —

радиус Земли. При $x = R$ сила $F(R) = \frac{\gamma}{R^2}$ равна весу P ракеты, поэтому

$\gamma = PR^2$ и $F(x) = \frac{PR^2}{x^2}$. Значит (§ 9.6, п. 2),

$$A = \int_R^{R+h} F(x) dx = PR^2 \int_R^{R+h} \frac{dx}{x^2} = -PR^2 \frac{1}{x} \Big|_R^{R+h} = \frac{PRh}{R+h}.$$

3. Определенный интеграл с переменным верхним пределом. Пусть функция $f(x)$ непрерывна на отрезке $[a; b]$. Рассмотрим интеграл

$$\int_a^x f(t) dt,$$

где x — любая точка из $[a; b]$.

Если $F(x)$ — первообразная функции $f(x)$, т.е. $F'(x) = f(x)$, то согласно формуле Ньютона — Лейбница имеем:

$$\int_a^x f(t) dt = F(x) - F(a).$$

Отсюда

$$\frac{d}{dx} \int_a^x f(t) dt = \frac{d}{dx} (F(x) - F(a)) = f(x).$$

Таким образом, *производная определенного интеграла с переменным верхним пределом по этому пределу равна значению подынтегральной функции для этого предела.*

4. Теорема о среднем.

Т е о р е м а. Если функция $f(x)$ непрерывна на сегменте $[a; b]$, то в интервале $(a; b)$ найдется такая точка c , что

$$\int_a^b f(x) dx = (b-a)f(c). \quad (3)$$

Д о к а з а т е л ь с т в о. По формуле Ньютона — Лейбница имеем:

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

где $F'(x) = f(x)$. Но в силу формулы Лагранжа (§ 8.6, п. 3)

$$F(b) - F(a) = (b-a) F'(c) = (b-a) f(c), \quad a < c < b,$$

что и приводит к искомой формуле (3).

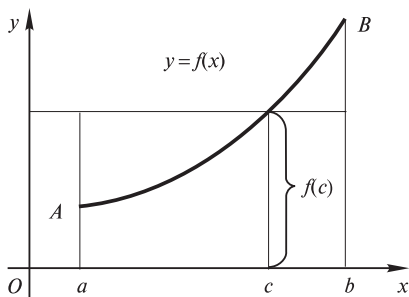


Рис. 111

Формула (3) при $f(x) \geq 0$ имеет простое геометрическое истолкование. Площадь криволинейной трапеции $aABb$ равна площади прямоугольника с тем же основанием и высотой, равной $f(c)$ (рис. 111).

5. Замена переменной в определенном интеграле. Предположим, что функция $f(x)$ непрерывна на сегменте $[a; b]$, функция $x = \varphi(t)$ имеет на сегменте $[\alpha; \beta]$ непре-

рывную производную, при этом $a \leq \varphi(t) \leq b$ и $\varphi(\alpha) = a$, $\varphi(\beta) = b$.

Пусть $F(x)$ — одна из первообразных функций $f(x)$. Тогда $F'(x) = f(x)$ и в силу формулы производной сложной функции (см. § 8.2, п., 1):

$$(F(\varphi(t)))' = F'_x(\varphi(t)) \cdot \varphi'(t) = f(\varphi(t)) \cdot \varphi'(t).$$

Теперь воспользуемся дважды формулой Ньютона — Лейбница:

$$\int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt = F(\varphi(\beta)) - F(\varphi(\alpha)) = F(b) - F(a) = \int_a^b f(x) dx.$$

Тем самым доказана формула замены переменной в определенном интеграле:

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

П р и м е р. Подстановка $x = \sin t$ дает

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

6. Интегрирование по частям. Пусть $u = u(x)$, $v = v(x)$ — непрерывно дифференцируемые на сегменте $[a; b]$ функции. Пользуясь формулой Ньютона — Лейбница, имеем:

$$\begin{aligned} [u(x)v(x)]_a^b &= \int_a^b (u(x)v(x))' dx = \\ &= \int_a^b (u(x)v'(x) + u'(x)v(x)) dx = \int_a^b u(x)v'(x) dx + \int_a^b v(x)u'(x) dx, \end{aligned}$$

откуда

$$\int_a^b u(x)v'(x) dx = [u(x)v(x)]_a^b - \int_a^b v(x)u'(x) dx.$$

Эта формула называется *формулой интегрирования по частям для определенного интеграла*.

Так как $v'(x) dx = dv$ и $u'(x) dx = du$, то эту формулу записывают еще в следующем, более компактном виде:

$$\int_a^b u dv = [uv]_a^b - \int_a^b v du. \quad (4)$$

При этом следует иметь в виду, что пределы интегрирования в формуле (4) относятся к независимой переменной x .

П р и м е р. $\int_0^{\pi} \underbrace{x}_{u} \underbrace{\cos x}_{dv} dx = [x \sin x]_0^{\pi} - \int_0^{\pi} \sin x dx = \cos x \Big|_0^{\pi} = -2.$

§ 9.8. Виды несобственных интегралов, их сходимость

При введении понятия определенного интеграла мы исходили из условий ограниченности подынтегральной функции и конечности пределов интегрирования. Такой интеграл называется *собственным* (слово «собственный» обычно опускается). Если хотя бы одно из этих двух условий не выполнено, то интеграл называется *несобственным*.

1. Интегралы с бесконечными пределами. Пусть функция $f(x)$ непрерывна при $a \leq x < +\infty$, т. е. для $x \geq a$. Тогда по определению полагают

$$\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} \int_a^b f(x) dx. \quad (1)$$

Если последний предел существует, то говорят, что интеграл

$$\int_a^{+\infty} f(x) dx \quad (2)$$

сходится, а если этот предел не существует, то интеграл (2) называют *расходящимся*. Такому интегралу не приписывают никакого значения.

Геометрически для неотрицательной при $x \geq a$ функции $f(x)$ несобственный интеграл (2) по аналогии с собственным интегралом (§ 9.6, п. 2) представляет собой площадь фигуры, ограниченной сверху графиком функции $y = f(x)$, слева отрезком прямой $x = a$, снизу осью Ox (рис. 112).

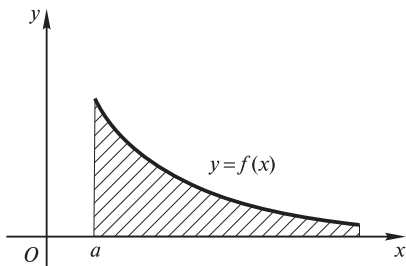


Рис. 112

Пр и м е р 1. $\int_0^{+\infty} e^{-x} dx = \lim_{b \rightarrow +\infty} \int_0^b e^{-x} dx = \lim_{b \rightarrow +\infty} (1 - e^{-b}) = 1.$

Заданный несобственный интеграл сходится.

Пр и м е р 2. $\int_1^{+\infty} \frac{dx}{x} = \lim_{b \rightarrow +\infty} \int_1^b \frac{dx}{x} = \lim_{b \rightarrow +\infty} \ln b = +\infty.$

Следовательно, данный интеграл расходится.

Пр и м е р 3. $\int_0^{+\infty} \cos x dx = \lim_{b \rightarrow +\infty} \int_0^b \cos x dx = \lim_{b \rightarrow +\infty} \sin b.$

Последний предел не существует, т. е. несобственный интеграл расходится.

Пр и м е р 4. Установить, при каких значениях α сходится интеграл $\int_1^{+\infty} \frac{dx}{x^\alpha}$. Случай

$\alpha = 1$ рассмотрен в примере 2. При $\alpha \neq 1$ имеем:

$$\int_1^{+\infty} \frac{dx}{x^\alpha} = \lim_{b \rightarrow +\infty} \frac{b^{1-\alpha} - 1}{1 - \alpha} = \begin{cases} +\infty & \text{при } \alpha < 1, \\ \frac{1}{\alpha - 1} & \text{при } \alpha > 1. \end{cases}$$

Значит, данный интеграл сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$.

На несобственные интегралы вида (2) непосредственно распространяются многие свойства собственных интегралов.

Пусть $F(x)$ — первообразная функция для подынтегральной функции $f(x)$ сходящегося интеграла (2). На основании формулы (1) и формулы Ньютона — Лейбница имеем:

$$\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} (F(b) - F(a)).$$

Если ввести условное обозначение

$$F(+\infty) = \lim_{b \rightarrow +\infty} F(b),$$

то получим для сходящегося несобственного интеграла (2) обобщенную формулу Ньютона — Лейбница

$$\int_a^{+\infty} f(x) dx = F(+\infty) - F(a),$$

которую записывают также в виде

$$\int_a^{+\infty} f(x) dx = F(x) \Big|_a^{+\infty} \text{ или } \int_a^{+\infty} f(x) dx = [F(x)]_a^{+\infty}.$$

Заметим еще, что для вычисления сходящихся интегралов вида (2) сохраняются методы подстановки и интегрирования по частям.

Самым простым признаком сходимости несобственных интегралов вида (2) является признак сравнения.

Рассмотрим его для случая неотрицательной подынтегральной функции.

Т е о р е м а с р а в н е н и я . Если в промежутке $[a; +\infty)$ функции $f(x)$, $g(x)$ непрерывны и удовлетворяют неравенствам $0 \leq g(x) \leq f(x)$, то из сходимости интеграла

$$\int_a^{+\infty} f(x) dx \quad (3)$$

следует сходимость интеграла

$$\int_a^{+\infty} g(x) dx. \quad (4)$$

Этот признак вытекает из геометрического смысла интеграла (2).

Из этого признака следует, что при том же условии из расходимости интеграла (4) следует расходимость интеграла (3).

П р и м е р 5. $\int_1^{+\infty} \frac{dx}{\sqrt[3]{1+x^5}}$ сходится, так как при $x \geq 1$ $\frac{1}{\sqrt[3]{1+x^5}} < \frac{1}{x^{\frac{5}{3}}}$ и интеграл $\int_1^{+\infty} \frac{dx}{x^{\frac{5}{3}}}$ сходится (см. пример 4).

П р и м е р 6. $\int_1^{+\infty} \frac{x+2}{x\sqrt{x}} dx$ расходится, так как при $x \geq 1$ $\frac{x+2}{x\sqrt{x}} > \frac{1}{\sqrt{x}}$ и интеграл $\int_1^{+\infty} \frac{dx}{x^{\frac{1}{2}}}$ расходится (см. пример 4).

Пусть теперь в интеграле (2) функция $f(x)$ может принимать значения разных знаков.

Т е о р е м а . Из сходимости интеграла

$$\int_a^{+\infty} |f(x)| dx \quad (5)$$

следует сходимость интеграла (2).

Д о к а з а т е л ь с т в о . Введя неотрицательные функции

$$\varphi(x) = \frac{f(x) + |f(x)|}{2}, \quad \psi(x) = \frac{|f(x)| - f(x)}{2},$$

имеем $\varphi(x) \leq |f(x)|$, $\psi(x) \leq |f(x)|$ и $f(x) = \varphi(x) - \psi(x)$. Так как интеграл (5) сходится, то по теореме сравнения сходятся интегралы

$$\int_a^{+\infty} \varphi(x) dx, \quad \int_a^{+\infty} \psi(x) dx$$

и, следовательно, сходится интеграл

$$\int_a^{+\infty} f(x) dx = \int_a^{+\infty} \varphi(x) dx - \int_a^{+\infty} \psi(x) dx.$$

Интеграл (2) называется *абсолютно сходящимся*, если сходится интеграл (5).

Пример 7. $\int_1^{+\infty} \frac{\cos x}{x^2} dx$ сходится абсолютно, потому что сходится по теореме сравнения $\int_1^{+\infty} \frac{|\cos x|}{x^2} dx$ (при $x \geq 1$ $\frac{|\cos x|}{x^2} \leq \frac{1}{x^2}$ и $\int_1^{+\infty} \frac{dx}{x^2}$ сходится; см. пример 4).

Все изложенное непосредственно переносится на интегралы вида

$$\int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx \quad (6)$$

(кстати заметим, что от интеграла (6) легко перейти к интегралу вида (2) с помощью подстановки $x = -y$).

Наконец, по определению

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^c f(x) dx + \int_c^{+\infty} f(x) dx, \quad (7)$$

где c — какое-нибудь число (выбор его безразличен). Последнее равенство следует понимать так: если каждый из интегралов, стоящих справа, сходится, то сходится и интеграл, стоящий слева. Если же расходится хотя бы один из интегралов, стоящих справа, то расходится и интеграл, стоящий слева.

2. Интегралы от неограниченных функций. Предположим, что функция $f(x)$ непрерывна при $a \leq x < b$. Предположим далее, что эта функция стремится к бесконечности, когда $x \rightarrow b$. Так что на отрезке $[a; b]$ функция $f(x)$ не ограничена. Положим

$$\int_a^b f(x) dx = \lim_{\varepsilon \rightarrow 0+0} \int_a^{b-\varepsilon} f(x) dx.$$

Если последний предел существует, то говорят, что интеграл

$$\int_a^b f(x) dx \quad (8)$$

сходится, а если этот предел не существует, то интеграл (8) называют *расходящимся*.

Подобным же образом равенством

$$\int_a^b f(x) dx = \lim_{\varepsilon \rightarrow 0+0} \int_{a+\varepsilon}^b f(x) dx$$

дается определение интеграла функции $f(x)$, стремящейся к бесконечности при $x \rightarrow a$.

Пр и м е р 1. $\int_a^b \frac{dx}{b-x} = - \lim_{\varepsilon \rightarrow 0+0} (\ln \varepsilon - \ln(b-a)) = +\infty$, интеграл расходится.

Пр и м е р 2. Выяснить, при каких значениях α сходится интеграл

$$\int_a^b \frac{dx}{(b-x)^\alpha}.$$

Случай $\alpha = 1$ рассмотрен в примере 1. При $\alpha \neq 1$ имеем:

$$\int_a^b \frac{dx}{(b-x)^\alpha} = - \lim_{\varepsilon \rightarrow 0+0} \frac{\varepsilon^{1-\alpha} - (b-a)^{1-\alpha}}{1-\alpha} = \begin{cases} +\infty & \text{при } \alpha > 1, \\ \frac{(b-a)^{1-\alpha}}{1-\alpha} & \text{при } \alpha < 1. \end{cases}$$

Значит, данный интеграл сходится при $\alpha < 1$ и расходится при $\alpha \geq 1$.

Все свойства для интегралов вида (2) и (6) переносятся и на только что введенные интегралы. В частности, имеет место теорема сравнения, аналогичная приведенной в п. 1.

Наконец, если функция $f(x)$ стремится к бесконечности при приближении к обоим концам промежутка $(a; b)$, то полагают:

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

При этом если сходятся оба интеграла в правой части последнего равенства, то сходится и интеграл слева.

3. Пример использования несобственного интеграла. В качестве примера использования несобственного интеграла проведем подсчет второй космической скорости.

Вычислим начальную скорость ракеты, при которой она способна выйти из поля притяжения Земли и уйти в межпланетное пространство.

Ранее (§ 9.7, п. 2, задача) с помощью определенного интеграла была подсчитана работа силы, необходимая для запуска ракеты массой m_p с поверхности Земли на высоту h , т. е. было получено:

$$A = \int_R^{R+h} F(x)dx = \frac{PRh}{R+h} = \frac{Ph}{1 + \frac{h}{R}},$$

откуда следует, что при R , несравнимо большем, чем h , $A \approx m_p gh$, где m_p — масса ракеты, g — ускорение свободного падения у поверхности Земли.

В силу закона сохранения энергии сумма кинетической и потенциальной энергии ракеты во все моменты времени одна и та же. Следовательно, чтобы полностью освободить ракету от земного притяжения,

необходимо, чтобы начальная кинетическая энергия была не меньше потенциальной энергии ракеты, бесконечно удаленной от Земли, т. е. надо подсчитать работу силы при $h \rightarrow \infty$:

$$\lim_{h \rightarrow \infty} \frac{PRh}{R+h} = \lim_{h \rightarrow \infty} \frac{PR}{1 + \frac{R}{h}} = PR = m_p g R$$

(движение Земли и притяжение других планет при этом не учитываются). Отсюда начальная скорость ракеты v должна быть такова, чтобы $\frac{m_p v^2}{2} \geq m_p g R$, или $v \geq \sqrt{2gR} = \sqrt{2 \cdot 10 \cdot 6\,400\,000} \approx 11,3 \cdot 10^3 \text{ м/с} = 11,3 \text{ км/с}$ (при вычислениях можно использовать калькулятор; g взято равным 10 м/с^2).

4. Гамма-функция. Интеграл

$$\int_0^{+\infty} e^{-x} x^{p-1} dx \quad (p = \text{const})$$

сходится при $p > 0$ (см. [9], т. 2).

Положим для $p > 0$

$$\Gamma(p) = \int_0^{+\infty} e^{-x} x^{p-1} dx.$$

Этот несобственный интеграл, зависящий от параметра p , называется *гамма-функцией*. (Эта функция введена Л. Эйлером в 1729 г.)

Установим следующие два свойства гамма-функции:

$$1. \Gamma(p+1) = p\Gamma(p).$$

Действительно, применяя формулу интегрирования по частям, получаем:

$$\Gamma(p+1) = \int_0^{+\infty} e^{-x} x^p dx = -[x^p e^{-x}]_0^{+\infty} + p \int_0^{+\infty} e^{-x} x^{p-1} dx = p\Gamma(p).$$

2. Если n — натуральное число, то

$$\Gamma(n+1) = n!.$$

В самом деле, применяя последовательно формулу $\Gamma(n+1) = n\Gamma(n)$, получим:

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = \dots = n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1\Gamma(1).$$

Но

$$\Gamma(1) = \int_0^{+\infty} e^{-x} dx = 1 \quad (\text{п. 1, пример 1}),$$

и поэтому

$$\Gamma(n+1) = n!.$$

§ 9.9. Геометрические приложения определенного интеграла

1. Вычисление площадей плоских фигур. Пусть функция $f(x)$ непрерывна на сегменте $[a; b]$. Известно (см. § 9.6), что если $f(x) \geq 0$ на $[a; b]$, то площадь S криволинейной трапеции, ограниченной линиями $y = f(x)$, $y = 0$, $x = a$, $x = b$, равна интегралу

$$S = \int_a^b f(x) dx. \quad (1)$$

(О понятии площади произвольной плоской фигуры, а также объема тела и площади поверхности см., например, в [9].)

Если же $f(x) \leq 0$ на $[a; b]$, то $-f(x) \geq 0$ на $[a; b]$. Поэтому площадь S соответствующей криволинейной трапеции выразится формулой

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

или

$$S = \left| \int_a^b f(x) dx \right|. \quad (2)$$

Если, наконец, линия $y = f(x)$ пересекает ось Ox , то сегмент $[a; b]$ надо разбить на части, в пределах которых $f(x)$ не меняет знака, и к каждой части применить ту из формул (1) или (2), которая ей соответствует.

П р и м е р. Найти площадь плоской фигуры, ограниченной графиком функции $y = \sin x$ и осью абсцисс при условии $0 \leq x \leq 2\pi$ (рис. 113). Разбиваем сегмент $[0; 2\pi]$ на два сегмента $[0; \pi]$ и $[\pi; 2\pi]$. На первом из них $\sin x \geq 0$, на втором $\sin x \leq 0$. Следовательно, используя формулы (1) и (2), имеем, что искомая площадь

$$S = \int_0^{\pi} \sin x dx + \left| \int_{\pi}^{2\pi} \sin x dx \right| = -\cos x \Big|_0^{\pi} + \left| -\cos x \Big|_{\pi}^{2\pi} \right| = 4.$$

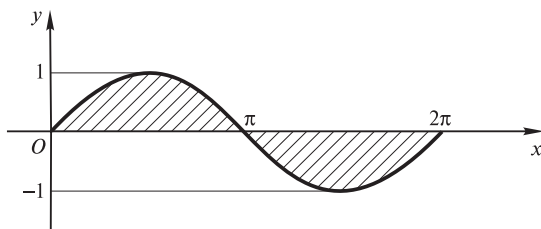


Рис. 113

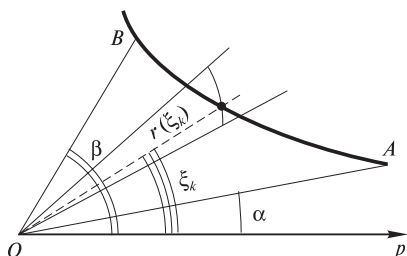


Рис. 114

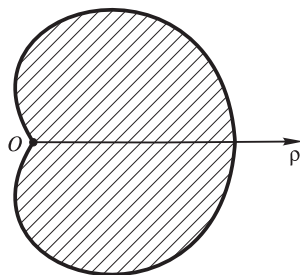


Рис. 115

2. Вычисление площади в полярных координатах. Пусть требуется определить площадь криволинейного сектора OAB (рис. 114), ограниченного лучами $\varphi = \alpha$, $\varphi = \beta$ и кривой AB , заданной в полярной системе координат уравнением $r = r(\varphi)$, где $r(\varphi)$ — функция, непрерывная на отрезке $[\alpha; \beta]$.

Разобьем отрезок $[\alpha; \beta]$ на n частей точками $\alpha = \varphi_0 < \varphi_1 < \dots < \varphi_{n-1} < \varphi_n = \beta$ и положим $\Delta\varphi_k = \varphi_k - \varphi_{k-1}$, $k = 1, \dots, n$. Наибольшую из этих разностей обозначим через λ : $\lambda = \max \Delta\varphi_k$. Разобьем данный сектор на n частей лучами $\varphi = \varphi_k$ ($k = 1, \dots, n-1$). Заменим k -й элементарный сектор круговым сектором радиуса $r(\xi_k)$, где $\xi_k \in [\varphi_{k-1}; \varphi_k]$. Тогда сумма $\sum_{k=1}^n \frac{1}{2} r^2(\xi_k) \Delta\varphi_k$ — это приближенно площадь сектора OAB . Отсюда

$$S = \frac{1}{2} \lim_{\lambda \rightarrow 0} \sum_{k=1}^n r^2(\xi_k) \Delta\varphi_k = \frac{1}{2} \int_{\alpha}^{\beta} r^2(\varphi) d\varphi. \quad (3)$$

П р и м е р. Найти площадь плоской фигуры, ограниченной кардиоидой $r = a(1 + \cos \varphi)$ (рис. 115). Учитывая симметричность кривой относительно полярной оси, по формуле (3) получаем:

$$\begin{aligned} S &= a^2 \int_0^{\pi} (1 + \cos \varphi)^2 d\varphi = a^2 \int_0^{\pi} (1 + 2 \cos \varphi + \cos^2 \varphi) d\varphi = \\ &= a^2 \varphi \Big|_0^{\pi} + 2a^2 \sin \varphi \Big|_0^{\pi} + \frac{a^2}{2} \int_0^{\pi} (1 + \cos 2\varphi) d\varphi = \\ &= a^2 \pi + \frac{a^2}{2} \varphi \Big|_0^{\pi} + \frac{a^2}{4} \sin 2\varphi \Big|_0^{\pi} = \frac{3}{2} \pi a^2. \end{aligned}$$

3. Вычисление длины дуги. Пусть дуга AB (рис. 116) задана уравнением $y = f(x)$, $x \in [a; b]$, где $f(x)$ — функция, имеющая на отрезке $[a; b]$ непрерывную производную.

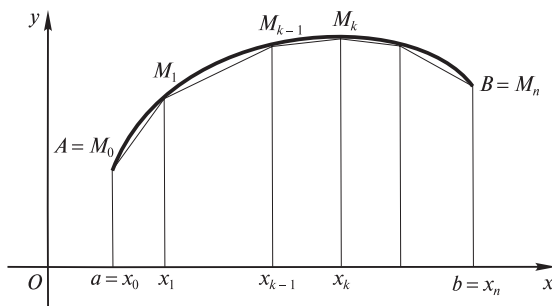


Рис. 116

Длиной дуги AB называется предел, к которому стремится длина ломаной линии, вписанной в эту дугу, когда длина наибольшего звена стремится к нулю.

Найдем длину дуги AB . Впишем в дугу AB ломаную линию $M_0M_1 \dots M_n$ ($M_0 = A$, $M_n = B$) (рис. 116). Пусть абсциссы точек $M_0, M_1, M_2, \dots, M_n$ соответственно $a = x_0, x_1, x_2, \dots, x_n = b$ (ординаты этих точек обозначим соответственно через $y_0, y_1, \dots, y_n$). Имеем разбиение отрезка $[a; b]$ на частичные отрезки $[x_{k-1}; x_k]$, $k = 1, 2, \dots, n$. Длина отрезка $[x_{k-1}; x_k]$ равна $\Delta x_k = x_k - x_{k-1}$. Пусть $\lambda = \max \Delta x_k$. Через Δy_k обозначим приращение функции $y = f(x)$ на отрезке $[x_{k-1}; x_k]$. По теореме Пифагора $M_{k-1}M_k = \sqrt{\Delta x_k^2 + \Delta y_k^2}$. Но в силу формулы Лагранжа (§ 8.6, п. 3) $\Delta y_k = f'(\xi_k)\Delta x_k$, где ξ_k — некоторая промежуточная точка отрезка $[x_{k-1}; x_k]$. Отсюда $M_{k-1}M_k = \sqrt{1 + f'^2(\xi_k)}\Delta x_k$, и, следовательно, длина ломаной линии $M_0M_1 \dots M_n$

$$\sum_{k=1}^n \sqrt{1 + f'^2(\xi_k)} \Delta x_k.$$

Переходя здесь к пределу при $\lambda \rightarrow 0$, получим:

$$l = \int_a^b \sqrt{1 + f'^2(x)} dx,$$

или

$$l = \int_a^b \sqrt{1 + y'^2} dx, \quad (4)$$

(l — длина дуги AB).

Отсюда длина дуги AM , где $M(x; y)$ — переменная точка дуги AB ,

$$l = l(x) = \int_a^x \sqrt{1 + y'^2} dx.$$

Поэтому (см. § 9.7, п. 3)

$$\frac{dl}{dx} = \sqrt{1 + y'^2},$$

откуда получаем формулу дифференциала дуги

$$dl = \sqrt{1 + y'^2} dx, \quad (5)$$

или

$$dl = \sqrt{dx^2 + dy^2}.$$

Если кривая AB задана параметрическими уравнениями

$$x = x(t), y = y(t) \quad (\alpha \leq t \leq \beta), \quad (6)$$

причем функции $x(t)$ и $y(t)$ имеют непрерывные производные $x'(t)$ и $y'(t)$ в $[\alpha; \beta]$, то путем замены переменной $x = x(t)$ в (4) получим:

$$l = \int_{\alpha}^{\beta} \sqrt{x'^2 + y'^2} dt. \quad (7)$$

Формула дифференциала дуги вместо (5) будет $dl = \sqrt{x'^2 + y'^2} dt$.

Если же кривая AB задана в полярных координатах уравнением

$$r = r(\varphi) \quad (\alpha \leq \varphi \leq \beta), \quad (8)$$

то, учитывая связь между прямоугольными и полярными координатами (см. § 1.1, п. 2), параметрические уравнения этой кривой будут

$$x = r \cos \varphi, y = r \sin \varphi \quad (\alpha \leq \varphi \leq \beta).$$

Поэтому

$$x' = r' \cos \varphi - r \sin \varphi, y' = r' \sin \varphi + r \cos \varphi,$$

и по формуле (7) получим:

$$l = \int_{\alpha}^{\beta} \sqrt{r^2 + r'^2} d\varphi. \quad (9)$$

Формула дифференциала дуги будет $dl = \sqrt{r^2 + r'^2} d\varphi$.

Пример 1. Найти длину дуги данной линии:

$$y = \frac{1}{2}(e^x + e^{-x}), 0 \leq x \leq 1 \text{ (рис. 117).}$$

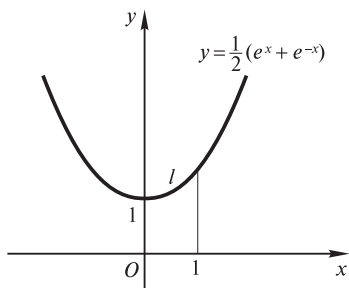


Рис. 117

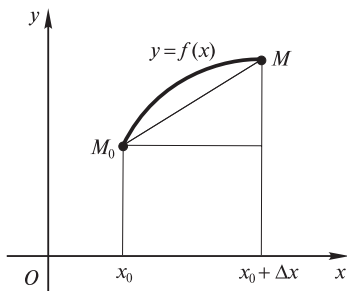


Рис. 118

Имеем $y' = \frac{1}{2}(e^x - e^{-x})$ и по формуле (4) находим

$$l = \int_0^1 \sqrt{1 + \frac{1}{4}(e^{2x} - 2 + e^{-2x})} dx = \frac{1}{2} \int_0^1 (e^x + e^{-x}) dx = \frac{1}{2} [e^x - e^{-x}]_0^1 = \frac{1}{2} \left(e - \frac{1}{e} \right).$$

Пример 2. Вычислить длину окружности радиуса R . Запишем уравнение окружности в полярных координатах: $r = R$ при $0 \leq \varphi \leq 2\pi$ — и по формуле (9) получим:

$$l = \int_0^{2\pi} R d\varphi = 2\pi R.$$

Укажем еще на одно применение формулы (4).

Теорема. Пусть дуга M_0M (рис. 118) задана уравнением $y = f(x)$ ($x_0 \leq x \leq x_0 + \Delta x$), причем $f(x)$ и $f'(x)$ — непрерывные функции. Тогда предел отношения длины этой дуги к длине стягивающей ее хорды при стремлении длины хорды к нулю равен единице.

Доказательство. Имеем:

$$M_0M = \sqrt{(\Delta x)^2 + (\Delta y)^2} = \sqrt{1 + \left(\frac{\Delta y}{\Delta x} \right)^2} \cdot \Delta x,$$

$$l = \int_{x_0}^{x_0 + \Delta x} \sqrt{1 + f'^2(x)} dx \quad (l — \text{длина дуги } M_0M), \text{ или в силу теоремы о среднем}$$

нем (см. § 9.7, п. 4) $l = \sqrt{1 + f'^2(c)} \cdot \Delta x$, где $x_0 < c < x_0 + \Delta x$. Отсюда, замечая, что если $\Delta x \rightarrow 0$, то $c \rightarrow x_0$, получим:

$$\lim_{M_0M \rightarrow 0} \frac{l}{MM_0} = \lim_{\Delta x \rightarrow 0} \frac{\sqrt{1 + f'^2(c)} \cdot \Delta x}{\sqrt{1 + \left(\frac{\Delta y}{\Delta x} \right)^2} \cdot \Delta x} = \frac{\sqrt{1 + f'^2(x_0)}}{\sqrt{1 + f'^2(x_0)}} = 1.$$

С л е д с т в и е. $\lim_{l \rightarrow 0} \frac{MM_0}{l} = 1$.

П р и м е ч а н и е. Если задана пространственная кривая параметрическими уравнениями (см. § 4.2)

$$x = x(t), y = y(t), z = z(t) \quad (\alpha \leq t \leq \beta),$$

где функции $x(t)$, $y(t)$, $z(t)$ имеют непрерывные производные, то ее длина (длина дуги для пространственной кривой определяется так же, как и для плоской) вычисляется по формуле

$$I = \int_{\alpha}^{\beta} \sqrt{x'^2 + y'^2 + z'^2} dx,$$

являющейся обобщением формулы (7). При этом длина дуги, отвечающая произвольному значению параметра t из $[\alpha; \beta]$, выразится формулой

$$I = \int_{\alpha}^t \sqrt{x'^2 + y'^2 + z'^2} dt,$$

откуда дифференциал дуги есть

$$dl = \sqrt{x'^2 + y'^2 + z'^2} dt.$$

4. Кривизна плоской кривой. Пусть к данной кривой $y = f(x)$ можно провести касательную в каждой точке. Проведем касательную к этой линии в точке $M_0(x_0, y_0)$. Обозначим через φ угол, образованный касательной с осью Ox (рис. 119). Пусть касательная в точке M образует с осью Ox угол $\varphi + \Delta\varphi$. Угол $\Delta\varphi$ между касательными в указанных точках называют *углом смежности*. Можно сказать, что при переходе из точки M_0 в точку M данной линии касательная к ней повернулась на угол $\Delta\varphi$, которому будем приписывать соответствующий знак в зависимости от направления поворота («плюс» при повороте против часовой стрелки, «минус» — по часовой стрелке). Величина угла поворота дает некоторое

представление о степени изогнутости линии на дуге M_0M . Однако сама по себе величина угла $\Delta\varphi$ еще не может служить мерой этой изогнутости; важно то, какая доля угла приходится в среднем на единицу длины дуги M_0M .

Средней кривизной дуги M_0M данной линии называется абсолютное значение отношения угла смежности $\Delta\varphi$ к длине Δl дуги M_0M :

$$K_{\text{ср}} = \left| \frac{\Delta\varphi}{\Delta l} \right|.$$

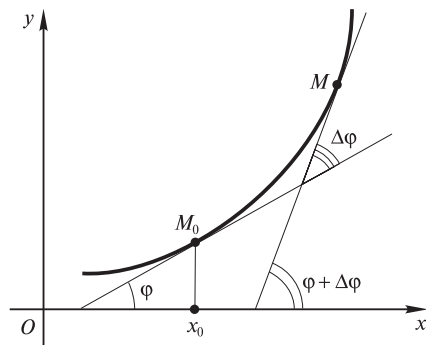


Рис. 119

Заметим, что разные дуги одной и той же линии могут иметь разную среднюю кривизну (на рисунке 120 средняя кривизна дуги M_0M значительно больше, чем дуги AB). Чтобы охарактеризовать степень изогнутости линии в данной точке, вводят понятие кривизны в этой точке.

Кривизной линии в данной точке M_0 называется предел средней кривизны дуги M_0M при $M \rightarrow M_0$:

$$K = \lim_{M \rightarrow M_0} \left| \frac{\Delta\varphi}{\Delta l} \right|.$$

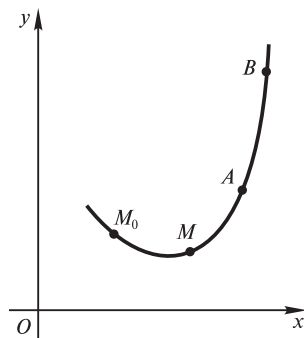


Рис. 120

Предположим теперь, что функция $y=f(x)$ имеет еще вторую производную в точке x_0 . Так как величины φ и l обе зависят от x (являются функциями от x), то, следовательно, φ можно рассматривать как функцию от l . Тогда

$$\lim_{\Delta l \rightarrow 0} \left| \frac{\Delta\varphi}{\Delta l} \right| = \left| \frac{d\varphi}{dl} \right|,$$

и, следовательно,

$$K = \left| \frac{d\varphi}{dl} \right|. \quad (10)$$

Из геометрического смысла производной следует $\operatorname{tg} \varphi = y'$, откуда $\varphi = \operatorname{arctg} y'$ и

$$d\varphi = \frac{y''}{1 + y'^2} dx. \quad (11)$$

Отсюда с учетом формул (5), (10), (11)

$$K = \frac{|y''|}{(1 + y'^2)^{\frac{3}{2}}}. \quad (12)$$

Заметим, что формула (12) содержит значения первой и второй производной данной функции $y=f(x)$, вычисленные при $x=x_0$.

Введем еще одно определение.

Радиусом кривизны данной линии и данной ее точке называется величина ρ , обратная кривизне K этой линии в рассматриваемой точке:

$$\rho = \frac{1}{K}. \quad (13)$$

Принимая во внимание формулу (12), из соотношения (13) получаем следующее выражение для радиуса кривизны линии $y=f(x)$ в точке $M_0(x_0; y_0)$:

$$\rho = \frac{(1 + y'^2)^{\frac{3}{2}}}{|y''|}$$

(здесь значения y' и y'' взяты при $x=x_0$).

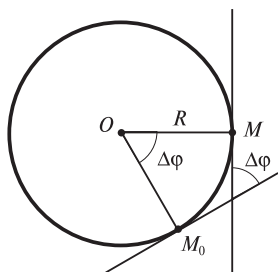


Рис. 121

Пример 1. В любой точке прямой линии касательная к ней совпадает с самой прямой, поэтому $\Delta\varphi = 0$ и средняя кривизна (и кривизна) равна нулю в любой ее точке.

Пример 2. Найти кривизну и радиус кривизны окружности. Для окружности угол смежности $\Delta\varphi$ равен углу между ее радиусами OM_0 и OM (рис. 121), длина дуги MM_0 выражается формулой $\Delta l = R\Delta\varphi$, где R — радиус данной окружности. Следовательно, $K_{\text{ср}} = \frac{\Delta\varphi}{\Delta l} = \frac{1}{R}$.

$K = \frac{1}{R}$, т. е. кривизна окружности постоянна и равна ве-

личине, обратной радиусу этой окружности. Из равенства (13) следует, что радиус кривизны окружности в любой ее точке равен радиусу окружности.

Пример 3. Найти кривизну параболы $y = px^2$ ($p > 0$) в точке $O(0; 0)$. Замечая, что $y' = 2px$, $y'' = 2p$, получим, что кривизна в точке O равна $2p$.

5. Площадь поверхности вращения. Переходя к площади поверхности вращения, предположим, что, как и в п. 3, дуга AB задана уравнением $y = f(x)$, $x \in [a; b]$, где $f(x)$ — функция, имеющая на $[a; b]$ непрерывную производную. Поверхность, образованная вращением k -го звена ломаной $M_0M_1\dots M_n$ вокруг оси Ox , есть боковая поверхность усеченного конуса с площадью

$$\pi(y_{k-1} + y_k) \cdot M_{k-1}M_k,$$

или

$$2\pi f(\eta_k) \sqrt{1 + f'^2(\xi_k)} \Delta x_k,$$

где

$$f(\eta_k) = \frac{f(x_{k-1}) + f(x_k)}{2}.$$

Следовательно, площадь поверхности вращения ломаной $M_0M_1\dots M_n$ вокруг оси Ox равна:

$$S_n = \sum_{k=1}^n 2\pi f(\eta_k) \sqrt{1 + f'^2(\xi_k)} \Delta x_k, \quad (14)$$

Площадь S поверхности вращения дуги AB вокруг оси Ox определим как $\lim_{n \rightarrow \infty} S_n$. Заметим при этом, что сумма (14) не является интегральной суммой для функции $2\pi f(x) \sqrt{1 + f'^2(x)}$, так как в слагаемом, соответствующем отрезку $[x_{k-1}; x_k]$, значения функций $f(x)$ и $f'^2(x)$ взяты в разных точках этого отрезка. Но можно доказать (см., например, в [8]), что

предел суммы (14) равняется пределу интегральной суммы для функции $2\pi f(x)\sqrt{1+f'^2(x)}$, т. е.

$$\lim_{\lambda \rightarrow 0} S_n = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n 2\pi f(\xi_k) \sqrt{1+f'^2(\xi_k)} \Delta x_k.$$

Поэтому

$$S = 2\pi \int_a^b y \sqrt{1+y'^2} dx. \quad (15)$$

Если данная кривая AB задана параметрическими уравнениями (6) или уравнением (8) в полярных координатах, то соответственно получим:

$$S = 2\pi \int_{\alpha}^{\beta} y \sqrt{x'^2 + y'^2} dt,$$

$$S = 2\pi \int_{\alpha}^{\beta} r \sin \varphi \sqrt{r'^2 + r'^2} d\varphi.$$

П р и м е р. Найти площадь поверхности шарового пояса, образованного вращением вокруг оси Ox дуги окружности $x^2 + y^2 = R^2$, соответствующей изменению x от a до b ($-R \leq a < 0$, $0 < b \leq R$). Здесь $y = \sqrt{R^2 - x^2}$, $y' = -\frac{x}{y}$, $1 + y'^2 = \frac{R^2}{y^2}$ и по формуле (15) $S = 2\pi R \int_a^b dx = 2\pi R(b-a)$. В частности, при $a = -R$, $b = R$ получаем площадь сферы $S = 4\pi R^2$.

6. Вычисление объема. Рассмотрим тело B , содержащееся между плоскостями $x = a$ и $x = b$ (рис. 122). Пусть для каждого x из сегмента $[a; b]$ дана площадь сечения этого тела $Q(x)$, перпендикулярного оси Ox . Найдем объем V данного тела при условии непрерывности $Q(x)$ в $[a; b]$. Разделим сегмент $[a; b]$ на n частей и через точки деления проведем плоскости, перпендикулярные оси Ox . Эти плоскости разобьют B на слои. Выделим k -й слой, ограниченный плоскостями $x = x_{k-1}$ и $x = x_k$. Объем этого слоя приближенно равен $Q(x_{k-1})\Delta x_k$. Образум сумму $V_n = \sum_{k=1}^n Q(x_{k-1})\Delta x_k$.

Объем V тела B определим как $\lim_{\lambda \rightarrow 0} V_n$. Этот предел существует в силу непрерывности $Q(x)$ на $[a; b]$ и равен $\int_a^b Q(x) dx$. Итак,

$$V = \int_a^b Q(x) dx.$$

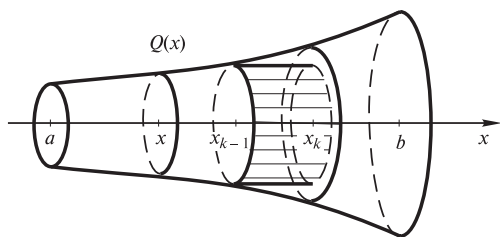


Рис. 122

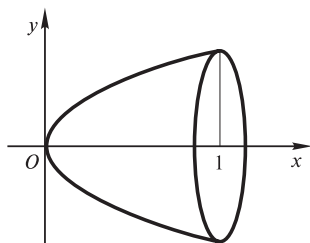


Рис. 123

В частности, если тело образовано поверхностью вращения линии $y=f(x)$ вокруг оси Ox в пределах изменения x от a до b , то $Q(x) = \pi f^2(x)$ и $V = \pi \int_a^b f^2(x) dx$, или, более кратко,

$$V = \pi \int_a^b y^2 dx. \quad (16)$$

П р и м е р. Найти объем тела, образованного вращением плоской фигуры, ограниченной линиями $y^2 = x$ и $x = 1$, вокруг оси Ox . Это тело называется *сегментом параболоида вращения* (рис. 123). Согласно формуле (16) имеем:

$$V = \pi \int_0^1 x dx = \pi \frac{x^2}{2} \Big|_0^1 = \frac{\pi}{2}.$$

§ 9.10. Физические приложения определенного интеграла

Статическим моментом материальной точки, находящейся в плоскости xOy , относительно координатной оси Ox (или Oy) называется произведение массы этой точки на ее ординату (соответственно абсциссу). Статическим моментом системы таких точек $M_1, M_2, \dots, M_n$ относительно координатной оси называется сумма статических моментов всех точек системы относительно этой оси.

Центром тяжести системы материальных точек с массами $m_1, m_2, \dots, m_n$ называется точка C , обладающая тем свойством, что если в ней сосредоточить всю массу системы $m = m_1 + m_2 + \dots + m_n$, то ее статический момент по отношению к любой оси равен статическому моменту системы точек относительно той же оси. Поэтому если обозначить через $M_x^{(n)}$ и $M_y^{(n)}$ статические моменты системы точек относительно координатных осей Ox и Oy , то координаты x_C и y_C центра тяжести C удовлетворяют соотношениям

$$mx_C = M_y^{(n)} = m_1 x_1 + \dots + m_n x_n, \quad my_C = M_x^{(n)} = m_1 y_1 + \dots + m_n y_n,$$

где x_k, y_k — декартовы координаты точки массой m_k . Следовательно, центр тяжести данной системы материальных точек имеет координаты

$$x_C = \frac{\sum_{k=1}^n m_k x_k}{m}, \quad y_C = \frac{\sum_{k=1}^n m_k y_k}{m}.$$

Статические моменты и координаты центра тяжести дуги плоской линии можно выразить через определенные интегралы. Пусть дуга AB задана уравнением $y = f(x)$, $x \in [a; b]$, где $f(x)$ — функция, имеющая на $[a; b]$ непрерывную производную, и на этой дуге распределено вещество с плотностью $\rho = \rho(x)$. Разделим дугу AB на n частичных дуг $M_{k-1}M_k$ ($k = 1, 2, \dots, n$). Сосредоточим массу каждого из элементов $M_{k-1}M_k$ в одной его точке $N_k(x_k; y_k)$. Тогда получим приближенные выражения элемента массы $\Delta m_k \approx \rho(x_k) M_{k-1}M_k$ и элементарных статических моментов относительно координатных осей $(\Delta M_x)_k \approx y_k \Delta m_k$, $(\Delta M_y)_k \approx x_k \Delta m_k$. Суммируя и переходя к пределу при $\lambda \rightarrow 0$, получим выражение массы материальной дуги

$$m = \int_a^b \rho \sqrt{1 + y'^2} dx$$

и ее статических моментов относительно координатных осей:

$$M_x = \int_a^b \rho y \sqrt{1 + y'^2} dx, \quad M_y = \int_a^b \rho x \sqrt{1 + y'^2} dx.$$

Для нахождения центра тяжести $C(x_C; y_C)$ материальной дуги AB в соответствии с определением этого понятия составим равенства

$$mx_C = M_y \text{ и } my_C = M_x,$$

из которых следует, что

$$x_C = \frac{M_y}{m}, \quad y_C = \frac{M_x}{m}. \quad (1)$$

Из формул (1) при $\rho(x) = \text{const}$ следует $\int_a^b y \sqrt{1 + y'^2} dx = y_C l$. Умножив обе части последнего равенства на 2π , получим:

$$2\pi \int_a^b y \sqrt{1 + y'^2} dx = 2\pi y_C l,$$

или

$$S = l \cdot 2\pi y_C, \quad (2)$$

где S — площадь поверхности, образуемой вращением дуги AB вокруг оси Ox . В правой части (2) $2\pi y_C$ — длина окружности, описываемой точкой $C(x_C; y_C)$ при вращении вокруг оси Ox .

Это приводит к следующей теореме.

Первая теорема Гульдина (Пауль Гульдин (1577—1643) — швейцарский математик). *Площадь поверхности, образуемой вращением дуги плоской кривой вокруг не пересекающей ее оси, лежащей в плоскости дуги, равна произведению длины вращающейся дуги на длину окружности, которую при этом вращении описывает центр тяжести дуги.*

Пример 1. Найти центр тяжести массы, распространенной вдоль полуокружности $y = \sqrt{R^2 - x^2}$ при условии $\rho = 1$. Из соображений симметрии заключаем, что $x_c = 0$. Далее, согласно формуле (2) $4\pi R^2 = \pi R 2\pi y_c$, откуда $y_c = \frac{2R}{\pi} \approx 0,637R$ (здесь можно использовать калькулятор).

Остановимся еще на статических моментах и координатах центра тяжести плоской фигуры. Пусть дана криволинейная трапеция, ограниченная линиями $x = a$, $x = b$, $y = f(x)$ ($f(x)$ непрерывна на $[a; b]$), $y = 0$, и на ней распределено вещество с плотностью $\rho = 1$. Разделим отрезок $[a; b]$ на n частичных отрезков, а криволинейную трапецию на n соответствующих частей. Заменяем каждую частичную трапецию прямоугольником с основанием Δx_k и высотой $y_{k-1} = f(x_{k-1})$. Тогда получим приближенные выражения элементов массы $\Delta m_k \approx y_{k-1} \Delta x_k$ и элементарных статических моментов относительно координатных осей $(\Delta M_x)_k \approx 1/2 y_{k-1} \Delta m_k$, $(M_y)_k \approx x_{k-1} \Delta m_k$. Суммируя и переходя к пределу при $\lambda \rightarrow 0$, получим выражения массы и статических моментов всей фигуры:

$$m = \int_a^b y dx, \quad M_x = \frac{1}{2} \int_a^b y^2 dx, \quad M_y = \frac{1}{2} \int_a^b x y dx. \quad (3)$$

Координаты центра тяжести x_c и y_c определяются так же, как и для материальной дуги, формулами (1), в которых m , M_x , M_y определяются по формулам (3).

Из равенства $y_c = M_x/m$ согласно формуле (3) имеем:

$$\frac{1}{2} \int_a^b y^2 dx = y_c \int_a^b y dx,$$

откуда

$$\pi \int_a^b y^2 dx = 2\pi y_c \int_a^b y dx.$$

Левая часть последнего равенства есть объем V тела, получаемого от вращения данной плоской фигуры вокруг оси Ox , правая — произведение площади S вращающейся фигуры на $2\pi y_c$ — длину ок-

ружности, описываемой при этом вращении центром тяжести фигуры. Итак,

$$V = S \cdot 2\pi y_c. \quad (4)$$

Это приводит ко второй теореме Гульдина.

Вторая теорема Гульдина. *Объем тела, образованного вращением данной плоской фигуры вокруг не пересекающей ее оси, лежащей в ее плоскости, равен произведению площади вращающейся фигуры на длину окружности, которую при этом вращении описывает центр тяжести этой фигуры.*

Пример 2. Найти центр тяжести полукруга, ограниченного осью Ox и полуокружностью $y = \sqrt{R^2 - x^2}$ (при условии $\rho = 1$). Из соображений симметрии следует, что $x_c = 0$. Далее, согласно формуле (4) $\frac{4}{3}\pi R^3 = \frac{\pi R^2}{2} \cdot 2\pi y_c$, откуда

$$y_c = \frac{4R}{3\pi} \approx 0,424R$$

(здесь также для вычислений можно использовать калькулятор).

§ 9.11. Вектор-функция скалярного аргумента

1. Векторная функция скалярного аргумента, ее график. До сих пор мы изучали функциональную зависимость между величинами, принимающими только числовые значения. Другими словами, если переменная величина x есть функция переменной величины t , т. е. $x = x(t)$, то как значения независимой переменной t , так и значения самой функции x суть числа. Таким образом, t и x являются в этом случае *скалярными величинами*, или просто *скалярами*; иными словами, мы имеем *скалярную функцию скалярного аргумента*. По аналогии со скалярной функцией можно ввести понятие векторной функции скалярного аргумента.

Определение. Если каждому значению параметра t из некоторого промежутка отвечает определенный вектор $\vec{r}$ (зависящий от t), то вектор $\vec{r}$ называется *векторной функцией* (кратко — *вектор-функция*) от скалярного аргумента t , и в этом случае пишут:

$$\vec{r} = \vec{r}(t). \quad (1)$$

При изменении аргумента t вектор $\vec{r}(t)$ изменяется, вообще говоря, как по величине, так и по направлению. В дальнейшем предполагается, что t изменяется в промежутке, конечном или бесконечном.

Как и в векторной алгебре (гл. 2), мы здесь рассматриваем свободные векторы, т. е. такие векторы, которые считаются равными, если

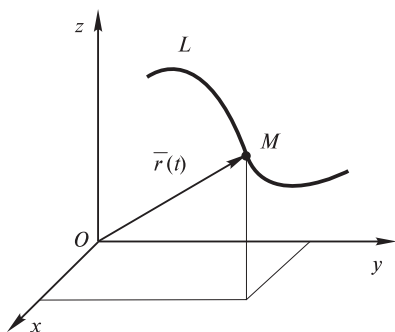


Рис. 124

они равны по величине и одинаково направлены, или, иначе говоря, если они имеют равные проекции на оси координат. Свободные векторы могут быть отложены от какой угодно точки. Будем считать, что вектор $\vec{r}(t)$ исходит из начала координат, т. е. вектор $\vec{r}$ — радиус-вектор некоторой точки M . В этом случае при изменении параметра t конец вектора $\vec{r}$ опишет некоторую линию L , называемую *годографом* векторной функции $\vec{r}(t)$

(рис. 124). При этом начало координат называют *полюсом* годографа. Уравнение (1) называют *векторным уравнением* этой кривой.

Если у вектора $\vec{r}(t)$ меняется только *модуль*, то годографом его будет луч, исходящий из полюса. Если модуль вектора $\vec{r}(t)$ постоянен ($|\vec{r}(t)| = \text{const}$) и меняется только его *направление*, то годограф есть линия, лежащая на сфере с центром в полюсе и радиусом, равным модулю вектора $\vec{r}(t)$.

С векторной функцией скалярного аргумента особенно часто приходится встречаться в кинематике при изучении движения точки, когда радиус-вектор $\vec{r}(t)$ движущейся точки является функцией времени t . Годографом такой функции является *траектория* движения точки. При этом уравнение $\vec{r} = \vec{r}(t)$ называют *уравнением движения*.

Если через x , y и z обозначить проекции вектора $\vec{r}(t)$ на оси прямоугольной системы координат в пространстве, то эти величины для каждого значения параметра t , в свою очередь, принимают определенные числовые значения и потому являются скалярными функциями скалярного аргумента t :

$$x = x(t), y = y(t), z = z(t). \quad (2)$$

В силу известного (§ 2.1, п. 8) разложения вектора по ортам прямоугольной системы координат имеем:

$$\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}.$$

Таким образом, задание векторной функции скалярного аргумента (т. е. функции (1)) равносильно заданию трех скалярных функций того же аргумента (т. е. функций (2)).

Так как уравнение (1) является уравнением некоторой кривой в пространстве, то ту же кривую задают и уравнения (2). Уравнения (2) — обычные *параметрические уравнения кривой в пространстве* (§ 4.2, п. 1).

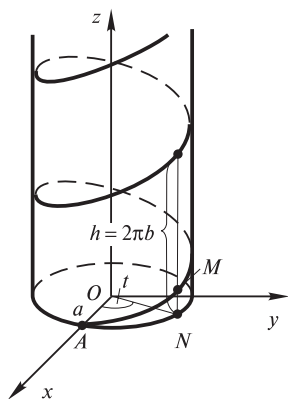
С помощью этих уравнений для каждого значения t определяются координаты x , y и z соответствующей точки кривой, а по координатам можно определить и радиус-вектор этой точки.

Рассмотрим кривую, заданную параметрически с помощью уравнений

$$x = a \cos t, y = a \sin t, z = bt.$$

Эта кривая называется *винтовой линией* (рис. 125). Ее векторное уравнение

$$\vec{r} = a \cos t \cdot \vec{i} + a \sin t \cdot \vec{j} + btk.$$



При любом значении параметра t

$$x^2 + y^2 = a^2(\cos^2 t + \sin^2 t) = a^2.$$

Это означает, что винтовая линия расположена на цилиндре $x^2 + y^2 = a^2$. Отсюда следует, что, когда точка M движется по винтовой линии, ее проекция N на плоскости xOy перемещается по окружности с радиусом a и центром в начале координат, причем t является полярным углом точки N . Когда точка N описывает полную окружность, аппликата z точки M винтовой линии увеличивается на $h = 2\pi b$. Эта величина называется *шагом* винтовой линии.

2. Предел, непрерывность и производная вектор-функции. Пусть функция $\vec{r}(t)$ определена в окрестности точки t_0 , кроме, быть может, самой точки t_0 .

Вектор $\vec{r}_0$ называется *пределом векторной функции* $\vec{r}(t)$ при стремлении t к t_0 (или, короче, в точке t_0), если

$$\lim_{t \rightarrow t_0} |\vec{r}(t) - \vec{r}_0| = 0. \quad (3)$$

Если $\vec{r}_0$ есть предел функции $\vec{r}(t)$ при $t \rightarrow t_0$, то это записывается так:

$$\lim_{t \rightarrow t_0} \vec{r}(t) = \vec{r}_0 \quad (4)$$

Итак, когда мы пишем соотношение (4), то подразумеваем под этим соотношением (3).

Если мы запишем векторную функцию $\vec{r}(t)$ и вектор $\vec{r}_0$ в проекциях (п. 1)

$$\begin{aligned} \vec{r}(t) &= x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}, \\ \vec{r}_0 &= x_0\vec{i} + y_0\vec{j} + z_0\vec{k}, \end{aligned}$$

Рис. 125

то получим:

$$\left| \vec{r}(t) - \vec{r}_0 \right| = \sqrt{(x(t) - x_0)^2 + (y(t) - y_0)^2 + (z(t) - z_0)^2}. \quad (5)$$

Тогда из условия (3) следует, что

$$\lim_{t \rightarrow t_0} x(t) = x_0, \lim_{t \rightarrow t_0} y(t) = y_0, \lim_{t \rightarrow t_0} z(t) = z_0. \quad (6)$$

Обратно: из формул (6) с помощью (5) сразу вытекает соотношение (3). Таким образом, соотношения (3) и (6) равносильны.

Из определения предела векторной функции с учетом очевидного неравенства $\left| \left| \vec{r}(t) \right| - \left| \vec{r}_0 \right| \right| \leq \left| \vec{r}(t) - \vec{r}_0 \right|$ непосредственно следует следующее свойство:

$$1. \text{ Если } \lim_{t \rightarrow t_0} \vec{r}(t) = \vec{r}_0, \text{ то и } \lim_{t \rightarrow t_0} \left| \vec{r}(t) \right| = \left| \vec{r}_0 \right|.$$

Отметим еще четыре свойства:

$$2. \lim_{t \rightarrow t_0} (\vec{r}_1(t) + \vec{r}_2(t)) = \lim_{t \rightarrow t_0} \vec{r}_1(t) + \lim_{t \rightarrow t_0} \vec{r}_2(t).$$

$$3. \lim_{t \rightarrow t_0} (f(t) \vec{r}(t)) = \lim_{t \rightarrow t_0} f(t) \lim_{t \rightarrow t_0} \vec{r}(t)$$

($f(t)$ — скалярная функция).

$$4. \lim_{t \rightarrow t_0} (\vec{r}_1(t) \vec{r}_2(t)) = \lim_{t \rightarrow t_0} \vec{r}_1(t) \lim_{t \rightarrow t_0} \vec{r}_2(t).$$

$$5. \lim_{t \rightarrow t_0} (\vec{r}_1(t) \times \vec{r}_2(t)) = \lim_{t \rightarrow t_0} \vec{r}_1(t) \times \lim_{t \rightarrow t_0} \vec{r}_2(t).$$

(Предполагается, что пределы в правых частях последних четырех равенств существуют.)

Свойства 2—5 с помощью формул (6) легко следуют из соответствующих свойств скалярных функций, если перейти к координатной записи векторов и их скалярных и векторных произведений.

Вектор-функция $\vec{r}(t)$, определенная в некоторой окрестности точки t_0 , называется *непрерывной* в точке t_0 , если $\lim_{t \rightarrow t_0} \vec{r}(t) = \vec{r}(t_0)$.

Из равносильности условий (3) и (6) следует, что, для того чтобы вектор-функция $\vec{r}(t)$ была непрерывной в точке t_0 , необходимо и достаточно, чтобы в этой точке были непрерывны функции x, y, z .

Из свойств пределов векторных функций следует, что сумма, скалярное и векторное произведения векторных функций, произведение скалярных функций на векторные будут непрерывны в точке t_0 , если в этой точке были непрерывны все слагаемые, соответственно сомножители. В дальнейшем мы будем рассматривать только непрерывные функции.

Перейдем теперь к вопросу о производной векторной функции скалярного аргумента

$$\vec{r}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}, \quad (7)$$

предполагая, что начало вектора $\vec{r}(t)$ находится в начале координат. Как уже отмечалось (п. 1), последнее уравнение является уравнением некоторой пространственной кривой.

Возьмем какое-нибудь фиксированное значение t , соответствующее какой-нибудь определенной точке M на кривой, заданной уравнением (7), и дадим t приращение Δt . Тогда получим вектор

$$\vec{r}(t + \Delta t) = x(t + \Delta t)\vec{i} + y(t + \Delta t)\vec{j} + z(t + \Delta t)\vec{k},$$

который определяет на кривой некоторую точку M' (рис. 126). Найдем приращение вектора:

$$\Delta \vec{r} = \vec{r}(t + \Delta t) - \vec{r}(t) = (x(t + \Delta t) - x(t))\vec{i} + (y(t + \Delta t) - y(t))\vec{j} + (z(t + \Delta t) - z(t))\vec{k}. \quad (8)$$

На рисунке 126, где $\overline{OM} = \vec{r}(t)$, $\overline{OM'} = \vec{r}(t + \Delta t)$, это приращение изображается вектором $\overline{MM'} = \Delta \vec{r}(t)$.

Рассмотрим отношение $\frac{\Delta \vec{r}(t)}{\Delta t}$ приращения вектор-функции к приращению скалярного аргумента; это есть вектор, коллинеарный с вектором $\Delta \vec{r}(t)$, так как получается из него умножением на скалярный множитель $\frac{1}{\Delta t}$. При этом вектор $\frac{\Delta \vec{r}(t)}{\Delta t}$ направлен в сторону, соответствующую возрастанию t (направление от M к M' на рисунке 126).

Действительно, если $\Delta t > 0$, то вектор $\frac{\Delta \vec{r}(t)}{\Delta t} = \overline{MQ'}$ имеет то же направление, что и вектор $\Delta \vec{r}(t) = \overline{MM'}$, т.е. направлен в ту сторону годографа, которая соответствует возрастанию параметра t . Если же $\Delta t < 0$, то вектор $\Delta \vec{r}(t) = \overline{MM''}$ направлен в противоположную сторону и при делении на отрицательное Δt мы получим вектор $\overline{MQ''}$, направленный снова в сторону возрастания t .

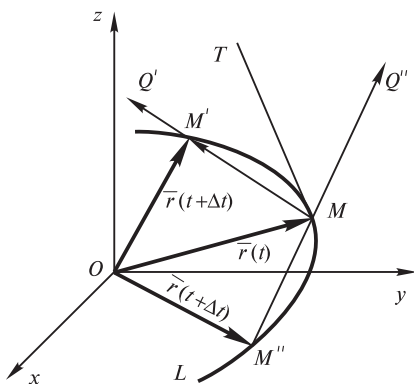


Рис. 126

Далее с учетом выражения (8) вектор $\frac{\Delta \bar{r}(t)}{\Delta t}$ можно представить в виде

$$\frac{\Delta \bar{r}(t)}{\Delta t} = \frac{x(t + \Delta t) - x(t)}{\Delta t} \bar{i} + \frac{y(t + \Delta t) - y(t)}{\Delta t} \bar{j} + \frac{z(t + \Delta t) - z(t)}{\Delta t} \bar{k}. \quad (9)$$

Если функции $x(t)$, $y(t)$, $z(t)$ имеют производные при выбранном значении t , то множители при $\bar{i}$, $\bar{j}$, $\bar{k}$ в равенстве (9) в пределе при $\Delta t \rightarrow 0$ обратятся в производные $x'(t)$, $y'(t)$, $z'(t)$. Следовательно, в этом случае существует предел

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta \bar{r}}{\Delta t} = x'(t) \bar{i} + y'(t) \bar{j} + z'(t) \bar{k}.$$

Вектор, определяемый последним равенством, называется *производной от вектора $\bar{r}(t)$ по скалярному аргументу t* , ее обозначают символом $\frac{d\bar{r}}{dt}$ или $\bar{r}'$. Итак,

$$\frac{d\bar{r}}{dt} = \bar{r}' = x'(t) \bar{i} + y'(t) \bar{j} + z'(t) \bar{k}, \quad (10)$$

или

$$\frac{d\bar{r}}{dt} = \frac{dx}{dt} \bar{i} + \frac{dy}{dt} \bar{j} + \frac{dz}{dt} \bar{k}.$$

Выясним направление вектора $\frac{d\bar{r}}{dt}$. Для этого заметим, что при $\Delta t \rightarrow 0$ точка $M'(M'')$ стремится к точке M и поэтому секущая $MM'(MM'')$ стремится к касательной в точке M . Отсюда производная $\bar{r}'(t)$ является вектором $\overline{MT}$, касательным к годографу вектор-функции $\bar{r}(t)$, направленным в сторону, соответствующую возрастанию параметра t .

Из разложения (10) следует, что

$$|\bar{r}'(t)| = \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}. \quad (11)$$

Но, как известно (§9.9, п. 3, примечание), дифференциал дуги

$$dl = \sqrt{x'^2(t) + y'^2(t) + z'^2(t)} dt,$$

откуда

$$\frac{dl}{dt} = \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}. \quad (12)$$

Из сопоставления формул (11) и (12) имеем:

$$|\bar{r}'(t)| = \frac{dl}{dt}. \quad (13)$$

Таким образом, модуль производной векторной функции $|\vec{r}'(t)|$ равен производной от длины годографа по аргументу t .

Если вектор-функция $\vec{r}(t)$ имеет постоянный модуль, но переменное направление, то ее производная $\frac{d\vec{r}}{dt}$ является вектором, перпендикулярным к вектору $\vec{r}(t)$. В самом деле, в этом случае годограф лежит на сфере, и поэтому производная $\vec{r}'(t)$ как вектор, касательный к годографу, перпендикулярна к вектору $\vec{r}(t)$.

Приведем правила дифференцирования вектор-функций (аргумент для краткости записи опустим):

$$1. (\vec{r}_1 + \vec{r}_2)' = \vec{r}'_1 + \vec{r}'_2.$$

$$2. (f\vec{r})' = f' \vec{r} + f \vec{r}'$$

(f — скалярная функция, в частности $(c\vec{r})' = c\vec{r}'$, где c — скалярная постоянная).

$$3. (\vec{r}_1 \vec{r}_2)' = \vec{r}'_1 \vec{r}_2 + \vec{r}_1 \vec{r}'_2.$$

$$4. (\vec{r}_1 \times \vec{r}_2)' = \vec{r}'_1 \times \vec{r}_2 + \vec{r}_1 \times \vec{r}'_2.$$

Все эти формулы доказываются аналогично формулам дифференцирования скалярных функций (см. § 8.2, п. 1), они могут быть получены и иначе — с использованием выражения (10).

Последовательным дифференцированием можно найти производные высших порядков от векторных функций. Так,

$$\vec{r}''(t) = x''(t)\vec{i} + y''(t)\vec{j} + z''(t)\vec{k} \text{ и т.д.}$$

3. Касательная, нормальная плоскость, главная нормаль. Пусть точка $M(x; y; z)$ описывает некоторую пространственную кривую L (рис. 127). Параметром, определяющим положение точки M на кривой, будем считать длину l дуги AM кривой, отсчитываемую от определенной точки A кривой до точки M . Так как положение точки $M(x; y; z)$ на кривой определяется значением дуги l , то координаты x , y и z , а также радиус-вектор $\vec{r}$ этой точки можно рассматривать как функции длины дуги l :

$$x = x(l), y = y(l), z = z(l), \vec{r} = \vec{r}(l),$$

или в проекциях:

$$\vec{r} = x(l)\vec{i} + y(l)\vec{j} + z(l)\vec{k}. \quad (14)$$

Продифференцируем вектор $\vec{r}(l)$ по l . Получим новый вектор

$$\vec{\tau} = \vec{r}'_l, \quad (15)$$

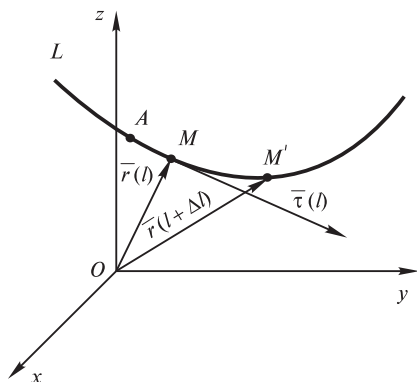


Рис. 127

направленный, как мы знаем (см. п. 2), по касательной к кривой в сторону возрастания параметра l (см. рис. 127). Модуль этого вектора

$$|\bar{\tau}| = |\bar{r}'| = \lim_{\Delta l \rightarrow 0} \left| \frac{\Delta \bar{r}}{\Delta l} \right|.$$

(Равенство $|\bar{r}'| = \lim_{\Delta l \rightarrow 0} \left| \frac{\Delta \bar{r}}{\Delta l} \right|$ записано на основе свойства 1 пределов (п. 2).)

Но, как ранее установлено (§9.9, п. 3), последний предел равен единице (он равен единице и в случае пространственной кривой).

Таким образом, вектор $\bar{\tau}$ есть единичный вектор, направленный по касательной к кривой в точке M в сторону возрастания l . Если вектор $\bar{r}$ задан в проекциях (14), то согласно формуле (10) имеем:

$$\bar{\tau} = \bar{r}' = x'_l \bar{i} + y'_l \bar{j} + z'_l \bar{k}.$$

На основе полученных результатов легко написать уравнения касательной к кривой L в точке $M(x; y; z)$. Так как вектор $\bar{\tau}$ на искомой касательной, то (§ 4.2, п. 3) ее уравнения будут

$$\frac{X - x}{x'_l} = \frac{Y - y}{y'_l} = \frac{Z - z}{z'_l}, \quad (16)$$

где X, Y, Z — текущие координаты точки касательной, а значения производных x', y', z' вычисляются в точке касания M . Если в уравнениях кривой, вместо длины дуги l взять любой другой параметр t , то, рассматривая длину дуги l как функцию параметра t , получим:

$$x'_l = x'_t \cdot l'_t, \quad y'_l = y'_t \cdot l'_t, \quad z'_l = z'_t \cdot l'_t.$$

Следовательно, в уравнениях (16) производные x'_l, y'_l, z'_l можно заменить производными x'_t, y'_t, z'_t соответственно. Тем самым приходим и к такой форме уравнений касательной:

$$\frac{X - x}{x'_t} = \frac{Y - y}{y'_t} = \frac{Z - z}{z'_t}. \quad (17)$$

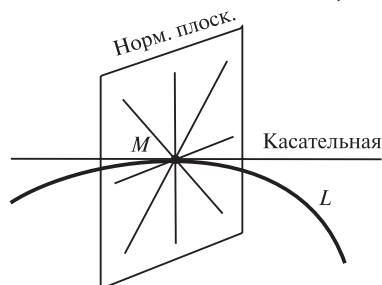


Рис. 128

Плоскость, перпендикулярная касательной (17) и проходящая через точку касания, называется *нормальной плоскостью* к кривой L в точке M (рис. 128). Всякая прямая, лежащая в этой плоскости и пересекающая кривую в точке касания M , называется *нормалью кривой L в точке M* , и, следовательно, пространственная кривая в каждой точке имеет

бесчисленное множество нормалей. Найдем уравнение нормальной плоскости к кривой L в точке $M(x; y; z)$.

Так как искомая плоскость перпендикулярна к касательной (17), то вектор (10) является нормальным этой плоскости и потому (§ 4.1, п. 2) ее уравнение имеет вид:

$$(X - x)x'_i + (Y - y)y'_i + (Z - z)z'_i = 0$$

(значения производных взяты в точке $M(x; y; z)$).

П р и м е р. Написать уравнения нормальной плоскости и касательной к винтовой линии

$$x = \cos t, y = \sin t, z = 3t$$

в точке P , для которой $t = \frac{\pi}{3}$.

Для любого t

$$x'_i = -\sin t, y'_i = \cos t, z'_i = 3,$$

поэтому при $t = \frac{\pi}{3}$ имеем:

$$x = \frac{1}{2}, y = \frac{\sqrt{3}}{2}, z = \pi,$$

$$x'_i = -\frac{\sqrt{3}}{2}, y'_i = \frac{1}{2}, z'_i = 3,$$

и, следовательно, для касательной получаем уравнения

$$\frac{X - \frac{1}{2}}{-\frac{\sqrt{3}}{2}} = \frac{Y - \frac{\sqrt{3}}{2}}{\frac{1}{2}} = \frac{Z - \pi}{3},$$

для нормальной плоскости уравнение

$$-\left(X - \frac{1}{2}\right)\sqrt{3} + \left(Y - \frac{\sqrt{3}}{2}\right) + (Z - \pi)6 = 0.$$

Рассмотрим теперь вторую производную $\bar{r}''_{i^2}$ векторной функции $\bar{r}(l)$. Из формулы (15) следует, что $\bar{r}''_{i^2} = [\bar{r}'_i]'_i = \bar{\tau}'_i$. Так как вектор $\bar{\tau}(l)$ единичный, то вектор производной $\bar{\tau}'_i$ перпендикулярен к нему (см. п. 2). Следовательно, вектор $\bar{\tau}'_i$ лежит в нормальной плоскости и определяет некоторую нормаль. Эту нормаль называют *главной*. Положительным направлением главной нормали будем считать то, которое совпадает с направлением вектора $\bar{\tau}'_i$. Итак, приходим к определению:

О п р е д е л е н и е. Главной нормалью кривой L в данной ее точке M называется прямая, проходящая через точку M и имеющая направление вектора $\bar{\tau}'_i$.

Так как

$$\bar{\tau}'_i = \bar{r}''_{i^2} = x''_{i^2}\bar{i} + y''_{i^2}\bar{j} + z''_{i^2}\bar{k},$$

то уравнения главной нормали кривой в точке $M(x; y; z)$ будут иметь вид:

$$\frac{X-x}{x''_{l^2}} = \frac{Y-y}{y''_{l^2}} = \frac{Z-z}{z''_{l^2}},$$

где X, Y, Z — текущие координаты точки на главной нормали.

4. Приложения к механике. Рассмотренное выше дифференцирование векторной функции скалярного аргумента имеет важное применение в механике при определении скорости и ускорения криволинейного движения.

Пусть годограф векторной функции $\vec{r} = \vec{r}(t)$ является траекторией движения точки, t — время. Тогда *вектор-производная*

$$\vec{v} = \frac{d\vec{r}}{dt}$$

называется *скоростью движения*.

Таким образом, *скорость движения есть вектор, касательный к траектории в соответствующей точке и направленный в сторону движения* (т. е. в сторону возрастания t).

Согласно формуле (13) *модуль скорости* равен:

$$|\vec{v}| = \left| \frac{d\vec{r}}{dt} \right| = \frac{dl}{dt},$$

т. е. равен *производной от пути по времени*.

Если движение прямолинейное, то скалярная величина $\frac{dl}{dt}$ вполне характеризует скорость движения (ее и называем *скоростью прямолинейного движения*).

Вектор

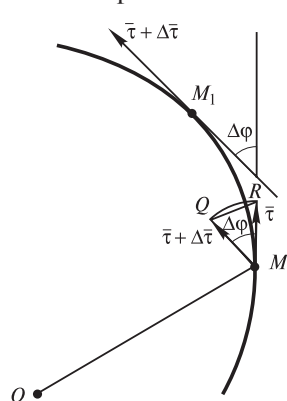


Рис. 129

$$\vec{w} = \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}}{dt^2}$$

называется *ускорением движения*.

Представим ускорение $\vec{w}$ в несколько ином виде. Пусть $\vec{\tau}$ — единичный вектор касательной к годографу в точке M (рис. 129). Тогда

$$\vec{v} = |\vec{v}| \vec{\tau} = \frac{dl}{dt} \vec{\tau}.$$

Дифференцируя это произведение (см. п. 2, правило 2), получим:

$$\vec{w} = \frac{d^2l}{dt^2} \vec{\tau} + \frac{dl}{dt} \cdot \frac{d\vec{\tau}}{dt}. \quad (18)$$

Рассмотрим теперь производную $\frac{d\bar{\tau}}{dt}$. Так как $\bar{\tau}$ — единичный вектор, то производная его будет к нему перпендикулярна, т. е. направлена по некоторой нормали к кривой (эту нормаль мы назвали главной нормалью (п. 3)). Обозначив единичный вектор этой нормали через $\bar{\nu}$, запишем:

$$\frac{d\bar{\tau}}{dt} = \left| \frac{d\bar{\tau}}{dt} \right| \bar{\nu}.$$

Из рисунка 129 ясно, что $|\Delta\bar{\tau}|$ есть хорда, стягивающая дугу окружности $QR = |\bar{\tau}|\Delta\phi$, где $\Delta\phi$ — угол смежности дуги MM_1 (см. § 9.9, п. 4). В силу эквивалентности бесконечно малой дуги и хорды (см. § 9.9, п. 3, теорема) можно записать:

$$\left| \frac{d\bar{\tau}}{dt} \right| = \left| \frac{d\phi}{dt} \right|.$$

Произведем еще следующее преобразование:

$$\left| \frac{d\phi}{dt} \right| = \left| \frac{d\phi}{dl} \cdot \frac{dl}{dt} \right| = \frac{dl}{dt} K,$$

где $K = \left| \frac{d\phi}{dl} \right|$. Эту величину так же, как и для плоской кривой, называют *кривизной* пространственной кривой в точке M , а величину $\rho = \frac{1}{K}$ — радиусом кривизны (формулу для кривизны пространственной кривой мы не выводим). Подставляя выражение для $\frac{d\bar{\tau}}{dt}$ в формулу (18), получим:

$$\bar{w} = \frac{d^2 l}{dt^2} \bar{\tau} + \frac{\left(\frac{dl}{dt} \right)^2}{\rho} \bar{\nu}.$$

Первый составляющий вектор $\bar{w}_t = \frac{d^2 l}{dt^2} \bar{\tau}$ называется *тангенциальным ускорением*; величина его равна второй производной от пути по времени, и он направлен по касательной к траектории движения. Вторым составляющий вектор $\bar{w}_n = \frac{1}{\rho} \left(\frac{dl}{dt} \right)^2 \bar{\nu} = \frac{|\bar{v}|^2}{\rho} \bar{\nu}$ называется *нормальным ускорением*; он направлен по главной нормали к траектории движения, величина его равна квадрату величины скорости деленному на радиус кривизны. Если движение *прямолинейное*, то вектор $\bar{\tau}$ постоянный, его производная равна нулю и ускорение полностью характеризуется величиной $\frac{d^2 l}{dt^2}$ — *ускорением прямолинейного движения*. Если дви-

жение *равномерное*, т. е. постоянна величина скорости $|\vec{v}| = \frac{dl}{dt}$, то первое слагаемое равно нулю и остается лишь нормальное ускорение $\vec{w}_n = \frac{|\vec{v}|^2}{\rho} \vec{v}$. В частности, когда точка равномерно движется по окружности радиуса R , то $\rho = R$ (§ 9.9, п. 4) и $|\vec{w}_n| = \frac{|\vec{v}|^2}{R}$. Это ускорение направлено к центру окружности и называется *центростремительным*.

Глава 10. РЯДЫ

§ 10.1. Числовые ряды

1. Основные понятия. Пусть дана бесконечная последовательность чисел $a_1, a_2, \dots, a_n, \dots$

О п р е д е л е н и е. Символ

$$a_1 + a_2 + a_n + \dots \quad (1)$$

называется *числовым рядом* или просто *рядом*, а числа $a_1, a_2, \dots, a_n, \dots$ называются *членами* ряда. Вместо (1), пользуясь знаком суммы, кратко пишут так:

$$\sum_{n=1}^{\infty} a_n.$$

Суммы конечного числа членов ряда (1) $S_1 = a_1$,

$$S_2 = a_1 + a_2, S_3 = a_1 + a_2 + a_3, \dots, S_n = a_1 + a_2 + a_3 + \dots + a_n, \dots$$

называются *частичными суммами* (или *отрезками*) ряда (1).

Рассмотрим последовательность

$$S_1, S_2, S_3, \dots, S_n, \dots \quad (2)$$

О п р е д е л е н и е. Если существует предел $S = \lim_{n \rightarrow \infty} S_n$, то ряд (1) называют *сходящимся*, а число S — *суммой* этого ряда. В этом случае пишут:

$$S = a_1 + a_2 + \dots + a_n + \dots = \sum_{n=1}^{\infty} a_n.$$

Если последовательность (2) не имеет предела, то ряд (1) называется *расходящимся*. Расходящийся ряд суммы не имеет.

Пример 1. Рассмотрим ряд, составленный из членов геометрической прогрессии $1, q, q^2, \dots, q^{n-1}, \dots$:

$$1 + q + q^2 + \dots + q^{n-1} + \dots \quad (3)$$

Если $q \neq 1$, то, как известно,

$$S_n = 1 + q + q^2 + \dots + q^{n-1} = \frac{q^{n-1} \cdot q - 1}{q - 1},$$

или

$$S_n = \frac{1 - q^n}{1 - q} = \frac{1}{1 - q} - \frac{q^n}{1 - q}.$$

При $|q| < 1 \lim_{n \rightarrow \infty} S_n = \frac{1}{1 - q}$, т. е. ряд (3) при $|q| < 1$ сходится и сумма его равна $\frac{1}{1 - q}$.

При $q = 1$ получаем ряд

$$1 + 1 + 1 + \dots + 1 + \dots$$

Следовательно, $S_n = n$ и $\lim_{n \rightarrow \infty} S_n = \infty$, т. е. ряд (3) при $q = 1$ расходится.

Пример 2. Исследовать на сходимость ряд

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)}. \quad (4)$$

Очевидно,

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1} \quad (n = 1, 2, 3, \dots).$$

Поэтому

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

Отсюда

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n+1}\right) = 1.$$

Следовательно, ряд (4) сходится и его сумма равна 1.

2. Основные свойства рядов. Если в ряде (1) отбросить конечное число первых членов, например m членов, то получим ряд

$$a_{m+1} + a_{m+2} + \dots + a_{m+k} + \dots, \quad (5)$$

который называется m -ым *остатком* ряда (1).

Теорема 1. Ряд (5) сходится (или расходится) одновременно с рядом (1).

Доказательство. Обозначим

$$S'_k = a_{m+1} + a_{m+2} + \dots + a_{m+k}.$$

Имеем:

$$S'_k = S_{m+k} - S_m. \quad (6)$$

Отсюда видно, что существование или отсутствие предела при $k \rightarrow \infty$ частичной суммы одного ряда влечет за собой существование или отсутствие предела частичной суммы другого ряда. Теорема доказана.

С л е д с т в и е 1. *При исследовании ряда на сходимость можно игнорировать конечное число его первых членов.*

Пусть ряд (1) сходится. Тогда согласно теореме 1 сходится и ряд (5), значит, существует его сумма. Обозначим ее через r_m . Перейдя к пределу в (6) при $k \rightarrow \infty$, получим:

$$r_m = S - S_m;$$

r_m есть та погрешность, которую мы допускаем, если вместо суммы S сходящегося ряда (1) берем сумму m первых его членов. Так как

$$\lim_{m \rightarrow \infty} r_m = \lim_{m \rightarrow \infty} (S - S_m) = S - S = 0, \quad (7)$$

то погрешность уменьшается с ростом m . Следовательно, абсолютная величина остатка

$$|r_m| = |S - S_m|$$

будет как угодно мала, если только число m взято достаточно большим. Таким образом, мы всегда имеем возможность подсчитать приближенно сумму сходящегося ряда, взяв достаточно большое число первых его членов.

Однако большую трудность представляет выяснение величины возникающей ошибки, т. е. оценки $|r_m|$. Задача состоит в том, чтобы по данному $\varepsilon > 0$ найти такое (наименьшее) m , чтобы выполнялось неравенство $|r_m| < \varepsilon$. В дальнейшем (см. п. 4) мы покажем, как иногда можно оценивать величину ошибки и тем самым устанавливать, сколько нужно брать первых членов ряда для получения его суммы с требуемой точностью.

Заметим, что полученное выше соотношение (7) выражает следующее предложение:

Т е о р е м а 2. *Предел суммы r_m m -го остатка сходящегося ряда (1) при $m \rightarrow \infty$ равен нулю.*

Т е о р е м а 3 (необходимый признак сходимости ряда). *Общий член a_n сходящегося ряда (1) стремится к нулю при неограниченном возрастании n , т. е.*

$$\lim_{n \rightarrow \infty} a_n = 0. \quad (8)$$

Д о к а з а т е л ь с т в о. Пусть ряд (1) сходится. Имеем $a_n = S_n - S_{n-1}$, откуда

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} S_n - \lim_{n \rightarrow \infty} S_{n-1} = S - S = 0.$$

С л е д с т в и е 2 . Если общий член a_n ряда (1) при $n \rightarrow \infty$ не стремится к нулю, то этот ряд расходится.

П р и м е р 1. Для ряда (3), у которого $|q| \geq 1$, имеем $|q|^{n-1} \geq 1$ для $n = 1, 2, \dots$, т. е. q^{n-1} не стремится к нулю при $n \rightarrow \infty$. Поэтому такой ряд расходится.

П р и м е ч а н и е 1. Отметим, что условие (8) не является достаточным для сходимости ряда. Действительно, для ряда

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots, \quad (9)$$

называемого *гармоническим* рядом,

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Однако этот ряд расходится, что можно установить рассуждениями от противного. Предположим, что ряд (9) сходится и его сумма равна S . Тогда

$$\lim_{n \rightarrow \infty} (S_{2n} - S_n) = \lim_{n \rightarrow \infty} S_{2n} - \lim_{n \rightarrow \infty} S_n = S - S = 0,$$

что противоречит неравенству

$$S_{2n} - S_n = \frac{1}{n+1} + \dots + \frac{1}{2n} > n \frac{1}{2n} = \frac{1}{2}.$$

Следовательно, гармонический ряд расходится.

Т е о р е м а 4. Если ряд (1) сходится и его сумма равна S , то ряд

$$\sum_{n=1}^{\infty} c a_n, \quad (10)$$

где c — произвольное число, также сходится и его сумма равна cS .

Д о к а з а т е л ь с т в о. Пусть S_n и σ_n — частичные суммы соответственно рядов (1) и (10). Тогда

$$\sigma_n = c a_1 + c a_2 + \dots + c a_n = c (a_1 + a_2 + \dots + a_n) = c S_n.$$

Отсюда

$$\lim_{n \rightarrow \infty} \sigma_n = \lim_{n \rightarrow \infty} c S_n = c \lim_{n \rightarrow \infty} S_n = c S.$$

Теорема доказана.

П р и м е р 2. С учетом примера 1 (п. 1) на основании теоремы 4 заключаем, что при $|q| < 1$ ряд

$$c + c q + c q^2 + \dots + c q^{n-1} + \dots,$$

где c — произвольное число, сходится.

Т е о р е м а 5. Если ряд (1) и ряд

$$b_1 + b_2 + \dots + b_n + \dots \quad (11)$$

сходятся и их суммы соответственно равны S и σ , то и ряд

$$\sum_{n=1}^{\infty} (a_n + b_n) \quad (12)$$

сходится и его сумма равна $S + \sigma$.

Д о к а з а т е л ь с т в о. Пусть S_n , σ_n и τ_n — частичные суммы соответственно рядов (1), (11) и (12). Тогда

$$\begin{aligned} \tau_n &= (a_1 + b_1) + (a_2 + b_2) + \dots + (a_n + b_n) = \\ &= (a_1 + a_2 + \dots + a_n) + (b_1 + b_2 + \dots + b_n) = S_n + \sigma_n. \end{aligned}$$

Отсюда с учетом теоремы 4 о пределах (§ 7.5, п. 1)

$$\lim_{n \rightarrow \infty} \tau_n = \lim_{n \rightarrow \infty} (S_n + \sigma_n) = \lim_{n \rightarrow \infty} S_n + \lim_{n \rightarrow \infty} \sigma_n = S + \sigma.$$

П р и м е ч а н и е 2. При условии теоремы (5) ряд

$$\sum_{n=1}^{\infty} (a_n - b_n)$$

также сходится и его сумма равна $S - \sigma$.

Наконец, заметим (см. [11]), что если ряд (1) сходится и имеет сумму S , то члены его можно любым образом сгруппировать скобками (однако не переставляя их), например, так:

$$a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6) + \dots,$$

образуя новый ряд, члены которого равны суммам чисел, стоящих в скобках. Новый ряд будет сходящимся и притом к S .

Однако раскрытие скобок в сходящемся ряде не всегда возможно. Так, ряд

$$(1 - 1) + (1 - 1) + \dots + (1 - 1) + \dots,$$

сходится, а ряд

$$1 - 1 + 1 - 1 + \dots + 1 - 1 + \dots \quad (13)$$

расходится (общий член ряда (13) не стремится к нулю).

3. Положительные ряды.

О п р е д е л е н и е. *Положительным рядом* называется ряд, члены которого неотрицательны.

Пусть дан положительный ряд

$$a_1 + a_2 + \dots + a_n + \dots, \quad (14)$$

т. е. $a_n \geq 0$ ($n = 1, 2, \dots$). Тогда очевидно,

$$S_{n+1} = S_n + a_{n+1} \geq S_n \quad (n = 1, 2, \dots),$$

т. е. последовательность $S_1, S_2, \dots, S_n, \dots$ является неубывающей.

Это с учетом приведенного в § 7.3 (п. 1) свойства 3 позволяет сформулировать следующее утверждение:

Т е о р е м а 1. *Для того чтобы положительный ряд (14) сходиллся, необходимо и достаточно, чтобы последовательность частичных сумм этого ряда была ограничена сверху.*

Т е о р е м а 2 (признак сравнения рядов). *Пусть даны два положительных ряда*

$$b_1 + b_2 + \dots + b_n + \dots, \quad (B)$$

$$c_1 + c_2 + \dots + c_n + \dots \quad (C)$$

Если члены ряда (B) не превосходят соответствующих членов ряда (C):

$$b_n \leq c_n \quad (n=1, 2, \dots), \quad (15)$$

то из сходимости ряда (C) следует сходимость ряда (B), а из расходимости ряда (B) следует расходимость ряда (C).

Д о к а з а т е л ь с т в о. Обозначив через B_n и C_n соответственно частичные суммы рядов (B) и (C), в силу неравенства (15) будем иметь:

$$B_n \leq C_n, \quad n=1, 2, 3, \dots \quad (16)$$

Если ряд (C) сходится, то по теореме 1 частичные суммы C_n ограничены сверху:

$$C_n \leq L \quad (L = \text{const}), \quad n=1, 2, 3, \dots \quad (17)$$

Из неравенств (16) и (17) имеем: $B_n \leq L, n=1, 2, 3, \dots$, и, значит, согласно той же теореме 1 ряд (B) сходится.

Пусть теперь ряд (B) расходится. Тогда расходится и ряд (C). В противном случае согласно только что доказанному сходил бы и ряд (B).

П р и м е ч а н и е 1. Ввиду следствия 1 (п. 2) теорема 2 остается справедливой, если условие (15) выполняется не для всех n , а лишь начиная с некоторого n .

П р и м е р 1. Исследовать на сходимость ряд

$$\sum_{n=1}^{\infty} \frac{1}{(n+1)^2}. \quad (18)$$

Сравниваем данный ряд со сходящимся рядом (см. п. 1) $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$. Имеем:

$$\frac{1}{(n+1)^2} < \frac{1}{n(n+1)} \quad (n=1, 2, 3, \dots).$$

Отсюда согласно теореме 2 получаем, что ряд (18) сходится. Попутно отметим, что тогда в силу следствия 1 (п. 2) сходится и ряд

$$\sum_{n=1}^{\infty} \frac{1}{n^2}. \quad (19)$$

Т е о р е м а 3 (признак Даламбера, Жан Лерон Даламбер (1717—1783) — французский математик). *Если члены положительного ряда (14) таковы, что существует предел*

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \rho,$$

то при $\rho < 1$ ряд (14) сходится, а при $\rho > 1$ ряд (14) расходится.

Д о к а з а т е л ь с т в о. В силу определения предела последовательности для любого $\varepsilon > 0$ существует такое натуральное число $N = N(\varepsilon)$, что для всех $n > N$ выполняется неравенство

$$\rho - \varepsilon < \frac{a_{n+1}}{a_n} < \rho + \varepsilon. \quad (20)$$

Если $\rho < 1$, то выберем ε столь малым, чтобы $\rho + \varepsilon$ было меньше единицы. Полагая $\rho + \varepsilon = q$, на основании правого неравенства (20) имеем:

$$\frac{a_{n+1}}{a_n} < q, \text{ или } a_{n+1} < a_n q,$$

для $n = N + 1, N + 2, \dots$. Давая n эти значения, из последнего неравенства получаем:

$$\begin{aligned} a_{N+2} &< a_{N+1} q, \\ a_{N+3} &< a_{N+2} q < a_{N+1} q^2, \\ a_{N+4} &< a_{N+3} q < a_{N+1} q^3, \\ &\dots \end{aligned}$$

т. е. члены ряда

$$a_{N+2} + a_{N+3} + a_{N+4} + \dots \quad (21)$$

меньше соответствующих членов сходящегося ряда $a_{N+1} q + a_{N+1} q^2 + a_{N+1} q^3 + \dots$ (см. п. 2, пример 2). Тогда по признаку сравнения ряд (21) сходится, и, следовательно, согласно теореме 1 (п. 2) сходится и ряд (14).

Пусть теперь $\rho > 1$. Возьмем ε столь малым, чтобы $\rho - \varepsilon > 1$. Тогда при $n > N$ в силу левой части неравенства (20) будет выполняться неравенство $\frac{a_{n+1}}{a_n} > 1$, или $a_{n+1} > a_n$. Таким образом, члены ряда (14), начиная с номера $n = N + 1$, возрастают с увеличением их номеров, т. е. общий член ряда (14) a_n не стремится к нулю при $n \rightarrow \infty$. Следовательно, согласно следствию 2 (п. 2) ряд (14) расходится.

Примечание 2. При $\rho = 1$ признак Даламбера на вопрос о том, сходится или расходится ряд, ответа не дает. В самом деле, для гармонического ряда $\rho = 1$, причем этот ряд расходится (см. п. 2, примечание 1). Вместе с тем для ряда (19) также $\rho = 1$, но этот ряд сходится (см. пример 1).

Примечание 3. Из доказательства признака Даламбера следует, что при $\rho > 1$ общий член a_n ряда (14) не стремится к нулю при $n \rightarrow \infty$.

Примечание 4. Ряд (14) будет расходиться и в том случае, когда $\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \infty$, так как тогда, начиная с некоторого номера $n = N$, будет $\frac{a_{n+1}}{a_n} > 1$, и, значит, a_n не стремится к нулю при $n \rightarrow \infty$.

Пример 2. Ряд $\sum_{n=1}^{\infty} \frac{1}{n!}$ сходится, так как

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!} = \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0 < 1.$$

Пример 3. Ряд $\sum_{n=1}^{\infty} \frac{n^n}{n!}$ расходится, так как

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \lim_{n \rightarrow \infty} \frac{(n+1)^{n+1} n!}{(n+1)! n^n} = \lim_{n \rightarrow \infty} \left(\frac{n+1}{n} \right)^n = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^n = e > 1.$$

Теорема 4 (интегральный признак Коши). Пусть члены положительного ряда (14) такие, что

$$a_1 = f(1), a_2 = f(2), \dots, a_n = f(n), \dots,$$

где функция $f(x)$ при $x \geq 1$ непрерывна, положительна и убывает. Тогда ряд (14) и несобственный интеграл

$$\int_1^{+\infty} f(x) dx \quad (22)$$

сходятся или расходятся одновременно.

Доказательство. Из рисунка 130 (имея в виду геометрический смысл определенного интеграла) видно, что

$$I_n = \int_1^{n+1} f(x) dx < f(1) \cdot 1 + f(2) \cdot 1 + \dots + f(n) \cdot 1 = a_1 + a_2 + \dots + a_n = S_n.$$

Отсюда, учитывая, что последовательности $\{I_n\}$ и $\{S_n\}$ монотонно возрастающие ($f(x) > 0$ и $a_n > 0, n = 1, 2, \dots$), с учетом теоремы 1 и свойства 3 из § 7.3 (п. 1) следует, что если ряд (14) сходится, то сходится и интеграл (22).

Точно так же из рисунка 131 ясно, что $S_{n+1} - a_1 < I_n$, откуда, подобно предыдущему, следует, что если интеграл (22) сходится, то сходится и ряд (14).

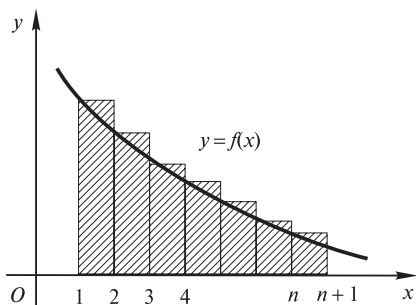


Рис. 130

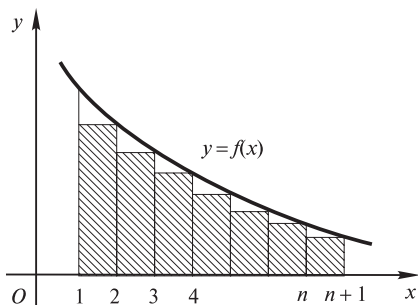


Рис. 131

Наконец, если ряд (14) расходится, то расходится и интеграл (22), ибо в противном случае в силу только что доказанного сошелся бы ряд (14). Аналогично если расходится интеграл (22), то расходится и ряд (14).

П р и м е ч а н и е 5. В интеграле (22) в качестве нижнего предела может быть фиксированное по произволу натуральное число $k > 1$. Это равносильно отбрасыванию $k - 1$ первых членов ряда (14), что на сходимость этого ряда не влияет (см. п. 2, следствие 1).

П р и м е р 4. Рассмотрим ряд

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}} \quad (\alpha > 0). \quad (23)$$

Функция $f(x) = \frac{1}{x^{\alpha}}$, где $x \geq 1$, удовлетворяет условиям теоремы 4. Члены ряда (23) равны значениям этой функции при $x = 1, 2, 3, \dots$. Как установлено ранее (см. § 9.8, п. 1), несобственный интеграл $\int_1^{+\infty} \frac{dx}{x^{\alpha}}$ при $\alpha > 1$ сходится, а при $\alpha \leq 1$ расходится. Следовательно, данный ряд сходится при $\alpha > 1$ и расходится при $0 < \alpha \leq 1$.

В частности, при $\alpha = 2$ получим сходящийся ряд (19), при $\alpha = 1$ — расходящийся гармонический ряд (9) (см. п. 2), при $\alpha = \frac{1}{2}$ — расходящийся ряд $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ и т. д.

4. Знакопередающие ряды. *Знакопередающим рядом* называется ряд вида

$$a_1 - a_2 + a_3 - a_4 + \dots + (-1)^{n-1} a_n + \dots, \quad (24)$$

где

$$a_n > 0 \quad (n = 1, 2, \dots).$$

Т е о р е м а 1 (теорема Лейбница). *Если члены ряда (24) по абсолютной величине монотонно убывают:*

$$a_{n+1} < a_n \quad (n = 1, 2, \dots) \quad (25)$$

и общий член стремится к нулю:

$$\lim_{n \rightarrow \infty} a_n = 0, \quad (26)$$

то ряд (24) сходится.

Д о к а з а т е л ь с т в о. Частичную сумму S_{2m} можно представить двояко:

$$S_{2m} = (a_1 - a_2) + (a_3 - a_4) + \dots + (a_{2m-1} - a_{2m}), \quad (27)$$

$$S_{2m} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \dots - (a_{2m-2} - a_{2m-1}) - a_{2m}. \quad (28)$$

Здесь в каждой круглой скобке разность положительна в силу условия (25). Из (27) следует, что $S_{2m} > 0$ и последовательность $\{S_{2m}\}$ монотонно возрастающая. Из (28) видно, что $S_{2m} < a_1$ т. е. $\{S_{2m}\}$ ограничена. Следовательно (§ 7.3, п. 1, свойство 3), эта последовательность имеет предел:

$$\lim_{m \rightarrow \infty} S_{2m} = S, \quad (29)$$

причем

$$0 < S < a_1.$$

Далее с учетом (29) и (26) имеем:

$$\lim_{m \rightarrow \infty} S_{2m+1} = \lim_{m \rightarrow \infty} S_{2m} + \lim_{m \rightarrow \infty} a_{2m+1} = S. \quad (30)$$

Из (29) и (30) следует, что $\lim_{n \rightarrow \infty} S_n = S$, т. е. ряд (24) сходится, причем

$$0 < S < a_1. \quad (31)$$

П р и м е р 1. Ряд

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + (-1)^{n-1} \frac{1}{n} + \dots$$

сходится, так как условия теоремы Лейбница здесь выполнены.

Т е о р е м а 2. *Остаток r_n знакопеременующегося ряда (24), удовлетворяющего условиям теоремы Лейбница, имеет знак своего первого члена и меньше его по абсолютной величине.*

Д о к а з а т е л ь с т в о. Если n четное, то

$$r_n = a_{n+1} - a_{n+2} + \dots$$

Так как этот ряд удовлетворяет теореме Лейбница, то согласно неравенству (31)

$$0 < r_n < a_{n+1}.$$

Если n нечетное, то

$$r_n = -a_{n+1} + a_{n+2} - \dots$$

Отсюда

$$-r_n = a_{n+1} - a_{n+2} + \dots,$$

и согласно (31)

$$0 < -r_n < a_{n+1}.$$

Отсюда

$$r_n < 0 \text{ и } |r_n| < a_{n+1}.$$

Теорема доказана.

П р и м е р 2. Вычислить с точностью до 0,1 сумму сходящегося ряда

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{(-1)^{n-1}}{n} + \dots \quad (32)$$

В качестве приближенного значения S ряда (32) мы должны взять ту частичную сумму S_n , для которой $|r_n| < 0,1$. Согласно теореме 2 $|r_n| < \frac{1}{n+1}$. Следовательно, достаточно положить $n+1 = 10$, т. е. $n = 9$. Тогда

$$S_9 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} \approx 0,74.$$

Отсюда $S \approx 0,7$ с точностью до 0,1.

5. Абсолютная и условная сходимость. Перейдем теперь к рядам с членами, имеющими любой знак. С каждым таким рядом

$$\sum_{n=1}^{\infty} a_n \quad (33)$$

связан ряд с неотрицательными членами, составленный из модулей членов данного ряда, т. е. ряд

$$\sum_{n=1}^{\infty} |a_n|. \quad (34)$$

Имеет место следующая теорема:

Т е о р е м а 1. Если сходится ряд (34), то сходится и ряд (33).

Д о к а з а т е л ь с т в о. Составим два положительных ряда

$$\sum_{n=1}^{\infty} p_n(p) \quad \text{и} \quad \sum_{n=1}^{\infty} q_n(q),$$

где

$$p_n = \begin{cases} a_n, & \text{если } a_n > 0, \\ 0, & \text{если } a_n \leq 0, \end{cases} \quad q_n = \begin{cases} 0, & \text{если } a_n \geq 0, \\ |a_n|, & \text{если } a_n < 0. \end{cases}$$

Так как при $n = 1, 2, \dots$ $p_n \leq |a_n|$ и $q_n \leq |a_n|$, то по теореме 2 (п. 3) ряды (p) и (q) сходятся. Обозначим их суммы соответственно через P и Q . Частичную сумму S_n данного ряда (33) можно с помощью обозначений для p_n и q_n переписать в виде

$$S_n = (p_1 + p_2 + \dots + p_n) - (q_1 + q_2 + \dots + q_n).$$

Переходя здесь к пределу при $n \rightarrow \infty$, получим:

$$\lim_{n \rightarrow \infty} S_n = P - Q. \quad (35)$$

Следовательно, ряд (33) сходится.

О п р е д е л е н и е. Ряд (33) называется *абсолютно сходящимся*, если ряд (34) сходится. Если же ряд (33) сходится, а ряд (34) расходится, то ряд (33) называется *неабсолютно сходящимся* или *условно сходящимся*.

П р и м е ч а н и е. Из равенства (35) имеем: $\left| \lim_{n \rightarrow \infty} S_n \right| \leq P + Q$. Но

$$P + Q = \sum_{n=1}^{\infty} p_n + \sum_{n=1}^{\infty} q_n = \sum_{n=1}^{\infty} |a_n|.$$

Следовательно, для абсолютно сходящегося ряда (33) имеет место неравенство

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n|.$$

П р и м е р. Ряд $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^\alpha}$ (36)

абсолютно сходится при $\alpha > 1$ и условно сходится при $0 < \alpha \leq 1$.

В самом деле, как установлено ранее (см. п. 3, пример 4), если $\alpha > 1$, то ряд (23) сходится, если же $0 < \alpha \leq 1$, то ряд (23) расходится, хотя сам ряд (36) сходится по теореме Лейбница.

П е р е м е с т и т е л ь н о е с в о й с т в о а б с о л ю т н о с х о д я щ и х с я р я д о в.

Пусть дан сходящийся ряд

$$S = a_1 + a_2 + \dots + a_n + \dots \quad (37)$$

Произвольным образом переставим в нем местами члены ряда, образовав ряд

$$a'_1 + a'_2 + \dots + a'_k + \dots, \quad (38)$$

причем члены ряда (38) связаны с членами ряда (37) следующим образом:

$$a'_1 = a_{n_1}, \quad a'_2 = a_{n_2}, \quad \dots, \quad a'_k = a_{n_k}, \quad \dots.$$

Т е о р е м а 2. Если ряд (37) абсолютно сходится, то ряд (38) также абсолютно сходится и сумма его равна сумме ряда (37), т. е. абсолютно сходящийся ряд (37) обладает переместительным свойством.

Д о к а з а т е л ь с т в о см., например, в [11].

Условно сходящиеся ряды переместительным свойством не обладают.

Абсолютно сходящиеся ряды обладают еще одним важным свойством: их можно перемножать.

Пусть даны два абсолютно сходящихся ряда

$$\begin{aligned} A &= a_1 + a_2 + a_3 + \dots + a_n + \dots, \\ B &= b_1 + b_2 + b_3 + \dots + b_n + \dots \end{aligned} \quad (39)$$

Образуем новый ряд:

$$a_1 b_1 + (a_1 b_2 + a_2 b_1) + (a_1 b_3 + a_2 b_2 + a_3 b_1) + \dots, \quad (40)$$

называемый *произведением двух рядов* (39). Оказывается (см. [9], т. 1), что ряд (40) также сходится абсолютно и имеет своей суммой число $S = AB$.

§ 10.2. Функциональные ряды

1. Область сходимости функционального ряда. Перейдем к рассмотрению рядов, членами которых являются не числа, а функции:

$$u_1(x) + u_2(x) + \dots + u_n(x) + \dots \quad (1)$$

Такие ряды называются *функциональными*.

Например, ряд

$$1 + x + x^2 + \dots + x^n + \dots$$

является функциональным.

Если в ряде (1) придать x какое-либо значение x_0 из области определения функций $u_n(x)$, $n = 1, 2, \dots$, то получим числовой ряд

$$u_1(x_0) + u_2(x_0) + \dots + u_n(x_0) + \dots \quad (2)$$

Этот ряд может сходиться или расходиться. Если он сходится, то точка x_0 называется *точкой сходимости* функционального ряда (1). Если ряд (2) расходится, то точка x_0 называется *точкой расходимости* функционального ряда (1). Для одних точек, взятых из области определения функций $u_n(x)$, ряд (1) может сходиться, а для других — расходиться.

О п р е д е л е н и е. Совокупность всех точек сходимости функционального ряда называется *областью его сходимости*.

Частичная сумма функционального ряда (1), т. е. сумма n первых его членов $S_n(x) = u_1(x) + u_2(x) + \dots + u_n(x)$, является функцией переменной x .

Из определения области сходимости функционального ряда следует, что для любой точки x этой области существует предел частичной суммы $S_n(x)$ при $n \rightarrow \infty$. В точках, не принадлежащих области сходимости, частичная сумма $S_n(x)$ не имеет предела. Ясно, что сумма $S(x)$ функционального ряда (1) является некоторой функцией от x , определенной в области сходимости ряда (1).

В этом случае пишут: $S(x) = u_1(x) + u_2(x) + \dots + u_n(x) + \dots$

Если функциональный ряд (1) сходится и имеет сумму $S(x)$, то разность $S(x) - S_n(x)$ называется, как и для числовых рядов, его n -м остатком. Этот остаток будем обозначать через $r_n(x)$: $r_n(x) = S(x) - S_n(x)$. Ясно, что

$$\lim_{n \rightarrow \infty} r_n(x) = 0.$$

П р и м е р. Ряд $\sum_{n=1}^{\infty} \left(\frac{x}{3}\right)^n$ сходится в промежутке $-3 < x < 3$, так как в этом промежутке $q = \left|\frac{x}{3}\right| < 1$. При $|x| \geq 3$ данный ряд расходится.

2. Равномерная сходимость функционального ряда.

О п р е д е л е н и е. Функциональный ряд (1) называется *равномерно сходящимся* на отрезке $[a; b]$, если для любого $\varepsilon > 0$ существует такое натуральное число $N = N(\varepsilon)$, не зависящее от x , что для $n > N$ неравенство $|S(x) - S_n(x)| < \varepsilon$ выполняется для всех x из отрезка $[a; b]$.

Т е о р е м а 1 (признак Вейерштрасса). *Если члены функционального ряда (1) удовлетворяют на отрезке $[a; b]$ условию*

$$|u_n(x)| \leq a_n \quad (n = 1, 2, \dots), \quad (3)$$

где a_n — члены сходящегося положительного ряда

$$\sum_{n=1}^{\infty} a_n, \quad (4)$$

то ряд (1) сходится равномерно (и абсолютно) на отрезке $[a; b]$.

Д о к а з а т е л ь с т в о. Возьмем произвольное число $\varepsilon > 0$. Так как ряд (4) сходится, то для этого ряда (см. § 10.1, п. 2, теорема 2) $\lim_{n \rightarrow \infty} r_n = 0$ (r_n — сумма n -го остатка ряда (4)). Поэтому существует такое натуральное число $N = N(\varepsilon)$, что при $n > N$ имеет место неравенство

$$\sum_{k=n+1}^{\infty} a_k < \varepsilon. \quad (5)$$

Пользуясь неравенством (3), получаем, что при любом натуральном p

$$\sum_{k=n+1}^{n+p} |u_k(x)| \leq \sum_{k=n+1}^{n+p} a_k, \quad (6)$$

откуда при $p \rightarrow \infty$ с учетом (5) следует, что

$$\sum_{k=n+1}^{\infty} |u_k(x)| \leq \sum_{k=n+1}^{\infty} a_k < \varepsilon \quad (n > N) \quad (7)$$

(предел при $p \rightarrow \infty$ у левой части неравенства (6) существует в силу теоремы 1 (§ 10.1, п. 3)). Это доказывает абсолютную сходимость ряда (1). Тогда (§ 10.1, п. 5, примечание) с учетом неравенства (7) имеем:

$$|r_n(x)| = \left| \sum_{k=n+1}^{\infty} u_k(x) \right| \leq \sum_{k=n+1}^{\infty} |u_k(x)| < \varepsilon$$

для $n > N$ и для всех x из $[a; b]$, т. е. ряд (1) равномерно сходится на $[a; b]$.

П р и м е р. Ряд $\sum_{n=1}^{\infty} \frac{x^n}{n^2}$ сходится равномерно на отрезке $[-1; 1]$, так как на этом отрезке $\left| \frac{x^n}{n^2} \right| \leq \frac{1}{n^2}$ ($n = 1, 2, \dots$) и ряд $\sum_{n=1}^{\infty} \frac{1}{n^2}$ сходится (см. § 10.1, п. 3, пример 1).

3. Свойства равномерно сходящихся рядов. Приведем без доказательства (доказательства см., например, в [15], т. I) некоторые теоремы о свойствах равномерно сходящихся рядов.

Т е о р е м а 1. Если члены равномерно сходящегося на отрезке $[a; b]$ функционального ряда (1) непрерывны на нем, то его сумма также непрерывна на отрезке $[a; b]$.

Т е о р е м а 2. Если члены равномерно сходящегося на отрезке $[a; b]$ функционального ряда (1) непрерывны на этом отрезке, то ряд (1) можно на отрезке $[a; b]$ почленно интегрировать.

Это значит, что если x_1 и x_2 — любые две точки отрезка $[a; b]$, то

$$\begin{aligned} & \int_{x_1}^{x_2} (u_1(x) + u_2(x) + \dots + u_n(x) + \dots) dx = \\ &= \int_{x_1}^{x_2} u_1(x) dx + \int_{x_1}^{x_2} u_2(x) dx + \dots + \int_{x_1}^{x_2} u_n(x) dx + \dots \end{aligned}$$

Т е о р е м а 3. Пусть функциональный ряд (1) сходится на отрезке $[a; b]$ и его члены имеют на нем непрерывные производные $u'_n(x)$ ($n = 1, 2, \dots$). Тогда если ряд $\sum_{n=1}^{\infty} u'_n(x)$ сходится равномерно на отрезке $[a; b]$, то на этом отрезке

$$(u_1(x) + u_2(x) + \dots + u_n(x) + \dots)' = u'_1(x) + u'_2(x) + \dots + u'_n(x) + \dots$$

Т е о р е м а 4. Если все члены равномерно сходящегося на отрезке $[a; b]$ ряда (1) умножить на ограниченную на этом отрезке функцию $\varphi(x)$, то полученный ряд

$$\varphi(x)u_1(x) + \varphi(x)u_2(x) + \dots + \varphi(x)u_n(x) + \dots$$

также будет равномерно сходящимся на отрезке $[a; b]$.

§ 10.3. Степенные ряды в действительной области

1. Степенной ряд и его область сходимости.

О п р е д е л е н и е. Степенным рядом называется функциональный ряд вида

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots, \quad (1)$$

где $a_0, a_1, a_2, \dots, a_n, \dots$ — постоянные вещественные числа, называемые *коэффициентами* степенного ряда.

Любой степенной ряд (1) сходится при $x=0$, так как в этой точке все члены ряда (1), кроме первого, — нули. Есть степенные ряды вида (1), которые сходятся лишь в точке $x=0$; такие ряды относят к рядам *первого класса*.

Например, ряд

$$1 + x + 2!x^2 + \dots + n!x^n + \dots \quad (2)$$

сходится лишь в точке $x=0$; в любой другой точке $x \neq 0$ этот ряд расходится. Действительно, при каждом $x \neq 0$ из числовой оси имеем числовой ряд. Исследуем его на сходимости. Образует ряд

$$1 + |x| + 2!|x|^2 + \dots + n!|x|^n + \dots \quad (3)$$

Применив к последнему ряду признак Даламбера (§ 10.1, п. 3), получим:

$$\lim_{n \rightarrow \infty} \frac{(n+1)!|x|^{n+1}}{n!|x|^n} = |x| \lim_{n \rightarrow \infty} (n+1) = \infty$$

при всех $x \neq 0$. Следовательно (см. § 10.1, п. 3, примечание 4), ряд (3) и ряд (2) расходятся при всех $x \neq 0$.

Есть степенные ряды вида (1), которые сходятся на всей числовой оси (такие ряды относят к рядам *второго класса*). Например, ряд

$$1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

сходится (притом абсолютно) на всей числовой оси, так как при каждом x из $(-\infty; +\infty)$

$$\lim_{n \rightarrow \infty} \frac{n!|x|^{n+1}}{(n+1)!|x|^n} = |x| \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0 < 1.$$

Ряды вида (1), не принадлежащие первому и второму классам, относят к рядам *третьего класса*.

Т е о р е м а 1 (теорема Абеля) (Нильс Хенрик Абель (1802 — 1829) — норвежский математик). Если степенной ряд (1) сходится при $x = x_0 \neq 0$, то он абсолютно сходится для любого x , удовлетворяющего условию

$$|x| < |x_0|;$$

если же при $x = x_0$ степенной ряд (1) расходится, то он расходится и при любом x , удовлетворяющем условию $|x| > |x_0|$.

Д о к а з а т е л ь с т в о. Пусть при $x = x_0 \neq 0$ степенной ряд (1) сходится, т. е. сходится числовой ряд

$$a_0 + a_1 x_0 + a_2 x_0^2 + \dots + a_n x_0^n + \dots \quad (4)$$

Тогда (см. § 10.1, п. 2, теорема 3)

$$\lim_{n \rightarrow \infty} a_n x_0^n = 0,$$

откуда следует, что члены ряда (4) ограничены по модулю:

$$|a_n x_0^n| \leq M, \quad n = 0, 1, 2, \dots \quad (5)$$

(M — постоянное положительное число).

Возьмем теперь любое x , для которого $|x| < |x_0|$, и рассмотрим ряд

$$|a_0| + |a_1 x| + |a_2 x^2| + \dots + |a_n x^n| + \dots \quad (6)$$

Имеем тождество $|a_n x^n| = |a_n x_0^n| \left| \frac{x}{x_0} \right|^n$ ($n = 0, 1, 2, \dots$).

Отсюда в силу (5) $|a_n x^n| \leq M \left| \frac{x}{x_0} \right|^n$ ($n = 0, 1, 2, \dots$)

или

$$|a_n x^n| \leq M q^n \quad (n = 0, 1, 2, \dots), \quad (7)$$

где

$$q = \left| \frac{x}{x_0} \right| < 1, \text{ так как } |x| < |x_0|.$$

Ряд

$$M + Mq + Mq^2 + \dots + Mq^n + \dots$$

сходится (см. § 10.1, п. 2, пример 2). Поэтому, имея в виду неравенства (7), согласно признаку сравнения рядов (см. § 10.1, п. 3) сходится ряд (6) для любого x , для которого $|x| < |x_0|$, а ряд (1), значит, для этого x сходится абсолютно.

Пусть теперь ряд (1) при $x = x_0$ расходится. Докажем, что он расходится и при любом x , удовлетворяющем неравенству $|x| > |x_0|$. Действительно, если бы при некотором значении $x = x_1$, удовлетворяющем неравенству $|x_1| > |x_0|$, ряд (1) был сходящимся, то по доказанному в первой части он был бы сходящимся и при $x = x_0$, что противоречит условию.

Теорема доказана.

С л е д с т в и е 1. Ряд второго класса (1) сходится абсолютно в интервале $(-\infty; +\infty)$.

С л е д с т в и е 2. Для каждого степенного ряда (1) третьего класса существует число $R > 0$, называемое радиусом сходимости этого ряда и обладающее следующими свойствами: при $|x| < R$ ряд (1) сходится абсолютно, при $|x| > R$ ряд (1) расходится.

Промежуток $(-R; R)$ называется *интервалом сходимости* степенного ряда (1).

Согласно следствию 1 для ряда (1) второго класса интервал сходимости $(-\infty; +\infty)$.

Областью сходимости степенного ряда (1) третьего класса является интервал $(-R; R)$, к которому в отдельных случаях добавляются один или оба конца этого интервала (это зависит от конкретного исследуемого ряда).

П р и м е ч а н и е 1. Для ряда (1) первого класса полагают $R=0$, для ряда (1) второго класса $R=\infty$.

Т е о р е м а 2. Пусть для ряда (1) существует и отличен от нуля предел

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L.$$

Тогда

$$R = \frac{1}{L}.$$

Д о к а з а т е л ь с т в о. Рассмотрим ряд, составленный из абсолютных величин членов ряда (1):

$$|a_0| + |a_1| |x| + |a_2| |x|^2 + \dots + |a_n| |x|^n + \dots \quad (8)$$

Применим признак Даламбера (§ 10.1, п. 3):

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| |x| = L |x|.$$

В соответствии с этим признаком ряд (8) сходится, если $|x| L < 1$, т. е. если $|x| < \frac{1}{L}$ и расходится, если $|x| L > 1$, т. е. если $|x| > \frac{1}{L}$ (в последнем случае общий член ряда (8), значит, и ряда (1) не стремится к нулю при $n \rightarrow \infty$).

Следовательно, ряд (1) сходится абсолютно при $|x| < \frac{1}{L}$ и расходится при $|x| > \frac{1}{L}$. Значит, радиусом сходимости ряда (1) является число

$$R = \frac{1}{L}.$$

П р и м е ч а н и е 2. Если $L=0$, то при любом x из числовой оси $|x| L=0 < 1$, т. е. ряд (8) сходится на всей числовой оси. Значит, ряд (1) абсолютно сходится на этой оси. Следовательно, $R=\infty$. Если $L=\infty$, то при любом $x \neq 0$ из числовой оси $|x| L=\infty$, и, значит, ряд (1) при любом $x \neq 0$ (§ 10.1, п. 3, примечание 4) расходится, т. е. $R=0$.

П р и м е р. Ряд $\sum_{n=1}^{\infty} \frac{1}{n} \left(\frac{x}{3}\right)^n$ имеет радиус сходимости $R=3$, так как

$$L = \lim_{n \rightarrow \infty} \frac{n3^n}{(n+1)3^{n+1}} = \frac{1}{3}.$$

Областью сходимости данного ряда является промежуток $-3 \leq x < 3$.

2. Свойства степенных рядов.

Т е о р е м а 1. *Всякий степенной ряд (1) с радиусом сходимости $R > 0$ сходится равномерно на всяком отрезке $[-\rho; \rho]$, содержащемся в интервале сходимости $(-R; R)$.*

Д о к а з а т е л ь с т в о. По условию $0 < \rho < R$. Поэтому положительный ряд

$$|a_0| + |a_1|\rho + |a_2|\rho^2 + \dots + |a_n|\rho^n + \dots$$

сходится, причем на отрезке $[-\rho; \rho]$:

$$|a_n x^n| \leq |a_n| \rho^n \quad (n = 0, 1, 2, \dots).$$

Следовательно, согласно признаку Вейерштрасса (см. § 10.2, п. 2) степенной ряд (1) сходится равномерно на отрезке $[-\rho; \rho]$.

Из установленной теоремы следует, что степенные ряды обладают всеми свойствами равномерно сходящихся рядов.

Сформулируем эти свойства (см., например, [15], т. 1):

Т е о р е м а 2. *Сумма степенного ряда (1) есть функция, непрерывная в каждой точке интервала сходимости ряда. Степенной ряд внутри его интервала сходимости можно почленно дифференцировать и интегрировать сколько угодно раз, причем в результате этих операций получаются степенные ряды, имеющие тот же радиус сходимости, что и исходный ряд.*

3. Ряды по степеням разности $x - a$. Степенным рядом называется также же функциональный ряд вида

$$a_0 + a_1(x - a) + a_2(x - a)^2 + \dots + a_n(x - a)^n + \dots \quad (9)$$

Интервалом сходимости степенного ряда (9) является интервал длиной $2R$ с центром в точке a .

Свойства степенных рядов по степеням x (п. 2) сохраняются и для рядов по степеням $x - a$.

4. Разложение функций в степенные ряды. Ряд Тейлора. Если функция $f(x)$ является суммой ряда (9), то в этом случае говорят, что функция $f(x)$ разлагается в ряд по степеням $x - a$.

Важность такого разложения видна хотя бы из того, что мы получаем возможность *приблизженно* заменить функцию суммой нескольких *первых членов степенного ряда*, т. е. *многочленом*.

Т е о р е м а 1. *Если функция $f(x)$ на интервале $(x_0 - R; x_0 + R)$ разлагается в степенной ряд*

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + \dots, \quad (10)$$

то это разложение единственно.

Д о к а з а т е л ь с т в о. Согласно теореме 2 (п. 2) имеем:

$$f'(x) = 1 \cdot a_1 + 2a_2(x - x_0) + 3a_3(x - x_0)^2 + \dots + na_n(x - x_0)^{n-1} + \dots,$$

$$f''(x) = 1 \cdot 2 \cdot a_2 + 2 \cdot 3a_3(x - x_0) + \dots + n(n-1)a_n(x - x_0)^{n-2} + \dots,$$

$$f'''(x) = 1 \cdot 2 \cdot 3a_3 + 2 \cdot 3 \cdot 4a_4(x - x_0) + \dots + (n-2)(n-1)na_n(x - x_0)^{n-3} + \dots,$$

.....

$$f^{(n)}(x) = 1 \cdot 2 \cdot 3 \dots (n-1)na_n + 2 \cdot 3 \dots n(n+1)a_{n+1}(x - x_0) + \dots$$

.....

Полагая в этих равенствах и в равенстве (10) $x = x_0$, получим:

$$a_0 = f(x_0), a_1 = \frac{f'(x_0)}{1!}, a_2 = \frac{f''(x_0)}{2!}, a_3 = \frac{f'''(x_0)}{3!}, \dots, a_n = \frac{f^{(n)}(x_0)}{n!}, \dots \quad (11)$$

Из (11) следует, что все коэффициенты ряда (10) определяются единственным образом формулами (11), что и доказывает теорему.

Подставляя полученные выражения коэффициентов в равенство (10), получим ряд

$$\begin{aligned} f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots \\ \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \dots, \end{aligned} \quad (12)$$

который называется *рядом Тейлора* функции $f(x)$; коэффициенты этого ряда

$$a_0 = f(x_0), a_1 = \frac{f'(x_0)}{1!}, a_2 = \frac{f''(x_0)}{2!}, \dots, a_n = \frac{f^{(n)}(x_0)}{n!}, \dots$$

называются *коэффициентами Тейлора* функции $f(x)$ в точке x_0 .

Таким образом, *если функция $f(x)$ разлагается в степенной ряд по степеням $x - x_0$, то этот ряд обязательно является рядом Тейлора этой функции.*

Если в ряде Тейлора положим $x_0 = 0$, то получим частный случай ряда Тейлора, который называют *рядом Маклорена*:

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \dots \quad (13)$$

Отметим, что все рассуждения были сделаны в предположении, что функция $f(x)$ *может быть разложена в степенной ряд*.

Если заранее этого не предполагать, а просто считать функцию $f(x)$ бесконечное число раз дифференцируемой и составить для нее ряд Тейлора, то ниоткуда не следует, что этот ряд сходится при значениях x , отличных от x_0 . Более того, могут быть даже такие случаи, что ряд Тейлора, составленный для функции $f(x)$, сходится, а сумма его вовсе не равна $f(x)$.

Пусть функция $f(x)$ в интервале $(x_0 - R; x_0 + R)$ имеет производные любого порядка. Тогда для любого x из этого интервала и любого n будет справедлива формула Тейлора (см. § 8.9, п. 1):

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + r_n(x), \quad (14)$$

где

$$r_n(x) = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{(n+1)!}(x - x_0)^{n+1}. \quad (15)$$

Перейдем теперь к выяснению условий, при которых можно утверждать, что ряд Тейлора, составленный для функции $f(x)$,

$$f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \dots \quad (16)$$

сходится в интервале $(x_0 - R; x_0 + R)$ и что его сумма равна $f(x)$.

Так как разность между $f(x)$ и суммой $(n+1)$ первых членов ряда (16) есть как раз $r_n(x)$, что видно из (14), то, очевидно, *для того чтобы ряд Тейлора (16) сходил к функции $f(x)$ (для которой он составлен) в интервале $(x_0 - R; x_0 + R)$, необходимо и достаточно, чтобы остаточный член в формуле Тейлора $r_n(x)$ для функции $f(x)$ в каждой точке этого интервала стремился к нулю при $n \rightarrow \infty$.*

Установим теперь достаточное условие сходимости ряда Тейлора функции $f(x)$ к этой функции.

Т е о р е м а 2. *Если в интервале $(x_0 - R; x_0 + R)$ функция $f(x)$ имеет производные любого порядка и все они по абсолютной величине ограничены одним и тем же числом*

$$|f^{(n)}(x)| \leq M \quad (n=1, 2, \dots), \quad (17)$$

то в этом интервале имеет место разложение (12).

Д о к а з а т е л ь с т в о. В силу (15) и (17) в интервале $(x_0 - R; x_0 + R)$ имеем:

$$|r_n(x)| = \frac{|f^{(n+1)}(x_0 + \theta(x - x_0))|}{(n+1)!} |x - x_0|^{n+1} \leq M \frac{|x - x_0|^{n+1}}{(n+1)!}. \quad (18)$$

Так как ряд

$$\sum_{n=0}^{\infty} \frac{|x - x_0|^{n+1}}{(n+1)!}$$

сходится при любом x по признаку Даламбера (§ 10.1, п. 3), то при любом x (§ 10.1, п. 2, теорема 3)

$$\lim_{n \rightarrow \infty} \frac{|x - x_0|^{n+1}}{(n+1)!} = 0.$$

Отсюда с учетом (18) $\lim_{n \rightarrow \infty} r_n(x) = 0$ в интервале $(x_0 - R; x_0 + R)$.

5. Разложение в степенные ряды основных элементарных функций.

1. $f(x) = e^x$. Здесь $f^{(n)}(x) = e^x$ и $f^{(n)}(0) = 1$. Условие (17) для данной функции выполнено в любом интервале $|x| < r$, так как $|f^{(n)}(x)| = e^x < e^r$. Поэтому согласно теореме 2 (п. 4) формула

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots \quad (19)$$

верна при всех x .

2. $f(x) = \sin x$. Здесь $f^{(k)}(x) = \sin\left(x + \frac{k\pi}{2}\right)$, $f^{(k)}(0) = \sin \frac{k\pi}{2} = 0$ при $k = 2n$, $f^{(k)}(0) = (-1)^n$ при $k = 2n + 1$. При этом $|f^{(n)}(x)| \leq 1$ на всей числовой оси. Поэтому согласно теореме 2 (п. 4) формула

$$\sin x = x - \frac{x^3}{3!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots$$

верна при всех x из $(-\infty; +\infty)$.

Аналогично формула

$$\cos x = 1 - \frac{x^2}{2!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!} + \dots$$

верна при всех x из $(-\infty; +\infty)$.

3. Рассмотрим функцию $f(x) = (1+x)^\alpha$, где α — любое вещественное число. Здесь

$$f^{(n)}(x) = \alpha(\alpha-1)\dots(\alpha-n+1)(1+x)^{\alpha-n},$$

$$f^{(n)}(0) = \alpha(\alpha-1)\dots(\alpha-n+1).$$

Можно доказать (см., например, в [9], т. 1), что равенство

$$(1+x)^{\alpha} = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n + \dots \quad (20)$$

верно при $|x| < 1$.

Ряд (20) называется *биномиальным*.

Если $\alpha = m$, где m — натуральное число, то равенство (20) обращается в формулу бинома Ньютона (см. § 8.9, п. 2):

$$(1+x)^m = 1 + \sum_{n=1}^m \frac{m(m-1)(m-2)\dots(m-n+1)}{n!}x^n.$$

Выделим следующие частные случаи биномиального ряда:

1) при $\alpha = -1$:

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n; \quad (21)$$

2) при $\alpha = \frac{1}{2}$:

$$\sqrt{1+x} = 1 + \frac{1}{2}x + \sum_{n=2}^{\infty} (-1)^{n-1} \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-3)}{2^n \cdot n!} x^n; \quad (22)$$

3) при $\alpha = -\frac{1}{2}$ имеем:

$$\frac{1}{\sqrt{1+x}} = 1 + \sum_{n=1}^{\infty} (-1)^n \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2^n \cdot n!} x^n.$$

4. Разложение $f(x) = \ln(1+x)$ получим путем интегрирования ряда (21) в промежутке от 0 до x (при $|x| < 1$):

$$\ln(1+x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1} \quad (-1 < x \leq 1). \quad (23)$$

5. Разложение $f(x) = \operatorname{arctg} x$ получим путем интегрирования ряда (21), если в нем предварительно заменить x на x^2 :

$$\operatorname{arctg} x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \quad (-1 \leq x \leq 1). \quad (24)$$

Области сходимости рядов (23), (24) указаны в скобках.

6. Приложения степенных рядов к приближенным вычислениям. Степенные ряды являются мощным вычислительным средством. С их помощью можно, например, вычислять приближенные значения функций, приближенно вычислять некоторые «неберущиеся» определенные интегралы.

Пример 1. Вычислить значение $e^{0,2}$ с точностью до 0,0001. Согласно формуле (19) имеем:

$$e^{0,2} = 1 + 0,2 + \frac{0,2^2}{2!} + \frac{0,2^3}{3!} + \frac{0,2^4}{4!} + \dots$$

Оценим погрешность, получаемую при отбрасывании всех членов этого ряда, начиная с пятого:

$$\begin{aligned} r_4 &= \frac{0,2^4}{4!} + \frac{0,2^5}{5!} + \frac{0,2^6}{6!} + \dots = \frac{0,2^4}{4!} \left(1 + \frac{0,2}{5} + \frac{0,2^2}{5 \cdot 6} + \dots \right) < \\ &< \frac{0,2^4}{4!} \left(1 + \frac{0,2}{5} + \left(\frac{0,2}{5} \right)^2 + \dots \right) = \frac{0,0016}{24} \cdot \frac{1}{1 - \frac{0,2}{5}} < 0,0001. \end{aligned}$$

Значит, с точностью до 0,0001 имеем:

$$e^{0,2} \approx 1 + 0,2 + \frac{0,2^2}{2!} + \frac{0,2^3}{3!} = 1,2 + \frac{0,04}{2} + \frac{0,008}{6} \approx 1,2213$$

(здесь можно использовать калькулятор).

Пример 2. Вычислить интеграл $\int_0^{\frac{1}{4}} e^{-x^2} dx$ с точностью до 0,0001. Заменяя x на $-x^2$ в формуле (19), получим:

$$e^{-x^2} = 1 - \frac{x^2}{1!} + \frac{x^4}{2!} - \frac{x^6}{3!} + \dots$$

Подставив этот ряд под знак данного интеграла и произведя почленное интегрирование, получаем:

$$\int_0^{\frac{1}{4}} e^{-x^2} dx = \int_0^{\frac{1}{4}} dx - \int_0^{\frac{1}{4}} x^2 dx + \frac{1}{2} \int_0^{\frac{1}{4}} x^4 dx - \frac{1}{3!} \int_0^{\frac{1}{4}} x^6 dx + \dots = 0,25 - \frac{1}{3 \cdot 4^3} + \frac{1}{10 \cdot 4^5} - \frac{1}{42 \cdot 4^7} + \dots \quad (25)$$

Это знакочередующийся ряд, удовлетворяющий теореме Лейбница (см. § 10.1, п. 4). Так как

$$\frac{1}{10 \cdot 4^5} = \frac{1}{10240} < \frac{1}{10000} = 0,0001,$$

то для получения нужной точности достаточно взять первые два члена ряда (25):

$$\int_0^{\frac{1}{4}} e^{-x^2} dx \approx 0,25 - \frac{1}{3 \cdot 4^3} \approx 0,25 - 0,0052 = 0,2448$$

(здесь также можно использовать калькулятор).

§ 10.4. Тригонометрические ряды

1. Тригонометрическая система функций, ее ортогональность. Функции $\varphi(x)$ и $\psi(x)$ называются *ортогональными* друг другу на отрезке $[a; b]$, если $\int_a^b \varphi(x)\psi(x)dx = 0$.

Система функций называется *ортгогональной* на отрезке $[a; b]$, если каждые две функции из этой системы ортгогональны друг другу на этом отрезке.

П р и м е р. Тригонометрическая система функций

$$1, \sin x, \cos x, \sin 2x, \cos 2x, \dots, \sin nx, \cos nx, \dots \quad (1)$$

ортгогональна на отрезке $[-\pi; \pi]$.

В самом деле, если $k \neq 0$ и целое, то

$$\int_{-\pi}^{\pi} \cos kx dx = \frac{1}{k} \sin xk \Big|_{-\pi}^{\pi} = 0, \quad (2)$$

$$\int_{-\pi}^{\pi} \sin kx dx = -\frac{1}{k} \cos xk \Big|_{-\pi}^{\pi} = 0. \quad (3)$$

Это значит, что единица ортгогональна ко всем остальным функциям системы (1).

Заметим теперь, что при натуральных m, n произведения $\sin nx \sin mx$ ($m \neq n$), $\cos nx \cos mx$ ($m \neq n$), $\sin nx \cos mx$ всегда можно представить суммой функций вида $\sin kx$ или $\cos kx$. Поэтому интеграл от $-\pi$ до π от этих произведений также равен нулю.

Укажем еще на одно свойство системы (1), заключающееся в том, что при любом натуральном n

$$\int_{-\pi}^{\pi} \cos^2 nx dx = \pi \text{ и } \int_{-\pi}^{\pi} \sin^2 nx dx = \pi. \quad (4)$$

2. Ряд Фурье. *Тригонометрическим рядом* называется функциональный ряд вида

$$\frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \dots + a_n \cos nx + b_n \sin nx + \dots,$$

или, более кратко, ряд вида

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx, \quad (5)$$

где a_0, a_n, b_n ($n=1, 2, \dots$) — постоянные числа, называемые *коэффициентами* тригонометрического ряда.

Свободный член ряда записан в виде $\frac{a_0}{2}$ для единообразия получающихся в дальнейшем формул.

Так как члены ряда (5) имеют общий период $T=2\pi$, то и сумма ряда, если он сходится, также является периодической функцией с периодом 2π .

Пусть периодическая функция $f(x)$ с периодом 2π является суммой ряда (5):

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx. \quad (6)$$

В таком случае говорят, что *функция $f(x)$ разлагается в тригонометрический ряд*. Предполагая, что этот ряд сходится равномерно на отрезке $[-\pi; \pi]$, покажем, как определить его коэффициенты.

Так как равномерно сходящийся на отрезке функциональный ряд можно почленно интегрировать на этом отрезке (см. § 10.2, п. 3, теорема 2), то с учетом формул (2), (3) имеем:

$$\int_{-\pi}^{\pi} f(x) dx = \frac{a_0}{2} \int_{-\pi}^{\pi} dx = a_0 \pi. \text{ Отсюда } a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx.$$

Умножим теперь ряд (6) на $\cos kx$ (k — фиксированное натуральное число). Получим опять равномерно сходящийся на отрезке $[-\pi; \pi]$ ряд (см. § 10.2, п. 3, теорема 4):

$$f(x) \cos kx = \frac{a_0}{2} \cos kx + \sum_{n=1}^{\infty} a_n \cos nx \cos kx + b_n \sin nx \cos kx.$$

Интегрируем его почленно на отрезке $[-\pi; \pi]$. И мы получим:

$$\int_{-\pi}^{\pi} f(x) \cos kx dx = \int_{-\pi}^{\pi} a_k \cos^2 kx dx$$

или согласно формуле (4)

$$\int_{-\pi}^{\pi} f(x) \cos kx dx = a_k \pi, \text{ откуда } a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx.$$

Аналогично, умножая обе части равенства (6) на $\sin kx$ и интегрируя в пределах от $-\pi$ до π , получим:

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx.$$

Таким образом, если периодическая функция $f(x)$ с периодом 2π является суммой равномерно сходящегося на отрезке $[-\pi; \pi]$ тригонометрического ряда (5), то коэффициенты этого ряда определяются по формулам

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nxdx, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nxdx \quad (7)$$

($n=1, 2, \dots$) и называются *коэффициентами Фурье* функции $f(x)$, а ряд (5) с этими коэффициентами называется *рядом Фурье* (Жан Батист Жозеф Фурье (1768—1830) — французский математик и физик) функции $f(x)$.

3. Сходимость ряда Фурье. При выводе формул (7) заранее предполагалось, что функция $f(x)$ разлагается в равномерно сходящийся тригонометрический ряд (6). Если такого предположения не делать, а допустить только, что для функции $f(x)$ существуют все интегралы, стоящие в правых частях формул (7), то по этим формулам можно вычислить коэффициенты a_0 , a_n , b_n и составить тригонометрический ряд (5), который представляет собой ряд Фурье, соответствующий данной функции.

Является ли построенный таким образом ряд Фурье сходящимся, и если он сходится, то имеем ли мы право утверждать, что он сходится именно к функции $f(x)$, с помощью которой вычислялись коэффициенты ряда?

Оказывается, что сходимость ряда Фурье к заданной функции имеет место для довольно широкого класса функций.

Сформулируем достаточные условия представимости функции рядом Фурье. Пусть функция $f(x)$ на отрезке $[-\pi; \pi]$ удовлетворяет условиям Дирихле (Петер Густав Дирихле (1805—1859) — немецкий математик). Это значит, что она на этом отрезке непрерывна или кусочно-непрерывна (т. е. имеет конечное число точек разрыва первого рода) и монотонна или кусочно-монотонна (т. е. отрезок $[-\pi, \pi]$ можно разделить на конечное число отрезков, внутри каждого из которых функция либо только возрастает, либо только убывает, либо постоянна).

Т е о р е м а Д и р и х л е. Если функция $f(x)$ удовлетворяет условиям Дирихле на отрезке $[-\pi; \pi]$, то ряд Фурье этой функции сходится на всем отрезке $[-\pi; \pi]$ и сумма этого ряда равна $f(x)$ в точках непрерывности функции, $\frac{1}{2}(f(x_0 - 0) + f(x_0 + 0))$ в точке x_0 разрыва функции, $\frac{1}{2}(f(\pi - 0) + f(-\pi + 0))$ на концах отрезка $[-\pi; \pi]$.

Доказательство см. в книге [15], т. II.

Сделаем некоторые замечания по поводу сформулированной теоремы Дирихле. Члены ряда (5) — периодические функции с периодом 2π . Поэтому если ряд (5) сходится на отрезке $[-\pi; \pi]$, то он сходится и при всех вещественных значениях x и сумма ряда (5) периодически повторяет с периодом 2π те значения, которые она давала на отрезке $[-\pi; \pi]$. Таким образом, если мы пользуемся рядом Фурье вне отрезка $[-\pi; \pi]$, то мы должны считать, что функция $f(x)$ продолжена вовне этого отрезка периодически с периодом 2π (рис. 132). С этой точки зрения его концы $x = -\pi$, $x = \pi$ явятся для продолженной таким образом функции точками разрыва, если $f(-\pi + 0) \neq f(\pi - 0)$.

На рисунке 132 изображена функция, непрерывная на отрезке $[-\pi; \pi]$, которая при периодическом продолжении дает разрывы в силу несовпадения значений $f(x)$ на концах отрезка $[-\pi; \pi]$.

П р и м е р. Функция $f(x) = x$ удовлетворяет условиям теоремы Дирихле и, следовательно, может быть разложена в ряд Фурье. По формулам (7) имеем:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} x dx = 0, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x \cos nx dx = \frac{1}{\pi n} x \sin nx \Big|_{-\pi}^{\pi} - \frac{1}{\pi n} \int_{-\pi}^{\pi} \sin nx dx = 0,$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin nx dx = -\frac{1}{\pi n} x \cos nx \Big|_{-\pi}^{\pi} + \frac{1}{\pi n} \int_{-\pi}^{\pi} \cos nx dx = -\frac{2}{n} \cos n\pi = -\frac{2}{n} (-1)^n = \frac{(-1)^{n+1} 2}{n}.$$

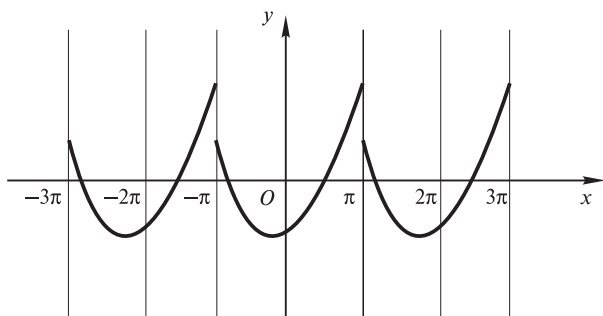


Рис. 132

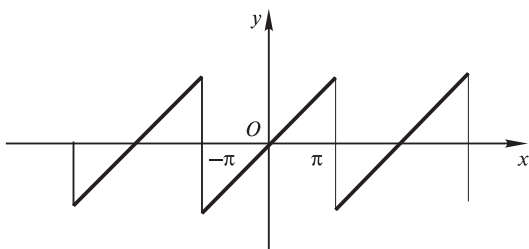


Рис. 133

Следовательно, согласно теореме Дирихле при $-\pi < x < \pi$

$$x = 2 \left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \frac{\sin 4x}{4} + \dots + (-1)^{n+1} \frac{\sin nx}{n} + \dots \right).$$

В точках $x = -\pi$ и $x = \pi$ сумма ряда Фурье по теореме Дирихле не совпадает со значениями функции $f(x) = x$, а равна

$$\frac{f(-\pi+0) + f(\pi-0)}{2} = \frac{-\pi + \pi}{2} = 0.$$

В силу периодичности суммы ряда график этой суммы имеет вид, изображенный на рисунке 133.

4. Ряды по косинусам и синусам. Если $f(x)$ — четная функция на отрезке $[-\pi; \pi]$, т. е. если $f(-x) = f(x)$, $x \in [-\pi; \pi]$, то ее коэффициенты Фурье b_n равны нулю. В самом деле,

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx = \frac{1}{\pi} \left(\int_{-\pi}^0 f(x) \sin nx dx + \int_0^{\pi} f(x) \sin nx dx \right).$$

Для разложения $f(x)$ в ряд по косинусам можно рассуждать следующим образом. Дополним определение данной функции $f(x)$ так, чтобы при $-\pi \leq x < 0$ было $f(x) = f(-x)$. В результате получится четная функция (в этом случае говорят: *функция $f(x)$ продолжена с отрезка $[0; \pi]$ на отрезок $[-\pi; 0]$ четным образом*; рис. 135). Тогда для «продолженной» четной функции справедливы все предыдущие рассуждения (п. 5), и, следовательно, коэффициенты Фурье могут быть вычислены по формулам

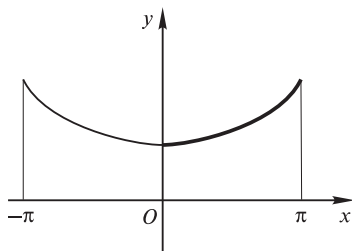


Рис. 135

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(x) \cos nx dx \quad (n = 0, 1, 2, \dots), \quad b_n = 0 \quad (n = 1, 2, \dots).$$

В этих формулах фигурируют лишь заданные на $[0; \pi]$ значения $f(x)$. Следовательно, при практических вычислениях *фактически* можно и не осуществлять указанное четное продолжение.

Если мы хотим разложить функцию $f(x)$ в ряд по синусам, то продолжаем ее с отрезка $[0; \pi]$ на отрезок $[-\pi; 0]$ нечетным образом.

Если мы дополним определение функции $f(x)$ так: $f(x) = -f(-x)$ при $-\pi \leq x < 0$, то получим нечетную функцию (говорят: *функция $f(x)$ продолжена нечетным образом*). При этом по смыслу, нечетности должны принять $f(0) = 0$. К «продолженной» нечетной функции опять применимы соображения, приведенные в п. 4, и, следовательно, для коэффициентов Фурье справедливы формулы

$$a_n = 0 \quad (n = 0, 1, 2, \dots), \\ b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx dx \quad (n = 1, 2, \dots).$$

Поскольку здесь участвуют лишь значения $f(x)$ на $[0; \pi]$, то, как и в случае ряда по косинусам, *фактически* продолжение функции $f(x)$ с отрезка $[0; \pi]$ на отрезок $[-\pi; 0]$ можно и не осуществлять.

6. Ряд Фурье с произвольным промежутком. Если надо разложить в тригонометрический ряд функцию $f(x)$ периода $2l$, которая на отрезке $[-l; l]$ удовлетворяет условиям Дирихле, то мы полагаем $x = \frac{lt}{\pi}$ и получаем функцию $\varphi(t) = f\left(\frac{lt}{\pi}\right)$ периода 2π :

$$\varphi(t + 2\pi) = f\left(\frac{l(t + 2\pi)}{\pi}\right) = f\left(\frac{lt}{\pi} + 2l\right) = f(x + 2l) = f(x) = f\left(\frac{lt}{\pi}\right) = \varphi(t),$$

которая на отрезке $[-\pi; \pi]$ удовлетворяет условиям Дирихле. Поэтому в интервале $(-\pi; \pi)$

$$\varphi(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \sin nt),$$

где

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \varphi(t) \cos ntdt \quad (n = 0, 1, 2, \dots),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \varphi(t) \sin ntdt \quad (n = 1, 2, \dots).$$

Возвратившись к прежней переменной x , получим, что в интервале $(-l; l)$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{\pi nx}{l} + b_n \sin \frac{\pi nx}{l} \right), \quad (8)$$

где

$$a_n = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{\pi nx}{l} dx \quad (n = 0, 1, 2, \dots), \quad (9)$$

$$b_n = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{\pi nx}{l} dx \quad (n = 1, 2, \dots). \quad (10)$$

Коэффициенты a_0, a_n, b_n ($n = 1, 2, \dots$) и здесь называются *коэффициентами Фурье для $f(x)$* , а ряд (8) — рядом Фурье для $f(x)$.

Если x — точка разрыва функции $f(x)$, то вместо $f(x)$ в равенстве (8) будет $\frac{f(x-0) + f(x+0)}{2}$.

Глава 11. ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 11.1. Функции нескольких переменных. Предел и непрерывность функции

1. Определение функции нескольких переменных. Понятие функции одной переменной не охватывает все зависимости, существующие в природе. Даже в самых простых задачах встречаются величины, значения которых определяются совокупностью значений нескольких величин.

Пример 1. Площадь S прямоугольника со сторонами, длины которых равны x и y , выражается формулой $S = xy$, т. е. значения S определяются совокупностью значений x и y .

Пр и м е р 2. Объем V прямоугольного параллелепипеда с ребрами, длины которых равны x , y и z , выражается формулой $V = xyz$, т. е. значения V определяются совокупностью значений x , y и z .

Пр и м е р 3. Абсолютная температура T , давление p и объем V данной массы газа связаны формулой Менделеева — Клапейрона $pV = RT$, где R — некоторая постоянная. Отсюда, например, $V = \frac{RT}{p}$, т. е. значения V определяются совокупностью значений T и p .

Пр и м е р 4. Количество тепла Q , выделяемое электрическим током, зависит от напряжения U , силы тока I и времени t , причем $Q = 0,24 UIt$, т. е. значения Q определяются совокупностью значений U , I и t .

Для изучения подобных зависимостей вводится понятие функций нескольких переменных.

Переменная z называется *функцией двух независимых переменных x и y* , если некоторым парам значений x и y по какому-либо правилу или закону ставится в соответствие определенное значение z .

Множество G пар значений x и y , которые могут принимать переменные x и y , называется *областью определения* или *областью существования функции*, а множество всех значений, принимаемых z в области определения, — *областью значений функции z* . Переменные x и y называются *аргументами функции z* . Символически функция двух переменных обозначается так: $z = f(x, y)$, $z = F(x, y)$, $z = \varphi(x, y)$, $z = z(x, y)$ и т. д.

Как известно (§ 1.1, п. 1), пара чисел x и y определяет положение точки M на плоскости xOy с координатами x и y (и, значит, радиус-вектор $\vec{r} = \vec{OM}$ (§ 2.1, п. 8)) и наоборот. Поэтому функцию двух переменных $z = f(x, y)$ можно рассматривать либо как функцию точки M и писать $z = f(M)$, либо как скалярную функцию векторного аргумента $\vec{r}$ и писать $z = f(\vec{r})$.

Способы задания функции двух переменных, как и в случае одной переменной, могут быть различными. В этой книге наиболее важным является аналитический способ задания, когда функция задается с помощью формулы. Областью определения функции в этом случае считается множество всех точек плоскости, для которых эта формула имеет смысл.

Пр и м е р 5. Областью определения функции $z = 1 - x - y$ является множество всех пар чисел (x, y) , т. е. вся плоскость xOy , а областью значений этой функции — промежуток $(-\infty; +\infty)$.

Пр и м е р 6. Для функции $z = \frac{x^2}{4} + \frac{y^2}{9}$ область определения — вся плоскость xOy , а область значений — промежуток $[0; +\infty)$.

Пр и м е р 7. Областью определения функции $z = \sqrt{1 - x^2 - y^2}$ является множество точек плоскости, координаты которых удовлетворяют неравенству $1 - x^2 - y^2 \geq 0$ (здесь

идет речь лишь о действительных значениях z) или неравенству $x^2 + y^2 \leq 1$, т. е. круг, ограниченный окружностью $x^2 + y^2 = 1$, включая эту окружность (замкнутый круг). Область значений этой функции — отрезок $[0; 1]$.

П р и м е р 8. Область определения функции $z = \frac{1}{\sqrt{1 - x^2 - y^2}}$ есть круг, ограниченный окружностью $x^2 + y^2 = 1$, исключая эту окружность (открытый круг), а область значений — промежуток $[1; +\infty)$.

Из рассмотренных примеров видно, что областью определения функции двух переменных может быть вся плоскость xOy или ее часть.

Известно (см. § 2.1, п. 8), что каждой тройке чисел $(x; y; z)$ в пространстве $Oxyz$ соответствует точка $M(x; y; z)$ (и, значит, радиус-вектор $\vec{r} = \overrightarrow{OM}$) и наоборот. Совершенно аналогично случаю двух переменных можно дать определение функции трех переменных $u = f(x; y; z) = f(M) = f(\vec{r})$. Областью определения функции трех переменных будет уже все пространство или его часть.

Аналогично можно дать определение функции четырех и более переменных.

В дальнейшем будем подробно рассматривать функции двух переменных, так как все важнейшие факты теории функций нескольких переменных наблюдаются уже на функциях двух переменных. Распространение определений и полученных результатов на функции трех и более переменных представляет собой, как правило, лишь технические трудности.

2. Геометрическое изображение функции двух переменных. Функции двух переменных допускают графическую иллюстрацию. *Графиком* функции $z = f(x, y)$, определенной в области G , называется множество точек $(x; y; z)$ пространства, у которых $(x; y)$ принадлежит G и $z = f(x, y)$. В наиболее простых случаях такой график представляет собой некоторую поверхность (см. также главу 4). Например, графиком функции из примера 5 (п. 1) является плоскость, проходящая через точки $(1; 0; 0)$, $(0; 1; 0)$ и $(0; 0; 1)$ (§ 4.1, п. 5); графиком функции из примера 6 (п. 1) — эллиптический параболоид (§ 5.2, п. 2).

Однако построение графиков функций двух переменных в большинстве случаев представляет значительные трудности. В связи с этим оказывается удобным геометрически описывать функции двух переменных, не выходя в трехмерное пространство. Средством такого описания являются *линии уровня*. Отметим на плоскости xOy все те точки $(x; y)$, в которых функция $f(x; y)$ принимает одно и то же значение, например значение, равное c . Иначе говоря, отмечаем те точки $(x; y)$, для которых

$$f(x; y) = c. \quad (1)$$

Множество этих точек и называется линией уровня функции $f(x, y)$. Понятно, что уравнение (1) есть уравнение этой линии. Придавая c

различные значения и каждый раз строя линию с уравнением (1), получим семейство линий уровня. Это семейство наглядно описывает функцию $f(x, y)$. Обычно рядом с линией уровня ставят то значение c , которому она соответствует.

П р и м е р. Построить линии уровня функции

$$z = x^2 + y^2. \quad (2)$$

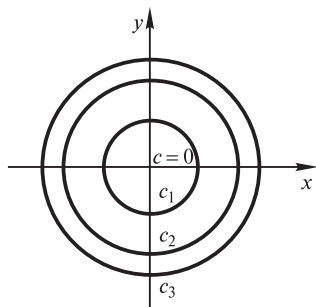


Рис. 136

Пересечем поверхность (2) плоскостью $z = c$ ($0 \leq c < +\infty$). Задавая c различные значения, например $c =$

$= 0, 1, 2, 3, \dots$, получим семейство линий уровня, представляющих собой окружности. При $c = 0$ окружность вырождается в точку $(0; 0)$ (рис. 136). Из того, что линиями уровня оказались окружности с центрами в начале координат, следует, что графиком данной функции должна быть поверхность вращения вокруг оси Oz . Действительно, как известно из аналитической геометрии (§ 5.2, п. 2), уравнение (2) определяет параболоид вращения.

Линиями уровня обозначают глубину морей и высоту гор на географических картах. Аналогичные линии описывают распределение тех или иных веществ в почве, распределение среднесуточной температуры и т. д.

3. Предел функции двух переменных. Множество точек $M(x; y)$, координаты которых x и y удовлетворяют неравенству $\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$, или, короче, $MM_0 < \delta$, называется δ -окрестностью точки $M_0(x_0; y_0)$. Другими словами, δ -окрестность точки M_0 — это все точки, лежащие внутри круга с центром M_0 радиуса δ .

О п р е д е л е н и е. Число A называется *пределом функции* $z = f(x, y) = f(M)$ при стремлении точки M к точке $M_0(x_0; y_0)$, что кратко записывается: $M \rightarrow M_0$, если для любого числа $\varepsilon > 0$ существует такое число $\delta = \delta(\varepsilon) > 0$, что для всех точек M из области определения этой функции, удовлетворяющих условию $0 < MM_0 < \delta$, имеет место неравенство $|f(M) - A| < \varepsilon$. Обозначают это так:

$$\lim_{M \rightarrow M_0} f(M) = A, \text{ или } \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y) = A.$$

Функция $z = f(M)$ называется *бесконечно малой* при $M \rightarrow M_0$, если $\lim_{M \rightarrow M_0} f(M) = 0$.

Все основные свойства о бесконечно малых и о пределах, установленные в главе 7 (§ 7.4, 7.5) для функций одной переменной, обобщаются и на случай функций двух и большего числа переменных.

4. Непрерывность функции двух переменных. Пусть точка $M_0(x_0; y_0)$ принадлежит области определения функции $z = f(x, y) = f(M)$.

О п р е д е л е н и е. Функция $z = f(x, y) = f(M)$ называется *непрерывной* в точке M_0 , если

$$\lim_{M \rightarrow M_0} f(M) = f(M_0), \text{ или } \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y) = f(x_0, y_0), \quad (3)$$

причем точка M стремится к точке M_0 произвольным образом, оставаясь в области определения функции.

Обозначим $x - x_0 = \Delta x$, $y - y_0 = \Delta y$. *Полным приращением функции* $z = f(x, y) = f(M)$ при переходе от точки M_0 к точке M называется разность значений функции в этих точках, а именно: $\Delta z = f(M) - f(M_0)$, т. е. $\Delta z = f(x, y) - f(x_0, y_0)$.

Условие (3) непрерывности функции в точке M_0 равносильно условию

$$\lim_{M \rightarrow M_0} \Delta z = 0, \text{ или } \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} \Delta z = 0.$$

П р и м е р. Функция $z = x^2 + y^2$ непрерывна в любой точке плоскости xOy , так как при любых значениях x и y величина $\Delta z = 2x\Delta x + 2y\Delta y + (\Delta x)^2 + (\Delta y)^2$ стремится к нулю при $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$.

Арифметические операции над непрерывными в точке функциями приводят к непрерывным функциям в той же точке (при условии, что деление производится на функцию, не обращающуюся в этой точке в нуль). Это устанавливается так же, как для функций одной переменной (§ 7.6, п. 1).

5. Понятие области. *Областью* или *открытой областью* называется множество точек плоскости, обладающее следующими двумя свойствами:

1) каждая точка области принадлежит ей вместе с некоторой окрестностью этой точки (свойство *открытости*);

2) всякие две точки области можно соединить непрерывной линией (непрерывная линия — множество точек $M(x; y)$ плоскости, координаты которых заданы как непрерывные функции $x = \varphi(t)$, $y = \psi(t)$, $a \leq t \leq b$), целиком лежащей в этой области (свойство *связности*).

Часть плоскости, лежащая внутри замкнутого контура L (рис. 137), является областью, так как, во-первых, для любой точки M , лежащей внутри L , существует окрестность, также лежащая внутри L ; во-вторых, две любые точки M и N , лежащие внутри L , можно соединить непрерывной линией, также лежащей внутри L .

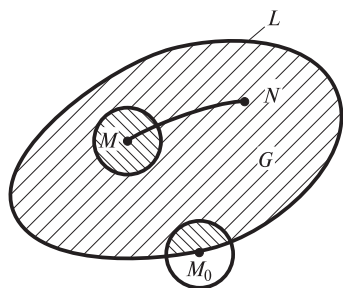


Рис. 137

Точка M_0 называется *граничной* точкой области G , если любая окрестность этой точки содержит как точки области G , так и точки, ей не принадлежащие (см. рис. 137).

Множество всех граничных точек области называется ее *границей*. На рисунке 137 любая точка контура L , очевидно, является граничной.

Если к открытой области присоединить ее границу, то полученное множество точек называется *замкнутой областью*. Так, область определения функции в примере 7 (п. 1) является замкнутой.

Если для данной области можно подобрать круг, полностью ее покрывающий, то такая область называется *ограниченной*. Например, области определения функций в примерах 7, 8 (п. 1) являются ограниченными.

Область G (открытая или замкнутая) называется *односвязной*, если для любого замкнутого контура, лежащего в этой области, ограниченная им часть плоскости целиком принадлежит области G . Например, области определения функций в примерах 5 — 8 (п. 1) являются односвязными. Область же, заключенная между окружностями $x^2 + y^2 = 2$ и $x^2 + y^2 = 4$, не является односвязной, так как окружность $x^2 + y^2 = 3$, лежащая в этой области, содержит внутри себя и точки, не принадлежащие области (например, начало координат). Это пример двусвязной области. Область, имеющая два выреза, называется *трахсвязной* и т.д.

П р и м е ч а н и е. Все введенные в этом пункте понятия почти без изменения переносятся на пространство трех и большего числа измерений.

6. Основные свойства непрерывных функций двух переменных. В § 7.6 (п. 3) были рассмотрены свойства функций, непрерывных на отрезке. Эти свойства распространяются и на случай функций двух и большего числа переменных, непрерывных в ограниченной замкнутой области.

Функция $z = f(x, y) = f(M)$ называется *непрерывной* в открытой или замкнутой области, если она непрерывна в каждой точке этой области. При этом функция $f(M)$ считается непрерывной в граничной точке M_0 , если в равенстве $\lim_{M \rightarrow M_0} f(M) = f(M_0)$ точка M стремится к точке M_0 вдоль любого пути, принадлежащего данной области.

Имеет место следующее предложение:

Если функция $z = f(M)$ непрерывна в ограниченной замкнутой области, то она в этой области:

- 1) имеет наибольшее и наименьшее значения;
- 2) ограничена: $|f(M)| \leq K$ (K — положительное число);
- 3) принимает все промежуточные значения между любыми двумя своими значениями, т. е. если $A < C < B$, где A и B — какие-то значения

функции $f(M)$ в данной области, то в этой области существует точка M_0 , в которой $f(M_0) = C$.

Из свойства 3, в частности, следует, что если M_1 и M_2 — точки данной области и $f(M_1) < 0$, а $f(M_2) > 0$, то в этой области существует точка M_0 , в которой $f(M_0) = 0$.

П р и м е ч а н и е. Заметим, что на случай функций двух и большего числа переменных распространяется (см. [11]) теорема 5 из п. 1 § 7.6.

§ 11.2. Частные производные. Полный дифференциал

1. Частные производные. Частной производной функции нескольких переменных по какой-нибудь переменной в рассматриваемой точке называется обычная производная по этой переменной, считая другие переменные фиксированными (постоянными). Например, для функции двух переменных $z = f(x, y)$ в точке $M_0(x_0, y_0)$ частные производные определяются так:

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta_x z}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x, y_0) - f(x_0, y_0)}{\Delta x},$$

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \rightarrow 0} \frac{\Delta_y z}{\Delta y} = \lim_{\Delta y \rightarrow 0} \frac{f(x_0, y_0 + \Delta y) - f(x_0, y_0)}{\Delta y},$$

если эти пределы существуют. Величина $\Delta_x z = f(x_0 + \Delta x, y_0) - f(x_0, y_0)$ ($\Delta_y z = f(x_0, y_0 + \Delta y) - f(x_0, y_0)$) называется *частным приращением* функции z в точке M_0 по аргументу x (y). Используются и другие обозначения частных производных:

$$z'_x, \frac{\partial f(x_0, y_0)}{\partial x}, \frac{\partial}{\partial x} f(x_0, y_0), f'_x(x_0, y_0), z'_y, \frac{\partial f(x_0, y_0)}{\partial y} \text{ и т. д.}$$

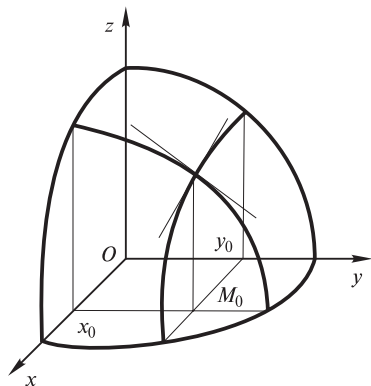


Рис. 138

Символы $\frac{\partial z}{\partial x}$, $\frac{\partial f(x_0, y_0)}{\partial x}$, $\frac{\partial z}{\partial y}$, $\frac{\partial f(x_0, y_0)}{\partial y}$ как дроби трактовать нельзя (в этом отличие от случая одной переменной).

Из определения следует геометрический смысл частной производной функции двух переменных: частная производная $f'_x(x_0, y_0)$ ($f'_y(x_0, y_0)$) — *угловой коэффициент касательной к линии пересечения поверхности $z = f(x, y)$ и плоскости $y = y_0$ ($x = x_0$) в соответствующей точке* (рис. 138).

Пользуясь понятием скорости изменения переменной (см. § 8.1, п.2), можно сказать, что частная производная $f'_x(x_0, y_0)$ ($f'_y(x_0, y_0)$) есть скорость изменения функции $f(x, y)$ относительно x (y) при постоянном y (x).

Из определения частных производных следует, что правила вычисления их остаются теми же, что и для функций одной переменной (§ 8.2), и только требуется помнить, по какой переменной находится производная.

Пример 1. Если $z = x^2 + y^2$, то $z'_x = 2x$, $z'_y = 2y$.

Пример 2. Если $p = \frac{RT}{V}$, то $p'_T = \frac{R}{V}$, $p'_V = -\frac{RT}{V^2}$. Величина p'_V называется *изотермическим коэффициентом упругости* идеального газа.

Аналогично определяются и обозначаются частные производные функции трех и большего числа независимых переменных.

2. Полный дифференциал. Как уже отмечалось (§ 11.1, п. 4), полным приращением функции $z = f(x, y)$ в точке $(x_0; y_0)$, соответствующим приращениям Δx и Δy переменных x и y , называется выражение

$$\Delta z = \Delta f(x_0, y_0) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0). \quad (1)$$

Если приращение (1) можно представить в виде

$$\Delta z = A\Delta x + B\Delta y + \alpha(\Delta x, \Delta y)\Delta x + \beta(\Delta x, \Delta y)\Delta y, \quad (2)$$

где A и B не зависят от Δx и Δy , а $\alpha(\Delta x, \Delta y)$ и $\beta(\Delta x, \Delta y)$ стремятся к нулю при стремлении к нулю Δx и Δy , то функция $z = f(x, y)$ называется *дифференцируемой* в точке $(x_0; y_0)$, а линейная часть $A\Delta x + B\Delta y$ приращения функции (т. е. та часть Δz , которая зависит от Δx и Δy линейно) называется *полным дифференциалом* (или просто *дифференциалом*) этой функции в точке $(x_0; y_0)$ и обозначается символом dz :

$$dz = A\Delta x + B\Delta y. \quad (3)$$

Из определения дифференцируемости функции следует, что если данная функция дифференцируема в точке $(x_0; y_0)$, то она в этой точке непрерывна.

Действительно, если в точке $(x_0; y_0)$ функция $z = f(x, y)$ дифференцируема, то для этой точки Δz представимо в форме (2), откуда следует, что

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \Delta z = 0,$$

а это и означает, что в точке $(x_0; y_0)$ функция $z = f(x, y)$ непрерывна.

Из дифференцируемости функции в данной точке следует существование ее частных производных в этой точке (*необходимое условие дифференцируемости*).

В самом деле, пусть функция $z = f(x, y)$ в точке $(x_0; y_0)$ дифференцируема. Тогда имеет место соотношение (2). Полагая в нем $\Delta y = 0$, имеем:

$$\Delta_x z = A\Delta x + \alpha(\Delta x, 0) \Delta x.$$

Деля на $\Delta x \neq 0$ и переходя к пределу при $\Delta x \rightarrow 0$, получаем:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta_x z}{\Delta x} = \lim_{\Delta x \rightarrow 0} (A + \alpha(\Delta x, 0)) = A.$$

Это означает, что в точке $(x_0; y_0)$ существует частная производная функции $z = f(x, y)$ по x и

$$\frac{\partial z}{\partial x} = A. \quad (4)$$

Аналогично доказывается, что в точке $(x_0; y_0)$ существует частная производная

$$\frac{\partial z}{\partial y} = B. \quad (5)$$

Используя формулы (4) и (5), можно переписать выражение (3) в виде $dz = \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y$.

Если положить $z = x$, то $dz = dx = 1 \cdot \Delta x + 0 \cdot \Delta y = \Delta x$, т. е. $dx = \Delta x$. Аналогично, полагая $z = y$, получим $dy = \Delta y$. Значит, дифференциалы независимых переменных совпадают с приращениями этих переменных, и мы можем записать дифференциал (3) в следующем виде: $dz = f'_x(x_0, y_0)dx + f'_y(x_0, y_0)dy$.

Т е о р е м а (достаточное условие дифференцируемости). *Если функция $z = f(x, y)$ имеет частные производные в некоторой окрестности точки $M_0(x_0; y_0)$ и эти производные непрерывны в самой точке M_0 , то эта функция дифференцируема в точке M_0 .*

Д о к а з а т е л ь с т в о. Дадим переменным x_0 и y_0 столь малые приращения Δx и Δy , чтобы точка $M(x_0 + \Delta x; y_0 + \Delta y)$ не вышла за пределы указанной окрестности точки M_0 . Полное приращение $\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ можно записать в виде

$$\Delta z = (f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)) + (f(x_0, y_0 + \Delta y) - f(x_0, y_0)).$$

Каждая из этих разностей представляет частное приращение функции. Преобразуем каждую из этих разностей по формуле Лагранжа (см. § 8.6, п. 3). Получим:

$$\Delta z = f'_x(x_0 + \theta \Delta x, y_0 + \Delta y) \Delta x + f'_y(x_0, y_0 + \theta_1 \Delta y) \Delta y \quad (6)$$

$$(0 < \theta, \theta_1 < 1).$$

Так как производные f'_x и f'_y непрерывны в точке M_0 , то

$$\begin{aligned}\lim_{\Delta x \rightarrow 0} f'_x(x_0 + \theta \Delta x, y_0 + \Delta y) &= f'_x(x_0, y_0), \\ \lim_{\Delta x \rightarrow 0} f'_y(x_0, y_0 + \theta_1 \Delta y) &= f'_y(x_0, y_0).\end{aligned}$$

Отсюда (см. § 11.1, п. 3, § 7.5, п. 1)

$$\begin{aligned}f'_x(x_0 + \theta \Delta x, y_0 + \Delta y) &= f'_x(x_0, y_0) + \alpha, \\ f'_y(x_0, y_0 + \theta_1 \Delta y) &= f'_y(x_0, y_0) + \beta,\end{aligned}$$

где α и β — бесконечно малые при $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$. Подставляя эти значения в равенство (6), находим:

$$\Delta z = f'_x(x_0, y_0) \Delta x + f'_y(x_0, y_0) \Delta y + \alpha \Delta x + \beta \Delta y,$$

а это и означает, что функция $z = f(x, y)$ дифференцируема в точке M_0 .

3. Производные и дифференциал сложной функции. Пусть $z = f(x, y)$, где $x = \varphi(t)$, $y = \psi(t)$. Тогда в конечном итоге z будет функцией одной переменной t . Предположим, что z'_x , z'_y непрерывны и $\frac{dx}{dt}$, $\frac{dy}{dt}$ существуют. Найдем $\frac{dz}{dt}$. Дадим переменной t приращение Δt . Тогда x , y , а следовательно, и z получат свои приращения Δx , Δy и Δz . В силу достаточного условия дифференцируемости

$$\Delta z = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y + \alpha \Delta x + \beta \Delta y,$$

откуда

$$\frac{\Delta z}{\Delta t} = f'_x(x, y) \frac{\Delta x}{\Delta t} + f'_y(x, y) \frac{\Delta y}{\Delta t} + \alpha \frac{\Delta x}{\Delta t} + \beta \frac{\Delta y}{\Delta t}.$$

Устремим теперь Δt к нулю. Тогда Δx и Δy будут стремиться к нулю, так как функции x и y непрерывны (мы предположили существование производных x'_t и y'_t), а потому α и β также будут стремиться к нулю. В пределе получим:

$$\frac{dz}{dt} = f'_x(x, y) \frac{dx}{dt} + f'_y(x, y) \frac{dy}{dt},$$

или, короче,

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}. \quad (7)$$

Формула (7) называется формулой *производной сложной функции*.

Пр и м е р 1. Пусть $z=f(x, y)$, $x=t^3+3$, $y=2t^4+1$. По формуле (7) имеем:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot 3t^2 + \frac{\partial z}{\partial y} \cdot 8t^3.$$

Предположим, в частности, что роль независимой переменной играет x , т. е. рассмотрим функцию $z=f(x, y)$, где $y=\psi(x)$. Согласно формуле (7) будем иметь:

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \frac{dy}{dx}, \quad (8)$$

так как $\frac{dx}{dx} = 1$. В формуле (8) $\frac{\partial z}{\partial x}$ — частная производная по первому аргументу функции двух переменных $f(x, y)$, а $\frac{dz}{dx}$ — обычная производная сложной функции одной переменной x : $z=f(x, \psi(x))$. Последнюю производную будем называть *полной производной* функции. В случае, когда $z=f(x, y)$, где $x=\varphi(y)$, аналогично получаем:

$$\frac{dz}{dy} = \frac{\partial z}{\partial x} \frac{dx}{dy} + \frac{\partial z}{\partial y}$$

($\frac{\partial z}{\partial y}$ — частная производная по второму аргументу функции $f(x, y)$, $\frac{dz}{dy}$ —

полная производная функции одной переменной y : $z=f(\varphi(y), y)$).

Пусть теперь $x=\varphi(t, \tau)$, $y=\psi(t, \tau)$ (здесь предполагается существование первых производных функций x, y по t и τ). В этом случае z будет функцией двух независимых переменных t и τ . Следовательно, для этого случая формулу (7) нужно переписать в виде

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t}. \quad (9)$$

Аналогично

$$\frac{\partial z}{\partial \tau} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial \tau} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial \tau}. \quad (10)$$

Пр и м е р 2. Если $z=xy$, где $x=t \cos 2\tau$, $y=t^2\tau$, то $z'_t = y \cos 2\tau + 2xt\tau$, $z'_\tau = -2yt \sin 2\tau + xt^2$.

Из формул (9) и (10) видно, что символ частной производной, как уже отмечалось выше, нельзя трактовать как дробь. В самом деле, если бы можно было сократить на dx и dy , то из формул (9) и (10) получили бы, что

$$\frac{\partial z}{\partial t} = 2 \frac{\partial z}{\partial t} \text{ и } \frac{\partial z}{\partial \tau} = 2 \frac{\partial z}{\partial \tau}.$$

Предположим еще, что x'_t, x'_τ, y'_t и y'_τ непрерывны.

Если бы x и y были независимыми переменными, то полный дифференциал функции z был бы равен $dz = z'_x dx + z'_y dy$. В данном случае z зависит через посредство x, y от переменных t и τ , следовательно, $dz = z'_t dt + z'_\tau d\tau$. Но

$$z'_t = z'_x x'_t + z'_y y'_t$$

и

$$z'_\tau = z'_x x'_\tau + z'_y y'_\tau.$$

Поэтому

$$dz = (z'_x x'_t + z'_y y'_t) dt + (z'_x x'_\tau + z'_y y'_\tau) d\tau = z'_x (x'_t dt + x'_\tau d\tau) + z'_y (y'_t dt + y'_\tau d\tau).$$

Нетрудно видеть, что выражения, стоящие в скобках, являются дифференциалами функций x, y , поэтому можно записать:

$$dz = z'_x dx + z'_y dy.$$

Мы пришли к той же форме записи дифференциала, что и в случае, когда x и y были независимыми переменными.

4. Неявные функции и их дифференцирование. Как уже отмечалось (см. § 7.1, п. 4), если уравнение, с помощью которого задается функция одной переменной x , не разрешено относительно y , то эта функция называется неявной. Разрешая это уравнение относительно y , мы получаем ту же функцию, но уже заданную в явной форме. Однако часто бывает, что разрешить такое уравнение относительно y невозможно (например, $2^y - 2y + x^2 - 1 = 0$) или нецелесообразно; в этом случае уравнение так и оставляют неразрешенным, в общем виде (когда все его члены перенесены в левую часть):

$$F(x, y) = 0. \quad (11)$$

В связи с этим встает вопрос о том, как найти производную неявной функции, не разрешая уравнения (11) относительно y .

Если в уравнении (11), определяющем неявную функцию $y = f(x)$, задавать значения независимой переменной x , то для нахождения соответствующего значения y надо решать уравнение. Затем, если в это уравнение подставить его решение, то получится тождество. Поэтому можно сказать также, что неявная функция $y = f(x)$, определенная уравнением (11), — это такая функция, которая, будучи подставлена в уравнение (11), обращает его в тождество. Дифференцируя это тождество по x согласно правилу дифференцирования сложной функции, получим:

$$F'_x(x, y) + F'_y(x, y) \frac{dy}{dx} = 0.$$

Отсюда при $F'_y(x, y) \neq 0$ вытекает формула для производной неявной функции

$$\frac{dy}{dx} = -\frac{F'_x(x, y)}{F'_y(x, y)}. \quad (12)$$

Пример 1. Пусть у как функция от x задана соотношением $e^{xy} - x - y = 0$. Найти $\frac{dy}{dx}$.
Для $F(x, y) = e^{xy} - x - y$ имеем: $F'_x = ye^{xy} - 1$, $F'_y = xe^{xy} - 1$ и согласно формуле (12)

$$\frac{dy}{dx} = \frac{1 - ye^{xy}}{xe^{xy} - 1}.$$

Пусть уравнение

$$F(x, y, z) = 0 \quad (13)$$

определяет z как неявную функцию $z = \varphi(x, y)$ независимых переменных x и y .

Пользуясь формулой (12), из равенства (13) имеем:

$$\frac{\partial z}{\partial x} = -\frac{F'_x}{F'_z}, \quad \frac{\partial z}{\partial y} = -\frac{F'_y}{F'_z}. \quad (14)$$

Пример 2. Найти частные производные неявной функции z , заданной уравнением $x^2 + y^2 + z^2 - 1 = 0$.

Согласно формулам (14)

$$\frac{\partial z}{\partial x} = -\frac{x}{z}, \quad \frac{\partial z}{\partial y} = -\frac{y}{z}.$$

5. Касательная плоскость и нормаль к поверхности. Рассмотрим некоторую поверхность

$$F(x, y, z) = 0 \quad (15)$$

и на ней точку $M_0(x_0; y_0; z_0)$. Проведем через точку M_0 какую-нибудь линию L , целиком лежащую на поверхности (15) (рис. 139). Пусть параметрические уравнения линии L :

$$x = \varphi(t), \quad y = \psi(t), \quad z = \omega(t),$$

причем при $t = t_0$ получаем координаты точки M_0 . Так как каждая точка линии L лежит на данной поверхности (15), то при любом значении параметра t будет

$$F(\varphi(t), \psi(t), \omega(t)) = 0.$$

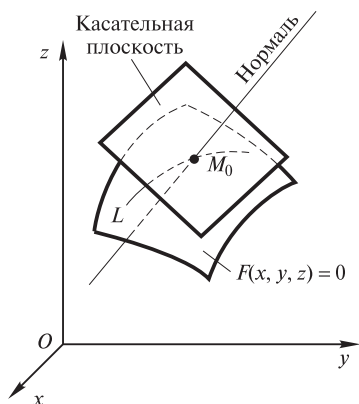


Рис. 139

Но тогда это соотношение есть тождество относительно t . Дифференцируя его, находим:

$$F'_x(\varphi(t), \psi(t), \omega(t))\varphi'(t) + F'_y(\dots)\psi'(t) + F'_z(\dots)\omega'(t) = 0.$$

Полагая здесь $t = t_0$, получим:

$$F'_x(M_0)\varphi'(t_0) + F'_y(M_0)\psi'(t_0) + F'_z(M_0)\omega'(t_0) = 0. \quad (16)$$

Касательная к линии L в точке M_0 определяется уравнениями

$$\frac{x - x_0}{\varphi'(t_0)} = \frac{y - y_0}{\psi'(t_0)} = \frac{z - z_0}{\omega'(t_0)}. \quad (17)$$

Введем в рассмотрение прямую, определив ее уравнениями

$$\frac{x - x_0}{\frac{\partial F(M_0)}{\partial x}} = \frac{y - y_0}{\frac{\partial F(M_0)}{\partial y}} = \frac{z - z_0}{\frac{\partial F(M_0)}{\partial z}}. \quad (18)$$

Соотношение (16) показывает (см. § 4.2, п. 5), что прямые (17), (18) перпендикулярны. Из уравнения (18) видно, что прямая (18) вполне определяется данной поверхностью (15) и выбранной на этой поверхности точкой M_0 , но не зависит от кривой L , которую мы провели на поверхности (15) через точку M_0 произвольно. Следовательно, касательные, проведенные в точке M_0 к всевозможным кривым, лежащим на поверхности (15) и проходящим через точку M_0 , перпендикулярны к одной и той же прямой (18) и потому лежат в одной плоскости, которую мы будем называть *касательной плоскостью* к данной поверхности в точке M_0 , а прямую (18) — *нормалью* к данной поверхности в точке M_0 .

Уравнение этой касательной плоскости как уравнение плоскости, проходящей через точку M_0 и перпендикулярной к прямой (18), есть (см. § 4.1, п. 2)

$$\frac{\partial F(M_0)}{\partial x}(x - x_0) + \frac{\partial F(M_0)}{\partial y}(y - y_0) + \frac{\partial F(M_0)}{\partial z}(z - z_0) = 0. \quad (19)$$

Если поверхность задана уравнением $z = f(x, y)$, то уравнение касательной плоскости будет получено как частный случай уравнения (19) при $F(x, y, z) = z - f(x, y)$. Тогда $F'_x = -f'_x$, $F'_y = -f'_y$, $F'_z = 1$ и уравнение (19) примет вид:

$$z - z_0 - f'_x(x_0, y_0)(x - x_0) - f'_y(x_0, y_0)(y - y_0) = 0.$$

Обозначив здесь $x - x_0 = \Delta x$, $y - y_0 = \Delta y$, $z - z_0 = \Delta z$, получим:

$$\Delta z = f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y,$$

или

$$df(x_0, y_0) = \Delta z,$$

т. е. *полный дифференциал функции двух переменных равен приращению аппликаты касательной плоскости*. Таково геометрическое истолкование полного дифференциала функции $f(x, y)$.

П р и м е р. Составить уравнение касательной плоскости и нормали к поверхности, заданной уравнением $z = x^2 + y^2$ в точке $M_0(1; 1; 2)$.

Здесь $F = z - x^2 - y^2$, $F'_x = -2x$, $F'_y = -2y$, $F'_z = 1$. В точке M_0 имеем $F'_x = F'_y = -2$, $F'_z = 1$. Поэтому согласно формулам (19) и (18) получим уравнение касательной плоскости $2x + 2y - z - 2 = 0$ и уравнение нормали

$$\frac{x-1}{2} = \frac{y-1}{2} = \frac{z-2}{-1}.$$

§ 11.3. Частные производные и дифференциалы высших порядков

1. Частные производные высших порядков. Частные производные функции нескольких переменных сами являются функциями этих переменных и могут иметь частные производные. Для исходной функции эти последние производные будут частными производными *второго порядка*. Так, для функции $z = f(x, y)$ двух независимых переменных можно определить (предполагается, что все производные существуют) четыре частные производные второго порядка, которые обозначаются символами

$$\begin{aligned} z''_{x^2} &= \frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \right), \quad z''_{xy} = \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} \right), \\ z''_{yx} &= \frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y} \right), \quad z''_{y^2} = \frac{\partial^2 z}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial y} \right). \end{aligned}$$

Частные производные z''_{xy} и z''_{yx} , отличающиеся порядком дифференцирования, называются *смешанными частными производными второго порядка*.

Аналогично определяются частные производные третьего, четвертого и старших порядков.

П р и м е р. Найти частные производные второго порядка функции $z = e^{x-2y}$.

Имеем:

$$\begin{aligned} z'_x &= e^{x-2y}, \quad z'_y = -2e^{x-2y}, \\ z''_{x^2} &= e^{x-2y}, \quad z''_{xy} = -2e^{x-2y}, \quad z''_{yx} = -2e^{x-2y}, \quad z''_{y^2} = 4e^{x-2y}. \end{aligned}$$

Здесь $z''_{xy} = z''_{yx}$. Оказывается, имеет место следующая теорема (см., например, [11]).

Т е о р е м а. Смешанные производные второго порядка равны, если они непрерывны: $f''_{xy}(x, y) = f''_{yx}(x, y)$.

С л е д с т в и е. Смешанные производные высших порядков равны, если они непрерывны и получены в результате дифференцирования по одним и тем же переменным одинаковое число раз, но может быть в разной последовательности.

Покажем это на примере:

$$z'''_{x^2y} = ((z'_x)'_x)'_y = ((z'_x)'_y)'_x = ((z'_y)'_x)'_x,$$

т. е.

$$z'''_{x^2y} = z'''_{xyx} = z'''_{yx^2}.$$

Здесь мы дважды пользовались только что отмеченной теоремой: первый раз применительно к функции z'_x (мы изменили порядок ее дифференцирования), второй раз использовали равенство $z''_{xy} = z''_{yx}$. В общем случае схема рассуждений аналогична.

2. Признак полного дифференцирования. Выясним, при каких условиях выражение

$$P(x, y)dx + Q(x, y)dy, \quad (1)$$

где $P(x, y)$ и $Q(x, y)$ непрерывны вместе со своими частными производными первого порядка в области D , является полным дифференциалом некоторой функции $u = u(x, y)$ в каждой точке области D , или, кратко, полным дифференциалом в области D .

Т е о р е м а. Выражение (1) есть полный дифференциал в области D тогда и только тогда, когда в каждой точке области D выполнено равенство

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

3. Дифференциалы высших порядков. Заметим прежде всего, что для функции нескольких переменных справедливы те же общие правила дифференцирования, что и для функции одной переменной (§ 8.3, п. 4):

$$\text{I. } d(u + v) = du + dv \quad (u = u(x, y), v = v(x, y)).$$

$$\text{II. } d(uv) = vdu + u dv.$$

$$\text{III. } d(Cu) = Cdu \quad (C - \text{const}).$$

$$\text{IV. } d\left(\frac{u}{v}\right) = \frac{vdu - u dv}{v^2} \quad (v \neq 0).$$

Например, имеем:

$$\begin{aligned} d(u + v) &= (u + v)'_x dx + (u + v)'_y dy = u'_x dx + v'_x dx + u'_y dy + v'_y dy = \\ &= (u'_x dx + u'_y dy) + (v'_x dx + v'_y dy) = du + dv. \end{aligned}$$

Пусть имеется функция $z=f(x, y)$ независимых переменных x и y , обладающая непрерывными частными производными второго порядка. Рассмотрим ее полный дифференциал

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

(dx и dy — произвольные приращения), который назовем *полным дифференциалом первого порядка* (или, кратко, *первым дифференциалом*).

Так как $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ по предположению имеют непрерывные частные производные первого порядка, то от функции dz , в свою очередь, можно взять полный дифференциал $d(dz)$. Так получим *полный дифференциал второго порядка* (или, кратко, *второй дифференциал*), который обозначается d^2z .

Аналогично, потребовав существования непрерывных частных производных третьего, четвертого, n -го порядков, можно получить полные дифференциалы соответственно третьего, четвертого, n -го порядков.

Найдем выражение для второго дифференциала через частные производные. Пользуясь правилами I, III (dx и dy не зависят от x и y , т. е. рассматриваются как постоянные) и приведенной в п. 1 теоремой, можно записать:

$$\begin{aligned} d^2z &= d\left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy\right) = d\left(\frac{\partial z}{\partial x}\right) dx + d\left(\frac{\partial z}{\partial y}\right) dy = \\ &= \frac{\partial^2 z}{\partial x^2} (dx)^2 + \frac{\partial^2 z}{\partial x \partial y} dx dy + \frac{\partial^2 z}{\partial y \partial x} dy dx + \frac{\partial^2 z}{\partial y^2} (dy)^2 = \\ &= \frac{\partial^2 z}{\partial x^2} dx^2 + 2 \frac{\partial^2 z}{\partial x \partial y} dx dy + \frac{\partial^2 z}{\partial y^2} dy^2 \end{aligned} \quad (2)$$

(здесь $dx^2 = (dx)^2$, $dy^2 = (dy)^2$).

Формула (2) обобщается на случай дифференциала n -го порядка.

4. Формула Тейлора для функции двух переменных. Пусть дана функция $z=f(x, y)$, имеющая непрерывные частные производные всех порядков до $(n+1)$ -го включительно в некоторой окрестности точки $M_0(x_0; y_0)$. Пусть точка $M(x_0 + \Delta x; y_0 + \Delta y)$ принадлежит этой окрестности. Вспомогательную функцию $F(t)$ определим на отрезке $0 \leq t \leq 1$ равенством

$$F(t) = f(x, y), \quad (3)$$

где $x = x_0 + t\Delta x$, $y = y_0 + t\Delta y$. Согласно известной (см. § 8.9, п. 2) формуле Тейлора имеем:

$$F(t) = F(0) + F'(0)t + \dots + \frac{F^{(n)}(0)}{n!} t^n + \frac{F^{(n+1)}(\theta t)}{(n+1)!} t^{n+1} \quad (0 < \theta < 1). \quad (4)$$

Вычислим коэффициенты формулы (4) с помощью равенства (3). При $t = 0$ имеем $F(0) = f(x_0, y_0)$. Дифференцируя сложную функцию $F(t)$ по t , получим:

$$\begin{aligned} F'(t) &= f'_x(x, y)\Delta x + f'_y(x, y)\Delta y = df(x, y), \\ F''(t) &= f''_{x^2}(x, y)(\Delta x)^2 + 2f''_{xy}(x, y)\Delta y\Delta x + f''_{y^2}(x, y)(\Delta y)^2 = d^2 f(x, y), \\ &\dots\dots\dots \\ F^{(n)}(t) &= d^n f(x, y), \\ F^{(n+1)}(t) &= d^{n+1} f(x, y). \end{aligned}$$

Заменяя в последнем равенстве t на θt , а в остальных положив $t = 0$, найдем:

$$\begin{aligned} F^{(k)}(0) &= d^k f(x_0, y_0), \quad (k = 1, 2, \dots, n), \\ F^{(n+1)}(\theta t) &= d^{n+1} f(x_0 + \theta t\Delta x, y_0 + \theta t\Delta y). \end{aligned}$$

Если подставить найденные выражения в равенство (4) и затем положить в нем $t = 1$, то получим для $f(x, y)$ формулу Тейлора:

$$\begin{aligned} f(x_0 + \Delta x, y_0 + \Delta y) &= f(x_0, y_0) + df(x_0, y_0) + \frac{d^2 f(x_0, y_0)}{2!} + \dots \\ &\dots + \frac{d^n f(x_0, y_0)}{n!} + \frac{d^{n+1} f(x_0 + \theta\Delta x, y_0 + \theta\Delta y)}{(n+1)!}. \end{aligned}$$

Такой же вид формула Тейлора имеет и в случае большего числа переменных.

§ 11.4. Экстремум функций двух переменных

1. Необходимые условия существования экстремума. Понятие максимума и минимума можно распространить и на функции нескольких переменных (здесь для случая двух переменных).

Говорят, что функция $z = f(x, y)$ имеет в точке $M_0(x_0; y_0)$ максимум (минимум), если существует такая окрестность точки M_0 , что для всех точек $M(x, y)$ из этой окрестности и отличных от M_0 выполняется неравенство

$$f(x_0, y_0) > f(x, y) \quad (f(x_0, y_0) < f(x, y))$$

или

$$\Delta z = f(x, y) - f(x_0, y_0) < 0 \quad (\Delta z = f(x, y) - f(x_0, y_0) > 0).$$

Т е о р е м а (необходимые условия существования экстремума). Если функция $z = f(x, y)$ имеет в точке $M_0(x_0; y_0)$ экстремум и в этой точке существуют частные производные z'_x и z'_y , то

$$f'_x(x_0, y_0) = 0, \quad f'_y(x_0, y_0) = 0, \quad (1)$$

Доказательство. Из определения экстремума следует, что функция $f(x, y_0)$, рассматриваемая как функция одной переменной x , при $x = x_0$ также имеет экстремум. Поэтому (см. § 8.7, п. 2, теорема 1) $f'_x(x_0, y_0) = 0$. Аналогично получаем равенство $f'_y(x_0, y_0) = 0$.

Примечание. Приведенные условия существования экстремума не являются достаточными, о чем свидетельствует следующий пример.

Пример. $z = x^3 + y^3$, $z'_x = 3x^2$, $z'_y = 3y^2$.

Частные производные равны нулю в точке $(0; 0)$, но экстремума эта функция в точке $(0; 0)$ не имеет, так как в любой окрестности этой точки она принимает значения разных знаков, а в самой точке $(0; 0)$ $z = 0$.

2. Достаточные условия существования экстремума.

Теорема (достаточные условия существования экстремума). Пусть функция $f(x, y)$, непрерывная вместе со своими частными производными первого и второго порядков в некоторой окрестности точки $M_0(x_0; y_0)$, удовлетворяет условиям (1).

Обозначим $A = f''_{x^2}(M_0)$, $B = f''_{xy}(M_0)$, $C = f''_{y^2}(M_0)$, $D = AC - B^2$. Тогда в точке M_0 функция $f(x, y)$: 1) имеет минимум, если $D > 0$ и $A > 0$; 2) имеет максимум, если $D > 0$ и $A < 0$; 3) не имеет экстремума, если $D < 0$.

Доказательство. Ради краткости доказательство проведем для случаев 1 и 2. Согласно формуле Тейлора, взятой для $n = 1$, с учетом условий (1) имеем:

$$\Delta f(x_0, y_0) = \frac{1}{2} (A'(\Delta x)^2 + 2B'\Delta x \Delta y + C'(\Delta y)^2), \quad (2)$$

где

$$A' = f''_{x^2}(M_*), B' = f''_{xy}(M_*), C' = f''_{y^2}(M_*)$$

$$(M_*(x_0 + \theta \Delta x, y_0 + \theta \Delta y), 0 < \theta < 1).$$

В силу непрерывности вторых частных производных в точке M_0 следует, что

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} A' = A, \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} (A'C' - (B')^2) = AC - B^2 = D.$$

Поэтому (§ 11.1, п. 6, примечание) для достаточно малых по модулю $\Delta x, \Delta y$ имеем:

$$A' > 0 \text{ (если } A > 0), \quad (3)$$

$$A' < 0 \text{ (если } A < 0), \quad (4)$$

$$A'C' - (B')^2 > 0 \text{ (если } D > 0). \quad (5)$$

В силу неравенств (3) и (4) равенство (2) можно представить в виде

$$\Delta f(x_0, y_0) = \frac{1}{2A'} ((A')^2 (\Delta x)^2 + 2A'B'\Delta x \Delta y + A'C'(\Delta y)^2)$$

или, дополняя до полного квадрата, в виде

$$\Delta f(x_0, y_0) = \frac{1}{2A'}((A'\Delta x + B'\Delta y)^2 + (A'C' - (B')^2)(\Delta y)^2).$$

Выражение во внешних скобках в силу неравенства (5) положительно. Поэтому: 1) если $A > 0$ (а тогда в силу неравенства (3) и $A' > 0$), то $\Delta f(x_0, y_0) > 0$ и, следовательно, в точке M_0 минимум; 2) если $A < 0$ (а тогда в силу неравенства (4) и $A' < 0$), то $\Delta f(x_0, y_0) < 0$ и, следовательно, в точке M_0 максимум.

Пример 1. Исследовать на экстремум функцию $z = x^3 + y^3 - 3xy$.

Ее частные производные $z'_x = 3x^2 - 3y$, $z'_y = 3y^2 - 3x$ обращаются в нуль в точках $M_0(0; 0)$ и $M_1(1; 1)$. Ее вторые производные равны $z''_{xx} = 6x$, $z''_{xy} = -3$, $z''_{yy} = 6y$. В точке M_0 имеем $D = -9 < 0$, следовательно, экстремума в этой точке нет. В точке M_1 $D = 27 > 0$, причем $A = 6 > 0$, следовательно, в точке M_1 минимум.

Примечание. Покажем на примерах, что в случае $D = 0$ экстремум может быть, но его может и не быть.

Пример 2. Функция $z = x^3 + y^3$ в точке $(0; 0)$, где $D = 0$, как показано выше (см. п. 1), экстремума не имеет.

Пример 3. Функция $z = x^4 + y^4$ в точке $(0; 0)$, где $D = 0$, имеет минимум, потому что в любой окрестности этой точки данная функция положительна, а в самой точке $(0; 0)$ равна нулю.

§ 11.5. Метод наименьших квадратов

В естествознании приходится пользоваться *эмпирическими* формулами, составленными на основе опыта и наблюдений. Один из наилучших методов получения таких формул — это способ наименьших квадратов. Изложим идею этого способа, ограничиваясь случаем линейной зависимости двух величин.

Пусть требуется установить зависимость между двумя величинами x и y (например, температурой и удлинением металлического стержня). Производим соответствующие измерения (например, n измерений) и результаты измерений сводим в таблицу:

x	x_1	x_2	x_3	...	x_n
y	y_1	y_2	y_3	...	y_n

(1)

Будем рассматривать x и y как прямоугольные координаты точек на плоскости. Предположим, что точки $(x_k; y_k)$, $k = 1, \dots, n$, группируются вдоль некоторой прямой линии (рис. 140). Естественно в этом случае считать, что между x и y существует приближенная линейная зависимость, т. е.

$$y = ax + b. \quad (2)$$

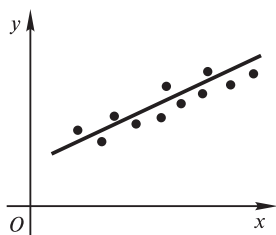


Рис. 140

Назовем *уклоном* (или *отклонением*) разность между точным значением функции (2) в точке x_k и соответствующим значением y_k из таблицы (1): $\varepsilon_k = ax_k + b - y_k$. Сумма квадратов уклонов — функция величин a и b :

$$u(a, b) = \sum_{k=1}^n \varepsilon_k^2 = \sum_{k=1}^n (ax_k + b - y_k)^2.$$

В методе наименьших квадратов лежит следующий принцип: искомыми значениями a и b являются те, при которых функция $u(a, b)$ имеет минимум.

Для этого (см. § 11.4, п. 1) необходимо, чтобы

$$\begin{aligned} \frac{\partial u}{\partial a} &= 2 \sum_{k=1}^n (ax_k + b - y_k)x_k = 2a \sum_{k=1}^n x_k^2 + 2b \sum_{k=1}^n x_k - 2 \sum_{k=1}^n x_k y_k = 0, \\ \frac{\partial u}{\partial b} &= 2 \sum_{k=1}^n (ax_k + b - y_k) = 2a \sum_{k=1}^n x_k + 2bn - 2 \sum_{k=1}^n y_k = 0. \end{aligned}$$

Отсюда

$$\begin{cases} a \sum_{k=1}^n x_k^2 + b \sum_{k=1}^n x_k = \sum_{k=1}^n x_k y_k, \\ a \sum_{k=1}^n x_k + bn = \sum_{k=1}^n y_k. \end{cases} \quad (3)$$

Это окончательный вид так называемой нормальной системы способа наименьших квадратов. Пусть $a = a_0$, $b = b_0$ — решение системы (3). Можно доказать, что a_0 и b_0 доставляют величине $u(a, b)$ минимум. Функция (2) при $a = a_0$ и $b = b_0$ дает эмпирическую формулу $y = a_0 x + b_0$.

Пример. Результаты измерения величин x и y даны в таблице:

x	-2	0	1	2	4
y	0,5	1	1,5	2	3

Предполагая, что между x и y существует линейная зависимость $y = ax + b$, способом наименьших квадратов определить коэффициенты a и b . Здесь, $n = 5$,

$$\sum_{k=1}^5 x_k^2 = 25, \quad \sum_{k=1}^5 x_k = 5, \quad \sum_{k=1}^5 x_k y_k = 16,5, \quad \sum_{k=1}^5 y_k = 8$$

и нормальная система (3) принимает вид:

$$\begin{cases} 25a + 5b = 16,5, \\ 5a + 5b = 8. \end{cases}$$

Решая эту систему, получим $a = 0,425$, $b = 1,175$. Поэтому $y = 0,425x + 1,175$.

Глава 12. ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

§ 12.1. Двойные интегралы

1. Задачи, приводящие к понятию двойного интеграла.

Задача о массе неоднородной пластины. Пусть плоская область G заполнена веществом с известной плотностью $\rho(M) = \rho(x, y)$. Найти массу (количество вещества) всей материальной области — «пластины». Под плотностью вещества в точке M понимается предел средней плотности бесконечно малой части G , содержащей точку M . Разобьем область G произвольным образом (рис. 141) на n частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$ без общих внутренних точек, площади которых обозначим соответственно через $\Delta w_1, \Delta w_2, \dots, \Delta w_n$. Предположим, что в пределах каждой частичной области плотность постоянна и равна $\rho(N_k)$ для Δ_k , где $N_k(\xi_k, \eta_k)$ — произвольная точка частичной области Δ_k . Тогда масса Δ_k приближенно равна:

$$\Delta m_k \approx \rho(\xi_k, \eta_k) \Delta w_k.$$

Для массы всей пластины получим приближенное выражение:

$$m = \sum_{k=1}^n \Delta m_k \approx \sum_{k=1}^n \rho(\xi_k, \eta_k) \Delta w_k. \quad (1)$$

Пусть λ — наибольший из диаметров частичных областей. Заметим, что диаметром области называется наибольшее из расстояний между точками ее границы. Например, диаметр прямоугольника равен его диагонали, диаметр эллипса — его большой оси. Для круга это определение диаметра совпадает с обычным. Сумма (1) будет тем точнее выражать искомую массу m , чем меньше будет каждый из диаметров частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$. Поэтому за массу m естественно принять $m = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n \rho(\xi_k, \eta_k) \Delta w_k$.

П р и м е ч а н и е. Точно так же, как эта задача, решаются задачи о суммарном заряде, распределенном в области G с заданной плотностью $\rho(M)$, о давлении жидкости на дно сосуда, о количестве световой энергии, падающей на площадку G , и многие другие.

Задача об объеме цилиндрида. Пусть дана функция $f(x, y)$, непрерывная и неотрицательная в области G . Найти объем тела, ограниченного сверху поверхностью $z = f(x, y)$, снизу областью G и с боков прямой цилиндрической поверхностью, направляющей которой

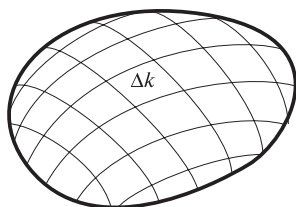


Рис. 141

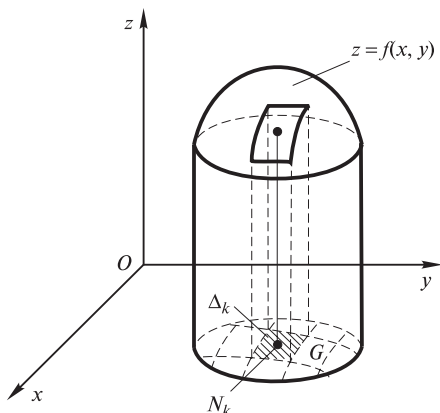


Рис. 142

служит замкнутый контур, ограничивающий область G (рис. 142). Такое тело для краткости будем называть *цилиндроидом*. В частности, когда верхнее основание цилиндроида есть плоскость, параллельная нижнему основанию, то цилиндرويد называется *цилиндром*. Примером цилиндра служит круговой цилиндр.

Для нахождения объема V данного цилиндроида разобьем область G произвольным образом на n частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$ без общих внутренних точек, площади которых обозначим соответственно через $\Delta w_1, \Delta w_2, \dots, \Delta w_n$. В каждой из этих частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$ выберем произвольную точку $N_k(\xi_k; \eta_k)$ и построим прямой цилиндрический столбик с основанием Δ_k и высотой $f(\xi_k, \eta_k)$. Объем такого столбика равен $f(\xi_k, \eta_k)\Delta w_k$. Сумма объемов этих цилиндрических столбиков представляет собой объем ступенчатого тела, приближенно заменяющего объем данного цилиндроида. Следовательно,

$$V \approx \sum_{k=1}^n f(\xi_k, \eta_k) \Delta w_k.$$

Эта сумма будет тем точнее выражать искомый объем V , чем меньше будет каждый из диаметров частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$. Поэтому за объем V естественно принять

$$V = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\xi_k, \eta_k) \Delta w_k,$$

где λ — наибольший из диаметров частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$.

2. Определение двойного интеграла. Из решения приведенных выше задач п. 1 видим, что, хотя эти задачи имеют различный смысл,

математический аппарат для их решения один и тот же. Во всех этих задачах получаем выражение одного и того же вида

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\xi_k, \eta_k) \Delta w_k, \quad (2)$$

где $f(x, y)$ — функция, заданная в области G .

О п р е д е л е н и е. Если существует предел (2), не зависящий от способа разбиения области G на частичные области Δ_k и выбора точек $N_k(\xi_k; \eta_k)$ в них, то он называется *двойным интегралом* от функции $f(x, y)$ по области G и обозначается символом

$$\iint_G f(x, y) dw = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\xi_k, \eta_k) \Delta w_k. \quad (3)$$

Функция $f(x, y)$ в этом случае называется *интегрируемой* в области G . При этом $f(x, y)$ называется *подынтегральной функцией*, dw — *элементом площади*, G — *областью интегрирования*, x и y — *переменными интегрирования*, сумма $\sum_{k=1}^n f(\xi_k, \eta_k) \Delta w_k$ — *интегральной суммой*.

П р и м е ч а н и е. Для двойного интеграла используется также обозначение $\iint_G f(x, y) dx dy$.

О п р е д е л е н и е. Кривая называется *гладкой*, если в каждой ее точке существует касательная и при переходе от точки к точке положение этой касательной меняется непрерывно. Поэтому кривая, заданная уравнениями $x = \varphi(t)$, $y = \psi(t)$, $\alpha \leq t \leq \beta$, будет гладкой, если функции $\varphi(t)$ и $\psi(t)$ непрерывны и имеют непрерывные производные $\varphi'(t)$ и $\psi'(t)$, не обращающиеся в нуль одновременно (тем самым кривая в каждой точке имеет касательную). Непрерывная кривая, составленная из конечного числа гладких кусков, называется *кусочно-гладкой*.

Справедлива следующая теорема (она следует из более общей теоремы, устанавливаемой в полных курсах математического анализа; см., например, [9], т. 2):

Т е о р е м а с у щ е с т в о в а н и я д в о й н о г о и н т е г р а л а. Если область G с кусочно-гладкой границей Γ ограничена и замкнута, а функция $f(x, y)$ непрерывна в области G , то эта функция интегрируема в области G .

В дальнейшем будем предполагать, что условия этой теоремы выполнены.

Из рассмотренных выше задач (п. 1) и определения двойного интеграла следует, что: 1) двойной интеграл (3) с положительной подынтегральной функцией может быть истолкован физически,

например как масса соответствующей пластины; 2) тот же интеграл с неотрицательной подынтегральной функцией может быть истолкован геометрически как объем соответствующего цилиндриоида. В частности, двойной интеграл от единичной функции ($f(x, y) \equiv 1$) по области G , т. е. интеграл $\iint_G dw$ численно равен площади области интегрирования $S_G = \iint_G dw$.

3. Свойства двойного интеграла. Эти свойства, как и их доказательства, аналогичны соответствующим свойствам определенного интеграла. Поэтому приведем их без доказательства.

1. Постоянный множитель можно выносить за знак двойного интеграла.

2. Двойной интеграл от суммы двух функций равен сумме двойных интегралов от этих функций.

Примечание. Свойство 2 распространяется на случай алгебраической суммы любого конечного числа функций.

3. Пусть область G разбита на две области G_1 и G_2 . Тогда

$$\iint_G f(x, y)dw = \iint_{G_1} f(x, y)dw + \iint_{G_2} f(x, y)dw.$$

4. Если функция $f(x, y) > 0$ в области G , то $\iint_G f(x, y)dw > 0$.

5. Двойной интеграл равен произведению значения подынтегральной функции в некоторой точке области интегрирования на площадь этой области (теорема о среднем).

4. Вычисление двойных интегралов. Пусть требуется вычислить двойной интеграл $\iint_G f(x, y)dw$ от непрерывной в области G функ-

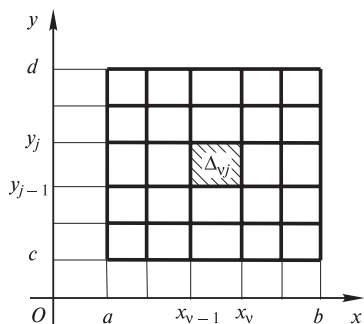


Рис. 143

ции $f(x, y)$.

Случай прямоугольной области. Пусть область G — прямоугольник $a \leq x \leq b$, $c \leq y \leq d$ (кратко $[a, b; c, d]$). Разобьем область G на частичные области прямыми, параллельными координатным осям (рис. 143) и проходящими через точки $x_0 = a, x_1, \dots, x_{m-1}, x_m = b$ оси Ox и точки $y_0 = c, y_1, \dots, y_p = d$ оси Oy . Тогда область G разо-

бьется на прямоугольники, наибольший из диаметров которых обозначим через λ . Пусть Δ_{vj} — прямоугольник, являющийся пересечением v столбца и j горизонтальной полосы. Площадь его будет $\Delta w_{vj} = \Delta x_v \Delta y_j$, где $\Delta x_v = x_v - x_{v-1}$, $\Delta y_j = y_j - y_{j-1}$. Выберем точку $(\xi_{vj}; \eta_{vj}) \in \Delta_{vj}$ так, чтобы $\xi_{vj} = x_{v-1}$, $j = 1, 2, \dots, p$. Тогда интегральная сумма будет

$$\sigma = \sum_{v,j} f(x_{v-1}, \eta_{vj}) \Delta x_v \Delta y_j, \quad (4)$$

где сумма распространена по всем прямоугольникам, т. е. по всем значениям v и j : $v = 1, 2, \dots, m$; $j = 1, 2, \dots, p$.

Сумма вида (4) с двумя индексами суммирования называется *двойной интегральной суммой*. Для ее вычисления можно сначала произвести суммирование по j при фиксированном v , т. е. сложить слагаемые, отвечающие одному (любому) столбцу, а затем результаты просуммировать по v . Тогда получим:

$$\sigma = \sum_{v=1}^m \left[\sum_{j=1}^p f(x_{v-1}, \eta_{vj}) \Delta x_v \Delta y_j \right] = \sum_{v=1}^m \left[\sum_{j=1}^p f(x_{v-1}, \eta_{vj}) \Delta y_j \right] \Delta x_v.$$

Разумеется, такой переход от двойной суммы к *повторной* можно было бы осуществить и вторым способом: первое, внутреннее, суммирование произвести по v , а второе, внешнее, — по j .

Используя одно из свойств определенного интеграла (§ 9.7, п. 1, свойство 4) и теорему о среднем (§ 9.7, п. 4), будем иметь:

$$\int_c^d f(x_{v-1}, y) dy = \sum_{j=1}^p \int_{y_{j-1}}^{y_j} f(x_{v-1}, y) dy = \sum_{j=1}^p f(x_{v-1}, \eta_{vj}) \Delta y_j.$$

Следовательно,

$$\sigma = \sum_{v=1}^m \Phi(x_{v-1}) \Delta x_v, \quad (5)$$

где

$$\Phi(x_{v-1}) = \int_c^d f(x_{v-1}, y) dy. \quad (6)$$

Перейдя в равенстве (5) к пределу при $\lambda \rightarrow 0$ (при $\lambda \rightarrow 0 \max \Delta x_v \rightarrow 0$; как и прежде, λ — наибольший из диаметров частичных областей), будем иметь:

$$\iint_G f(x, y) dw = \int_a^b \Phi(x) dx,$$

или с учетом равенства (6):

$$\iint_G f(x, y) dw = \int_a^b \left[\int_c^d f(x, y) dy \right] dx. \quad (7)$$

Обычно формулу (7) записывают в виде

$$\iint_G f(x, y) dw = \int_a^b dx \int_c^d f(x, y) dy. \quad (8)$$

Выражение, стоящее в правой части последней формулы, называется *повторным интегралом*. Для его вычисления надо последовательно взять два обычных определенных интеграла: сначала *внутренний интеграл*

$$\int_c^d f(x, y) dy,$$

в котором x считается постоянной, а затем полученное выражение (оно зависит от x) проинтегрировать по x от a до b — *внешний интеграл*.

Аналогично при втором способе перехода от двойной интегральной суммы к повторной получили бы

$$\iint_G f(x, y) dw = \int_c^d dy \int_a^b f(x, y) dx. \quad (9)$$

Пример. Вычислить двойной интеграл

$$I = \iint_G (x^2 + y^2) dw,$$

где G — квадрат $0 \leq x \leq 1, 0 \leq y \leq 1$.

По формуле (8) имеем:

$$I = \int_0^1 dx \int_0^1 (x^2 + y^2) dy = \int_0^1 \left(x^2 y + \frac{y^3}{3} \right) \bigg|_0^1 dx = \int_0^1 \left(x^2 + \frac{1}{3} \right) dx = \left(\frac{x^3}{3} + \frac{1}{3} x \right) \bigg|_0^1 = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

Этот же двойной интеграл можно вычислить и по формуле (9).

Случай произвольной области. Пусть теперь G — область на плоскости xOy , изображенная на рисунке 144. Тогда вместо интеграла (случай прямоугольной области)

$$\int_c^d f(x_{v-1}, y) dy$$

будем иметь интеграл

$$\int_{\varphi_1(x_{v-1})}^{\varphi_2(x_{v-1})} f(x_{v-1}, y) dy.$$

Здесь $y = \varphi_1(x)$ и $y = \varphi_2(x)$ — уравнения нижней и верхней частей кон-

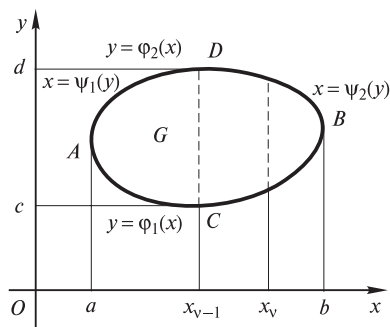


Рис. 144

тура области G , на которые он делится точками A и B . Соответственно и окончательный результат взамен формулы (8) запишется в виде

$$\iint_G f(x, y)dw = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)dy. \quad (10)$$

Можно интегрировать и в другом порядке — сначала по x , а затем по y . Тогда получается формула

$$\iint_G f(x, y)dw = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x, y)dx, \quad (11)$$

где $x = \psi_1(y)$ и $x = \psi_2(y)$ — уравнения левой и правой частей контура области G (см. рис. 144), на которые он делится точками C и D .

Формулы (10), (11) получены при условии, что граница области G пересекается прямыми, параллельными оси Oy (Ox), не более чем в двух точках. Если это условие нарушено, то область G разбивают на части.

Пример. Найти

$$\iint_G (x + y)dx dy$$

по области G , ограниченной линиями $y = x$, $y = x^2$ (рис. 145). Интегрируя сначала по y , а потом по x , получаем:

$$\begin{aligned} \iint_G (x + y)dx dy &= \int_0^1 dx \int_{x^2}^x (x + y)dy = \int_0^1 \left[xy + \frac{y^2}{2} \right]_{x^2}^x dx = \\ &= \int_0^1 \left(x^2 + \frac{x^2}{2} - x^3 - \frac{x^4}{2} \right) dx = \left(\frac{x^3}{2} - \frac{x^4}{4} - \frac{x^5}{10} \right) \Big|_0^1 = \frac{3}{20}. \end{aligned}$$

Проверить результат можно, изменив порядок интегрирования.

5. Двойной интеграл в полярных координатах. Пусть рассматривается двойной интеграл

$$\iint_G f(x, y)dw,$$

где G — область на плоскости xOy , изображенная на рисунке 146. Как известно, $x = r \cos \varphi$, $y = r \sin \varphi$. Разобьем область G на частичные области посредством координатных линий полярной системы, т. е. линий $r = \text{const}$ и $\varphi = \text{const}$ (рис. 146). Выделим частичную об-

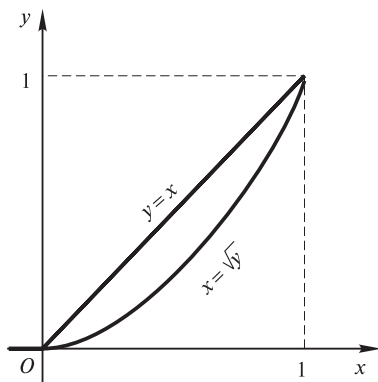


Рис. 145

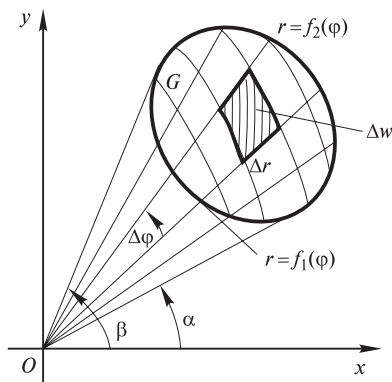


Рис. 146

ласть, для которой центральный угол $\Delta\varphi$ и боковая сторона Δr , а радиус, соответствующий нижнему основанию этой области, r (значит, нижнее основание $r\Delta\varphi$). Эту частичную область, представляющую собой криволинейный четырехугольник, можно принять приближенно за прямоугольник со сторонами Δr и $r\Delta\varphi$. (Это замена с точностью до малых высшего порядка, так как площадь криволинейного четырехугольника будет

$$\frac{(r + \Delta r)^2 \Delta\varphi}{2} - \frac{r^2 \Delta\varphi}{2} = r\Delta r\Delta\varphi + \frac{(\Delta r)^2 \Delta\varphi}{2}.$$

Тогда $\Delta w = r\Delta r\Delta\varphi$, и мы будем иметь:

$$\lim_{\lambda \rightarrow 0} \sum_G f(x_k, y_k) \Delta w_k = \lim_{\lambda \rightarrow 0} \sum f(r_k \cos \varphi_k, r_k \sin \varphi_k) r_k \Delta r_k \Delta \varphi_k$$

$(\sum_G f(x_k, y_k) \Delta w_k)$ — интегральная сумма для $f(x, y)$ по области G) или

$$\iint_G f(x, y) dw = \iint_G f(r \cos \varphi, r \sin \varphi) r dr d\varphi. \quad (12)$$

Переходя к повторному интегралу, получим:

$$\iint_G f(r \cos \varphi, r \sin \varphi) r dr d\varphi = \int_{\alpha}^{\beta} d\varphi \int_{f_1(\varphi)}^{f_2(\varphi)} f(r \cos \varphi, r \sin \varphi) r dr, \quad (13)$$

где смысл пределов интегрирования показан на рисунке 146.

З а м е ч а н и е. Если подынтегральная функция или уравнение границы области интегрирования содержит сумму $x^2 + y^2$, то в большинстве случаев упрощение интеграла достигается преобразованием его к полярным координатам, так как данная сумма в полярных координатах получает весьма простой вид

$$(r \cos \varphi)^2 + (r \sin \varphi)^2 = r^2.$$

П р и м е р. Переходя к полярным координатам, вычислить двойной интеграл

$$I = \iint_G \frac{dxdy}{\sqrt{x^2 + y^2}},$$

где G — первая четверть круга радиуса $R = 1$ с центром в начале координат (рис. 147).

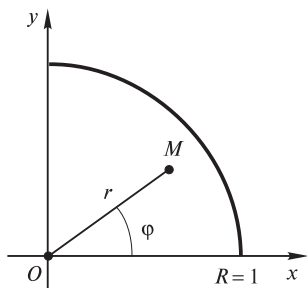


Рис. 147

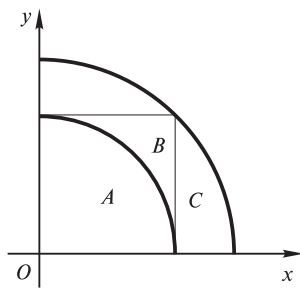


Рис. 148

Имеем:

$$\sqrt{x^2 + y^2} = r.$$

В области G r меняется в пределах от 0 до 1, а φ — от 0 до $\frac{\pi}{2}$. Таким образом, по формулам (12) и (13) получаем:

$$I = \int_0^{\frac{\pi}{2}} d\varphi \int_0^1 dr = \frac{\pi}{2}.$$

6. Интеграл Эйлера — Пуассона. (Симеон Дени Пуассон (1781—1840) — французский математик и физик.) Так называется следующий интеграл:

$$\int_0^{+\infty} e^{-x^2} dx = \lim_{R \rightarrow +\infty} \int_0^R e^{-x^2} dx. \quad (14)$$

Он встречается, например, в теории вероятностей и статистической физике. Докажем, что он сходится, и найдем его величину.

Пусть A — четверть круга радиуса R , B — квадрат со стороной R , содержащий A , и C — четверть круга радиуса $R\sqrt{2}$, содержащая B (рис. 148). Согласно свойствам двойного интеграла имеем:

$$\iint_A e^{-x^2-y^2} dw < \iint_B e^{-x^2-y^2} dw < \iint_C e^{-x^2-y^2} dw. \quad (15)$$

Интегралы по областям A и C вычислим в полярных координатах:

$$\begin{aligned} \iint_A e^{-x^2-y^2} dw &= \int_0^{\frac{\pi}{2}} d\varphi \int_0^R e^{-r^2} r dr = \frac{\pi}{4} (1 - e^{-R^2}), \\ \iint_C e^{-x^2-y^2} dw &= \frac{\pi}{4} (1 - e^{-2R^2}). \end{aligned}$$

Интеграл по области B сведем к квадрату определенного интеграла:

$$\iint_B e^{-x^2-y^2} dw = \int_0^R e^{-x^2} dx \int_0^R e^{-y^2} dy = \left(\int_0^R e^{-x^2} dx \right)^2.$$

Соотношение (15) теперь примет вид:

$$\frac{\pi}{4}(1-e^{-R^2}) < \left(\int_0^R e^{-x^2} dx \right)^2 < \frac{\pi}{4}(1-e^{-2R^2}),$$

откуда

$$\frac{\sqrt{\pi}}{2}\sqrt{1-e^{-R^2}} < \int_0^R e^{-x^2} dx < \frac{\sqrt{\pi}}{2}\sqrt{1-e^{-2R^2}}.$$

Крайние члены этого неравенства при $R \rightarrow +\infty$ стремятся к $\frac{\sqrt{\pi}}{2}$. Поэтому в силу известной теоремы (см. § 7.5, п. 1, теорема 5) $\lim_{R \rightarrow +\infty} \int_0^R e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$.

Следовательно, интеграл (14) сходится, и он равен: $\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$.

7. Замена переменных в двойном интеграле. Пусть в некоторой области G для функции $f(x, y)$ существует двойной интеграл

$$\iint_G f(x, y) dx dy. \quad (16)$$

Предположим, далее, что с помощью формул

$$x = x(u, v), \quad y = y(u, v) \quad (17)$$

мы переходим к новым переменным u и v . Предполагается, что u и v определяются из (17) единственным образом:

$$u = u(x, y), \quad v = v(x, y). \quad (18)$$

Обозначим через G^* множество точек $M^*(u, v)$. С помощью формул (18) каждой точке $M(x, y)$ из области G ставится в соответствие некоторая точка $M^*(u, v)$ из области G^* . Формулы (18) называют *формулами преобразования координат*, а формулы (17) — *формулами обратного преобразования*.

При сделанных предположениях доказывается (см. [2]), что если функции (17) имеют в области G^* непрерывные частные производные первого порядка и если выражение

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \quad (19)$$

отлично в G^* от нуля, то для интеграла (16) справедлива формула замены переменных:

$$\iint_G f(x, y) dx dy = \iint_{G^*} f(x(u, v), y(u, v)) |J| du dv. \quad (20)$$

Выражение (19) называется *функциональным определителем* функций $x = x(u, v)$ и $y = y(u, v)$ или *якобианом* (но имени немецкого математика Карла Якоби (1804—1851)).

В частности, при переходе к полярным координатам имеем: $x = r \cos \varphi$, $y = r \sin \varphi$,

$$J = \begin{vmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{vmatrix} = r(\cos^2 \varphi + \sin^2 \varphi) = r.$$

Таким образом,

$$\iint_G f(x, y) dx dy = \iint_{G^*} f(r \cos \varphi, r \sin \varphi) r dr d\varphi.$$

Эта формула и была, по существу, установлена в п. 5 непосредственно для полярных координат. Однако там несколько упростив рассуждения, мы не рассматривали новой области интегрирования G^* и поэтому формулу преобразования интеграла записали в условном виде (см. формулу (12) из п. 5). Практически, переходя к полярным координатам, не строя область G^* , а находят нужные пределы интегрирования по r и φ , совмещая декартову и полярную системы координат и используя вид области G , а также геометрический смысл полярных координат. Так мы и поступили в примере из п. 5.

В двойном интеграле, как и в определенном, замена переменных — важнейший способ приведения двойного интеграла к виду, более удобному для его вычисления. Один из примеров, показывающий это, был рассмотрен в п. 5. Приведем еще такой пример.

Пусть $I = \iint_G (x + y) dx dy$, где G — параллелограмм, ограниченный прямыми

$$x + y = 1, \quad x + y = 2, \quad 2x - y = 1, \quad 2x - y = 3 \quad (\text{рис. 149, а}),$$

было бы затруднительным, так как для его вычисления необходимо область G разбить на три области и вычислять соответственно три интеграла.

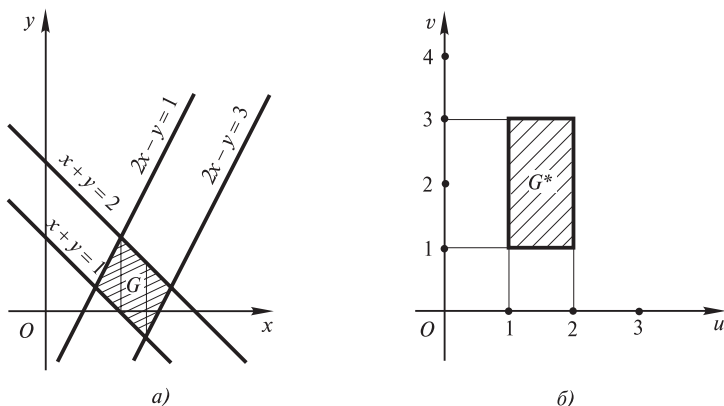


Рис. 149

Положим

$$x + y = u, \quad 2x - y = v,$$

откуда

$$x = \frac{u+v}{3}, \quad y = \frac{2u-v}{3}.$$

Следовательно,

$$J = \frac{1}{3}.$$

При этом данный нам параллелограмм соответствует прямоугольнику G^* в плоскости uOv , ограниченному прямыми $u = 1$, $u = 2$, $v = 1$, $v = 3$ (рис. 149, б).

По формуле (20) получаем:

$$I = \frac{1}{3} \int_1^2 du \int_1^3 u dv = \frac{2}{3} \int_1^2 u du = 1.$$

8. Вычисление площади кривой поверхности. Пусть σ — участок поверхности $z = f(x, y)$, а область G — его проекция на координатную плоскость xOy (рис. 150). Предположим, что функция $f(x, y)$ и ее первые частные производные непрерывны в G вплоть до границы. Требуется найти площадь S поверхности σ (здесь и в § 12.4 так кратко называем участок поверхности).

Разобьем область G на n частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$ без общих внутренних точек, площади которых обозначим соответственно через $\Delta w_1, \Delta w_2, \dots, \Delta w_n$. В каждой из этих частичных областей Δ_k выберем произвольную точку $N_k(\xi_k; \eta_k)$. Этой точке будет соответствовать на поверхности σ точка $M_k(\xi_k; \eta_k; \theta_k)$, где $\theta_k = f(\xi_k, \eta_k)$. Построим в точке M_k касательную плоскость к поверхности σ и нормаль к этой поверхности (рис. 151).

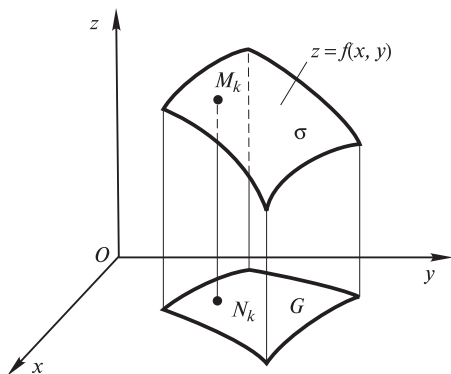


Рис. 150

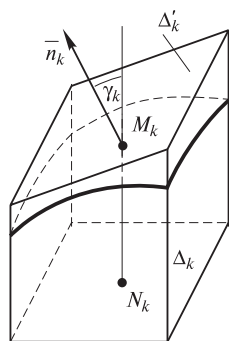


Рис. 151

На касательной плоскости рассмотрим фигуру Δ'_k (площадь ее обозначим через $\Delta w'_k$), которую вырезает из этой плоскости прямой цилиндр с основанием Δ_k . Таким образом, поверхность σ покрывается плоскими пластинками Δ'_k , сумма площадей которых дает приближенное значение площади S поверхности σ , т. е.

$$S \approx \sum_{k=1}^n \Delta w'_k.$$

Известно, что

$$\Delta w_k = \Delta w'_k \cos \gamma_k$$

(для многоугольников это соотношение доказывается в элементарной геометрии; оно остается верным и для произвольной плоской фигуры; [15], т. II), или

$$\Delta w'_k = \frac{\Delta w_k}{\cos \gamma_k},$$

где γ_k — острый угол между касательной плоскостью и плоскостью xOy . Этот угол равен углу между перпендикулярами к ним, т. е. углу между нормалью к поверхности σ в точке M_k и осью Oz . Поэтому (см. § 2.2, п. 3; § 11.2, п. 5)

$$\cos \gamma_k = \frac{1}{\sqrt{1 + f_x'^2(\xi_k, \eta_k) + f_y'^2(\xi_k, \eta_k)}}.$$

Следовательно,

$$\Delta w'_k = \sqrt{1 + f_x'^2(\xi_k, \eta_k) + f_y'^2(\xi_k, \eta_k)} \Delta w_k,$$

и потому

$$\sum_{k=1}^n \Delta w'_k = \sum_{k=1}^n \sqrt{1 + f_x'^2(\xi_k, \eta_k) + f_y'^2(\xi_k, \eta_k)} \Delta w_k. \quad (21)$$

По определению площадью S поверхности σ называется предел последней суммы, когда наибольший из диаметров частичных областей Δ_k стремится к нулю. Так как в правой части равенства (21) стоит интегральная сумма для непрерывной функции

$$\sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)},$$

то предел суммы (21) существует и, значит,

$$S = \iint_G \sqrt{1 + f_x'^2(x, y) + f_y'^2(x, y)} dw.$$

9. Приложения двойного интеграла в механике. Выше (см. п. 1) уже отмечались приложения двойного интеграла в геометрии и некоторых разделах физики, укажем еще на его приложения в механике.

Статические моменты и центр тяжести пластинки. Начнем с вычисления статических моментов рассматривавшейся выше пластинки (см. п. 1) относительно осей координат. Для этого сосредоточим в точках $N_k(\xi_k; \eta_k)$ массы соответствующих частичных областей и найдем статические моменты полученной системы материальных точек (см. § 9.10):

$$M_x^{(n)} = \sum_{k=1}^n \eta_k \rho(\xi_k, \eta_k) \Delta w_k, \quad M_y^{(n)} = \sum_{k=1}^n \xi_k \rho(\xi_k, \eta_k) \Delta w_k.$$

Переходя здесь к пределу при обычных условиях, получим подобно случаю, рассмотренному в § 9.10,

$$M_x = \iint_G y \rho(x, y) dw, \quad M_y = \iint_G x \rho(x, y) dw.$$

Наконец, по определению центра тяжести имеем:

$$x_c = \frac{M_y}{m}, \quad y_c = \frac{M_x}{m},$$

где m — масса пластинки.

Отсюда в случае однородной пластинки

$$x_c = \frac{1}{S} \iint_G x dw, \quad y_c = \frac{1}{S} \iint_G y dw.$$

Моменты инерции пластины. Моментом инерции материальной точки P с массой m относительно какой-либо оси называется произведение массы на квадрат расстояния от точки P до этой оси.

Метод составления выражений для моментов инерции пластины относительно осей координат такой же, какой применялся для вычисления статических моментов. Поэтому приведем лишь окончательные результаты:

$$J_x = \iint_G y^2 \rho(x, y) dw, \quad J_y = \iint_G x^2 \rho(x, y) dw.$$

Отметим еще, что интеграл $\iint_G x y \rho(x, y) dw$ называется *центробежным моментом инерции* и обозначается J_{xy} .

В механике часто рассматривают *полярный момент инерции* точки, равный произведению массы точки на квадрат расстояния от нее до данной точки — полюса. Полярный момент инерции пластины относительно начала координат будет равен:

$$J_0 = \iint_G (x^2 + y^2) \rho(x, y) dw = J_x + J_y.$$

§ 12.2. Тройные интегралы

Тройной интеграл является полным аналогом двойного интеграла. Поэтому изложение этого параграфа будем вести по возможности кратко.

1. Задача о массе неоднородного тела. Тройной интеграл. Пусть пространственная (трехмерная) область Ω заполнена веществом с известной плотностью $\rho(M) = \rho(x, y, z)$. Найти массу всей материальной области — «тела». Разобьем область Ω произвольным образом на n частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$ без общих внутренних точек, объемы которых обозначим соответственно через $\Delta v_1, \Delta v_2, \dots, \Delta v_n$. Предположим, что в пределах каждой частичной области плотность постоянна и равна $\rho(N_k)$ для частичной области Δ_k , где $N_k(\xi_k, \eta_k, \theta_k)$ — произвольная точка этой частичной области. Тогда масса Δ_k приближенно равна $\rho(\xi_k, \eta_k, \theta_k)\Delta v_k$. Для массы всего тела получим приближенное выражение

$$m \approx \sum_{k=1}^n \rho(\xi_k, \eta_k, \theta_k) \Delta v_k.$$

Эта сумма будет тем точнее выражать искомую массу m , чем меньше будет каждый из диаметров частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$. (Определение диаметра трехмерной области такое же, как и определение диаметра плоской области (см. § 12.1, п. 1).) За массу m естественно принять

$$m = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n \rho(\xi_k, \eta_k, \theta_k) \Delta v_k,$$

где λ — наибольший из диаметров частичных областей $\Delta_1, \Delta_2, \dots, \Delta_n$.

К вычислению подобного рода пределов приводят и другие задачи. Поэтому будем рассматривать выражение

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\xi_k, \eta_k, \theta_k) \Delta v_k, \quad (1)$$

где $f(x, y, z)$ — функция, заданная в области Ω .

О п р е д е л е н и е. Если существует предел (1), не зависящий от способа разбиения области Ω на частичные области Δ_k и выбора точек $N_k(\xi_k, \eta_k, \theta_k)$ в них, то он называется *тройным интегралом* от функции $f(x, y, z)$ по области Ω и обозначается символом

$$\iiint_{\Omega} f(x, y, z) dv = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(\xi_k, \eta_k, \theta_k) \Delta v_k. \quad (2)$$

Функция $f(x, y, z)$ в этом случае называется *интегрируемой* в области Ω .

Терминология для тройных интегралов аналогична соответствующей терминологии для двойных интегралов. Так же формулируется и теорема существования тройного интеграла.

Из рассмотренной выше задачи о массе неоднородного тела и определения тройного интеграла следует, что тройной интеграл (2) с положительной подынтегральной функцией может быть истолкован *физически* как масса соответствующего тела. В частности, тройной интеграл от единичной функции ($f(x, y, z) \equiv 1$) численно равен объему области интегрирования:

$$\iiint_{\Omega} dv = V_{\Omega}. \quad (3)$$

Свойства двойных интегралов, перечисленные в п. 3 § 12.1, полностью переносятся на тройные интегралы. Заметим лишь, что

$$\iiint_{\Omega} f(x, y, z) dv = f(\xi, \eta, \theta) V_{\Omega},$$

где (ξ, η, θ) — некоторая точка области Ω (*теорема о среднем*).

2. Вычисление тройных интегралов. Вычисление тройного интеграла так же, как и двойного, может быть сведено к ряду однократных интегрирований. Пусть для простоты область Ω есть тело, ограниченное сверху поверхностью $z = z_2(x, y)$, снизу поверхностью $z = z_1(x, y)$, а с боков цилиндрической поверхностью с образующими, параллельными оси Oz (рис. 152). Тогда, подобно формуле (8) (для двойного интеграла) из § 12.1, имеем следующую формулу:

$$\iiint_{\Omega} f(x, y, z) dv = \iint_G \left(\int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z) dz \right) dw,$$

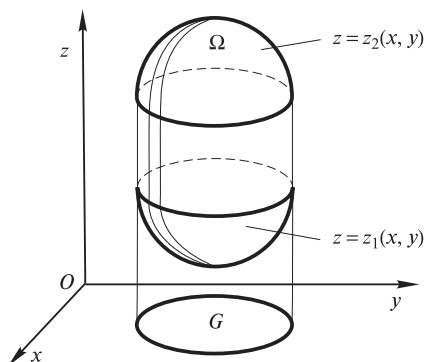


Рис. 152

где G — проекция области Ω на плоскость xOy . Здесь во внутреннем определенном интеграле x и y считаются постоянными. После того как этот внутренний интеграл будет вычислен, получим выражение, зависящее от x и y . Эту функцию от двух переменных надо затем проинтегрировать по плоской области G . Двойной же интеграл, как установлено в § 12.1 (п. 4), сводится к двум определенным интегралам.

Пр и м е р. Вычислить тройной интеграл

$$I = \iiint_{\Omega} (x + y + z) dv,$$

где Ω — область, ограниченная плоскостями $x=0$, $y=0$, $z=0$, $x+y+z=1$ (рис. 153).

Интегрирование по z совершается от $z=0$ до $z=1-x-y$. Поэтому, обозначая проекцию области Ω на плоскость xOy через G , получим:

$$\begin{aligned} I &= \iint_G \left(\int_0^{1-x-y} (x+y+z) dz \right) dw = \iint_G \left((x+y)z + \frac{z^2}{2} \right) \bigg|_0^{1-x-y} dw = \\ &= \iint_G \left((x+y) - (x+y)^2 + \frac{(1-x-y)^2}{2} \right) dw. \end{aligned}$$

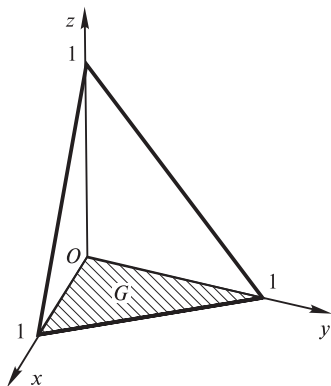


Рис. 153

Теперь, учитывая, что G — треугольник, ограниченный прямыми $x=0$, $y=0$, $x+y=1$, имеем:

$$I = \int_0^1 dx \int_0^{1-x} \left((x+y) - (x+y)^2 + \frac{(1-x-y)^2}{2} \right) dy = \frac{1}{8}.$$

3. Тройной интеграл в цилиндрических и сферических координатах.

Как и в двумерном случае, для тройных интегралов имеют место формулы преобразования интеграла от прямоугольных координат к новым системам координат. Наиболее употребительные из них — цилиндрические и сферические координаты.

Ц и л и н д р и ч е с к и е к о о р д и н а т ы. В этой системе координат положение точки M пространства определяется полярными координатами r и φ точки M' — проекции точки M на плоскость xOy — и аппликатой самой точки M (рис. 154). Числа r , φ и z называются *цилиндрическими координатами* точки M , причем $r \geq 0$, $0 \leq \varphi < 2\pi$ и z — любое действительное число. Из рисунка 154 видно, что цилиндрические координаты r , φ и z связаны с прямоугольными соотношением

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z.$$

Вопрос о преобразовании тройного интеграла к цилиндрическим координатам решается таким же путем, как и преобразование двойного интеграла к полярным координатам. Формула перехода для тройного интеграла от прямоугольных координат к цилиндрическим координатам имеет вид:

$$\iiint_{\Omega} f(x, y, z) dv = \iiint_{\Omega} f(r \cos \varphi, r \sin \varphi, z) r dr d\varphi dz.$$

Вычисление тройного интеграла в цилиндрических координатах приводится к трем однократным интегрированиям по z , r и φ на основании тех же принципов, что и в случае прямоугольных координат.

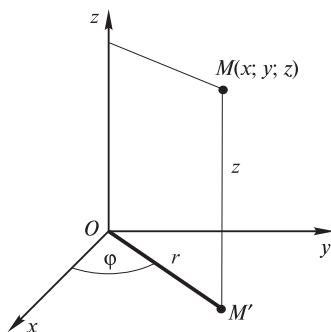


Рис. 154

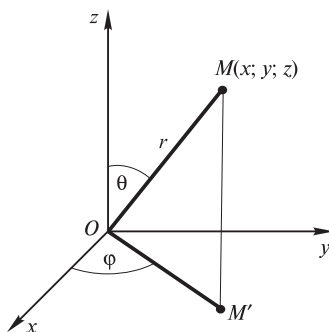


Рис. 155

П р и м е р 1. Найти объем кругового цилиндра высотой H с радиусом основания R .

Используя формулу (3), получим: $V = \int_0^{2\pi} d\varphi \int_0^R r dr \int_0^H dz = \pi R^2 H$.

С ф е р и ч е с к и е к о о р д и н а т ы. В этом случае положение точки M в пространстве определяется ее расстоянием r от начала O , углом φ между положительным направлением оси Ox и отрезком OM' — проекцией отрезка OM на плоскость xOy , углом θ между положительным направлением оси Oz и отрезком OM (рис. 155). Числа r , φ и θ называются *сферическими координатами* точки M или *полярными координатами* в пространстве, при этом $r \geq 0$, $0 \leq \varphi < 2\pi$, $0 \leq \theta \leq \pi$. Из рисунка 155 видно, что сферические координаты r , φ и θ связаны с прямоугольными координатами соотношениями

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta.$$

Подобно тому как была установлена формула перехода в двойном интеграле от прямоугольных координат к полярным, устанавливается формула перехода в тройном интеграле от прямоугольных координат к сферическим:

$$\iiint_{\Omega} f(x, y, z) dv = \iiint_{\Omega} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta dr d\theta d\varphi. \quad (4)$$

Вычисление последнего интеграла также приводится к трем однократным интегрированиям по r , φ и θ .

Если в формуле (4) $f(x, y, z) = 1$, то в силу (3) получаем формулу для объема тела Ω в сферических координатах:

$$V_{\Omega} = \iiint_{\Omega} r^2 \sin \theta dr d\theta d\varphi; \quad (5)$$

при этом выражение $r^2 \sin \theta dr d\theta d\varphi$ называется *элементом объема в сферических координатах*.

Пример 2. Найти объем шара радиуса R .

Используя формулу (5), получим: $V = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta \int_0^R r^2 dr = 2\pi \cdot 2 \cdot \frac{R^3}{3} = \frac{4}{3} \pi R^3$.

4. Приложения тройного интеграла в механике. Выше (см. п. 1) уже отмечались применения тройного интеграла в геометрии и физике. Укажем еще на его приложения в механике.

Для вычисления координат центра тяжести тела нужны статические моменты относительно координатных плоскостей xOy , xOz , yOz (обозначим их соответственно через M_{xy} , M_{xz} , M_{yz}). Повторяя рассуждения п. 9 § 12.1, получим следующие формулы для координат x_C , y_C , z_C центра тяжести тела, занимающего область Ω :

$$x_C = \frac{M_{yz}}{m}, \quad y_C = \frac{M_{xz}}{m}, \quad z_C = \frac{M_{xy}}{m},$$

где

$$M_{yz} = \iiint_{\Omega} x \rho dv, \quad M_{xz} = \iiint_{\Omega} y \rho dv, \quad M_{xy} = \iiint_{\Omega} z \rho dv,$$

m — масса тела, $\rho = \rho(x, y, z)$ — плотность тела.

Отсюда в случае однородного тела

$$x_C = \frac{1}{V} \iiint_{\Omega} x dv, \quad y_C = \frac{1}{V} \iiint_{\Omega} y dv, \quad z_C = \frac{1}{V} \iiint_{\Omega} z dv,$$

где V — объем тела.

Пример 1. Найти центр тяжести однородного ($\rho=1$) полушара Ω :

$$x^2 + y^2 + z^2 \leq R^2, \quad z \geq 0.$$

В силу симметрии заключаем, что $x_C = y_C = 0$. Далее имеем:

$$m = \frac{2}{3} \pi R^3, \quad \iiint_{\Omega} z dv = \int_0^{\frac{\pi}{2}} \sin\theta \cos\theta d\theta \int_0^{2\pi} d\varphi \int_0^R r^3 dr = \frac{1}{2} \cdot 2\pi \cdot \frac{R^4}{4} = \frac{1}{4} \pi R^4.$$

Поэтому

$$z_C = \frac{3}{8} R.$$

Перейдем к вычислению моментов инерции тела относительно координатных осей. Так как квадраты расстояний от точки $M(x, y, z)$ до осей Ox , Oy , Oz соответственно равны $y^2 + z^2$, $x^2 + z^2$, $x^2 + y^2$, то получим следующие формулы:

$$J_x = \iiint_{\Omega} (y^2 + z^2) \rho dv, \quad J_y = \iiint_{\Omega} (x^2 + z^2) \rho dv, \quad J_z = \iiint_{\Omega} (x^2 + y^2) \rho dv.$$

Аналогично плоскому случаю интегралы

$$J_{xy} = \iiint_{\Omega} xypdv, \quad J_{yz} = \iiint_{\Omega} yzpdv, \quad J_{zx} = \iiint_{\Omega} zxpdv$$

называются *центробежными моментами инерции*.

Для полярного момента инерции формула имеет вид:

$$J_0 = \iiint_{\Omega} (x^2 + y^2 + z^2)pdv.$$

Следовательно,

$$2J_0 = J_x + J_y + J_z.$$

Пример 2. Для однородного ($\rho = 1$) шара $x^2 + y^2 + z^2 \leq R^2$ имеем:

$$J_0 = \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta \int_0^R r^2 \cdot r^2 dr = \frac{4\pi R^5}{5}.$$

§ 12.3. Криволинейные интегралы

1. Задачи, приводящие к криволинейным интегралам.

Задача о массе материальной линии. Пусть вдоль некоторой гладкой кривой AB распределена масса с переменной линейной плотностью $\rho(M)$, где M — любая точка кривой AB ($\rho(M)$ — предел средней плотности распределения вещества на бесконечно малой дуге, содержащей точку M). Требуется определить массу m дуги AB .

Для решения задачи раздробим дугу AB на n произвольных частей и вычислим приближенно массу каждой части, предполагая, что на каждой из них плотность постоянна и равна $\rho(N_k)$ для k -й части, где N_k — одна из точек этой части, безразлично какая. Тогда масса k -й части приближенно равна $\Delta m_k \approx \rho(N_k)\Delta l_k$ ($k = 1, 2, \dots, n$), а масса всей дуги приближенно равна $m \approx \sum_{k=1}^n \rho(N_k)\Delta l_k$, где Δl_k — длина k -й части.

В пределе при $\lambda \rightarrow 0$ ($\lambda = \max \Delta l_k$) получим точное значение массы всей дуги AB , т. е.

$$m = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n \rho(N_k)\Delta l_k.$$

Задача о площади цилиндрической поверхности. Пусть в плоскости xOy дана некоторая гладкая кривая AB и на этой кривой определена непрерывная функция $f(M) = f(x, y) \geq 0$. (Непрерывность $f(M)$ вдоль кривой AB означает, что в любой точке M_0 этой кривой $\lim_{M \rightarrow M_0} f(M) = f(M_0)$, где M также точка этой кривой.) Тогда точки пространства $(x, y, f(x, y))$ в совокупности составят некоторую

кривую, лежащую на цилиндрической поверхности, для которой кривая AB — направляющая, а образующая перпендикулярна к плоскости xOy . Требуется определить площадь части поверхности, которая ограничена сверху кривой $z = f(x, y)$, снизу кривой AB , а с боков прямыми AA' и BB' (рис. 156).

Произвольным образом разобьем дугу AB на n частей точками $A = M_0, M_1, M_2, \dots, M_n = B$.

Из каждой точки дробления M_k ($k = 1, 2, \dots, n-1$) проведем перпендикуляры к плоскости xOy высотой $f(M_k)$. В результате вся цилиндрическая поверхность разобьется на n полосок.

Каждую такую полоску заменим прямоугольником с основанием Δl_k , где Δl_k — длина дуги $M_{k-1}M_k$ ($k = 1, 2, \dots, n$), и высотой, равной значению функции $f(N_k)$, где N_k — одна из точек дуги $M_{k-1}M_k$, безразлично какая. На рисунке 156 в целях его упрощения в качестве такой точки взята точка M_{k-1} . Тогда площадь k -й полоски будет приближенно равна: $S_k \approx f(N_k)\Delta l_k$, а площадь всей поверхности $AA'B'B$

$$S \approx \sum_{k=1}^n f(N_k) \Delta l_k.$$

При $\lambda \rightarrow 0$ в пределе получим точное значение искомой площади:

$$S = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(N_k) \Delta l_k.$$

Задача о работе силы. В § 9.6 была рассмотрена задача о работе переменной силы при движении материальной точки по прямой линии, причем направление силы совпадало с направлением движения. Сейчас мы рассмотрим более общую задачу.

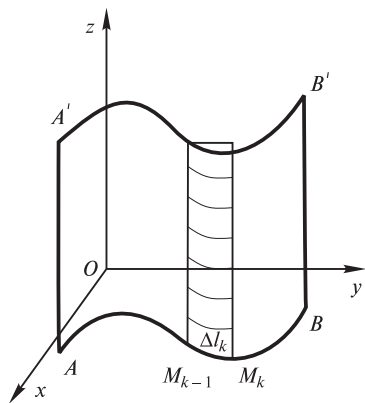


Рис. 156

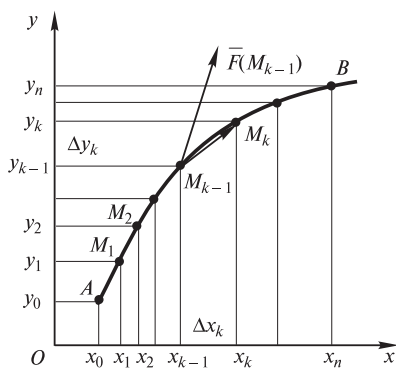


Рис. 157

Пусть материальная точка под действием силы $\vec{F}$ перемещается вдоль непрерывной плоской кривой AB в направлении от A к B . Сила $\vec{F}$ предполагается переменной, зависящей от положения точки на кривой AB . Вычислим работу силы $\vec{F}$, затраченную на перемещение точки из A в B . С этой целью разобьем произвольно точками $A = M_0, M_1, M_2, \dots, M_n = B$ дугу AB на n частичных дуг $M_0M_1, M_1M_2, \dots, M_{n-1}M_n$ с длинами $\Delta l_1, \Delta l_2, \dots, \Delta l_n$ (рис. 157). Наибольшую из длин Δl_k ($k=1, 2, \dots, n$) обозначим λ . Ввиду малости Δl_k можно приближенно принять, что: а) вектор силы $\vec{F}$ сохраняет на дуге $M_{k-1}M_k$ постоянное значение, равное $\vec{F}(N_k)$, где N_k — одна из точек элемента $M_{k-1}M_k$, безразлично какая (на рисунке 157 в целях его упрощения в качестве такой точки взята точка M_{k-1}); б) дуга $M_{k-1}M_k$ может быть заменена хордой $M_{k-1}M_k$, стягивающей концы этого элемента. Вектор $\overline{M_{k-1}M_k}$ равен приращению радиуса-вектора $\vec{r}(M)$: $\Delta \vec{r}_k = \vec{r}(M_k) - \vec{r}(M_{k-1}) (\Delta \vec{r}_k(x_k - x_{k-1}; y_k - y_{k-1}))$. Тогда на элементе дуги $M_{k-1}M_k$ работа силы $\vec{F}$ приближенно равна

$$\vec{F}(N_k) \Delta \vec{r}_k.$$

Пусть вектор $\vec{F}(M)$ имеет проекции $P(M), Q(M)$ соответственно на оси Ox, Oy . Тогда работа силы $\vec{F}$ вдоль всей дуги AB будет приближенно равна

$$\sum_{k=1}^n (P(N_k) \Delta x_k + Q(N_k) \Delta y_k), \quad (1)$$

где $\Delta x_k = x_k - x_{k-1}, \Delta y_k = y_k - y_{k-1}$.

Перейдя в сумме (1) к пределу при $\lambda \rightarrow 0$, получим точное выражение работы силы $\vec{F}$ вдоль всей дуги AB :

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n (P(N_k) \Delta x_k + Q(N_k) \Delta y_k). \quad (2)$$

2. Определение криволинейных интегралов, их свойства. Из решения первых двух задач (п. 1) видно, что хотя они имеют различный смысл, но математический аппарат для их решения один и тот же. В этих двух задачах получаем выражение одного и того же вида

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(N_k) \Delta l_k. \quad (3)$$

О п р е д е л е н и е. Если существует предел (3), не зависящий от способа деления дуги AB на частичные дуги и выбора точек N_k , то он называется *криволинейным интегралом первого рода* от функции $f(M)$ по дуге AB и обозначается так:

$$\int_{AB} f(M) dl, \text{ или } \int_{AB} f(x, y) dl.$$

Дуга AB называется *путем интегрирования*, точка A — *начальной*, а точка B — *конечной* точками интегрирования, сумма $\sum_{k=1}^n f(N_k) \Delta l_k$ — *интегральной суммой*.

Рассмотренные в п. 1 первые две задачи показывают: 1) криволинейный интеграл первого рода при $f(M) \geq 0$ ($f(M)$ на дуге AB непрерывна) численно равен площади участка цилиндрической поверхности с образующей, параллельной оси Oz . Снизу этот участок ограничен дугой AB , а сверху — кривой, изображающей подынтегральную функцию $z = f(M)$. В этом состоит геометрический смысл криволинейного интеграла первого рода; 2) криволинейный интеграл

$$\int_{AB} \rho(M) dl$$

($\rho(M)$ — линейная плотность) равен массе m материальной дуги AB . В этом состоит его физический смысл. Из 1) следует, что $\int_{AB} dl$ численно равен длине дуги AB .

Хотя, как будет показано в п. 3, криволинейный интеграл первого рода непосредственно сводится к определенному, между этими понятиями есть и следующее различие. В выражении (3) величины Δl_k обязательно положительны независимо от того, какую точку кривой AB считать начальной, а какую — конечной. Поэтому

$$\int_{AB} f(x, y) dl = \int_{BA} f(x, y) dl.$$

Как установлено в п. 1, решение задачи о работе силы сводится к вычислению предела вида (2). К вычислению подобного рода пределов приводят и другие задачи. Поэтому будем рассматривать выражение

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n (P(N_k) \Delta x_k + Q(N_k) \Delta y_k). \quad (4)$$

(Здесь $P(M)$ и $Q(M)$ — проекции вектор-функции $\vec{a}(M)$, определенной на дуге AB , на оси координат Ox и Oy .)

О п р е д е л е н и е. Если существует предел (4), не зависящий от способа деления дуги AB на частичные дуги и выбора точек N_k , то он называется *криволинейным интегралом второго рода от векторной функции* $\vec{a}(M) = P(M)\vec{i} + Q(M)\vec{j}$ по дуге AB и обозначается

$$\int_{AB} \vec{a} d\vec{r} \text{ или } \int_{AB} Pdx + Qdy.$$

Сумма

$$\sum_{k=1}^n (P(N_k) \Delta x_k + Q(N_k) \Delta y_k) \quad (5)$$

называется *интегральной суммой*.

Физическое истолкование криволинейного интеграла второго рода, например, как следует из рассмотренной задачи в предыдущем пункте, — это работа силы $\vec{a}(M)$ вдоль дуги AB .

Если $Q(x, y) \equiv 0$ ($P(x, y) \equiv 0$), то интеграл второго рода имеет вид:

$$\int_{AB} P(x, y) dx \left(\int_{AB} Q(x, y) dy \right) \quad (6)$$

и называется *криволинейным интегралом по координате $x(y)$* .

В отличие от криволинейного интеграла первого рода криволинейный интеграл второго рода (как непосредственно следует из его определения) зависит от того, в каком направлении (от A к B или от B к A) пробегается кривая AB (кратко L), и меняет знак при изменении направления обхода кривой.

В случае, когда L — замкнутая кривая, т. е. когда точка B совпадает с точкой A , из двух возможных направлений обхода замкнутого контура L условимся называть *положительным* то, при котором область, лежащая внутри этого контура, остается слева по отношению к точке, совершающей обход. Противоположное направление обхода контура L условимся называть *отрицательным*.

Криволинейный интеграл второго рода по замкнутому контуру L , пробегаемому в положительном направлении, часто обозначают символом

$$\oint_L P(x, y) dx + Q(x, y) dy.$$

Рассмотрим радиус-вектор точки M : $\vec{r}(M) = x\vec{i} + y\vec{j}$. Для общего члена суммы (5) имеем:

$$P(N_k) \Delta x_k + Q(N_k) \Delta y_k = \vec{a}(N_k) \Delta \vec{r}_k, \text{ где } \Delta \vec{r}_k = \Delta x_k \cdot \vec{i} + \Delta y_k \cdot \vec{j}.$$

Но $\vec{a}(N_k) \Delta \vec{r}_k = |\vec{a}(N_k)| |\Delta \vec{r}_k| \cos(\vec{a}, \hat{\Delta \vec{r}_k})$. Поэтому $P(N_k) \Delta x_k + Q(N_k) \Delta y_k = |a(N_k)| |\Delta \vec{r}_k| \cos(\vec{a}, \hat{\Delta \vec{r}_k})$. Пользуясь последним соотношением, можно доказать (см. [2]), что имеет место следующая зависимость между криволинейными интегралами первого и второго рода:

$$\int_{AB} \vec{a}(M) d\vec{r} = \int_{AB} a_\tau(M) dl,$$

$\vec{\tau} = \vec{\tau}(M)$ — единичный вектор касательной к дуге AB в точке M и соответствующий направлению дуги от A к B : $a_\tau(M) = |\vec{a}(M)| \cos(\vec{a}, \hat{\vec{\tau}})$ — проекция вектора $\vec{a}(M)$ на эту касательную.

Так же, как в случае определенных интегралов, из определения криволинейных интегралов двух родов устанавливаются следующие

щие три свойства, общие криволинейным интегралам первого и второго рода:

1. Постоянный множитель можно выносить за знак криволинейного интеграла.

2. Криволинейный интеграл от алгебраической суммы конечного числа функций равен такой же сумме криволинейных интегралов от этих функций.

3. Если путь интегрирования разбит на конечное число частей, то криволинейный интеграл по всему пути равен сумме криволинейных интегралов по всем его частям.

Заметим, что кривая AB может быть и замкнутой.

Справедливо и еще одно свойство, общее для криволинейных интегралов первого и второго рода:

Криволинейный интеграл вдоль замкнутого контура не зависит от выбора начальной точки на этом контуре.

Действительно, если принять за начальную точку A (рис. 158), то в силу свойства 3 получим:

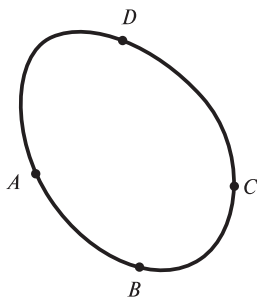


Рис. 158

$$\int_{ABCD} = \int_{ABC} + \int_{CDA} \quad (7)$$

(ради краткости здесь и иногда в дальнейшем подынтегральное выражение не пишем).

Если же за начальную точку принять C , то получим:

$$\int_{CDABC} = \int_{CDA} + \int_{ABC} \quad (8)$$

Из равенств (7) и (8) следует, что $\int_{ABCD} = \int_{CDABC}$.

3. Вычисление криволинейных интегралов первого и второго рода.

Пусть гладкая дуга AB задана параметрически уравнениями $x = x(t)$, $y = y(t)$ ($\alpha \leq t \leq \beta$) и функции $f(x, y)$, $P(x, y)$, $Q(x, y)$ определены и непрерывны на этой дуге.

Для вычисления криволинейного интеграла первого рода представим приращение длины дуги Δl_k в виде интеграла (см. § 9.9, п. 3, формула (7)) и с помощью теоремы о среднем (§ 9.7, п. 4) получим:

$$\Delta l_k = \int_{t_{k-1}}^{t_k} \sqrt{x'^2(t) + y'^2(t)} dt = \sqrt{x'^2(t_k^*) + y'^2(t_k^*)} \Delta t_k,$$

где среднее значение аргумента t_k^* принадлежит промежутку $(t_{k-1}; t_k)$. Выберем в качестве точки N_k дуги $M_{k-1}M_k$ точку N_k^* , соответствующую значению параметра t_k^* . Получим:

$$\sum_{k=1}^n f(N_k^*) \Delta l_k = \sum_{k=1}^n f(x(t_k^*), y(t_k^*)) \sqrt{x'^2(t_k^*) + y'^2(t_k^*)} \Delta l_k.$$

Правая часть этого равенства есть интегральная сумма для функции $f(x(t), y(t)) \sqrt{x'^2(t) + y'^2(t)}$ на сегменте $[\alpha; \beta]$ (§ 9.6, п. 2). Поэтому в результате предельного перехода при $\lambda \rightarrow 0$ получим:

$$\int_{AB} f(x, y) dl = \int_{\alpha}^{\beta} f(x(t), y(t)) \sqrt{x'^2(t) + y'^2(t)} dt. \quad (9)$$

Эта формула одновременно доказывает существование криволинейного интеграла первого рода от непрерывной функции $f(x, y)$ по гладкой дуге AB , если считать существование определенного интеграла от непрерывной функции известным (см. § 9.6, п. 2).

В частности, если дуга AB задана уравнением $y = y(x)$ на отрезке $[a; b]$ (роль параметра t играет величина x), то согласно формуле (9)

$$\int_{AB} f(x, y) dl = \int_a^b f(x, y(x)) \sqrt{1 + y'^2} dx. \quad (10)$$

Для вычисления криволинейного интеграла второго рода (6) представим величину Δx_k с помощью формулы Лагранжа (§ 8.6, п. 3) в виде произведения $\Delta x_k = x'(t_k^*) \Delta t_k$, где t_k^* принадлежит интервалу $(t_{k-1}; t_k)$. Выберем в качестве точки N_k дуги $M_{k-1}M_k$ точку N_k^* , соответствующую значению t_k^* . Интегралу (6) соответствует интегральная сумма

$$\sum_{k=1}^n P(N_k^*) \Delta x_k = \sum_{k=1}^n P(x(t_k^*), y(t_k^*)) x'(t_k^*) \Delta t_k.$$

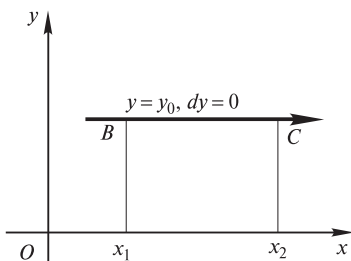
Правая часть этого равенства есть вместе с тем интегральная сумма для функции $P(x(t), y(t)) x'(t)$ на отрезке $[\alpha, \beta]$, (см. § 9.6, п. 2). Поэтому в результате предельного перехода при $\lambda \rightarrow 0$ получим:

$$\int_{AB} P(x, y) dx = \int_{\alpha}^{\beta} P(x(t), y(t)) x'(t) dt. \quad (11)$$

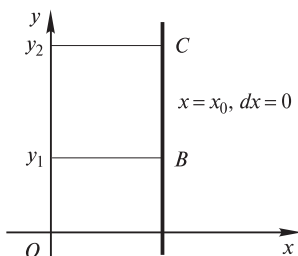
Эта формула одновременно доказывает существование криволинейного интеграла второго рода (6) от непрерывной функции $f(x, y)$ по гладкой дуге AB .

Аналогично выводится формула

$$\int_{AB} Q(x, y) dy = \int_{\alpha}^{\beta} Q(x(t), y(t)) y'(t) dt. \quad (12)$$



а)



б)

Рис. 159

Сложив почленно равенства (11) и (12), получим:

$$\int_{AB} Pdx + Qdy = \int_{\alpha}^{\beta} (Px' + Qy')dt. \quad (13)$$

В частности, если дуга AB задана уравнением $y = y(x)$ на отрезке $[a; b]$, то аналогично формуле (10) из формулы (13) имеем:

$$\int_{AB} P(x, y)dx + Q(x, y)dy = \int_a^b (P(x, y(x)) + Q(x, y(x))y'(x))dx. \quad (14)$$

П р и м е ч а н и е. Если путь интегрирования L — отрезок прямой, параллельной оси абсцисс (рис. 159, а), то криволинейный интеграл второго рода сразу превращается в обыкновенный. Действительно, так как $y = y_0$, и значит, $dy = 0$, то

$$\int_L P(x, y)dx + Q(x, y)dy = \int_{x_1}^{x_2} P(x, y_0)dx.$$

Аналогично если L — отрезок прямой, параллельной оси ординат (рис. 159, б).

П р и м е р 1. Найти массу четверти окружности $x^2 + y^2 = R^2$, $x \geq 0$, $y \geq 0$, если плотность в каждой точке равна ее ординате.

Параметрические уравнения окружности $x^2 + y^2 = R^2$: $x = R \cos t$, $y = R \sin t$. Следовательно, $x^2 + y^2 = R^2$ и в силу формулы (9)

$$m = \int_{AB} \rho dl = \int_{AB} y dl = \int_0^{\frac{\pi}{2}} R^2 \sin t dt = -R^2 \cos t \Big|_0^{\frac{\pi}{2}} = R^2.$$

П р и м е р 2. Вычислить работу силы $\vec{F} = 2xy\vec{i} + x\vec{j}$ при перемещении точки M из положения $A(2; 0)$ в положение $B(-1; 3)$ вдоль прямой AB и вдоль ломаной ACB (рис. 160).

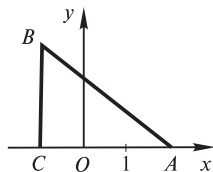


Рис. 160

Задача сводится к вычислению криволинейного интеграла

$$I = \int_{AB} 2xydx + xdy.$$

1) Вдоль прямой AB имеем $y = 2 - x$, $dy = -dx$, и потому

$$I = \int_2^{-1} (2x(2 - x) - x)dx = \frac{3}{2}.$$

2) Вдоль ломаной ACB на участке AC имеем $y = 0$ и $dy = 0$; на участке CB имеем $x = -1$, $dx = 0$. Поэтому

$$I = \int_{AC} + \int_{CB} = 0 - \int_0^3 dy = -3.$$

В заключение заметим, что мы рассмотрели криволинейные интегралы для плоских кривых. Однако все сказанное о них может быть перенесено и на пространственные кривые.

По аналогии со случаем плоской кривой можно определить криволинейный интеграл первого рода

$$\int_{AB} f(x, y, z)dl$$

и криволинейные интегралы второго рода

$$\begin{aligned} & \int_{AB} P(x, y, z)dx, \int_{AB} Q(x, y, z)dy, \int_{AB} R(x, y, z)dz, \\ & \int_{AB} P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz. \end{aligned}$$

Техника вычисления таких интегралов, по существу, ничем не отличается от техники вычисления соответствующих интегралов по плоской кривой.

4. Формула Римана — Грина (Георг Риман (1826—1866) — немецкий математик, Джордж Грин (1793—1841) — известный английский математик и физик).

Пусть функции $P(x, y)$, $Q(x, y)$ (кратко P , Q) непрерывны вместе со своими частными производными P'_y , Q'_x в замкнутой области G , граница L которой пересекается прямыми, параллельными осям координат, не более чем в двух точках (для краткости такие области будем называть *простыми*). Предположим, что контур L гладкий или кусочно-гладкий и может быть задан как уравнениями $y = y_1(x)$, $y = y_2(x)$ ($a \leq x \leq b$), так и уравнениями $x = x_1(y)$, $x = x_2(y)$ ($c \leq y \leq d$) (рис. 161).

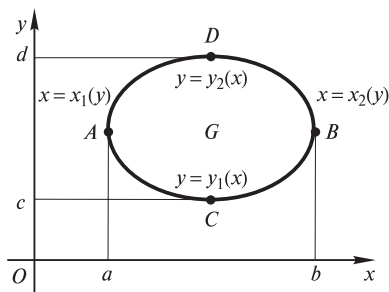


Рис. 161

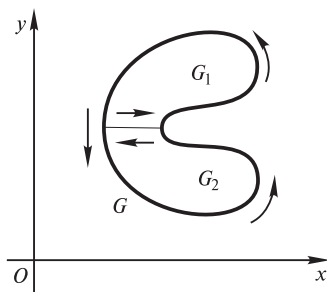


Рис. 162

Рассмотрим интеграл

$$\iint_G \frac{\partial P}{\partial y} dx dy.$$

Представляя его в виде повторного, выполним во внутреннем интеграле по формуле Ньютона— Лейбница (§ 9.7, п. 2) интегрирование по y . Получим:

$$\iint_G \frac{\partial P}{\partial y} dx dy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y} dy = \int_a^b (P(x, y_2(x)) - P(x, y_1(x))) dx.$$

С другой стороны, используя свойство 3 для криволинейного интеграла (п. 2) и формулу (14), имеем:

$$\oint_L P(x, y) dx = \int_a^b (P(x, y_1(x)) - P(x, y_2(x))) dx.$$

Таким образом,

$$\iint_G \frac{\partial P}{\partial y} dx dy = - \oint_L P(x, y) dx. \quad (15)$$

Аналогично устанавливается формула

$$\iint_G \frac{\partial Q}{\partial x} dx dy = \oint_L Q(x, y) dy. \quad (16)$$

Вычитая равенство (15) из равенства (16), получим формулу

$$\iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \oint_L P(x, y) dx + Q(x, y) dy, \quad (17)$$

называемую формулой *Римана — Грина*.

Эта формула устанавливает связь между двойным и криволинейным интегралами. Она имеет широкое применение в математическом анализе и его приложениях.

П р и м е ч а н и е. Формула Римана — Грина остается справедливой и для замкнутой области G , которую можно разбить на конечное число простых областей проведением дополнительных линий (рис. 162).

5. Условия независимости криволинейного интеграла второго рода от пути интегрирования. Пусть функции $P(x, y)$ и $Q(x, y)$ непрерывны вместе со своими производными $\frac{\partial P}{\partial y}$ и $\frac{\partial Q}{\partial x}$ в некоторой односвязной области G (§ 11.1, п. 5) плоскости xOy .

Рассмотрим в области G две произвольные точки A и B . Эти точки можно соединить различными линиями, лежащими в G . Если интеграл

$$\int_{AB} P(x, y)dx + Q(x, y)dy \quad (18)$$

по любому из путей, лежащих в G и соединяющих точки A и B , принимает одно и то же значение, то говорят, что он *не зависит от пути интегрирования* в области G .

В следующих теоремах приводятся условия, при которых криволинейный интеграл (18) не зависит от пути интегрирования.

Т е о р е м а 1. *Для того чтобы криволинейный интеграл (18) в области G не зависел от пути интегрирования, необходимо и достаточно, чтобы*

$$\oint_L P(x, y)dx + Q(x, y)dy = 0, \quad (19)$$

где L — любой замкнутый контур, лежащий в этой области.

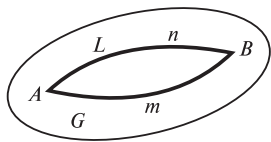


Рис. 163

Д о к а з а т е л ь с т в о. Пусть выполнено условие (19), где L — любой замкнутый контур, лежащий в области G . Соединим произвольные две точки A и B , принадлежащие области G , двумя различными произвольно выбранными кривыми AmB и AnB , лежащими в области G (рис. 163). Дуги AmB и AnB образуют замкнутый контур $AmBnA$. Учитывая свойство 3 криволинейного интеграла (п. 2), получим:

$$\oint_{AmBnA} = \int_{AmB} + \int_{BnA} = \int_{AmB} - \int_{AnB}.$$

В силу формулы (19) $\oint_{AmBnA} = 0$. Следовательно, $\int_{AmB} = \int_{AnB}$, т. е. криволинейный интеграл (18)

не зависит от пути интегрирования.

Обратно: пусть в области G криволинейный интеграл (18) не зависит от пути интегрирования. Рассмотрим произвольный замкнутый контур L , лежащий в области G , и возьмем на нем две произвольные точки A и B (рис. 163). Тогда

$$\oint_{ABnBA} = \int_{AB} + \int_{BA} = \int_{AB} - \int_{AB} = 0,$$

т. е. получаем равенство (19).

Т е о р е м а 2. Для того чтобы криволинейный интеграл (18) не зависел от пути интегрирования в области G , необходимо и достаточно, чтобы в каждой точке этой области выполнялось условие

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}. \quad (20)$$

В главе 11 (см. § 11.3, п. 2) было отмечено, что выполнение условия (20) равносильно тому, что выражение $P(x, y)dx + Q(x, y)dy$ есть полный дифференциал. Отсюда следует:

Т е о р е м а 3. Для того чтобы криволинейный интеграл (18) в области G не зависел от пути интегрирования, необходимо и достаточно, чтобы выражение $P(x, y)dx + Q(x, y)dy$ было в этой области полным дифференциалом.

6. Интегрирование полных дифференциалов. Пусть подынтегральное выражение в криволинейном интеграле

$$\oint_L P(x, y)dx + Q(x, y)dy \quad (21)$$

является полным дифференциалом. Тогда (см. п. 5, теорема 3) величина этого интеграла зависит лишь от начальной и конечной точек линии интегрирования L . Поэтому интеграл (21) записывается обычно в виде

$$\int_{(x_0; y_0)}^{(x_1; y_1)} P(x, y)dx + Q(x, y)dy.$$

Для вычисления такого интеграла воспользуемся примечанием, отмеченным в п. 3, и будем интегрировать по ломаной $M_0M_2M_1$, звенья которой параллельны осям координат (рис. 164). Тогда

$$\int_{(x_0; y_0)}^{(x_1; y_1)} P(x, y)dx + Q(x, y)dy = \int_{x_0}^{x_1} P(x, y_0)dx + \int_{y_0}^{y_1} Q(x_1, y)dy. \quad (22)$$

Если окажется, что ломаные $M_0M_2M_1$, и $M_0M_3M_1$, выходят из области G (рис. 165), то можно производить интегрирование, например, по ломаной $M_0M_4M_5M_1$.

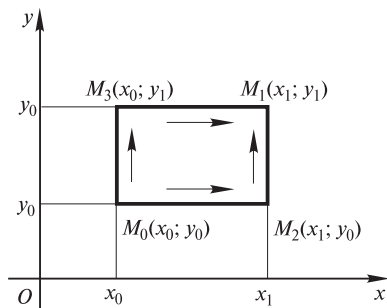


Рис. 164

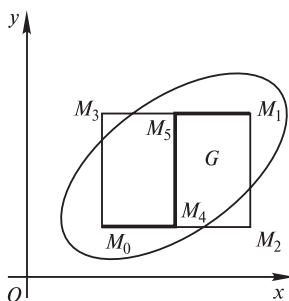


Рис. 165

§ 12.4. Поверхностные интегралы

1. Определение поверхностного интеграла первого рода.

Задача о массе изогнутой пластины. Пусть на поверхности σ непрерывно распределено вещество с известной плотностью $\rho(M)$. При этом под плотностью вещества в точке M поверхности σ понимается предел средней плотности на бесконечно малом элементе, содержащем точку M . Требуется определить всю массу материальной поверхности σ .

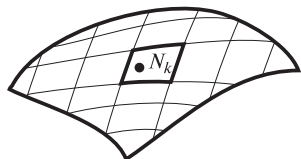


Рис. 166

Разделим поверхность σ (рис. 166) произвольными гладкими линиями на n частей $\sigma_1, \sigma_2, \dots, \sigma_n$ без общих внутренних точек с площадями $\Delta s_1, \Delta s_2, \dots, \Delta s_n$; наибольшую из этих площадей обозначим через λ . Предположим, что в каждой части σ_k плотность постоянна и равна $\rho(N_k)$, где N_k — одна из точек σ_k , безразлично какая. Тогда масса k -го

элемента будет приближенно равна: $\Delta m_k \approx \rho(N_k) \Delta s_k$. Для массы всей поверхности получим приближенное выражение

$$m \approx \sum_{k=1}^n \rho(N_k) \Delta s_k.$$

За массу материальной поверхности (изогнутой пластины) естественно принять предел полученной суммы при стремлении λ к нулю:

$$m = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n \rho(N_k) \Delta s_k.$$

Сформулируем определение поверхностного интеграла первого рода в общем случае. Пусть функция $f(M) = f(x, y, z)$ определена на гладкой

или кусочно-гладкой поверхности σ . (Поверхность называется *гладкой*, если в каждой ее точке существует касательная плоскость и при переходе от точки к точке положение этой касательной плоскости меняется непрерывно; поверхность, состоящая из конечного числа гладких кусков, которые соединены непрерывно, называется *кусочно-гладкой*.) Разделим, как и выше, σ на n частей. Выберем на каждой частичной поверхности σ_k произвольную точку N_k и составим интегральную сумму

$$\sum_{k=1}^n f(N_k) \Delta s_k. \quad (1)$$

О п р е д е л е н и е. Предел интегральной суммы (1) при стремлении λ к нулю (если он существует и не зависит от способа деления σ на частичные поверхности σ_k и от выбора точек N_k) называется *поверхностным интегралом первого рода* от функции $f(M)$ по поверхности σ и обозначается символом

$$\iint_{\sigma} f(M) ds \quad \text{или} \quad \iint_{\sigma} f(x, y, z) ds.$$

Его физическое истолкование при $f(M) > 0$: например, это масса материальной поверхности с плотностью распределения вещества $f(M)$.

Данное определение, по сути дела, аналогично определению двойного интеграла. Поэтому теорема существования двойного интеграла и его свойства (§ 12.1, пп. 2 и 3) без особых изменений переносятся на поверхностные интегралы первого рода.

В частности, если на поверхности σ $f(x, y, z) \equiv 1$, то

$$\iint_{\sigma} ds = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n \Delta s_k = S_{\sigma} \quad (S_{\sigma} \text{ — площадь поверхности } \sigma).$$

2. Вычисление поверхностных интегралов первого рода. Пусть поверхность σ задана уравнением $z = z(x, y)$, где функция $z(x, y)$ вместе с производными $z'_x(x, y)$ и $z'_y(x, y)$ непрерывна в замкнутой области G — проекции σ на плоскость xOy (рис. 167) и пусть функция $f(x, y, z)$ непрерывна на поверхности σ и, следовательно, интегрируема по поверхности σ .

Как и выше (п. 1), разобьем поверхность σ произвольно на n частей $\sigma_1, \sigma_2, \dots, \sigma_n$, не имеющих общих внутренних точек, с площадями $\Delta s_1, \Delta s_2, \dots, \Delta s_n$ и спроектируем это разбиение на плоскость xOy . Получим соответственно разбиение области G на части $\Delta_1, \Delta_2, \dots, \Delta_n$, площади которых обозначим через $\Delta w_1, \Delta w_2, \dots, \Delta w_n$.

Площадь Δs_k каждой части поверхности может быть представлена в виде (см. § 12.1, п. 7):

$$\Delta s_k = \iint_{\Delta_k} \sqrt{1 + z'^2_x(x, y) + z'^2_y(x, y)} dw.$$

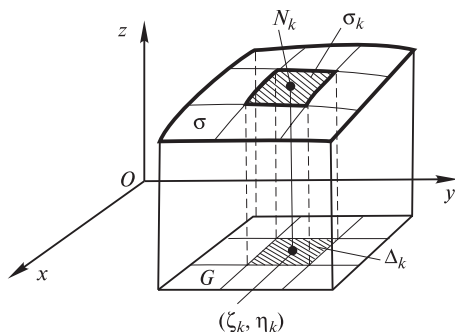


Рис. 167

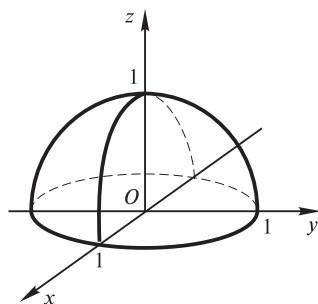


Рис. 168

Применяя здесь теорему о среднем (§ 12.1, п. 3, свойство 5), получим:

$$\Delta s_k = \sqrt{1 + z_x'^2(\xi_k, \eta_k) + z_y'^2(\xi_k, \eta_k)} \Delta w_k. \quad (2)$$

Обозначим через N_k точку на k -й части поверхности σ_k с координатами $(\xi_k; \eta_k; \theta_k)$, где $\theta_k = z(\xi_k, \eta_k)$. Составим интегральную сумму:

$$\begin{aligned} & \sum_{k=1}^n f(\xi_k, \eta_k, \theta_k) \Delta s_k = \\ & = \sum_{k=1}^n f(\xi_k, \eta_k, z(\xi_k, \eta_k)) \sqrt{1 + z_x'^2(\xi_k, \eta_k) + z_y'^2(\xi_k, \eta_k)} \Delta w_k. \end{aligned}$$

Переходя здесь к пределу при $\lambda \rightarrow 0$ (очевидно, при $\lambda \rightarrow 0 \max \Delta w_k \rightarrow 0$), получим (в силу определений двойного и поверхностного интеграла первого рода) формулу

$$\iint_{\sigma} f(x, y, z) ds_k = \iint_G f(x, y, z(x, y)) \sqrt{1 + z_x'^2(x, y) + z_y'^2(x, y)} dx dy. \quad (3)$$

Аналогично получаются формулы, выражающие поверхностный интеграл первого рода по поверхности σ через двойные по ее проекциям на плоскости yOz и xOz .

П р и м е р. Вычислить интеграл

$$I = \iint_{\sigma} \sqrt{1 + 4x^2 + 4y^2} ds,$$

где σ — часть поверхности параболоида вращения $z = 1 - x^2 - y^2$, отсеченная плоскостью $z = 0$ (рис. 168).

Поверхность σ , заданная уравнением $z = 1 - x^2 - y^2$, проектируется на плоскость xOy в область G , ограниченную окружностью $x^2 + y^2 = 1$. Следовательно, область G является круг $x^2 + y^2 \leq 1$. В нем функции

$$z = 1 - x^2 - y^2, \quad z'_x(x, y) = -2x, \quad z'_y(x, y) = -2y$$

непрерывны. По формуле (3) получаем:

$$I = \iint_G \sqrt{1 + 4x^2 + 4y^2} \sqrt{1 + 4x^2 + 4y^2} dx dy = \iint_G (1 + 4x^2 + 4y^2) dx dy.$$

Переходя в последнем интеграле к полярным координатам, находим:

$$I = \int_0^{2\pi} d\varphi \int_0^1 (1 + 4r^2) r dr = 3\pi.$$

3. Определение поверхностного интеграла второго рода. Для того чтобы определить поверхностный интеграл второго рода, необходимо сначала ввести понятие стороны поверхности.

Возьмем на гладкой поверхности σ произвольную точку M , проведем через нее нормаль к поверхности (вектор $\vec{n}$). Проведем теперь на поверхности σ через точку M какой-нибудь замкнутый контур, не имеющий общих точек с границей поверхности σ , и начнем перемещать точку M по замкнутому контуру так, чтобы вектор $\vec{n}$ все время оставался нормальным к σ и чтобы его направление менялось при этом непрерывно (рис. 169). В прежнее положение точка M вернется либо с тем же направлением нормали, либо с прямо противоположным.

Если обход по любому замкнутому контуру, лежащему на поверхности σ и не пересекающему ее границы, при возвращении в исходную точку не меняет направления нормали к поверхности, то поверхность называется *двусторонней* (в противном случае — *односторонней*). Примерами двусторонних поверхностей могут служить плоскость, сфера и т. д.

Мы будем рассматривать лишь двусторонние поверхности.

Для двусторонней поверхности совокупность всех ее точек с выбранным в них направлением нормали, изменяющимся непрерывно при переходе от точки к точке, называется *стороной поверхности*, а

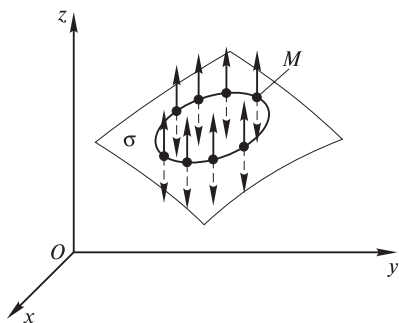


Рис. 169

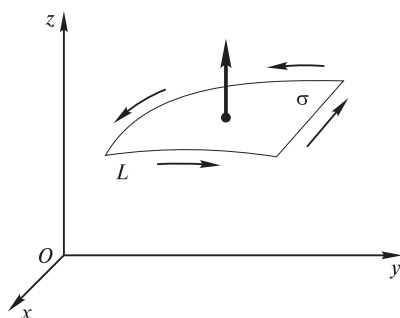


Рис. 170

выбор определенной ее стороны — *ориентацией поверхности*. Двустороннюю поверхность называют также *ориентируемой*, а одностороннюю — *неориентируемой*.

С понятием стороны поверхности тесно связано понятие ориентации ее границы.

Пусть σ — ориентированная (сторона уже выбрана) поверхность, ограниченная контуром L , не имеющим точек самопересечения. Будем считать *положительным* направление обхода контура L (*согласованным с ориентацией σ*) то, при движении по которому сама поверхность остается слева по отношению к точке, совершающей обход (рис. 170). Противоположное направление будем считать *отрицательным*. Если изменить ориентацию поверхности, то положительное и отрицательное направления обхода контура L поменяются ролями.

Задача о потоке жидкости. Пусть пространство заполнено движущейся жидкостью, скорость которой в каждой точке $M(x; y; z)$ задана вектором

$$\vec{v}(M) = P(x, y, z)\vec{i} + Q(x, y, z)\vec{j} + R(x, y, z)\vec{k},$$

где P, Q, R — проекции скорости на координатные оси. (В § 9.11 рассматривалась векторная функция одного скалярного аргумента.)

Пусть P, Q, R — непрерывные функции координат. Вычислим количество жидкости Π , протекающей за единицу времени через некоторую ориентированную поверхность σ (в потоке жидкости σ надо мыслить как воображаемую поверхность, не препятствующую течению), ограниченную пространственной кривой L , считая плотность жидкости $\rho = 1$.

Пусть $\vec{n} = \vec{n}(M) = \vec{i} \cos \alpha + \vec{j} \cos \beta + \vec{k} \cos \gamma$ — единичный вектор нормали к поверхности σ в текущей точке M , и пусть его направляющие косинусы являются непрерывными функциями координат x, y, z точек данной поверхности.

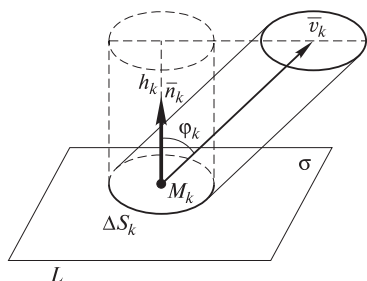


Рис. 171

Разобьем поверхность σ произвольно на n частей, не имеющих общих внутренних точек, с площадями $\Delta s_1, \Delta s_2, \dots, \Delta s_n$ и в каждой из них выберем произвольную точку $M_k(\zeta; \eta_k, \theta_k)$. Подсчитаем количество $\Delta \Pi_k$ жидкости, протекающей за единицу времени через k -ю часть поверхности (рис. 171).

Обозначим через φ_k угол между векторами $\vec{n}_k = \vec{n}(M_k)$ и $\vec{v}_k = \vec{v}(M_k)$. Приближенно можно считать, что

при достаточно мелком разбиении поверхности σ скорость $\bar{v}$ во всех точках k -й части постоянна и равна $\bar{v}(M_k)$, а частичные поверхности плоские. Тогда количество $\Delta\Pi_k$ жидкости, протекающей через k -ю часть за единицу времени в направлении нормали $\bar{n}_k$, приближенно равно объему цилиндра с основанием Δs_k и высотой h_k , равной проекции вектора на нормаль $\bar{n}_k$, т. е. $\Delta\Pi_k \approx \Delta s_k \cdot h_k$. А так как

$$h_k = |\bar{v}_k| \cos \varphi_k = |\bar{v}_k| |\bar{n}_k| \cos \varphi_k = (\bar{v}_k, \bar{n}_k),$$

то

$$\Delta\Pi_k \approx (\bar{v}_k, \bar{n}_k) \Delta s_k.$$

Для количества жидкости, протекающей через поверхность σ за единицу времени, получим приближенное выражение

$$\Pi = \sum_{k=1}^n \Delta\Pi_k \approx \Pi_n, \text{ где } \Pi_n = \sum_{k=1}^n (\bar{v}_k, \bar{n}_k) \Delta s_k$$

или

$$\Pi_n = \sum_{k=1}^n (P(M_k) \cos \alpha_k + Q(M_k) \cos \beta_k + R(M_k) \cos \gamma_k) \Delta s_k. \quad (4)$$

Точное значение этого количества получаем при переходе к пределу в (4) при $\lambda \rightarrow 0$, где λ — наибольший из диаметров частей поверхности σ :

$$\Pi = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n (P(M_k) \cos \alpha_k + Q(M_k) \cos \beta_k + R(M_k) \cos \gamma_k) \Delta s_k. \quad (5)$$

Преобразуем сумму (4). Пусть $(\Delta w_k)_{y_k}$ — площадь проекции k -й части поверхности на плоскость yOz , взятая со знаком «плюс» или «минус» в зависимости от того, образует нормаль $\bar{n}_k$ с осью Ox острый или тупой угол. Имеем (см. § 12.1, п. 8):

$$(\Delta w_k)_{yz} = \Delta s_k \cos \alpha_k. \quad (6)$$

Аналогично

$$(\Delta w_k)_{xz} = \Delta s_k \cos \beta_k, \quad (7)$$

$$(\Delta w_k)_{xy} = \Delta s_k \cos \gamma_k. \quad (8)$$

С учетом равенств (6), (7), (8) сумма (4) и предел (5) принимают соответственно вид:

$$\Pi_n = \sum_{k=1}^n (P(M_k) (\Delta w_k)_{yz} + Q(M_k) (\Delta w_k)_{xz} + R(M_k) (\Delta w_k)_{xy}), \quad (9)$$

$$\Pi = \lim_{\lambda \rightarrow 0} \sum_{k=1}^n (P(M_k) (\Delta w_k)_{yz} + Q(M_k) (\Delta w_k)_{xz} + R(M_k) (\Delta w_k)_{xy}). \quad (10)$$

П р и м е ч а н и е. Отметим, что пределы в правых частях равенств (5) и (10) равны (предполагается, что эти пределы существуют), так как формулы (4) и (9) выражают одну и ту же сумму, но записанную в разных формах.

Перейдем теперь к определению поверхностного интеграла второго рода.

Пусть σ — некоторая ориентируемая поверхность, заданная уравнением $z = f(x, y)$, и пусть $R(x, y, z)$ — функция, определенная в точках поверхности σ . Выберем одну из двух сторон поверхности, т. е. выберем одно из двух возможных направлений векторов нормали в точках поверхности (тем самым мы ориентировали поверхность). Если векторы нормалей составляют острые углы с осью Oz , то будем говорить, что выбрана *верхняя сторона* поверхности $z = f(x, y)$, если тупые углы, то *нижняя сторона* поверхности. Разобьем поверхность σ произвольно на n частей, не имеющих общих внутренних точек. Обозначим через Δ_k проекцию k -й части поверхности на плоскость xOy . Выбрав на каждой частичной поверхности σ_k произвольную точку $N_k(\zeta_k, \eta_k, \theta_k)$, составим сумму

$$\sum_{k=1}^n R(\zeta_k, \eta_k, \theta_k) \Delta w_k, \quad (11)$$

где Δw_k — площадь Δ_k , взятая со знаком «плюс», если выбрана верхняя сторона поверхности σ , и со знаком «минус», если выбрана нижняя сторона поверхности σ . Обозначим через λ наибольший из диаметров частей поверхности σ и дадим следующее определение.

О п р е д е л е н и е. Предел интегральной суммы (11) при $\lambda \rightarrow 0$ (если он существует и не зависит от способа деления σ на частичные поверхности σ_k и выбора точек N_k) называется *поверхностным интегралом второго рода* от функции $R(x, y, z)$ по выбранной стороне поверхности σ и обозначается одним из символов

$$\iint_{\sigma} R(M) dx dy \text{ или } \iint_{\sigma} R(x, y, z) dx dy.$$

Аналогичным образом определяется поверхностный интеграл второго рода:

$$\iint_{\sigma} P(x, y, z) dy dz \left(\iint_{\sigma} Q(x, y, z) dz dx \right).$$

Сумму

$$\iint_{\sigma} P(x, y, z) dy dz + \iint_{\sigma} Q(x, y, z) dz dx + \iint_{\sigma} R(x, y, z) dx dy$$

принято называть *общим поверхностным интегралом второго рода* и обозначать символом

$$\iint_{\sigma} P(x, y, z) dy dz + Q(x, y, z) dz dx + R(x, y, z) dx dy. \quad (12)$$

Поверхностный интеграл второго рода обладает всеми свойствами поверхностного интеграла первого рода, за исключением одного: при изменении стороны поверхности (переориентации) интеграл меняет знак.

Из рассмотренной выше задачи о потоке жидкости следует (см. (10)), что в этой задаче поверхностный интеграл (12) может быть истолкован физически как количество жидкости, протекающее за единицу времени через указанную в этой задаче поверхность σ .

4. Вычисление поверхностных интегралов второго рода. Пусть ориентированная (выберем верхнюю сторону) гладкая поверхность σ задана уравнением $z=f(x, y)$, где $f(x, y)$ — функция, определенная в замкнутой области G — проекции поверхности σ на плоскость xOy , а $R(x, y, z)$ — непрерывная функция на поверхности σ .

Разобьем произвольно поверхность σ на n частей, не имеющих общих внутренних точек, и спроектируем это разбиение на плоскость xOy (рис. 172).

Составим интегральную сумму $\sum_{k=1}^n R(\zeta_k, \eta_k, \theta_k) \Delta w_k$, где Δw_k — площадь Δ_k .

Так как $\theta_k = f(\zeta_k, \eta_k)$, то

$$\sum_{k=1}^n R(\zeta_k, \eta_k, \theta_k) \Delta w_k = \sum_{k=1}^n R(\zeta_k, \eta_k, f(\zeta_k, \eta_k)) \Delta w_k. \quad (13)$$

В правой части равенства (13) стоит интегральная сумма для двойного интеграла от непрерывной в области G функции $R(x, y, f(x, y))$. Переходя к пределу в (13) при $\lambda \rightarrow 0$, получаем:

$$\iint_{\sigma} R(x, y, z) dx dy = \iint_{\sigma} R(x, y, f(x, y)) dx dy. \quad (14)$$

Если выбрать нижнюю сторону поверхности, то перед интегралом в правой части (14) появится знак «минус».

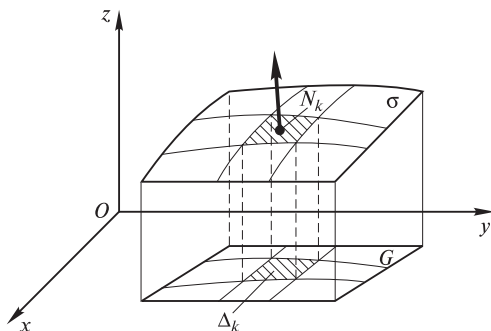


Рис. 172

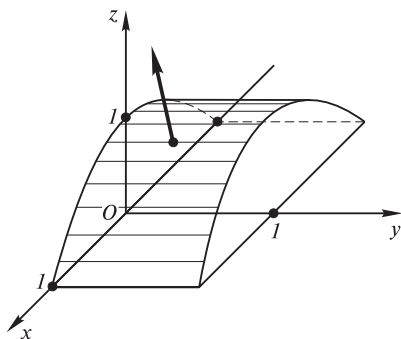


Рис. 173

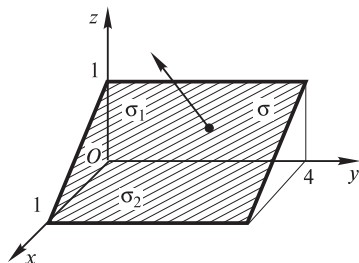


Рис. 174

Аналогично вычисляются интегралы

$$\iint_{\sigma} P(x, y, z) dy dz, \iint_{\sigma} Q(x, y, z) dz dx.$$

Для вычисления интеграла общего вида (12) используются формулы указанных трех интегралов, если поверхность σ однозначно проектируется на все три координатные плоскости.

Пример 1. Вычислить интеграл

$$\iint_{\sigma} (y^2 + z^2) dx dy,$$

где σ — верхняя сторона поверхности $z = \sqrt{1-x^2}$ отсеченная плоскостями $y = 0, y = 1$ (рис. 173).

Проекцией данной поверхности на плоскость xOy является прямоугольник $G: -1 \leq x \leq 1, 0 \leq y \leq 1$. По формуле (14) находим:

$$\iint_{\sigma} (y^2 + z^2) dx dy = \iint_G (y^2 + (\sqrt{1-x^2})^2) dx dy = \int_{-1}^1 dx \int_0^1 (y^2 + 1 - x^2) dy = 2.$$

Пример 2. Вычислить интеграл

$$\iint_{\sigma} x dy dz + y dz dx + z dx dy,$$

где σ — верхняя сторона части плоскости $x + z - 1 = 0$, отсеченной плоскостями $y = 0, y = 4$ и лежащей в первом октанте (рис. 174).

Так как плоскость σ параллельна оси Oy , то $\iint_{\sigma} y dz dx = 0$. Поэтому

$$\begin{aligned} \iint_{\sigma} x dy dz + y dz dx + z dx dy &= \iint_{G_1} (1-z) dy dz + \iint_{G_2} (1-x) dx dy = \\ &= \int_0^4 dy \int_0^1 (1-z) dz + \int_0^4 dy \int_0^1 (1-x) dx = 2 + 2 = 4. \end{aligned}$$

5. Связь между поверхностными интегралами первого и второго рода. Из примечания, отмеченного в п. 3, и определений поверхностных интегралов первого и второго рода следует, что

$$\begin{aligned} \iint_{\sigma} [P(x, y, z)\cos\alpha + Q(x, y, z)\cos\beta + R(x, y, z)\cos\gamma] ds = \\ = \iint_{\sigma} P(x, y, z) dydz + Q(x, y, z) dzdx + R(x, y, z) dxdy. \end{aligned} \quad (15)$$

Глава 13. Дифференциальные уравнения

§ 13.1. Основные понятия о дифференциальных уравнениях

Дифференциальным уравнением называется соотношение, связывающее независимую переменную x , искомую функцию $y=f(x)$ и ее производные. Если искомая функция есть функция одной независимой переменной, то дифференциальное уравнение называется *обыкновенным*. В этой главе мы будем заниматься только обыкновенными дифференциальными уравнениями. Порядок старшей производной, входящей в дифференциальное уравнение, называется *порядком* данного уравнения. Следовательно, общий вид дифференциального уравнения n -го порядка следующий:

$$F(x, y, y', \dots, y^{(n)}) = 0, \quad (1)$$

причем в частных случаях в это уравнение могут и не входить x , y и отдельные производные порядка ниже, чем n . Например, уравнения $y' - \frac{y}{x} = x$, $y'' + y' = 1$ имеют соответственно первый и второй порядок.

Всякая функция $y=f(x)$, которая, будучи подставлена в уравнение (1), обращает его в тождество, называется *решением* этого уравнения.

§ 13.2. Дифференциальные уравнения первого порядка

1. Дифференциальное уравнение первого порядка, его геометрическое истолкование, общее решение и начальные условия. Дифференциальное уравнение первого порядка имеет общий вид

$$F(x, y, y') = 0$$

или (если это уравнение можно разрешить относительно y') вид

$$y' = f(x, y). \quad (1)$$

Будем рассматривать в уравнении (1) переменные x и y как прямоугольные координаты точки на плоскости xOy . Пусть $y = \varphi(x)$ — решение уравнения (1); тогда кривая, определяемая уравнением $y = \varphi(x)$, называется *интегральной кривой* дифференциального уравнения (1). Рассмотрим касательную к интегральной кривой в произвольной точке $M(x; y)$. Согласно геометрическому смыслу производной в этой точке имеем:

$$\frac{dy}{dx} = \operatorname{tg} \alpha,$$

где α — угол наклона этой касательной к оси Ox . Из последнего равенства и из (1) получаем:

$$\operatorname{tg} \alpha = f(x, y),$$

где x, y — координаты точки M . Таким образом, угловой коэффициент касательной к интегральной кривой в каждой ее точке равен значению в этой точке правой части уравнения (1). Итак, уравнение (1) определяет в каждой точке интегральной кривой направление касательной к этой кривой.

Каждой точке $M(x; y)$ той области, где определена функция $f(x, y)$ (правая часть уравнения (1)), сопоставим отрезок с угловым коэффициентом $k = f(x, y)$, где $(x; y)$ — координаты точки M . Мы получаем совокупность направлений, или, как говорят, поле направлений данного дифференциального уравнения.

Таким образом, уравнению (1) соответствует его поле направлений. В этом состоит геометрический смысл дифференциального уравнения первого порядка (1). Проведя указанные выше отрезки для достаточно большого числа точек области, получим наглядное изображение поля направлений. Так как касательная в точке интегральной кривой имеет то же направление, что и отрезок в этой точке, то задачу решения (интегрирования) уравнения (1) геометрически можно истолковать следующим образом: найти такую кривую, чтобы ее касательная в каждой точке имела направление, совпадающее с направлением поля в этой точке.

Приведенные рассуждения хорошо иллюстрировать на известном опыте с железными опилками, помещенными в магнитное поле. Сами опилки образуют поле направлений, а интегральной кривой служит одна из магнитных силовых линий.

Решение уравнения (1), содержащее произвольную постоянную C , т. е. имеющее вид

$$y = \varphi(x, C), \quad (2)$$

называется *общим решением* этого уравнения. Иногда, впрочем, это решение получается в неявной форме $\Phi(x, y, C) = 0$ или $\Psi(x, y) = C$. В этом случае соотношение $\Phi(x, y, C) = 0$ или $\Psi(x, y) = C$ называется *общим интегралом* уравнения (1).

Решить, или *проинтегрировать*, данное дифференциальное уравнение — значит найти его общее решение в той или иной форме.

Решение, которое получается из общего решения при некотором фиксированном значении произвольной постоянной C , называется *частным решением*.

Для уравнения (1) справедлива следующая теорема, называемая теоремой о существовании и единственности решения дифференциального уравнения (1).

Т е о р е м а. Если в уравнении (1) функция $f(x, y)$ и ее частная производная $f'_y(x, y)$ непрерывны в некоторой области D на плоскости xOy , содержащей некоторую точку $(x_0; y_0)$, то существует единственное решение этого уравнения $y = \varphi(x)$, удовлетворяющее условию: при $x = x_0$, $y = y_0$.

Геометрический смысл этой теоремы состоит в том, что существует, и притом единственная, функция $y = \varphi(x)$, график которой проходит через точку $(x_0; y_0)$. (Доказательство ее выходит за рамки настоящей книги; читатель может найти его, например, в книге [12].)

Условие, что при $x = x_0$ функция y должна равняться заданному числу y_0 , называется *начальным условием*. Начальное условие дает возможность выделить из общего решения (2) частное решение.

2. Уравнения с разделяющимися переменными. Запишем уравнение (1) в виде

$$\frac{dy}{dx} = f(x, y), \text{ или } dy = f(x, y)dx.$$

Такому уравнению можно придать следующую форму:

$$M(x, y)dx + N(x, y)dy = 0. \quad (3)$$

Форма (3) удобна тем, что здесь переменные x и y равноправны, т. е. каждую из них можно рассматривать как функцию другой. Предположим, что функции $M(x, y)$ и $N(x, y)$ можно представить произведениями

$$M(x, y) = M_1(x)M_2(y), \quad N(x, y) = N_1(x)N_2(y),$$

в которых сомножители зависят только от одной переменной. Тогда уравнение (3) можно переписать в виде

$$M_1(x)M_2(y)dx + N_1(x)N_2(y)dy = 0, \quad (4)$$

откуда, деля почленно на произведение $M_2(y)N_1(x)$ (предполагаем, что оно не равно нулю), имеем:

$$\frac{M_1(x)}{N_1(x)}dx + \frac{N_2(y)}{M_2(y)}dy = 0. \quad (5)$$

Заметим, что в уравнении (5) множитель перед dx — функция только одной переменной x , а множитель перед dy — функция только одной переменной y .

Уравнение (5) называется *уравнением с разделенными переменными*, а уравнение (4) — *уравнением с разделяющимися переменными*. Итак, уравнение с разделяющимися переменными (4) сводится к уравнению с разделенными переменными путем деления обеих частей уравнения (4) на произведение $M_2(y)N_1(x)$. Эта операция называется *разделением переменных*.

Покажем, что соотношение

$$F(x, y) = C, \quad (6)$$

где

$$F(x, y) = \int \frac{M_1(x)}{N_1(x)}dx + \int \frac{N_2(y)}{M_2(y)}dy,$$

есть общий интеграл уравнения (5) и уравнения (4). Действительно, пусть $y = \varphi(x, C)$ (или, кратко, $y = \varphi$) — функция, определяемая уравнением (6). Тогда имеем тождество

$$F(x, \varphi) \equiv C.$$

Дифференцируя это тождество по x , получим тождество

$$F'_x(x, \varphi) + F'_y(x, \varphi)\varphi' \equiv 0, \text{ или } \frac{M_1(x)}{N_1(x)}dx + \frac{N_2(y)}{M_2(y)}dy \equiv 0.$$

Следовательно, функция $y = y(x, C)$ оказывается общим (поскольку зависит от C) решением уравнения (5), а следовательно, и уравнения (4). Значит, соотношение (6) или соотношение

$$\int \frac{M_1(x)}{N_1(x)}dx + \int \frac{N_2(y)}{M_2(y)}dy = C$$

есть общий интеграл уравнения (5) и уравнения (4).

П р и м е ч а н и е. В общем случае, деля на произведение $M_2(y)N_1(x)$, мы рискуем потерять те решения уравнения (4), которые обращают это произведение в нуль. Непо-

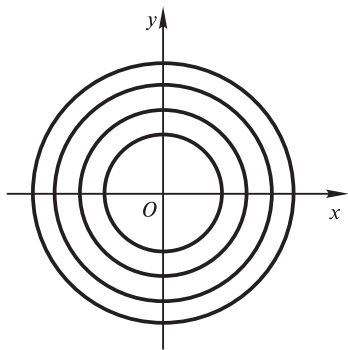


Рис. 175

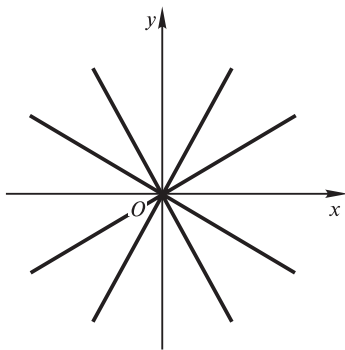


Рис. 176

средственной подстановкой легко убедиться, что функция $y = b$, где b — корень уравнения $M_2(y) = 0$, есть решение уравнения (4). Аналогично функция $x = a$, где a — корень уравнения $N_1(x) = 0$, также является решением уравнения (4).

П р и м е р 1. Решить уравнение $xdx + ydy = 0$.

Интегрируя, находим $\frac{x^2}{2} + \frac{y^2}{2} = C_1$. Так как левая часть последнего равенства неотрицательна, то и правая часть тоже неотрицательна. Обозначив $2C_1$ через C^2 , будем иметь $x^2 + y^2 = C^2$. Это уравнение семейства концентрических окружностей (рис. 175) с центром в начале координат и радиусом C .

П р и м е р 2. Решить уравнение $xdy = ydx$. Разделяя переменные, получим $\frac{dy}{y} = \frac{dx}{x}$. Интегрируя последнее уравнение, будем иметь:

$$\ln y = \ln x + \ln C. \quad (7)$$

Строго говоря, мы должны писать $\ln |y| = \ln |x| + \ln C$, где $C > 0$. Однако допущенная в (7) вольность не отразится на окончательном результате, если после потенцирования произвольную постоянную C считать действительным числом. Это следует иметь в виду для дальнейшего.

В равенстве (7) произвольная постоянная взята в логарифмической форме, что законно, так как всякое положительное или отрицательное число C_1 может быть представлено как логарифм другого числа $C_1 = \ln C$, где $C = e^{C_1}$.

Потенцируя равенство (7), получим общее решение данного дифференциального уравнения: $y = Cx$. Это семейство прямых, проходящих через начало координат (рис. 176).

3. Однородные уравнения. Функция $f(x, y)$ называется *однородной измерения t* , если имеет место тождество

$$f(tx, ty) = t^m f(x, y).$$

Пример 1. Функция $f(x, y) = x^2 + 2y^2 - xy$ является однородной функцией измерения 2, так как

$$(tx)^2 + 2(ty)^2 - (txty) = t^2(x^2 + 2y^2 - xy).$$

С понятием однородной функции связано понятие однородного дифференциального уравнения.

Уравнение

$$M(x, y)dx + N(x, y)dy = 0 \quad (8)$$

называется *однородным дифференциальным уравнением первого порядка*, если функции $M(x, y)$ и $N(x, y)$ являются однородными функциями одного и того же измерения.

Можно показать, что с помощью подстановки $y = ux$, где u — новая искомая функция от x , однородное уравнение (8) легко приводится к уравнению с разделяющимися переменными. Заметим, что $dy = udx + xdu$.

Иногда целесообразно вместо подстановки $y = ux$ использовать подстановку $x = uy$.

Пример 2. Решить уравнение $(y^2 - 3x^2)dx + 2xydy = 0$, если $y = 0$ при $x = 0$. Применяя подстановку $y = ux$, имеем $(u^2x^2 - 3x^2)dx + 2x^2u(udx + xdu) = 0$, откуда $3(u^2 - 1)dx + 2xudu = 0$. Разделяя переменные и интегрируя, получаем:

$$\frac{3dx}{x} + \frac{2udu}{u^2 - 1} = 0, \quad 3 \ln x + \ln(u^2 - 1) = \ln C,$$

что после потенцирования дает $x^3(u^2 - 1) = C$. Так как $u = \frac{y}{x}$, то $x^3\left(\frac{y^2}{x^2} - 1\right) = C$ и общий интеграл $x(y^2 - x^2) = C$. Используя начальное условие, имеем: $C = 0$. Поэтому искомыми частными решениями являются $y = \pm x$.

4. Линейные дифференциальные уравнения первого порядка. Уравнение

$$y' + py = q, \quad (9)$$

где $p = p(x)$ и $q = q(x)$ — заданные непрерывные в интервале $(a; b)$ функции, называется *линейным дифференциальным уравнением первого порядка*. Для решения уравнения (9) применим подстановку $y = uv$, причем функцию $u = u(x)$ будем считать новой неизвестной функцией, а функцию $v = v(x)$ мы выбираем произвольно. Эта подстановка дает $u'v + uv' + puv = q$ или $v \frac{du}{dx} + \left(\frac{dv}{dx} + pv\right)u = q$. Используя произволь-

ный выбор функции v , подчиним ее условию $\frac{dv}{dx} + pv = 0$.

Разделяя переменные и интегрируя, получаем:

$$\frac{dv}{v} = -pdx, \ln v = -\int pdx,$$

откуда:

$$v = e^{-\int pdx}.$$

Поэтому имеем уравнение

$$e^{-\int pdx} \frac{du}{dx} = q.$$

Решая его, получаем:

$$u = \int qe^{\int pdx} dx + C.$$

Возвращаясь к переменной y , находим общее решение уравнения (9):

$$y = e^{-\int pdx} \left[\int qe^{\int pdx} dx + C \right]. \quad (10)$$

П р и м е ч а н и е. Если в уравнении (9) $q(x) \equiv 0$, то оно называется *линейным однородным уравнением первого порядка*, в противном случае — *линейным неоднородным уравнением первого порядка*. Следовательно, линейное однородное уравнение первого порядка имеет вид:

$$y' + py = 0. \quad (11)$$

Из формулы (10) следует формула общего решения уравнения (11):

$$y = Ce^{-\int pdx}.$$

П р и м е р. Решить уравнение $y' - \frac{y}{x} = x$.

Согласно формуле (10) имеем:

$$y = e^{\int \frac{dx}{x}} \left(C + \int x e^{-\int \frac{dx}{x}} dx \right) = e^{\ln x} \left(C + \int x e^{-\ln x} dx \right) = x \left(C + \int dx \right) = (Cx + x^2).$$

§ 13.3. Уравнения высших порядков

1. Основные понятия. Общий вид дифференциального уравнения n -го порядка есть

$$F(x, y, y', \dots, y^{(n)}) = 0. \quad (1)$$

Здесь $F(x, y, y', \dots, y^{(n)})$ может не зависеть от некоторых из величин $x, y, y', \dots$. Однако если (1) есть уравнение именно n -го порядка, то от $y^{(n)}$ функция F обязательно зависит. Наиболее простым дифференциальное уравнение (1) оказывается тогда, когда оно имеет вид:

$$y^{(n)} = f(x), \quad (2)$$

где $f(x)$ — заданная функция.

Примером такого простейшего уравнения служит хотя бы дифференциальное уравнение

$$y'' = -\frac{1}{x^2}. \quad (3)$$

Из этого уравнения сразу видно, что

$$y' = -\int \frac{dx}{x^2} + C_1 = \frac{1}{x} + C_1, \quad (4)$$

где C_1 — произвольная постоянная. В свою очередь из уравнения (4) следует, что

$$y = \int \left(\frac{1}{x} + C_1 \right) dx + C_2 = \ln|x| + C_1 x + C_2,$$

где C_2 — произвольная постоянная, никак не связанная с постоянной C_1 .

Найденное решение зависит от *двух* произвольных постоянных, при этом исходное дифференциальное уравнение (3) было уравнением *второго* порядка. Такое решение называется *общим решением* этого уравнения.

Аналогично посредством *n* последовательных интегрирований решается любое уравнение вида (2). В связи с этим вводится определение.

О п р е д е л е н и е. *Общим решением* дифференциального уравнения *n*-го порядка (1) называется функция

$$y = \varphi(x, C_1, C_2, \dots, C_n),$$

существенно зависящая от *n* произвольных постоянных и обращающая уравнение (1) в тождество при любых значениях этих постоянных. Решения, получаемые из общего при закреплении постоянных $C_1, C_2, \dots, C_n$, называются *частными*.

З а м е ч а н и е. В данном определении употреблено выражение «существенно зависящая». Это означает, что число постоянных нельзя снизить за счет введения новых обозначений. Например, функция

$$y = (C_1^2 + 2C_2 + C_3)x + C_4 + 6C_5$$

существенно зависит лишь от двух постоянных

$$C_1^* = C_1^2 + 2C_2 + C_3 \text{ и } C_2^* = C_4 + 6C_5$$

и может быть записана в виде

$$y = C_1^* x + C_2^*.$$

В прикладных вопросах часто приходится искать такое решение дифференциального уравнения *n*-го порядка, которое удовлетворяет

n условиям: при заданном значении $x = x_0$ сама функция y и ее первые $n - 1$ производных $y', y'', \dots, y^{(n-1)}$ должны принимать заданные значения

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y'_0, \quad \dots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}. \quad (5)$$

Вообще говоря, условия (5), называемые *начальными*, выделяют из общего решения

$$y = \varphi(x, C_1, C_2, \dots, C_n)$$

единственное частное решение.

2. Случай понижения порядка. Рассмотрим три типа дифференциальных уравнений n -го порядка, допускающих понижение порядка.

I. Уравнение вида (2). Общее решение этого уравнения мы получим, произведя последовательно n интегрирований; при каждом таком интегрировании будет появляться новая произвольная постоянная.

II. Уравнение

$$y^{(n)} = f(x, y^{(k)}, \dots, y^{(n-1)}) \quad (k \leq n-1), \quad (6)$$

не содержащее явно y и младших производных до порядка $k - 1$ включительно, допускает понижение порядка на k единиц. Для этого введем новую искомую функцию $z = y^{(k)}$. Тогда

$$y^{(k+1)} = z', \quad \dots, \quad y^{(n)} = z^{(n-k)}$$

и уравнение относительно z будет порядка $n - k$:

$$z^{(n-k)} = f(x, z, z', \dots, z^{(n-k-1)}).$$

Если найдено общее решение этого уравнения $z = \varphi(x, C_1, \dots, C_{n-k})$, то для y имеем уравнение $y^{(k)} = \varphi(x, C_1, \dots, C_{n-k})$. Интегрируя его, найдем общее решение уравнения (6).

III. Уравнение

$$y^{(n)} = f(y, y', \dots, y^{(n-1)}), \quad (7)$$

не содержащее явно независимой переменной. Здесь порядок уравнения понижается на единицу путем замены обеих переменных. В качестве новой искомой функции мы выбираем $y' = p$, а за новую независимую переменную принимаем y .

З а д а ч а. Материальная точка массой m движется по прямой линии к центру O (рис. 177), притягивающему ее с силой $\frac{km}{r^3}$, где r — расстояние от этой точки до центра.

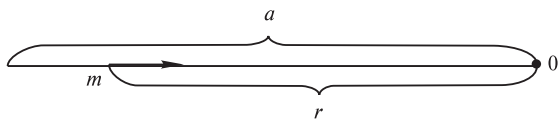


Рис. 177

Движение начинается с состояния покоя при $r = a$. Найти время, за которое материальная точка достигнет центра O .

Решение. По условию задачи в любой момент времени t на материальную точку действует сила $F = \frac{-km}{r^3}$. Отсюда получаем дифференциальное уравнение

$$m \frac{d^2 r}{dt^2} = \frac{-km}{r^3}$$

или

$$\frac{d^2 r}{dt^2} = -\frac{k}{r^3}, \quad (8)$$

Обозначим $\frac{dr}{dt} = p$. Тогда $\frac{d^2 r}{dt^2} = \frac{dp}{dt} = \frac{dp}{dr} \frac{dr}{dt} = p \frac{dp}{dr}$ и уравнение (8) перепишется в виде

$$\frac{p dp}{dr} = -\frac{k}{r^3}.$$

Разделяя переменные и интегрируя, будем иметь:

$$p dp = -\frac{k}{r^3} dr, \quad p^2 = \frac{k}{r^2} + C_1,$$

откуда

$$p = \frac{dr}{dt} = -\sqrt{\frac{k}{r^2} + C_1}$$

(перед радикалом ставится знак «минус», так как по смыслу задачи функция r убывает и $\frac{dr}{dt} < 0$), или

$$\frac{dr}{dt} = -\frac{\sqrt{k + C_1 r^2}}{r}.$$

Разделяя переменные в последнем уравнении и затем интегрируя, получаем:

$$-\frac{r dr}{\sqrt{k + C_1 r^2}} = dt, \quad -\frac{\sqrt{k + C_1 r^2}}{C_1} = t + C_2.$$

Используя начальные условия (при $t = 0$ $r = a$, $dr/dt = 0$), получаем:

$$C_2 = -\frac{\sqrt{k + C_1 a^2}}{C_1}, \quad 0 = -\frac{\sqrt{k + C_1 a^2}}{r},$$

откуда

$$C_1 = -\frac{k}{a^2}, \quad C_2 = 0.$$

Поэтому

$$t = \frac{a\sqrt{a^2 - r^2}}{\sqrt{k}}.$$

Когда точка достигнет центра O , будем иметь $r = 0$ и $t = \frac{a^2}{\sqrt{k}}$.

§ 13.4. Линейные уравнения второго порядка

1. Общие сведения о линейных дифференциальных уравнениях второго порядка. Уравнение

$$y'' + py' + qy = f(x), \quad (1)$$

где $p = p(x)$, $q = q(x)$ и $f(x)$ — непрерывные функции в интервале $(a; b)$, называется *неоднородным линейным дифференциальным уравнением второго порядка*, функции p , q — его *коэффициентами*. Если $f(x) \equiv 0$ в этом интервале, то уравнение принимает вид

$$y'' + py' + qy = 0 \quad (2)$$

и называется *однородным линейным дифференциальным уравнением второго порядка*. Если уравнение (2) имеет те же коэффициенты, как (1), то оно называется *однородным уравнением, соответствующим неоднородному уравнению (1)*.

Функции $y_1 = y_1(x)$ и $y_2 = y_2(x)$, определенные и непрерывные в интервале $(a; b)$, называются *линейно зависимыми* в этом интервале, если существуют постоянные числа α_1 и α_2 (причем по крайней мере одно из них не равно нулю), такие, что для всех значений x в рассматриваемом интервале выполняется тождество

$$\alpha_1 y_1 + \alpha_2 y_2 = 0. \quad (3)$$

Функции y_1 и y_2 называются *линейно независимыми* в интервале $(a; b)$, если тождество (3) может иметь место только при $\alpha_1 = \alpha_2 = 0$.

Т е о р е м а 1. Если y_1 и y_2 — линейно независимые частные решения линейного однородного уравнения второго порядка (2), то общее решение этого уравнения имеет вид:

$$y = C_1 y_1 + C_2 y_2, \quad (4)$$

где C_1 и C_2 — произвольные постоянные.

Д о к а з а т е л ь с т в о. Так как y_1 и y_2 — решения уравнения (2), то имеем тождества $y_1'' + py_1' + qy_1 = 0$, $y_2'' + py_2' + qy_2 = 0$. Используя их, получаем тождество

$$(C_1y_1 + C_2y_2)'' + p(C_1y_1 + C_2y_2)' + q(C_1y_1 + C_2y_2) = \\ = C_1(y_1'' + py_1' + qy_1) + C_2(y_2'' + py_2' + qy_2) = 0.$$

Следовательно, выражение (4) является решением уравнения (2), и поскольку это решение содержит две произвольные постоянные, то оно является общим решением однородного уравнения (2). Теорема доказана.

Пусть y_1 и y_2 — два линейно зависимых решения уравнения (2). Тогда выполняется тождество (3), где либо $\alpha_1 \neq 0$, либо $\alpha_2 \neq 0$. Предположим для определенности, что $\alpha_2 \neq 0$. Тогда из тождества (3) имеем $y_2 = -\frac{\alpha_1}{\alpha_2}y_1$, или $y_2 = ay_1$ $\left(a = -\frac{\alpha_1}{\alpha_2}\right)$. Подставляя это выражение в уравнение (4), получаем:

$$y = C_1y_1 + C_2ay_1 = (C_1 + aC_2)y_1 = Cy_1,$$

где $C = C_1 + aC_2$. Отсюда видно, что если y_1 и y_2 — линейно зависимые решения однородного уравнения (2), то решение (4) содержит только одну произвольную постоянную C и, следовательно, не является общим.

П р и м е ч а н и е. Отметим, что при условии теоремы 1 определитель

$$\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix},$$

называемый *определителем Вронского* или *вронскианом* (Юзеф Вронский (1776—1853) — польский математик) для функций y_1 и y_2 , не равен нулю ни при одном значении x из $(a; b)$. (Доказательство этого факта см., например, в [16].)

Для общего решения неоднородного уравнения (1) справедлива следующая теорема:

Т е о р е м а 2. *Общее решение неоднородного уравнения (1) равно сумме общего решения соответствующего однородного уравнения (2) и любого частного решения данного неоднородного уравнения.*

Д о к а з а т е л ь с т в о. Пусть $Y = C_1y_1 + C_2y_2$ — общее решение уравнения (2), соответствующего уравнению (1), и z — любое частное решение уравнения (1). Имеем тождества $Y'' + pY' + qY = 0$, $z'' + pz' + qz = f(x)$. Складывая почленно эти два тождества, получим тождество $(Y + z)'' + p(Y + z)' + q(Y + z) = f(x)$. Следовательно, функция $y = Y + z = C_1y_1 + C_2y_2 + z$ — решение уравнения (1) и при этом общее, так как в эту функцию входят две произвольные постоянные C_1 и C_2 .

2. Линейные однородные уравнения второго порядка с постоянными коэффициентами. Пусть в линейном уравнении

$$y'' + py' + qy = 0 \quad (5)$$

p и q — постоянные действительные числа.

Частное решение уравнения (5) будем искать в виде функции

$$y = e^{kx}, \quad (6)$$

где k — действительное или комплексное число, подлежащее определению. Дифференцируя по x выражение (6), получим:

$$y' = ke^{kx}, \quad y'' = k^2 e^{kx}. \quad (7)$$

Подставляя выражения (6) и (7) в уравнение (5), будем иметь

$$e^{kx}(k^2 + pk + q) = 0.$$

Отсюда, учитывая, что $e^{kx} \neq 0$, имеем:

$$k^2 + pk + q = 0. \quad (8)$$

Алгебраическое уравнение (8) называется *характеристическим уравнением* однородного уравнения (5). Характеристическое уравнение и дает возможность найти k . Уравнение (8) есть уравнение второй степени и потому имеет два корня. Обозначим их через k_1 и k_2 . Возможны три случая.

1) Корни k_1 и k_2 действительные и разные ($k_1 \neq k_2$). В этом случае по формуле (6) получим два частных решения уравнения (5) $y_1 = e^{k_1 x}$, $y_2 = e^{k_2 x}$, которые являются линейно независимыми. Действительно, если бы эти решения были линейно зависимы, то в интервале $(a; b)$ должно было бы выполняться тождество $\alpha_1 e^{k_1 x} + \alpha_2 e^{k_2 x} = 0$, где α_1 и α_2 одновременно не нули, или тождество $\alpha_1 e^{k_1 x} = -\alpha_2 e^{k_2 x}$. Отсюда $\alpha_1 e^{(k_1 - k_2)x} = -\alpha_2$, что невозможно, так как справа в последнем тождестве постоянное число, а слева функция от x . По теореме 1 (п. 1) следует, что общее решение уравнения (5) будет $y = C_1 e^{k_1 x} + C_2 e^{k_2 x}$.

П р и м е р 1. Решить уравнение $y'' - 3y' + 2y = 0$.

Его характеристическое уравнение $k^2 - 3k + 2 = 0$ имеет два различных действительных корня $k_1 = 1$ и $k_2 = 2$. Поэтому общее решение есть $y = C_1 e^x + C_2 e^{2x}$.

2) Корни k_1 и k_2 действительные и равные ($k_1 = k_2$). В этом случае одно частное решение уравнения (5) выразится функцией $y_1 = e^{k_1 x}$. Частным решением уравнения (5) в этом случае будет также функция $y_2 = xe^{k_1 x}$.

Действительно, $y_2'' + py_2' + qy_2 = (2 + k_1 x)k_1 e^{k_1 x} + p(1 + k_1 x)e^{k_1 x} + qxe^{k_1 x} = e^{k_1 x}[(k_1^2 + pk_1 + q)x + 2k_1 + p] = e^{k_1 x}(-p + p) = 0$.

Заметим, что решения $e^{k_1 x}$ и $xe^{k_1 x}$ линейно независимы, так как если бы функции $e^{k_1 x}$ и $xe^{k_1 x}$ были линейно зависимы, то в интервале $(a; b)$ выполнялось бы тождество $\alpha_1 e^{k_1 x} + \alpha_2 xe^{k_1 x} = 0$ (α_1 и α_2 одновременно не нули) и, значит, тождество $\alpha_1 + \alpha_2 x = 0$, что невозможно. Следовательно, общее решение уравнения (5) в данном случае

$$y = C_1 e^{k_1 x} + C_2 x e^{k_1 x} = e^{k_1 x} (C_1 + C_2 x).$$

Пример 2. Уравнению $y'' - 6y' + 9y = 0$ соответствует характеристическое уравнение $k^2 - 6k + 9 = 0$, имеющее равные корни $k_1 = k_2 = 3$. Поэтому общее решение будет

$$y = (C_1 + C_2 x)e^{3x}.$$

3) Корни $k_1 = \alpha + \beta i$ и $k_2 = \alpha - \beta i$ комплексные. Можно показать (см. [16]), что общее решение уравнения (5) в этом случае есть

$$y = e^{\alpha x} (C_1 \cos \beta x + C_2 \sin \beta x).$$

Пример 3. Уравнению $y'' - 2y' + 2y = 0$ соответствует характеристическое уравнение $k^2 - 2k + 2 = 0$, имеющее комплексные корни $k_1 = 1 + i$, $k_2 = 1 - i$. Следовательно, общим решением будет функция $y = e^x (C_1 \cos x + C_2 \sin x)$.

3. Линейные неоднородные уравнения второго порядка с постоянными коэффициентами. Рассмотрим теперь решение некоторых типов линейного неоднородного уравнения второго порядка с постоянными коэффициентами

$$y'' + py' + qy = f(x), \quad (9)$$

где p, q — постоянные действительные числа, $f(x)$ — известная непрерывная функция в интервале $(a; b)$. По теореме 2 (п. 1) для нахождения общего решения уравнения (9) надо знать общее решение Y соответствующего однородного уравнения $y'' + py' + qy = 0$ (для этого используются результаты п. 2 настоящего параграфа) и частное решение z уравнения (9).

Вид частного решения уравнения (9) зависит от вида правой части этого уравнения. Рассмотрим некоторые случаи:

а) $f(x) = a_2 x^2 + a_1 x + a_0$ ($a_2 \neq 0$). Если $q \neq 0$, то частное решение уравнения (9) ищем также в форме квадратного трехчлена: $z = A_2 x^2 + A_1 x + A_0$, где A_2, A_1 и A_0 — неопределенные коэффициенты. Отсюда $z' = 2A_2 x + A_1$,

$z'' = 2A_2$. Подставляя эти выражения в уравнение (9), в котором $f(x) = a_2x^2 + a_1x + a_0$, получим тождество

$$A_2qx^2 + (2A_2p + A_1q)x + 2A_2 + A_1p + A_0q = a_2x^2 + a_1x + a_0,$$

откуда

$$A_2q = a_2, 2A_2p + A_1q = a_1, 2A_2 + A_1p + A_0q = a_0. \quad (10)$$

Так как $q \neq 0$, то из равенств (10) для коэффициентов A_2 , A_1 и A_0 получаются определенные числовые значения. Тем самым частное решение z будет вполне определено. Если $q = 0$, то частное решение z уравнения (9) ищем в виде $z = x(A_2x^2 + A_1x + A_0)$, когда 0 — однократный корень характеристического уравнения (8), и в виде $z = x^2(A_2x^2 + A_1x + A_0)$, когда 0 — двукратный корень характеристического уравнения (8). Аналогично обстоит дело, если $f(x)$ — многочлен $P(x)$ произвольной степени.

Пример 1. Решить уравнение $y'' + y' = 2x + 1$. Имеем:

$$k^2 + k = 0, k_1 = 0, k_2 = -1, Y = C_1 + C_2e^{-x}.$$

Так как 0 — однократный корень характеристического уравнения, то частное решение данного уравнения ищем в виде $z = x(A_1x + A_0)$. Далее имеем: $z' = 2A_1x + A_0$, $z'' = 2A_1$, $2A_1x + 2A_1x + A_0 = 2x + 1$, $A_1 = 1$, $A_0 = -1$, $z = x^2 - x$, $y = C_1 + C_2e^{-x} + x^2 - x$.

б) $f(x) = ae^{bx}$ ($a \neq 0$). Частное решение ищем в виде $z = Ae^{bx}$, где A — неопределенный коэффициент. Отсюда $z' = Abe^{bx}$, $z'' = Ab^2e^{bx}$. Подставляя эти выражения в уравнение (9), в котором $f(x) = ae^{bx}$, после сокращения на e^{bx} будем иметь $A(b^2 + pb + q) = a$. Отсюда видно, что если b не является корнем характеристического уравнения, то

$$z = \frac{ae^{bx}}{b^2 + pb + q}.$$

Если b — корень характеристического уравнения, то частное решение уравнения (9) ищем в виде $z = Axe^{bx}$, когда b — однократный корень, и в виде $z = Ax^2e^{bx}$, когда b — двукратный корень. Аналогично будет, если $f(x) = P(x)e^{bx}$.

Пример 2. Решить уравнение $y'' - 2y' + y = 2e^x$. Имеем: $k^2 - 2k + 1 = 0$, $k_1 = k_2 = 1$, $Y = (C_1 + C_2x)e^x$. Так как в данном уравнении $b = 1$ — корень кратности 2 характеристического уравнения, то частное решение данного уравнения ищем в виде $z = Ax^2e^x$. Далее имеем:

$$\begin{aligned} z' &= Ax(x+2)e^x, z'' = A(x^2 + 4x + 2)e^x, \\ Ae^x(x^2 + 4x + 2) - 2Axe^x(x+2) + Ax^2e^x &= 2e^x, A = 1, \\ z &= x^2e^x, y = (C_1 + C_2x)e^x + x^2e^x. \end{aligned}$$

в) $f(x) = a \cos \omega x + b \sin \omega x$ (a и b не нули одновременно). В этом случае частное решение z ищем также в форме тригонометрического двучлена

$$z = A \cos \omega x + B \sin \omega x,$$

где A и B — неопределенные коэффициенты. Отсюда $z' = -A\omega \sin \omega x + B\omega \cos \omega x$, $z'' = -A\omega^2 \cos \omega x - B\omega^2 \sin \omega x$. Подставляя эти выражения в уравнение (9), в котором $f(x) = a \cos \omega x + b \sin \omega x$, получим:

$$(-A\omega^2 + Bp\omega + Aq)\cos \omega x + (-B\omega^2 - Ap\omega + Bq)\sin \omega x = a \cos \omega x + b \sin \omega x.$$

Так как последнее равенство представляет собой тождество, то коэффициенты при $\cos \omega x$ и $\sin \omega x$ в левой и правой частях этого равенства должны быть соответственно равны друг другу. Поэтому

$$A(q - \omega^2) + Bp\omega = a, -Ap\omega + B(q - \omega^2) = b.$$

Эти уравнения определяют коэффициенты A и B , кроме случая, когда $p = 0$, $q = \omega^2$ (или когда $\pm \omega i$ — корни характеристического уравнения). В последнем случае частное решение уравнения (9) ищется в виде: $z = x(A \cos \omega x + B \sin \omega x)$.

Пример 3. Решить уравнение $y'' + y = \cos x$. Имеем: $k^2 + 1 = 0$, $k_1 = i$, $k_2 = -i$, $Y = C_1 \cos x + C_2 \sin x$. Так как $\pm i$ — корни характеристического уравнения, то частное решение данного уравнения ищем в виде: $z = x(A \cos x + B \sin x)$. Далее имеем:

$$\begin{aligned} z' &= A \cos x + B \sin x + x(-A \sin x + B \cos x), \\ z'' &= -2A \sin x + 2B \cos x - x(A \cos x + B \sin x), \\ -2A \sin x + 2B \cos x &= \cos x, \quad A = 0, \quad B = \frac{1}{2}, \quad z = \frac{x}{2} \sin x, \\ y &= C_1 \cos x + C_2 \sin x + \frac{x}{2} \sin x. \end{aligned}$$

4. Гармонический осциллятор. Резонанс. Пусть на идеально гладком столе лежит шарик массы m , прикрепленный к пружине с коэффициентом жесткости $\lambda > 0$ (рис. 178). Направим ось Ox вдоль пружины, а за начало координат примем ту точку, в которой шарик находится в положении равновесия (пружина не растянута). Отведем теперь шарик от положения равновесия на расстояние x_0 и отпустим его. Тогда со стороны пружины на шарик будет действовать сила F , стремящаяся вернуть



Рис. 178

его в положение равновесия. Из физики известно, что эта сила равна (для малых x)

$$F(x) = -\lambda x \quad (11)$$

(знак минус поставлен потому, что направление действующей силы обратно по знаку смещению x).

Запишем второй закон Ньютона для шарика

$$F = ma, \quad (12)$$

где ускорение $a = \frac{d^2x}{dt^2}$ (см. § 8.4, п. 2). Из (11), (12) имеем:

$$ma = -\lambda x$$

или

$$m \frac{d^2x}{dt^2} = -\lambda x.$$

Отсюда

$$\frac{d^2x}{dt^2} + \omega^2 x = 0, \quad (13)$$

где

$$\omega = \sqrt{\frac{\lambda}{m}}.$$

Функция:

$$x = C_1 \cos \omega t + C_2 \sin \omega t, \quad (14)$$

является *общим решением уравнения* (13) (пп. 2; 3).

По условию задачи в момент времени $t = 0$ $x = x_0$, а скорость центра шарика равна нулю. Значит,

$$\begin{cases} x(0) = x_0, \\ x'(0) = 0. \end{cases} \quad (15)$$

Из выражения (14) следует, что

$$\frac{dx}{dt} = -C_1 \omega \sin \omega t + C_2 \omega \cos \omega t.$$

Поэтому (15) перепишется в виде:

$$\begin{cases} C_1 + 0 = x_0, \\ 0 + C_2 \omega = 0. \end{cases}$$

Отсюда $C_1 = x_0$, $C_2 = 0$ и

$$x = x_0 \cos \omega t. \quad (16)$$

Другими словами, шарик будет совершать гармонические колебания с амплитудой $|x_0|$ и периодом $T = \frac{2\pi}{\omega}$. Как говорят в физике, мы имеем здесь *гармонический осциллятор*.

В действительности мы знаем, что шарик не может колебаться бесконечно долго, и амплитуда колебаний стремится к нулю. Это происходит потому, что в любом реальном опыте присутствует сила трения, которой мы пренебрегли при выводе уравнения (13). Однако если сила трения очень мала, а промежуток времени не слишком большой, то (13) и (16) описывают процесс с хорошим приближением.

Пусть теперь на шарик действует еще одна сила F_1 , направленная вдоль оси Ox и изменяющаяся по закону $F_1 = F_0 \sin pt$, где F_0 и p — положительные постоянные.

Тогда второй закон Ньютона примет вид: $m \frac{d^2 x}{dt^2} = -\lambda x + F_0 \sin pt$ или

$$\frac{d^2 x}{dt^2} + \omega^2 x = \alpha_0 \sin pt, \quad (17)$$

где $\alpha_0 = \frac{F_0}{m}$.

Уравнение (17) называется уравнением *вынужденных колебаний*, а уравнение (13) — уравнением *свободных колебаний*.

Найдем частное решение неоднородного уравнения (17).

1) Пусть $p \neq \omega$, т. е. частота внешней силы не совпадает с частотой свободных колебаний. Частное решение ищем в виде:

$$z = A \cos pt + B \sin pt \quad (\text{см. п. 3, в}),$$

и общее решение уравнения (17) для случая $p \neq \omega$ будет

$$x = C_1 \cos \omega t + C_2 \sin \omega t + \frac{\alpha_0}{\omega^2 - p^2} \sin pt.$$

Это решение представляет собой наложение двух гармонических колебаний с частотами ω и p , причем колебания ограничены.

2) $p = \omega$, т. е. частота внешней силы совпадает с частотой свободных колебаний. В этом случае частное решение уравнения (17) ищем в виде

$$z = t(A \cos \omega t + B \sin \omega t) \quad (\text{см. п. 3, в}).$$

и общее решение уравнения (17) в случае $p = \omega$ будет

$$x = C_1 \cos \omega t + C_2 \sin \omega t - \frac{\alpha_0 t}{2\omega} \cos \omega t.$$

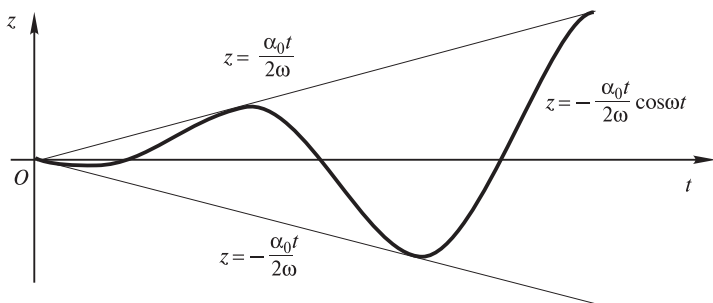


Рис. 179

Последняя формула показывает, что размах колебаний x неограниченно растет вместе со временем t (рис. 179). Это явление носит название *резонанса*. При работе многих механизмов резонанс крайне нежелателен, так как он приводит к нарушению их правильной работы и даже к разрушению.

5. Метод вариации произвольных постоянных. Лагранжу принадлежит общий метод нахождения частных решений неоднородного линейного уравнения. Этот метод применим как к уравнениям с постоянными коэффициентами, так и к уравнениям, в которых коэффициенты являются функциями от x .

Пусть y_1 и y_2 — два линейно независимых частных решения однородного уравнения (2), соответствующего неоднократному уравнению (1). Тогда общее решение этого однородного уравнения (см. п. 1, теорема 1) есть

$$y = C_1 y_1 + C_2 y_2,$$

где C_1 и C_2 — произвольные постоянные.

Будем искать решение неоднородного уравнения (1) в виде

$$y = C_1(x) y_1 + C_2(x) y_2, \quad (18)$$

где $C_1(x)$ и $C_2(x)$ — неизвестные функции, подлежащие определению.

Продифференцируем равенство (18):

$$y' = C_1(x) y_1' + C_2(x) y_2' + C_1'(x) y_1 + C_2'(x) y_2.$$

Подберем искомые функции $C_1(x)$ и $C_2(x)$ так, чтобы выполнялось равенство $C_1'(x) y_1 + C_2'(x) y_2 = 0$. Тогда

$$y' = C_1(x) y_1' + C_2(x) y_2'.$$

Дифференцируя последнее выражение, получим:

$$y'' = C_1(x) y_1'' + C_2(x) y_2'' + C_1'(x) y_1' + C_2'(x) y_2'.$$

Подставляя выражения для y , y' и y'' в уравнение (1), получим:

$$C_1(x)(y_1'' + p(x)y_1' + q(x)y_1) + \\ + C_2(x)(y_2'' + p(x)y_2' + q(x)y_2) + C_1'(x)y_1' + C_2'(x)y_2' = f(x).$$

Так как y_1 и y_2 — решения уравнения (2), то выражения, стоящие в скобках, тождественно равны нулю. Следовательно,

$$C_1'(x)y_1' + C_2'(x)y_2' = f(x).$$

Таким образом, приходим к системе уравнений

$$\begin{cases} C_1'(x)y_1 + C_2'(x)y_2 = 0, \\ C_1'(x)y_1' + C_2'(x)y_2' = f(x). \end{cases}$$

Определитель этой системы, как уже отмечалось ранее (см. п. 1, примечание), не обращается в нуль. Следовательно, мы можем найти $C_1(x)$ и $C_2(x)$ как определенные функции от x :

$$C_1'(x) = \varphi_1(x), C_2'(x) = \varphi_2(x).$$

Интегрируя, получим: $C_1(x) = \int \varphi_1(x)dx + C_1$, $C_2(x) = \int \varphi_2(x)dx + C_2$.

Подставляя найденные выражения для $C_1(x)$ и $C_2(x)$ в равенство (18), найдем решение уравнения (1), зависящее от двух произвольных постоянных, т. е. его общее решение. Если положить $C_1 = C_2 = 0$, то получим частное решение уравнения (1).

§ 13.5. Системы линейных дифференциальных уравнений

1. Общие вопросы. О п р е д е л е н и е. Система дифференциальных уравнений вида

$$\begin{cases} \frac{dy}{dx} = f(x, y, z), \\ \frac{dz}{dx} = \varphi(x, y, z). \end{cases} \quad (1)$$

где y и z — неизвестные функции независимой переменной x , называется *нормальной системой*.

Решением системы (1) называется совокупность функций $y(x)$ и $z(x)$, удовлетворяющих каждому из уравнений этой системы.

Если правые части нормальной системы дифференциальных уравнений являются линейными функциями относительно y и z , то эта система дифференциальных уравнений называется *линейной*.

Иногда нормальную систему (1) дифференциальных уравнений удается свести к одному уравнению 2-го порядка, содержащему одну неизвестную функцию. Сведение нормальной системы к одному урав-

нению может быть достигнуто дифференцированием одного из уравнений системы и исключением одной неизвестной функции y или z (так называемый *метод исключения*).

П р и м е р. Решить систему дифференциальных уравнений

$$\begin{cases} \frac{dy}{dx} = y + z, \\ \frac{dz}{dx} = z - y. \end{cases}$$

Продифференцируем по x первое уравнение системы:

$$\frac{d^2 y}{dx^2} = \frac{dy}{dx} + \frac{dz}{dx}.$$

Исключая из этого уравнения $\frac{dz}{dx}$, имеем

$$\frac{d^2 y}{dx^2} = \frac{dy}{dx} + z - y.$$

Исключая отсюда z , получаем:

$$\frac{d^2 y}{dx^2} = \frac{dy}{dx} + \frac{dy}{dx} - y - y \quad \text{или} \quad \frac{d^2 y}{dx^2} = 2 \frac{dy}{dx} - 2y$$

и, наконец

$$\frac{d^2 y}{dx^2} - 2 \frac{dy}{dx} + 2y = 0.$$

Корни характеристического уравнения $k^2 - 2k + 2 = 0$ будут $k_1 = 1 + i$, $k_2 = 1 - i$. Следовательно, общее решение для y будет

$$y = e^x (C_1 \cos x + C_2 \sin x).$$

Общее решение для z находим из первого уравнения системы

$$\begin{aligned} z = \frac{dy}{dx} - y &= e^x (C_1 \cos x + C_2 \sin x) + e^x (-C_1 \sin x + C_2 \cos x) - e^x (C_1 \cos x + C_2 \sin x) = \\ &= e^x (C_2 \cos x - C_1 \sin x). \end{aligned}$$

Итак, функции

$$y = e^x (C_1 \cos x + C_2 \sin x),$$

$$z = e^x (C_2 \cos x - C_1 \sin x)$$

являются решением данной системы.

Итак, интегрируя нормальную систему двух дифференциальных уравнений, мы получим ее решение, зависящее от двух произвольных постоянных C_1 и C_2 . Можно показать, что в общем случае для нормальной системы, состоящей из n уравнений, ее общее решение зависит от n произвольных постоянных.

2. Системы линейных дифференциальных уравнений с постоянными коэффициентами. Кроме рассмотренного метода интегрирования нормальной системы уравнений, мы укажем сейчас еще один метод (метод Эйлера), применимый только к нормальным системам линейных уравнений с постоянными коэффициентами.

Пусть дана нормальная система линейных дифференциальных уравнений с постоянными коэффициентами. Для простоты ограничимся системой двух уравнений с двумя неизвестными функциями

$$\begin{cases} \frac{dy}{dx} = ay + bz, \\ \frac{dz}{dx} = cy + dz. \end{cases} \quad (2)$$

Будем искать частное решение этой системы в виде

$$y = \alpha e^{kx}, \quad z = \beta e^{kx}. \quad (3)$$

Мы должны определить коэффициенты α , β и показатель степени k так, чтобы функции (3) были решением системы (2). Подставляя эти функции в равенства (2) и сокращая на множитель $e^{kx} \neq 0$, получим

$$\begin{cases} k\alpha = a\alpha + b\beta, \\ k\beta = c\alpha + d\beta. \end{cases}$$

Переносим все члены в одну сторону, получаем следующую систему линейных алгебраических уравнений относительно неизвестных α , β :

$$\begin{cases} (a - k)\alpha + b\beta = 0, \\ c\alpha + (d - k)\beta = 0. \end{cases} \quad (4)$$

Система (4) является однородной системой уравнений. Как известно, для того чтобы однородная система имела решения, отличные от нулевого, необходимо и достаточно, чтобы определитель системы равнялся нулю. Таким образом, для того чтобы система (4) имела решения, отличные от нулевого, должно выполняться равенство

$$\begin{vmatrix} a - k & b \\ c & d - k \end{vmatrix} = 0. \quad (5)$$

Равенство (5) представляет собой уравнение второй степени относительно k и называется *характеристическим* уравнением для системы (2). Ограничимся случаем, когда характеристическое уравнение имеет различные действительные корни k_1 , k_2 . Для каждого из этих корней напишем соответствующую систему уравнений (4) и определим коэффициенты α_1 , β_1 ; α_2 , β_2 . Если обозначить частные решения системы (2), соответствующие корню характеристического уравнения k_1 , через y_1 , z_1 , а

соответствующие корню k_2 — через y_2, z_2 , то, как можно показать, общее решение системы дифференциальных уравнений (2) запишется в виде

$$y = C_1 y_1 + C_2 y_2, \quad z = C_1 z_1 + C_2 z_2$$

или

$$y = C_1 \alpha_1 e^{k_1 x} + C_2 \alpha_2 e^{k_2 x}, \quad z = C_1 \beta_1 e^{k_1 x} + C_2 \beta_2 e^{k_2 x}. \quad (6)$$

П р и м е р. Решить систему дифференциальных уравнений

$$\begin{cases} \frac{dy}{dx} = -2y - 3z, \\ \frac{dz}{dx} = -y. \end{cases} \quad (7)$$

Р е ш е н и е. Характеристическое уравнение для системы (7) имеет вид

$$\begin{vmatrix} -2 - k & -3 \\ -1 & 0 - k \end{vmatrix} = 0$$

или

$$k^2 + 2k - 3 = 0.$$

Его корни $k_1 = -3, k_2 = 1$. Частные решения системы (7) будем искать в виде

$$y_1 = \alpha_1 e^{k_1 x}, \quad z_1 = \beta_1 e^{k_1 x};$$

$$y_2 = \alpha_2 e^{k_2 x}, \quad z_2 = \beta_2 e^{k_2 x}.$$

Система уравнений для определения α и β при $k_1 = -3$ запишется таким образом

$$\begin{cases} [-2 - (-3)]\alpha_1 - 3\beta_1 = 0, \\ -\alpha_1 + [0 - (-3)]\beta_1 = 0 \end{cases}$$

или

$$\begin{cases} \alpha_1 - 3\beta_1 = 0, \\ -\alpha_1 + 3\beta_1 = 0. \end{cases}$$

Эта система имеет бесчисленное множество решений, так как второе уравнение — следствие первого. Полагая, например, $\beta_1 = 1$, находим $\alpha_1 = 3$. Итак, корню характеристического уравнения $k_1 = -3$ соответствуют частные решения $y_1 = 3e^{-3x}$ и $z_1 = e^{-3x}$. Система уравнений для определения α и β при $k_2 = 1$ имеет вид

$$\begin{cases} -3\alpha_2 - 3\beta_2 = 0, \\ -\alpha_2 - \beta_2 = 0. \end{cases}$$

В качестве решений этой системы можно взять $\alpha_2 = 1, \beta_2 = -1$. Тогда корню характеристического уравнения $k_2 = 1$ соответствуют частные решения $y_2 = e^x$ и $z_2 = -e^x$.

Общее решение данной системы согласно формуле (6) таково:

$$y = 3C_1 e^{-3x} + C_2 e^x, \quad z = C_1 e^{-3x} - C_2 e^x.$$

Если среди корней характеристического уравнения (5) имеются комплексные, то соответствующие им частные решения преобразуются по формулам Эйлера подобно тому, как это делалось для одного линейного уравнения.

Раздел III

СПЕЦИАЛЬНЫЕ ГЛАВЫ

Глава 14. ТЕОРИЯ ПОЛЯ

§ 14.1. Скалярные поля

1. Понятие скалярного поля. Пусть Ω — область в пространстве.

Если каждой точке M из Ω ставится в соответствие по известному закону некоторое число $u = u(M)$, то говорят, что в области Ω задано *скалярное поле u* . Иными словами, задать скалярное поле — это означает задать скалярную функцию $u = u(M)$, называемую *функцией поля*. Запись $u(M)$ означает, что величина u является, как говорят (§ 11.1, п. 1), *функцией точки*. Заметим, что областью Ω может быть и все пространство.

Если величина $u = u(M)$ не зависит от времени t , то скалярное поле u называется *стационарным* (или *установившимся*). Мы будем рассматривать только стационарные скалярные поля.

Поле температуры внутри нагретого тела, поле плотности массы, поле распределения потенциала в электрическом поле — примеры скалярных полей.

Запись $u = u(M)$ не предполагает введения в пространстве никакой системы координат. Если же пространство отнесено к некоторой системе координат, например к прямоугольной системе координат $Oxyz$, то задание точки M равносильно заданию ее координат x, y, z в этой системе и функция поля $u(M)$ превращается в обычную функцию трех переменных $u(x, y, z)$. Мы всегда будем предполагать, что эта функция имеет непрерывные частные производные.

Можно ввести в пространстве и другие системы координат, например цилиндрическую и сферическую (§ 2.1, п. 9). В цилиндрической системе координат функция поля $u = u(r, \varphi, z)$, в сферической $u = u(r, \varphi, \theta)$.

2. Поверхности уровня. Скалярные поля часто изображаются геометрически с помощью так называемых поверхностей уровня.

Поверхностью уровня скалярного поля $u(M)$ называется множество точек пространства, в которых функция поля u имеет постоянное значение.

Уравнение поверхности уровня в прямоугольной системе координат *Охуз* имеет вид $u(x, y, z) = C$, где C — некоторая постоянная.

Заметим, что в курсе физики при рассмотрении поля потенциала поверхности уровня называют обычно *эквипотенциальными поверхностями* (т. е. поверхностями равного потенциала).

Придавая C различные значения, получим семейство поверхностей уровня.

Указанный способ изображения скалярного поля удобен, если речь идет о *плоском* скалярном поле, т. е. о поле, заданном в плоской области. Функция u этого поля зависит только от двух переменных x и y . Поэтому плоские скалярные поля геометрически изображают с помощью *линий уровня* (см. § 11.1, п. 2).

В случае поля температур на плоскости линии уровня называются *изотермами*, в случае поля давлений — *изобарами* и т. д.

3. Производная поля по направлению.

О п р е д е л е н и е. *Производной скалярного поля $u(M)$ в точке M по направлению $\vec{n}$* называется предел (если он существует) отношения приращения Δu функции $u(M)$ при смещении точки M в направлении вектора $\vec{n}$ (рис. 180) к величине этого смещения $d = MM_1$, когда последнее стремится к нулю; она



Рис. 180

обозначается символом $\frac{\partial u}{\partial n}$:

$$\frac{\partial u}{\partial n} = \lim_{d \rightarrow 0} \frac{\Delta u}{d} = \lim_{d \rightarrow 0} \frac{u(M_1) - u(M)}{d}. \quad (1)$$

Выведем для $\frac{\partial u}{\partial n}$ формулу, удобную в вычислительном отношении.

Пусть x, y, z — координаты фиксированной точки M , а $\cos \alpha, \cos \beta, \cos \gamma$ — направляющие косинусы вектора $\vec{n}$. Тогда точка M_1 будет иметь координаты $x + d \cos \alpha, y + d \cos \beta, z + d \cos \gamma$. Величины $x, y, z, \cos \alpha, \cos \beta, \cos \gamma$ фиксированы, поэтому $u(M_1)$ является функцией только смещения d . Обозначим эту функцию через

$$\psi(d) = u(x + d \cos \alpha, y + d \cos \beta, z + d \cos \gamma).$$

Имеем $\psi(0) = u(x, y, z)$. Следовательно, в силу равенства (1)

$$\frac{\partial u}{\partial n} = \lim_{d \rightarrow 0} \frac{\psi(d) - \psi(0)}{d} = \psi'(0).$$

Формула производной сложной функции (§ 11.2, п. 3) дает

$$\psi'(d) = u'_x(M_1) \cos \alpha + u'_y(M_1) \cos \beta + u'_z(M_1) \cos \gamma,$$

откуда

$$\psi'(0) = u'_x(M) \cos \alpha + u'_y(M) \cos \beta + u'_z(M) \cos \gamma,$$

или

$$\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma. \quad (2)$$

Из формулы (2) видно, что если направление $\bar{n}$ совпадает с положительным направлением оси Ox , т. е. $\alpha = 0, \beta = \gamma = \frac{\pi}{2}$, то $\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}$. Аналогично, если направление $\bar{n}$ будет совпадать с направлением оси Oy (Oz).

Подобно тому как частные производные $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}$ и $\frac{\partial u}{\partial z}$ характеризуют скорость изменения функции u в направлении осей координат (§ 11.2, п. 1), так и производная по направлению $\frac{\partial u}{\partial n}$ будет являться *скоростью* изменения функции $u(x, y, z)$ в точке M по направлению $\bar{n}$. *Абсолютная величина производной* $\frac{\partial u}{\partial n}$ по направлению $\bar{n}$ определяет величину скорости, а *знак производной* — характер изменения функции u (возрастание или убывание). В этом состоит *физический смысл* производной по направлению.

Из формулы (2) следует, что производная по направлению $\bar{n}'$, *противоположному направлению $\bar{n}$* , равна *производной по направлению $\bar{n}$* , взятой с *обратным знаком*. Действительно, при перемене направления углы α, β, γ изменятся на π .

Если поле u плоское, то направление $\bar{n}$ вполне определено углом α его наклона к оси абсцисс. Формулу для производной по направлению в случае такого поля можно получить из общей формулы (2), положив $\gamma = \frac{\pi}{2}$ и

$$\beta = \frac{\pi}{2} - \alpha. \text{ При этом если } \alpha = 0, \text{ то } \frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}, \text{ а если } \alpha = \frac{\pi}{2}, \text{ то } \frac{\partial u}{\partial n} = \frac{\partial u}{\partial y}.$$

4. Градиент скалярного поля. Пусть дано скалярное поле $u = u(M) = u(x, y, z)$.

О п р е д е л е н и е. Вектор $\frac{\partial u}{\partial x} \bar{i} + \frac{\partial u}{\partial y} \bar{j} + \frac{\partial u}{\partial z} \bar{k}$ называется *градиентом скалярного поля $u = u(M)$ в точке M* и обозначается $\text{grad } u$:

$$\text{grad } u = \frac{\partial u}{\partial x} \bar{i} + \frac{\partial u}{\partial y} \bar{j} + \frac{\partial u}{\partial z} \bar{k}. \quad (3)$$

Обозначим через $\bar{n}_0 (\cos \alpha; \cos \beta; \cos \gamma)$ единичный вектор направления $\bar{n}$. Тогда

$$\frac{\partial u}{\partial n} = \bar{n}_0 \text{grad } u = |\text{grad } u| \cos \varphi = \text{пр}_{\bar{n}} \text{grad } u,$$

где φ — угол между векторами $\text{grad } u$ и $\bar{n}$.

Значит, производная скалярного поля $u(M)$ в точке M в данном направлении равна проекции градиента поля в точке M на это направление. Отсюда следует, что $\frac{\partial u}{\partial n}$ в точке M имеет наибольшее значение в направлении градиента (для этого направления $\varphi = 0$ и $\cos 0 = 1$). В этом случае

$$\frac{\partial u}{\partial n} = |\operatorname{grad} u| = \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2}.$$

Таким образом, $\operatorname{grad} u$ и есть вектор, указывающий направление наибольшего возрастания скалярного поля в данной точке и имеющий модуль, равный скорости этого возрастания. В этом состоит физический смысл градиента.

В плоском скалярном поле $u = u(M) = u(x, y)$ градиент определяется равенством

$$\operatorname{grad} u = \frac{\partial u}{\partial x} \bar{i} + \frac{\partial u}{\partial y} \bar{j}.$$

П р и м е р. Определить градиент скалярного поля потенциала электростатического поля, образованного точечным зарядом величины q , помещенным в начало координат.

Из электростатики известно, что упомянутый потенциал в точке $M(x; y; z)$ равен $u = \frac{q}{r}$, где $r = \sqrt{x^2 + y^2 + z^2}$. По формуле (3) находим:

$$\operatorname{grad} u = -\frac{q}{r^2} \frac{x}{r} \bar{i} - \frac{q}{r^2} \frac{y}{r} \bar{j} - \frac{q}{r^2} \frac{z}{r} \bar{k} = -\frac{q}{r^2} \frac{\bar{r}}{r} = -\frac{q}{r^2} \bar{r}_0,$$

где $\bar{r} = x\bar{i} + y\bar{j} + z\bar{k}$ — радиус-вектор точки M , а $\bar{r}_0 = \frac{\bar{r}}{r}$ — единичный вектор в направлении радиуса-вектора. Вектор $\bar{E} = \frac{q}{r^2} \bar{r}_0$ называется *вектором напряженности* рассматриваемого электростатического поля в точке M . Таким образом, $\operatorname{grad} u = -\bar{E}$.

5. Оператор набла и исчисление градиентов. Английским математиком Уильямом Гамильтоном (1805 — 1865) был введен векторный дифференциальный оператор

$$\nabla = \bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z},$$

называемый *оператором набла*. (Набла — греческое слово, означающее «арфа» — музыкальный инструмент, по форме напоминающий перевернутый треугольник.) В этой формуле величина, к которой прилагается оператор ∇ , должна быть поставлена под знаком частных производных (операторов дифференцирования) $\frac{\partial}{\partial x}$ и т. д. При этом сущест-

венно отметить, что единичные векторы следует писать до операторов дифференцирования, так как эти операторы действуют на выражения, стоящие справа от них.

Поэтому

$$\nabla u = \bar{i} \frac{\partial u}{\partial x} + \bar{j} \frac{\partial u}{\partial y} + \bar{k} \frac{\partial u}{\partial z}, \text{ или } \nabla u = \text{grad } u.$$

Из известных (см. § 11.2) правил дифференциального исчисления функций нескольких переменных вытекают следующие простые правила:

$$1. \nabla(C_1 u + C_2 v) = C_1 \nabla u + C_2 \nabla v$$

(C_1, C_2 — постоянные, u, v — функции переменных x, y, z), т. е.

$$\text{grad}(C_1 u + C_2 v) = C_1 \text{grad } u + C_2 \text{grad } v.$$

$$2. \nabla(uv) = v \nabla u + u \nabla v,$$

т. е. $\text{grad}(uv) = v \text{grad } u + u \text{grad } v.$

$$3. \nabla f(u) = f'(u) \nabla u,$$

т. е. $\text{grad } f(u) = f'(u) \text{grad } u.$

$$4. \nabla f(u, v) = f'_u \nabla u + f'_v \nabla v,$$

т. е. $\text{grad } f(u, v) = f'_u \text{grad } u + f'_v \text{grad } v.$

Аналогично в случае $f(u, v, w).$

§ 14.2. Векторные поля

1. Понятие векторного поля. Пусть Ω — область в пространстве.

Если каждой точке M из Ω сопоставлен вполне определенный вектор $\vec{a}(M)$, то говорят, что в области Ω задано *векторное поле* $\vec{a}$.

Заметим, что областью Ω может быть и все пространство. Мы будем рассматривать стационарные векторные поля, в которых вектор $\vec{a}(M)$ зависит только от точки M и не зависит от времени.

Поле силы тяжести, поле скорости частиц текущей жидкости, поле электрической и магнитной индукции, поле плотности электрического тока — примеры векторных полей.

В прямоугольной системе координат $Oxyz$ вектор $\vec{a}(M)$ запишется в виде (§ 2.1, п. 8)

$$\vec{a}(M) = a_x(x, y, z) \bar{i} + a_y(x, y, z) \bar{j} + a_z(x, y, z) \bar{k},$$

где $a_x(x, y, z), a_y(x, y, z), a_z(x, y, z)$ (кратко a_x, a_y, a_z) — проекции вектора, заданного в точке $M(x, y, z)$ соответственно на оси координат Ox, Oy, Oz .

В § 9.11 рассматривалась векторная функция одного скалярного аргумента. Здесь изучаем векторную функцию *трех* скалярных аргументов.

Далее всюду предполагается, что функции a_x , a_y , a_z непрерывны вместе со своими частными производными.

Укажем на частные случаи векторных полей.

Векторное поле называется *однородным*, если $\vec{a}(M)$ — постоянный вектор, т. е. a_x , a_y , a_z — постоянные величины. Примером однородного поля может служить, например, поле силы тяжести.

Векторное поле называется *плоским*, если проекции вектора $\vec{a}(M)$ не зависят от одной из трех переменных x , y , z и одна из проекций равна нулю, например:

$$\vec{a}(M) = a_x(x, y)\vec{i} + a_y(x, y)\vec{j}.$$

С плоскими полями очень часто приходится встречаться в гидродинамике при изучении *плоских течений жидкости*, т. е. таких течений, когда все частицы жидкости движутся параллельно некоторой плоскости, причем скорости частиц, расположенных на одной и той же прямой, перпендикулярной к этой плоскости, одинаковы. Рассмотрим следующий физический пример.

П р и м е р. Пусть твердое тело вращается с постоянной угловой скоростью $\vec{\omega}$. Найдем поле линейных скоростей точек этого тела.

Как известно из кинематики, линейная скорость $\vec{v}$ равна векторному произведению $\vec{v} = [\vec{\omega}, \vec{r}]$, где $\vec{\omega}$ — вектор угловой скорости (т. е. вектор, отложенный на оси вращения и численно равный величине угловой скорости; этот вектор направлен так, что если смотреть из его конца, вращение кажется происходящим против часовой стрелки), а $\vec{r}$ — радиус-вектор точки M вращающегося тела относительно какой-либо точки оси вращения. Выбрав эту неподвижную точку за начало координат и направив ось вращения по оси Oz (рис. 181), найдем проекции вектора $\vec{v}$. Имеем:

$$\vec{\omega} = \omega\vec{k}, \quad \vec{r} = x\vec{i} + y\vec{j} + z\vec{k}, \quad \vec{v} = -\omega y\vec{i} + \omega x\vec{j}.$$

Таким образом, $v_x = -\omega y$, $v_y = \omega x$, $v_z = 0$, т. е. поле является плоским.

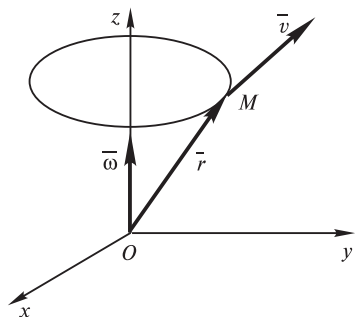


Рис. 181

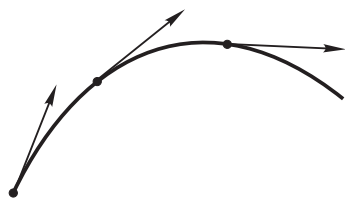


Рис. 182

Векторной линией векторного поля называется линия, в каждой точке которой касательная совпадает с вектором, соответствующим этой точке (рис. 182).

Векторные линии в конкретных полях имеют ясный физический смысл. Так, если мы рассматриваем поле скоростей текущей жидкости, то векторные линии суть *линии тока* этой жидкости, т. е. линии, по которым движутся частицы жидкости.

В электрическом поле векторные линии — это *силовые линии* этого поля. Например, в поле точечного заряда такими линиями будут лучи, выходящие из заряда. Для магнитного поля векторными (силовыми) линиями будут линии, выходящие из северного полюса и оканчивающиеся в южном.

2. Поток векторного поля через поверхность. Как уже отмечалось (см. § 12.4, п. 3), количество жидкости, протекающей в единицу времени через поверхность σ , выражается формулой

$$\Pi = \iint_{\sigma} P(x, y, z) dydz + Q(x, y, z) dzdx + R(x, y, z) dxdy$$

или в силу формулы (15) из § 12.4 формулой

$$\Pi = \iint_{\sigma} [P(x, y, z) \cos \alpha + Q(x, y, z) \cos \beta + R(x, y, z) \cos \gamma] ds. \quad (1)$$

Величина Π называется *поток*ом *жидкости* через поверхность σ . Поскольку P, Q, R — координаты вектора-скорости $\vec{v}$, $\cos \alpha, \cos \beta, \cos \gamma$ — координаты единичного вектора $\vec{n}_0$ нормали $\vec{n}$ к поверхности σ и $P \cos \alpha + Q \cos \beta + R \cos \gamma = \vec{v} \cdot \vec{n}_0$, то формулу (1) можно записать в виде

$$\Pi = \iint_{\sigma} \vec{v} \cdot \vec{n}_0 ds, \text{ или } \Pi = \iint_{\sigma} v_n ds,$$

где v_n — проекция вектора $\vec{v}$ на нормаль $\vec{n}$.

О п р е д е л е н и е. *Потоком* Π *вектора* $\vec{a}(M)$ *или потоком векторного поля* $\vec{a}(M)$ *через поверхность* σ *называется поверхностный интеграл*

$$\Pi = \iint_{\sigma} \vec{a}(M) \cdot \vec{n}_0 ds, \quad (2)$$

где $\vec{n}_0$ — единичный вектор нормали к поверхности σ в ее текущей точке M . Поверхность должна быть ориентированной, это означает, что в каждой ее точке выбрано одно из двух направлений нормали так, чтобы $\vec{n}(M)$ непрерывно менялся по поверхности σ . В случае замкнутой поверхности σ в качестве $\vec{n}(M)$ выбирается вектор внешней нормали.

Пусть

$$\vec{a}(M) = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}, \quad \vec{n}_0 = \vec{i} \cos \alpha + \vec{j} \cos \beta + \vec{k} \cos \gamma.$$

Тогда

$$\Pi = \iint_{\sigma} (a_x \cos \alpha + a_y \cos \beta + a_z \cos \gamma) ds.$$

Отсюда согласно формуле (15) из § 12.4 (п. 5) интеграл (2) можно представить в виде

$$\Pi = \iint_{\sigma} a_x dydz + a_y dzdx + a_z dxdy.$$

Таким образом, вычисление потока вектора сводится к вычислению интегралов по поверхности. Из самого определения следует, что поток вектора Π — величина скалярная. Если изменить направление нормали $\bar{n}$ на противоположное, т. е. переменить сторону поверхности σ , то (§ 12.4, п. 3) поток Π изменит знак. Так как скалярное произведение вектора $\bar{a}(M)$ на единичный вектор нормали $\bar{n}$ равно $a_n(M)$ — проекции вектора $\bar{a}(M)$ на направление $\bar{n}$ (§ 2.2, п. 1), то поток Π можно представить в виде

$$\Pi = \iint_{\sigma} a_n(M) ds. \quad (3)$$

Отсюда, в частности, следует, что если на некотором участке поверхности проекция вектора $\bar{a}(M)$ на нормаль постоянна, т. е. $a_n(M) = C = \text{const}$, то поток через такой участок просто равен $C \cdot Q$, где Q — площадь этого участка поверхности.

П р и м е р. Найти поток радиуса-вектора $\bar{r}$ через боковую поверхность σ_1 , верхнее основание σ_2 и нижнее основание σ_3 прямого цилиндра радиуса R и высотой H , если начало координат лежит в центре нижнего основания цилиндра, а ось цилиндра совпадает с осью Oz (рис. 183).

На всех поверхностях $\bar{n}$ имеет направление внешней нормали. На боковой поверхности σ_1 внешняя нормаль $\bar{n}$ параллельна плоскости xOy и проекция r_n равна R . Поэтому, используя формулу (3) и формулу $\iint_{\sigma} ds = S_{\sigma}$ (§ 12.4, п. 1), получаем:

$$\Pi_1 = \iint_{\sigma_1} R ds = R \cdot 2\pi RH = 2\pi R^2 H.$$

На верхнем основании σ_2 нормаль $\bar{n}$ направлена параллельно оси Oz и $r_n = H$. Следовательно,

$$\Pi_2 = \iint_{\sigma_2} H ds = H\pi R^2 = \pi R^2 H.$$

Наконец, на нижнем основании σ_3 проекция $r_n = 0$ и $\Pi_3 = 0$.

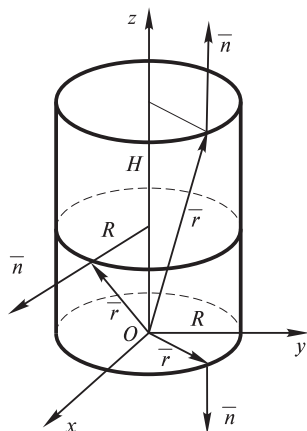


Рис. 183

Особый интерес представляет случай, когда σ — замкнутая поверхность. Если берется внешняя нормаль, то мы будем говорить о *потоке изнутри* поверхности σ . Он обозначается так:

$$\Pi = \oiint_{\sigma} a_n(M) ds.$$

(В векторном анализе интегралы по замкнутым поверхностям обозначаются обычно $\oiint_{\sigma}$, часто также употребляют и символ $\oint$.) Область, которую ограничивает поверхность σ , будем обозначать Ω .

Когда векторное поле $\vec{a}(M)$ представляет поле скоростей жидкости, величина потока Π дает разность между количеством жидкости, вытекающей из области Ω , и количеством жидкости, втекающей в эту область. (Действительно, так как интеграл $\oiint_{\sigma} a_n ds$ берется по внешней стороне поверхности, то если вектор скорости $\vec{v}$ в данной точке поверхности σ направлен наружу, то $\vec{v} \cdot \vec{n}_0 > 0$; если же вектор $\vec{v}$ направлен внутрь, то $\vec{v} \cdot \vec{n}_0 < 0$.)

Если $\Pi = 0$, то в область Ω жидкости втекает столько же, сколько и вытекает. Так, например, будет для любой области, расположенной в потоке воды, текущей в реке.

Если же величина Π отлична от нуля, например положительна, то из области Ω жидкости вытекает больше, чем втекает. Это означает, что в области Ω имеются *источники*, питающие поток жидкости. Наоборот, если величина Π отрицательна, то это указывает на наличие *стоков* — мест, где жидкость удаляется из потока.

Равенство нулю потока Π может означать либо отсутствие источников и стоков, либо, по крайней мере, наличие такого распределения источников и стоков, что их общая мощность равна нулю.

3. Формула Остроградского — Гаусса. Формула Остроградского — Гаусса (М. В. Остроградский (1801—1861) — выдающийся русский математик) связывает поверхностный интеграл по замкнутой поверхности и тройной интеграл по пространственной области, ограниченной этой поверхностью. Эта формула является аналогом формулы Римана — Грина (§ 12.3, п. 4), связывающей криволинейный интеграл по замкнутой кривой с двойным интегралом по плоской области, ограниченной этой кривой.

Пусть функции $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ (кратко P , Q , R) непрерывны вместе со своими частными производными P'_x , Q'_y , R'_z в замкнутой пространственной области Ω , граница которой пересекается с любой прямой, параллельной осям координат, не более чем в двух точках. Назовем для краткости такие области *простыми*.

При этом будем рассматривать внешнюю сторону поверхности σ , ограничивающей эту область. Предполагается, что поверхность σ гладкая или кусочно-гладкая (§ 12.4, п. 1).

Пусть область G есть проекция поверхности σ (и области Ω) на плоскость xOy (рис. 184), а $z = z_1(x, y)$ и $z = z_2(x, y)$ — уравнения соответствующих частей σ — нижней части σ_1 и верхней σ_2 , где $z_1(x, y)$ и $z_2(x, y)$ непрерывны в области G . Преобразу-

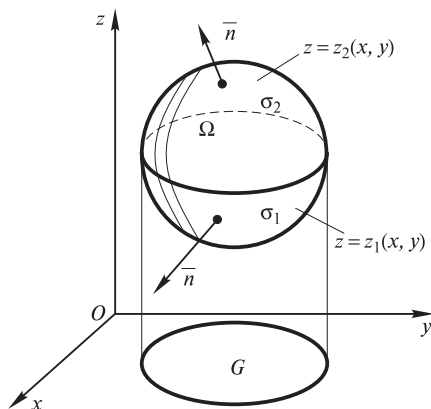


Рис. 184

ем, как указано ниже, интеграл $\iiint_{\Omega} \frac{\partial R}{\partial z} dx dy dz$:

$$\begin{aligned} \iiint_{\Omega} \frac{\partial R}{\partial z} dx dy dz &= \iint_G dx dy \int_{z_1(x, y)}^{z_2(x, y)} \frac{\partial R}{\partial z} dz = \\ &= \iint_G (R(x, y, z_2(x, y)) - R(x, y, z_1(x, y))) dx dy = \\ &= \iint_{\sigma_2} R(x, y, z) dx dy - \iint_{\sigma_1} R(x, y, z) dx dy. \end{aligned}$$

Преобразование I основано на правиле вычисления тройного интеграла (§ 12.2), II — на формуле Ньютона — Лейбница (§ 9.7, п. 2), III — на формуле (14) из п. 4 § 12.4. В результате получаются интегралы по верхним сторонам поверхностей σ_1 и σ_2 . Меняя в интеграле по σ_1 сторону поверхности, с учетом свойств поверхностного интеграла второго рода (§ 12.4, п. 3) получаем:

$$\iiint_{\Omega} \frac{\partial R}{\partial z} dx dy dz = \iint_{\sigma} R(x, y, z) dx dy, \quad (4)$$

где интеграл справа берется по внешней стороне поверхности σ .

Аналогичным образом устанавливаются формулы

$$\iiint_{\Omega} \frac{\partial P}{\partial x} dx dy dz = \iint_{\sigma} P dy dz, \quad (5)$$

$$\iiint_{\Omega} \frac{\partial Q}{\partial y} dx dy dz = \iint_{\sigma} Q dz dx. \quad (6)$$

Складывая почленно равенства (4), (5), (6), приходим к формуле

$$\iiint_{\Omega} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx dy dz = \iint_{\sigma} P dy dz + Q dz dx + R dx dy, \quad (7)$$

называемой *формулой Остроградского — Гаусса* (в координатной форме), где поверхностный интеграл берется по внешней стороне поверхности σ .

З а м е ч а н и е. Формула Остроградского—Гаусса остается верной и для любой замкнутой пространственной области Ω , которую можно разбить на конечное число простых областей.

4. Дивергенция. Рассмотрим некоторую точку M_0 векторного поля $\vec{a}(M)$ и окружим ее замкнутой поверхностью σ , целиком содержащейся в поле. Вычислим поток вектора через поверхность σ и возьмем отношение этого потока к объему V области Ω , ограниченной поверхностью σ :

$$\frac{\oiint_{\sigma} a_n(M) ds}{V}. \quad (8)$$

В поле скоростей жидкости это отношение определяет количество жидкости, возникающее в единицу времени в области Ω , отнесенное к единице объема, т. е., как говорят, *среднюю объемную мощность источника*; если поток изнутри поверхности σ меньше нуля, то соответственно говорят σ *мощности стока*.

Найдем теперь предел отношения (8) при условии, что область Ω стягивается в точку M_0 (т. е. что V стремится к нулю):

$$\lim_{V \rightarrow 0} \frac{\oiint_{\sigma} a_n(M) ds}{V}. \quad (9)$$

Если этот предел существует и положителен, то точка M_0 называется *источником* поля, а если отрицателен, то *стоком* поля. Сама величина этого предела характеризует *мощность* источника или стока. В первом случае в любом бесконечно малом объеме, окружающем точку M_0 , жидкость возникает, а во втором — исчезает.

Предел (9) (если он существует) называется *дивергенцией* или *расходимостью* векторного поля в точке M_0 ; обозначение $\operatorname{div} \vec{a}(M_0)$. Таким образом,

$$\operatorname{div} \vec{a}(M_0) = \lim_{V \rightarrow 0} \frac{\oiint_{\sigma} a_n(M) ds}{V}.$$

Из определения следует, что дивергенция векторного поля есть величина скалярная.

Следующая теорема решает вопрос о существовании предела (9) и нахождении его величины в декартовых координатах:

Т е о р е м а. Дивергенция векторного поля $\vec{a}(M) = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}$ (a_x, a_y, a_z непрерывны вместе со своими частными производными $\frac{\partial a_x}{\partial x}, \frac{\partial a_y}{\partial y}, \frac{\partial a_z}{\partial z}$) выражается формулой

$$\operatorname{div} \vec{a}(M) = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}, \quad (10)$$

где значения частных производных берутся в точке M .

Д о к а з а т е л ь с т в о. По формуле Остроградского — Гаусса (п. 3) поток вектора $\vec{a}$ можно представить в виде

$$\begin{aligned} \oiint_{\sigma} a_n(M) ds &= \iiint_{\Omega} a_x(M) dydz + a_y(M) dzdx + a_z(M) dxdy = \\ &= \iiint_{\Omega} \left(\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right) dxdydz. \end{aligned}$$

По теореме о среднем (см. § 12.2, п. 1)

$$\iiint_{\Omega} \left(\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right) dxdydz = \left(\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right) \bigg|_{M_1} V.$$

Если область Ω стягивается в точку M , то точка M_1 стремится к точке M , и мы получаем:

$$\operatorname{div} \vec{a}(M) = \lim_{M_1 \rightarrow M} \frac{\left(\frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \right) \bigg|_{M_1} V}{V} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z},$$

что и требовалось доказать.

Укажем теперь на векторную запись формулы Остроградского—Гаусса. Если в формуле (7) функции P, Q, R рассматривать как проекции некоторого вектора $\vec{a}(M)$, то правая часть ее равна потоку вектора $\vec{a}(M)$ (п. 2) через замкнутую поверхность σ (в направлении внешней нормали к σ), а

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = \operatorname{div} \vec{a}(M),$$

и формула Остроградского—Гаусса может быть переписана в виде

$$\oiint_{\sigma} a_n(M) ds = \iiint_{\Omega} \operatorname{div} \vec{a}(M) dv. \quad (11)$$

Она выражает следующий результат: *поток векторного поля через замкнутую поверхность* (в направлении внешней нормали к ней) *равен тройному интегралу от дивергенции этого поля, взятому по области, ограниченной этой поверхностью.*

Предположим, что в некоторой точке M дивергенция скорости $\bar{v}$ положительна: $\operatorname{div} \bar{v} > 0$. В силу непрерывности частных производных она является положительной и в точках достаточно малого шара w , ограниченного сферой σ с центром в точке M . Но тогда $\iiint_w \operatorname{div} \bar{v} dx dy dz > 0$, следовательно, на основании формулы (11) $\iint_{\sigma} v_n ds > 0$, т. е. из области w че-

рез ее границу σ жидкости вытекает больше, чем втекает. По этой причине точку M называют *источником*. Если же $\operatorname{div} \bar{v} < 0$ в точке M , то через достаточно малую сферу с центром в этой точке втекает жидкости больше, чем вытекает. Поэтому точку M в этом случае называют *стоком* (то же и в случае произвольного векторного поля $\bar{a}(M)$).

Векторная форма формулы Остроградского — Гаусса выражает в поле текущей жидкости тот факт, что поток жидкости через поверхность равен суммарной мощности всех источников и стоков, т. е. количеству жидкости, возникающей в рассматриваемой области за единицу времени. (Если мощность стоков больше, чем источников, то жидкость в области исчезает.) Если, в частности, дивергенция во всех точках равна нулю, то равен нулю и поток через любую замкнутую поверхность.

П р и м е р 1. Вычислить дивергенцию поля скоростей жидкости $\bar{v} = a\bar{i} + b\bar{j} + c\bar{k}$, где a, b, c — постоянные. Это однородное векторное поле. Ясно, что $\operatorname{div} \bar{v} = 0$.

Таким образом, в данном поле нет ни источников, ни стоков. Все частицы жидкости имеют одну и ту же скорость, жидкость движется поступательно как твердое тело. Поток такой жидкости через любую замкнутую поверхность равен нулю.

П р и м е р 2. Вычислить поток вектора $\bar{a} = \bar{r} = x\bar{i} + y\bar{j} + z\bar{k}$ через замкнутую поверхность. Так как $\operatorname{div} \bar{r} = 3$, то по формуле (11) получаем, что $\Pi = 3 \iiint_{\Omega} dv = 3V_{\Omega}$. Таким образом, поток радиуса-вектора $\bar{r}$ через замкнутую поверхность равен утроенному объему области, ограниченной этой поверхностью.

П р и м е р 3. Вычислим дивергенцию поля линейных скоростей вращающегося тела (см. п. 1, пример) $\bar{v} = -\omega y\bar{i} + \omega x\bar{j}$. Здесь

$$\operatorname{div} \bar{v} = \frac{\partial(-\omega y)}{\partial x} + \frac{\partial(\omega x)}{\partial y} = 0.$$

Если представить себе жидкость, вращающуюся как твердое тело, то ясно, что в таком потоке нет ни источников, ни стоков.

Примечание 1. Заметим, что формулу (10) можно записать в виде

$$\operatorname{div} \bar{a} = (\nabla, \bar{a}),$$

если под $(\nabla, \bar{a})$ понимать формально образованное скалярное произведение векторов

$$\nabla = \bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z} \text{ и } \bar{a} = \bar{a}(M) = a_x \bar{i} + a_y \bar{j} + a_z \bar{k},$$

обращаясь с операторами дифференцирования $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$ как со скалярными величинами.

Примечание 2. Рассмотрим некоторые формулы, облегчающие вычисление дивергенции более сложных векторов полей.

1. Если C_1 и C_2 — скалярные постоянные, то согласно известным свойствам скалярного произведения (см. § 2.2, п. 1) имеем:

$$(\nabla, C_1 \bar{a}_1 + C_2 \bar{a}_2) = C_1 (\nabla, \bar{a}_1) + C_2 (\nabla, \bar{a}_2),$$

т. е.

$$\operatorname{div}(C_1 \bar{a}_1 + C_2 \bar{a}_2) = C_1 \operatorname{div} \bar{a}_1 + C_2 \operatorname{div} \bar{a}_2.$$

2. Если $\bar{a}$ — векторное поле постоянного направления, $\bar{a} = u(M) \bar{c}$, где $\bar{c} = c_x \bar{i} + c_y \bar{j} + c_z \bar{k}$ — постоянный вектор, а $u = u(M)$ — скалярное поле, то

$$(\nabla, \bar{a}) = (\nabla, u \bar{c}) = c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} + c_z \frac{\partial u}{\partial z} = (\bar{c}, \nabla u),$$

т. е.

$$\operatorname{div}(u \bar{c}) = (\operatorname{grad} u, \bar{c}),$$

3. Если $\bar{c} = \bar{c}(M)$ — переменный вектор, $u = u(M)$ — скалярное поле, то

$$(\nabla, u \bar{c}) = c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} + c_z \frac{\partial u}{\partial z} + u \left(\frac{\partial c_x}{\partial x} + \frac{\partial c_y}{\partial y} + \frac{\partial c_z}{\partial z} \right) = (\bar{c}, \nabla u) + u(\nabla, \bar{c}),$$

т. е.

$$\operatorname{div}(u \bar{c}) = (\operatorname{grad} u, \bar{c}) + u \operatorname{div} \bar{c},$$

5. Циркуляция, завихренность и ротор векторного поля. Под *циркуляцией* векторного поля $\bar{a}(M)$ по контуру L понимается следующий криволинейный интеграл по замкнутому пути L , снабженному направлением обхода:

$$\Pi_L = \oint_L a_\tau dl,$$

где a_τ — проекция вектора поля $\bar{a}(M)$ на касательную к пути L в точке M , причем на этой касательной положительным считается то направление, которое совпадает с направлением обхода контура (рис. 185).

Очевидно, что циркуляция тем больше, чем ближе направление вектора $\bar{a}(M)$ в каждой точке M пути L к направлению указанной касательной. Очевидно, что изменение направления обхода контура L влечет за собой изменение знака циркуляции на обратный.

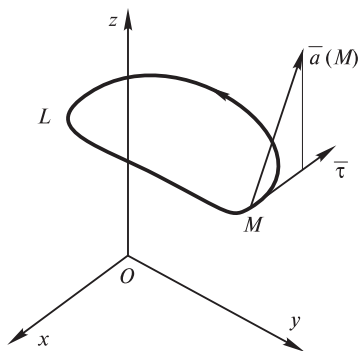


Рис. 185

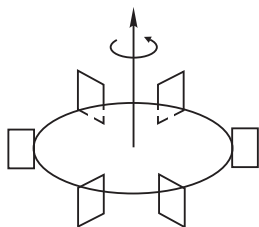


Рис. 186

Выясним физический смысл циркуляции. Пусть $\vec{a}$ — поле скоростей текущей жидкости, L — окружность, и представим себе, что эта окружность — периферия колесика с лопастями, расположенными по окружности этого колеса (рис. 186), могущего вращаться вокруг оси, проходящей через его центр перпендикулярно к его плоскости. Тогда в стационарном потоке чем больше циркуляция Π_L при данном радиусе колесика, тем быстрее будет оно вращаться. Например, в однородном поле, где вектор скорости $\vec{a}$ — постоянный вектор, такое колесико не будет вращаться ни при одном положении в потоке, так как на каждую его пару диаметрально противоположных лопаток будет действовать одна и та же сила (по величине и направлению). В этом случае, как мы увидим дальше (см. п. 6, примечание), действительно $\Pi_L = 0$. Но в потоке, образуемом вращением жидкости как твердого тела вокруг некоторой оси, колесико будет вращаться и притом с разной угловой скоростью в зависимости от угла наклона его оси к оси вращения жидкости. Быстрее всего оно будет вращаться тогда, когда его ось параллельна оси вращения жидкости (в этом случае на каждую его пару диаметрально противоположных лопаток будет действовать пара сил).

Величина циркуляции Π_L по плоскому контуру L , отнесенная к единице площади, ограниченной этим контуром, т. е. отношение $\frac{\Pi_L}{S}$, где S — площадь, ограниченная контуром L , характеризует *среднюю завихренность* потока вокруг оси, перпендикулярной плоскости контура L и проходящей через некоторую внутреннюю точку плоской области, ограниченной этим контуром.

Пусть дана точка M_0 поля $\vec{a}(M)$ и проходящая через нее ось направления $\vec{n}$ (рис. 187). В плоскости, содержащей точку M_0 и перпендику-

лярной $\vec{n}$, проведем контур L , обходящий точку M_0 в том направлении, в котором надо вращать правый винт, чтобы он имел поступательное движение в направлении $\vec{n}$. Такое направление обхода контура L будем называть *согласованным с $\vec{n}$* . Образуем

$$\lim_{\substack{L \rightarrow M_0 \\ (S \rightarrow 0)}} \frac{1}{S} \int_L a_\tau dl \quad (12)$$

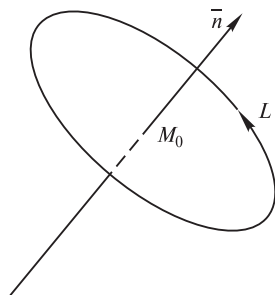


Рис. 187

при условии стягивания контура L в точку M_0 (тогда S стремится к нулю). Если существует предел (12), не зависящий от того, какие формы принимает контур L в своей плоскости в процессе стягивания в точку M_0 , то он называется *завихренностью* поля $\vec{a}(M)$ в точке M_0 вокруг направления $\vec{n}$ и обозначается через $w_{\vec{n}}(M_0)$:

$$\lim_{\substack{L \rightarrow M_0 \\ (S \rightarrow 0)}} \frac{1}{S} \int_L a_\tau dl = w_{\vec{n}}(M_0).$$

Рассмотрим вектор $\vec{b}$ такой, что

$$\text{Пр}_{\vec{n}} \vec{b}(M_0) = w_{\vec{n}}(M_0).$$

Последняя формула вполне определяет вектор $\vec{b}$ в данной точке поля M_0 , так как известна проекция этого вектора на любое направление $\vec{n}$. Вектор $\vec{b}$ называется *вихрем* или *ротором* данного векторного поля $\vec{a}$ в точке M_0 и обозначается $\text{rot } \vec{a}(M_0)$. Следовательно,

$$\text{Пр}_{\vec{n}} \text{rot } \vec{a}(M_0) = w_{\vec{n}}(M_0), \quad (13)$$

короче,

$$\text{rot } \vec{a}(M_0) = w_{\vec{n}}(M_0).$$

Из формулы (13) следует, что завихренность $w_{\vec{n}}(M_0)$ будет наибольшей в том случае, когда направление нормали $\vec{n}$ совпадает с направлением $\text{rot } \vec{a}(M_0)$.

Найдем формулу вычисления ротора в прямоугольной системе координат. Для этого вычислим, например, проекцию ротора на ось Oz . Пусть γ — замкнутый контур, лежащий в плоскости xOy , и S_G — площадь ограниченной им области G (рис. 188). Используя связь между криволинейными интегралами первого и

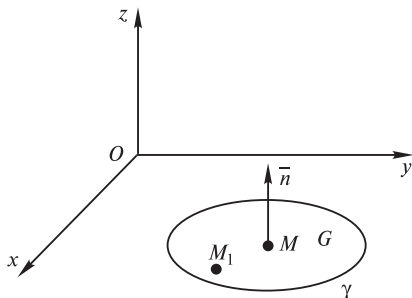


Рис. 188

второго рода (§ 12.3) и формулу Римана — Грина (§ 12.3, п. 4), можем записать, что

$$\oint_{\gamma} a_{\gamma} dl = \oint_{\gamma} a_x dx + a_y dy = \iint_G \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) dx dy.$$

Применяя к последнему интегралу теорему о среднем (§ 12.1, п. 3), находим:

$$\oint_{\gamma} a_{\gamma} dl = \left(\frac{\partial a_y(M_1)}{\partial x} - \frac{\partial a_x(M_1)}{\partial y} \right) S_G.$$

Здесь M_1 — некоторая точка области G . Отсюда в силу формулы (13) получаем:

$$\text{Pr}_{O_z} \text{rot } \bar{a} = \lim_{\substack{M_1 \rightarrow M \\ (\gamma \rightarrow M)}} \left(\frac{\partial a_y(M_1)}{\partial x} - \frac{\partial a_x(M_1)}{\partial y} \right) = \frac{\partial a_y(M)}{\partial x} - \frac{\partial a_x(M)}{\partial y}.$$

Таким же образом находим проекции ротора на оси Ox и Oy :

$$\text{Pr}_{O_x} \text{rot } \bar{a} = \frac{\partial a_z(M)}{\partial y} - \frac{\partial a_y(M)}{\partial z}, \quad \text{Pr}_{O_y} \text{rot } \bar{a} = \frac{\partial a_x(M)}{\partial z} - \frac{\partial a_z(M)}{\partial x}.$$

Таким образом,

$$\text{rot } \bar{a} = \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) \bar{i} + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right) \bar{j} + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) \bar{k}, \quad (14)$$

или для удобства запоминания

$$\text{rot } \bar{a} = \begin{pmatrix} \bar{i} & \bar{j} & \bar{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{pmatrix}.$$

Следовательно (см. § 3.3, п. 1), $\text{rot } \bar{a} = [\nabla, \bar{a}]$.

Пусть $\bar{n}_0 = \bar{i} \cos \alpha + \bar{j} \cos \beta + \bar{k} \cos \gamma$ — единичный вектор направления $\bar{n}$. Тогда с учетом формул (13) и (14) имеем:

$$w_{\bar{n}}(M_0) = \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) \cos \alpha + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right) \cos \beta + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) \cos \gamma.$$

Точно так же, как для градиента и дивергенции,

$$\text{rot}(C_1 \bar{a}_1 + C_2 \bar{a}_2) = C_1 \text{rot } \bar{a}_1 + C_2 \text{rot } \bar{a}_2.$$

Для ротора поля постоянного направления $\bar{a} = u \bar{c}$, где $\bar{c} = c_x \bar{i} + c_y \bar{j} + c_z \bar{k}$ — постоянный вектор, а $u = u(M)$ — скалярное поле, имеем:

$$\text{rot}(u \bar{c}) = -[\bar{c}, \text{grad } u]. \quad (15)$$

Формула (15) является частным случаем более общей формулы. Если $\vec{c} = \vec{c}(M)$ — переменный вектор, то $[\nabla, u\vec{c}] = -[\vec{c}, \nabla u] + u[\nabla, \vec{c}]$.

Пр и м е р 1. Если $\vec{c}$ — постоянный вектор, то $\text{rot } \vec{c} = 0$ и $w_n = 0$.

Пр и м е р 2. Если $\vec{a} = \vec{r}$ — радиус-вектор, то

$$\text{rot } \vec{a} = \left(\frac{\partial z}{\partial y} - \frac{\partial y}{\partial z} \right) \vec{i} + \left(\frac{\partial x}{\partial z} - \frac{\partial z}{\partial x} \right) \vec{j} + \left(\frac{\partial y}{\partial x} - \frac{\partial x}{\partial y} \right) \vec{k} = 0 \text{ и } w_n = 0.$$

Приведем теперь пример, поясняющий физический смысл ротора.

Пр и м е р 3. Вычислим ротор поля линейных скоростей вращающегося тела (см. § 14.2, п. 1, пример) $\vec{v} = -\omega y \vec{i} + \omega x \vec{j}$.

Имеем: $\text{rot } \vec{v} = 2\omega \vec{k}$.

Значит, $|\text{rot } \vec{v}| = 2\omega$, т. е. модуль вектора $\text{rot } \vec{v}$ равен удвоенной угловой скорости, с которой твердое тело вращается вокруг оси Oz . Отсюда происхождение названия «ротор», т. е. «вращение». Установленный механический смысл ротора распространяется и на поле скоростей текущей жидкости.

Пр и м е ч а н и е. Справедливо следующее соотношение:

$$(\nabla, \vec{a}_1 \times \vec{a}_2) = \vec{a}_2[\nabla, \vec{a}_1] - \vec{a}_1[\nabla, \vec{a}_2], \text{ т. е. } \text{div}(\vec{a}_1 \times \vec{a}_2) = \vec{a}_2 \text{ rot } \vec{a}_1 - \vec{a}_1 \text{ rot } \vec{a}_2,$$

что доказывается переходом к координатам входящих сюда векторов.

6. Формула Стокса (Джордж Габриель Стокс (1819—1903) — английский физик и математик). Для поверхностных интегралов имеет место формула, аналогичная формуле Римана—Грина (§ 12.3, п. 4), позволяющая свести вычисление интеграла по поверхности σ к вычислению криволинейного интеграла по контуру L , ограничивающему эту поверхность (рис. 189). Будем считать, что поверхность σ пересекается с любой прямой, параллельной оси Oz , не более чем в одной точке.

Пусть $z = z(x, y)$ — уравнение поверхности σ , где функции $z(x, y)$, $z'_x(x, y)$, $z'_y(x, y)$ непрерывны в замкнутой области G — проекции σ на плоскость xOy , L — контур, ограничивающий σ , а Γ — его проекция на плоскость xOy , являющаяся контуром, ограничивающим область G . Выберем для ориентации верхнюю сторону поверхности σ (рис. 189). Пусть функции $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ (кратко P , Q , R) непрерыв-

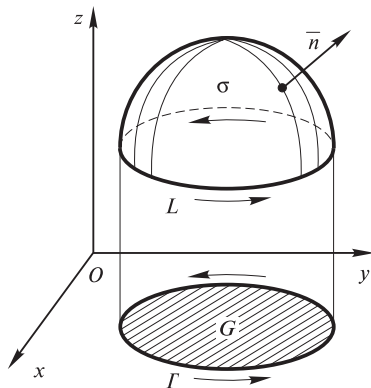


Рис. 189

ны вместе со своими частными производными первого порядка на поверхности σ .

Рассмотрим интеграл

$$\oint_L P(x, y, z) dx,$$

где направление обхода контура L согласовано с выбором стороны поверхности (см. рис. 189). В случае выбора нижней стороны поверхности σ направление обхода контура L также должно быть согласовано с таким выбором.

Так как контур L лежит на поверхности σ , то координаты его точек удовлетворяют уравнению $z = z(x, y)$, и поэтому значения функции P в точках контура L равны значениям функции $P(x, y, z(x, y))$ в соответствующих точках контура Γ . Проекции же соответствующих участков разбиения контуров L и Γ на ось Ox совпадают. Значит, совпадают интегральные суммы для криволинейных интегралов второго рода от функции P по контурам L и Γ , а значит, равны и интегралы

$$\oint_L P(x, y, z) dx = \oint_{\Gamma} P(x, y, z(x, y)) dx.$$

Далее, применяя формулу Римана—Грина (§ 12.3, п. 4), перейдем к двойному интегралу по области G . Получаем:

$$\oint_{\Gamma} P(x, y, z(x, y)) dx = - \iint_G \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} z'_y \right) dx dy.$$

Так как выбрана верхняя сторона поверхности σ (см. § 12.4, п. 3), то направляющие косинусы нормали $\vec{n}$ к ней выражаются формулами (§ 2.2, п. 3; § 11.2, п. 5):

$$\begin{aligned} \cos \alpha &= - \frac{z'_x}{\sqrt{1 + z'^2_x + z'^2_y}}, \\ \cos \beta &= - \frac{z'_y}{\sqrt{1 + z'^2_x + z'^2_y}}, \\ \cos \gamma &= \frac{1}{\sqrt{1 + z'^2_x + z'^2_y}}. \end{aligned}$$

Отсюда $\frac{\cos \beta}{\cos \gamma} = -z'_y$. Поэтому

$$- \iint_G \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} z'_y \right) dx dy = - \iint_G \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z} \frac{\cos \beta}{\cos \gamma} \right) dx dy.$$

Теперь, воспользовавшись формулами (14) и (15) § 12.4, можно этот двойной интеграл преобразовать в поверхностный. Получаем:

$$-\iint_{\sigma} \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z} \frac{\cos \beta}{\cos \gamma} \right) dx dy = -\iint_{\sigma} \left(\frac{\partial P}{\partial y} \cos \gamma - \frac{\partial P}{\partial z} \cos \beta \right) ds.$$

Итак,

$$\oint_L P dx = \iint_{\sigma} \left(\frac{\partial P}{\partial y} \cos \gamma - \frac{\partial P}{\partial z} \cos \beta \right) ds. \quad (16)$$

(Можно показать, что формула (16) верна и для поверхностей более сложного вида.)

Аналогично при соответствующих условиях доказывается:

$$\oint_L Q dy = \iint_{\sigma} \left(\frac{\partial Q}{\partial x} \cos \gamma - \frac{\partial Q}{\partial z} \cos \alpha \right) ds, \quad (17)$$

$$\oint_L R dz = \iint_{\sigma} \left(\frac{\partial R}{\partial y} \cos \alpha - \frac{\partial R}{\partial x} \cos \beta \right) ds. \quad (18)$$

Складывая почленно равенства (16), (17), (18), получаем формулу

$$\begin{aligned} & \oint_L P dx + Q dy + R dz = \\ & = \iint_{\sigma} \left[\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \cos \gamma + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \cos \alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \cos \beta \right] ds, \end{aligned} \quad (19)$$

называемую *формулой Стокса*.

Интеграл в левой части формулы (19) равен $\oint_L a_{\tau} dl$ (см. § 12.3) — циркуляции вектора $\vec{a} = P\vec{i} + Q\vec{j} + R\vec{k}$ вдоль контура L . Интеграл в правой части представляет поток вектора $\text{rot } \vec{a}$ через поверхность σ , ограниченную контуром L .

Следовательно, формулу Стокса в векторной форме можно записать так:

$$\oint_L a_{\tau} dl = \iint_{\sigma} \text{rot } \vec{a} \cdot \vec{n} ds, \quad (20)$$

Таким образом, *циркуляция вектора $\vec{a}$ вдоль замкнутого контура L равна потоку вихря этого вектора $\vec{a}$ через поверхность σ , лежащую в поле вектора $\vec{a}$ и ограниченную контуром L .*

Так как $\text{rot } \vec{a}$ есть завихренность $w_{\vec{n}}(M)$ в точке M вокруг направления $\vec{n}$, то равенство (20) может быть переписано в виде

$$\oint_L a_{\tau} dl = \iint_{\sigma} w_{\vec{n}}(M) ds.$$

Эта формула означает, что *интеграл по поверхности σ от завихренности поля вокруг направления нормали равен циркуляции вектора поля по контуру L , ограничивающему поверхность σ , направление обхода которого согласованно с направлением нормали.*

П р и м е ч а н и е. Из формулы Стокса следует, что циркуляция постоянного вектора вдоль любого контура равна нулю, так как $\operatorname{rot} \vec{c} = 0$ (п. 5, пример 1), и по формуле (20)

$$\oint_L a_{\zeta} dl = 0.$$

На этот факт мы ссылались в начале п. 5.

Формулу Стокса с помощью формулы связи поверхностных интегралов (§ 12.4, п. 5, формула (15)) можно переписать в виде

$$\begin{aligned} \oint_L Pdx + Qdy + Rdz = \\ = \iint_{\sigma} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) dzdx. \end{aligned} \quad (21)$$

Формулу (21) легко запомнить, заметив, что первое слагаемое в правой ее части — это то же выражение, которое стоит под знаком двойного интеграла в формуле Римана—Грина, а второе и третье получаются из него циклической перестановкой координат x, y, z и функций P, Q, R .

В частности, если поверхность σ — область плоскости xOy , ограниченная контуром L , то интегралы по $dzdx$ и $dydz$ обращаются в нуль и формула Стокса переходит в формулу Римана—Грина.

Из формулы Стокса следует, что если

$$\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}, \quad \frac{\partial R}{\partial y} = \frac{\partial Q}{\partial z}, \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \quad (22)$$

то криволинейный интеграл по любой пространственной замкнутой кривой L равен нулю:

$$\oint_L Pdx + Qdy + Rdz = 0. \quad (23)$$

А это значит, что в данном случае криволинейный интеграл $\int_{AB}$ не зависит от пути интегрирования.

Как и в случае плоской кривой, условия (22) являются необходимыми и достаточными для существования равенства (23).

При выполнении этих условий подынтегральное выражение $Pdx + Qdy + Rdz$ представляет собой полный дифференциал некоторой функции $u(x, y, z)$. Это устанавливается так же, как в § 12.3 для функции двух переменных.

§ 14.3. Дифференциальные операции второго порядка и их приложения

1. Дифференциальные операции второго порядка. Применение оператора ∇ к скалярным и векторным полям носит название *дифференциальных операций первого порядка*, так как ∇ является дифференциальным оператором первого порядка (его применение требует образования производных первого порядка). Будучи векторным оператором, ∇ может быть единственным образом применен к скалярному полю (ввиду наличия только одного вида произведения вектора на скаляр) и двумя путями применен к векторному полю (скалярным или векторным умножением). Это дает три уже известные нам дифференциальные операции первого порядка: градиент скалярного поля, дивергенцию и ротор векторного поля.

Важную роль играют и результаты двукратного применения к полям оператора ∇ , которые называются *дифференциальными операциями второго порядка*.

Рассмотрим все возможные дифференциальные операции второго порядка. Начнем со скалярного поля $u = u(M)$. Так как $\nabla u = \text{grad } u$ — векторное поле (см. § 14.1, п. 4), то к нему оператор ∇ может быть применен скалярным или векторным умножением, что дает первые две операции второго порядка:

$$(\nabla, \nabla u) = \text{div grad } u, \quad (1)$$

$$[\nabla, \nabla u] = \text{rot grad } u. \quad (2)$$

Пусть теперь имеем векторное поле $\bar{a} = \bar{a}(M)$. Так как $(\nabla, \bar{a}) = \text{div } \bar{a}$ есть скалярное поле, то к нему оператор ∇ может быть применен единственным образом, именно образованием

$$\nabla(\nabla, \bar{a}) = \text{grad div } \bar{a}. \quad (3)$$

Это третья операция второго порядка.

Наконец, так как $[\nabla, \bar{a}] = \text{rot } \bar{a}$ есть векторное поле, то к нему мы имеем возможность применить оператор ∇ скалярным и векторным умножением, что дает четвертую и пятую операции второго порядка:

$$(\nabla, [\nabla, \bar{a}]) = \text{div rot } \bar{a} \quad (4)$$

и

$$[\nabla, [\nabla, \bar{a}]] = \text{rot rot } \bar{a}. \quad (5)$$

Итак, получаются всего пять дифференциальных операций второго порядка. Две из этих операций, именно (2) и (4), тождественно равны нулю, так как в произведениях $[\nabla, \nabla]u$ и $(\nabla, [\nabla, \bar{a}])$ два «вектора» равны (см. § 2.2, п. 4 и 5).

Далее, переходя к операции (1), с использованием формул (3) из § 14.1 и (10) из § 14.2 получаем:

$$\operatorname{div} \operatorname{grad} u = \operatorname{div} \left(\frac{\partial u}{\partial x} \bar{i} + \frac{\partial u}{\partial y} \bar{j} + \frac{\partial u}{\partial z} \bar{k} \right) = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

Введем в рассмотрение следующий дифференциальный оператор второго порядка:

$$\nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

называемый *оператором Лапласа* (Пьер Симон Лаплас (1749—1827) — французский астроном, математик и физик) или *лапласианом*. Тогда

$$\begin{aligned} \Delta u &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}, \\ \Delta \bar{a} &= \frac{\partial^2 \bar{a}}{\partial x^2} + \frac{\partial^2 \bar{a}}{\partial y^2} + \frac{\partial^2 \bar{a}}{\partial z^2} \end{aligned} \quad (6)$$

(Δu — скаляр, $\Delta \bar{a}$ — вектор).

Из равенства (6) следует, что если $\bar{a} = a_x \bar{i} + a_y \bar{j} + a_z \bar{k}$, то

$$\Delta \bar{a} = \bar{i} \Delta a_x + \bar{j} \Delta a_y + \bar{k} \Delta a_z.$$

Формально умножив скалярно вектор ∇ сам на себя (взяв скалярный квадрат вектора ∇), найдем:

$$\nabla \nabla = \nabla^2 = \left(\frac{\partial}{\partial x} \right)^2 + \left(\frac{\partial}{\partial y} \right)^2 + \left(\frac{\partial}{\partial z} \right)^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \Delta,$$

поэтому оператор Лапласа часто обозначают еще символом ∇^2 . Итак,

$$\operatorname{div} \operatorname{grad} u = \nabla(\nabla u) = \nabla^2 u = \Delta u, \quad (7)$$

т. е. *дивергенция градиента скалярной функции равна лапласиану этой функции*.

Наконец, можно показать, что операции (5) и (3) связаны между собой соотношением

$$\operatorname{rot} \operatorname{rot} \bar{a} = \operatorname{grad} \operatorname{div} \bar{a} - \Delta \bar{a}. \quad (8)$$

(Доказывается с помощью перехода к координатам.)

Приведем пример применения дифференциальных операций второго порядка к выводу уравнения электромагнитных колебаний из уравнений Максвелла (Джеймс Клерк Максвелл (1831 — 1879) — английский физик) электромагнитного поля. *Уравнения*

Максвелла при отсутствии пространственных зарядов и соответствующем выборе системы единиц имеют вид:

$$\operatorname{rot} \bar{E} = -\mu \frac{\partial \bar{H}}{\partial t}, \quad \operatorname{rot} \bar{H} = 4\pi\gamma \bar{E} + \epsilon \frac{\partial \bar{E}}{\partial t}, \quad \operatorname{div}(\epsilon \bar{E}) = 0, \quad \operatorname{div}(\mu \bar{H}) = 0,$$

где ϵ — диэлектрическая постоянная, μ — магнитная проницаемость и γ — удельная проводимость (предполагаемые постоянными). Исключим теперь из этих уравнений вектор магнитной напряженности $\bar{H}$ для того, чтобы получить уравнение для одного только вектора $\bar{E}$ электрической напряженности. С этой целью образуем роторы обеих частей первого уравнения Максвелла:

$$\operatorname{rot} \operatorname{rot} \bar{E} = -\mu \operatorname{rot} \frac{\partial \bar{H}}{\partial t} = -\mu \frac{\partial}{\partial t} \operatorname{rot} \bar{H}$$

(последнее равенство справедливо потому, что дифференциальные операторы $\frac{\partial}{\partial t}$ и rot перестановочны в силу равенства смешанных производных). Поэтому в силу второго уравнения Максвелла

$$\operatorname{rot} \operatorname{rot} \bar{E} = -\mu \frac{\partial}{\partial t} \left(4\pi\gamma \bar{E} + \epsilon \frac{\partial \bar{E}}{\partial t} \right) = -4\pi\mu\gamma \frac{\partial \bar{E}}{\partial t} - \mu\epsilon \frac{\partial^2 \bar{E}}{\partial t^2}.$$

Но так как по формуле (8)

$$\operatorname{rot} \operatorname{rot} \bar{E} = \operatorname{grad} \operatorname{div} \bar{E} - \Delta \bar{E},$$

то, используя третье уравнение Максвелла, получим, что

$$\mu\epsilon \frac{\partial^2 \bar{E}}{\partial t^2} + 4\pi\mu\gamma \frac{\partial \bar{E}}{\partial t} - \Delta \bar{E} = 0. \quad (9)$$

(Так как четвертое уравнение Максвелла относится только к вектору магнитной напряженности, то оно при выводе уравнения (9) не используется.)

Поскольку $\bar{E}$ — вектор, то этому уравнению удовлетворяет каждая из его проекций, т. е. если обозначить через E любую из этих проекций, то

$$\mu\epsilon \frac{\partial^2 E}{\partial t^2} + 4\pi\mu\gamma \frac{\partial E}{\partial t} - \Delta E = 0.$$

Это уравнение называют *уравнением электромагнитных колебаний*.

2. Ортогональные криволинейные координаты. Коэффициенты Ламе. Прямоугольная система координат в пространстве и на плоскости чаще всего применяется при решении различных задач физики и техники. Тем не менее многие из этих задач проще решаются в других системах координат, из которых наиболее употребительны уже встречавшиеся нам ранее полярная система координат на плоскости, цилиндрическая и сферическая системы координат в пространстве.

Пусть x, y, z — прямоугольные координаты точек пространства. Рассмотрим три уравнения

$$x = x(u, v, w), \quad y = y(u, v, w), \quad z = z(u, v, w), \quad (10)$$

выражающие эти переменные через некоторую другую тройку переменных u, v, w . Предполагая, что уравнения (10) однозначно разрешимы относительно u, v, w , и решая эти уравнения относительно указанных переменных, получим:

$$u = u(x, y, z), v = v(x, y, z), w = w(x, y, z). \quad (11)$$

В дальнейшем предполагаем выполненными два условия:

1) функции u, v, w непрерывны и обладают непрерывными частными производными до второго порядка включительно;

2) разным тройкам (u, v, w) из области D^* изменения этих переменных отвечают различные тройки чисел (x, y, z) из соответствующей области D .

При этих условиях каждой тройке чисел (u, v, w) будет отвечать в пространстве точка M с координатами (x, y, z) , причем это соответствие будет и взаимно однозначным, ибо, наоборот, каждой точке $M(x, y, z)$ пространства будет отвечать определенная тройка чисел (u, v, w) . В силу сказанного переменные (u, v, w) можно принять в качестве координат точек пространства, причем если по крайней мере одно из уравнений (10) нелинейно, то эти переменные называются *криволинейными координатами*.

Рассмотрим поверхность

$$u(x, y, z) = C_1 \quad (12)$$

(кратко $u = C_1$).

На этой поверхности координата u сохраняет постоянное значение C_1 , такая поверхность называется *координатной поверхностью*. Аналогично рассматриваются координатные поверхности

$$v(x, y, z) = C_2 \quad (13)$$

(кратко $v = C_2$),

$$w(x, y, z) = C_3 \quad (14)$$

(кратко $w = C_3$).

Уравнения (12), (13), (14) с произвольными параметрами C_1, C_2, C_3 , дают три семейства координатных поверхностей рассматриваемой системы координат. Например, в цилиндрической системе координат это цилиндры с осью Oz , полуплоскости, ограниченные осью Oz , и плоскости, перпендикулярные к оси Oz , в сферической системе координат это сферы с центром в начале координат, конусы с осью Oz , полуплоскости, ограниченные осью Oz . Через каждую точку пространства $M_0(u_0, v_0, w_0)$ проходят три пересекающиеся координатные поверхности $u = u_0, v = v_0, w = w_0$.

Линия пересечения любых двух координатных поверхностей называется *координатной линией*. Так, координатная линия u есть линия пересечения поверхностей $v = v_0$ и $w = w_0$, на этой линии изменяется лишь величина u , а v и w сохраняют постоянные значения. Аналогично вдоль каждой координатной линии изменяется только одна из трех координат u, v, w . Параметрические уравнения координатных линий получаются из уравнений (10), если в них зафиксировать какие-нибудь две из трех криволинейных координат u, v, w . Всего имеем три семейства координатных линий. В цилиндрической системе координат это лучи, перпендикулярные оси Oz , окружности с центром на оси Oz в плоскостях, перпендикулярных к оси Oz , и прямые, параллельные оси Oz ; в сферической

системе координат это лучи, исходящие из полюса, дуги окружностей с центром в начале координат и окружности с центром на оси Oz .

Через каждую точку пространства проходят три пересекающиеся координатные линии системы (u, v, w) .

Обозначим через $\bar{e}_u$, $\bar{e}_v$, $\bar{e}_w$ единичные векторы касательных к соответствующим координатным линиям в рассматриваемой точке M_0 , направленные в сторону возрастания координат.

Система координат называется *ортогональной*, если ее координатные линии пересекаются только под прямым углом. Ортогональными являются, например прямоугольная, цилиндрическая и сферическая системы координат.

Система координат называется *криволинейной*, если среди ее координатных линий есть не прямые линии. Например, цилиндрическая и сферическая системы — ортогональные криволинейные координатные системы. Будем рассматривать только ортогональные криволинейные координатные системы. Поэтому указанные выше векторы $\bar{e}_u$, $\bar{e}_v$ и $\bar{e}_w$ будут попарно перпендикулярны. Их в этом случае называют *ортами* системы (u, v, w) .

Рассмотрим координатные линии системы (u, v, w) , проходящие через точку M_0 . Параметрические уравнения, например, координатной линии u получим, положив в (10) $v = v_0$, $w = w_0$:

$$x = x(u, v_0, w_0), y = y(u, v_0, w_0), z = z(u, v_0, w_0).$$

Под *элементами дуг координатных линий* u , v , w понимают дифференциалы длин дуг этих линий обозначения, соответственно dl_u , dl_v , dl_w . Как известно (см. § 9.9, п. 3),

$$dl_u = \sqrt{x_u'^2 + y_u'^2 + z_u'^2} du, dl_v = \sqrt{x_v'^2 + y_v'^2 + z_v'^2} dv, dl_w = \sqrt{x_w'^2 + y_w'^2 + z_w'^2} dw.$$

Величины

$$H_u = \sqrt{x_u'^2 + y_u'^2 + z_u'^2}, H_v = \sqrt{x_v'^2 + y_v'^2 + z_v'^2}, H_w = \sqrt{x_w'^2 + y_w'^2 + z_w'^2} \quad (15)$$

называются *коэффициентами Ламе* (Габриель Ламе (1795—1870) — французский математик) системы координат (u, v, w) .

П р и м е р 1. В случае цилиндрической системы координат имеем (см. § 12.2, п. 3) $x = r \cos \varphi$, $y = r \sin \varphi$, $z = z$ и, полагая $u = r$, $v = \varphi$, $w = z$, по формулам (15) получим:

$$H_r = 1, H_\varphi = r, H_z = 1. \quad (16)$$

П р и м е р 2. В случае сферической системы координат (см. § 12.2, п. 3)

$$x = r \sin \theta \cos \varphi; y = r \sin \theta \sin \varphi, z = r \cos \theta,$$

и, полагая $u = r$, $v = \varphi$, $w = \theta$, по формулам (15) получим:

$$H_r = 1, H_\varphi = r \sin \theta, H_\theta = r. \quad (17)$$

3. Градиент, дивергенция и лапласиан в ортогональных криволинейных координатах. Вычислим градиент скалярной функции $f = f(M) = f(u, v, w)$.

Подставляя в первое из уравнений (11) вместо x, y и z их выражения из (10), получим тождество

$$u(x(u, v, w), y(u, v, w), z(u, v, w)) = u.$$

Затем, продифференцировав это тождество по u , будем иметь:

$$u'_x x'_u + u'_y y'_u + u'_z z'_u = 1. \quad (18)$$

Аналогичные формулы справедливы для v и w .

Формула градиента сложной функции (§ 14.1, п. 5, правило 4) дает

$$\text{grad } f = f'_u \text{grad } u + f'_v \text{grad } v + f'_w \text{grad } w. \quad (19)$$

Известно (см. [15], т. II), что градиент скалярной функции $u(x, y, z)$ есть вектор, направленный перпендикулярно поверхности уровня $u = u_0$. Поэтому он коллинеарен вектору $\bar{e}_u$. Следовательно, имеем:

$$\text{grad } u = u'_x \bar{i} + u'_y \bar{j} + u'_z \bar{k} = h \bar{e}_u, \quad (20)$$

где h — скалярный множитель, подлежащий определению.

Производная векторной функции $\bar{r} = x\bar{i} + y\bar{j} + z\bar{k}$ по скалярному аргументу, например u , согласно п. 2 § 9.11 есть вектор, направленный по касательной к годографу вектора $\bar{r}(u)$ (рассматриваемого как функция u), т. е. по касательной к координатной линии u . Поэтому

$$r'_u = x'_u \bar{i} + y'_u \bar{j} + z'_u \bar{k} = H_u \bar{e}_u, \quad (21)$$

где H_u — соответствующий коэффициент Ламе (п. 2) системы (u, v, w) .

В результате скалярного умножения равенств (20) и (21) получим равенство

$$u'_x x'_u + u'_y y'_u + u'_z z'_u = h H_u, \quad (22)$$

левая часть которого равна единице по формуле (18). Поэтому $h = \frac{1}{H_u}$ и согласно формуле (20) имеем:

$$\text{grad } u = \frac{1}{H_u} \bar{e}_u. \quad (23)$$

Аналогично получим:

$$\text{grad } v = \frac{1}{H_v} \bar{e}_v, \quad (24)$$

$$\text{grad } w = \frac{1}{H_w} \bar{e}_w. \quad (25)$$

Теперь из формул (19), (23), (24), (25) находим выражение градиента скалярной функции точки в криволинейных ортогональных координатах:

$$\text{grad } f = \frac{1}{H_u} \frac{\partial f}{\partial u} \bar{e}_u + \frac{1}{H_v} \frac{\partial f}{\partial v} \bar{e}_v + \frac{1}{H_w} \frac{\partial f}{\partial w} \bar{e}_w. \quad (26)$$

Пример 1. Из формулы (26), используя выражения (16) для коэффициентов Ламе в цилиндрической системе координат (п. 2, пример. 1), найдем, что в этих координатах

$$\text{grad } f = \frac{\partial f}{\partial r} \bar{e}_r + \frac{1}{r} \frac{\partial f}{\partial \varphi} \bar{e}_\varphi + \frac{\partial f}{\partial z} \bar{e}_z.$$

Пример 2. Аналогично с помощью формул (17) (п. 2, пример 2) получим выражение для градиента в сферических координатах:

$$\text{grad } f = \frac{\partial f}{\partial r} \bar{e}_r + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \varphi} \bar{e}_\varphi + \frac{\partial f}{\partial \theta} \bar{e}_\theta.$$

Можно показать (см., например, [15], т. II), что выражения дивергенции вектора $\bar{a} = \bar{a}(M)$ в ортогональных криволинейных координатах имеет вид

$$\text{div} \bar{a} = \frac{1}{H_u H_v H_w} \left(\frac{\partial}{\partial u} (a_u H_v H_w) + \frac{\partial}{\partial v} (a_v H_u H_w) + \frac{\partial}{\partial w} (a_w H_u H_v) \right), \quad (27)$$

где a_u, a_v, a_w — проекции вектора $\bar{a}$ соответственно на направления $\bar{e}_u, \bar{e}_v, \bar{e}_w$.

Оператор Лапласа скалярной функции $f(u, v, w)$ согласно формуле (7) есть дивергенция от ее градиента. Поэтому, применяя формулу (27) к вектору $\bar{a} = \text{grad } f$ в форме (26), т. е. полагая в (27)

$$a_u = \frac{1}{H_u} \frac{\partial f}{\partial u}, \quad a_v = \frac{1}{H_v} \frac{\partial f}{\partial v}, \quad a_w = \frac{1}{H_w} \frac{\partial f}{\partial w},$$

получим выражение оператора Лапласа в любых ортогональных криволинейных координатах:

$$\Delta f = \frac{1}{H_u H_v H_w} \left(\frac{\partial}{\partial u} \left(\frac{H_v H_w}{H_u} \frac{\partial f}{\partial u} \right) + \frac{\partial}{\partial v} \left(\frac{H_u H_w}{H_v} \frac{\partial f}{\partial v} \right) + \frac{\partial}{\partial w} \left(\frac{H_u H_v}{H_w} \frac{\partial f}{\partial w} \right) \right).$$

Пример 3. Найдите выражения для лапласиана в цилиндрической и сферической системах координат.

С помощью формул (16) и (17) соответственно имеем:

$$\Delta f = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \varphi^2} + \frac{\partial^2 f}{\partial z^2}$$

и

$$\Delta f = \frac{1}{r^2} \left(\frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial^2 f}{\partial \varphi^2} \right).$$

4. Простейшие векторные поля и их основные свойства. Безвихревые поля. Векторное поле $\bar{a} = \bar{a}(M)$, в каждой точке которого

$$\text{rot} \bar{a} = 0, \quad (28)$$

называется *безвихревым* или *потенциальным*.

Так как для любого скалярного поля u $\text{rot grad } u = 0$ (§ 14.2), то *поле градиента любого скалярного поля безвихревое*.

Справедливо обратное предложение.

Вектор $\vec{a}$ безвихревого поля является градиентом скалярного поля u : $\vec{a} = \text{grad } u$.

Скаляр u называется *скалярным потенциалом* этого безвихревого поля.

В силу (28) и формулы Стокса (см. § 14.2, п. 6) в безвихревом поле $\vec{a}$ циркуляция по любому контуру равна нулю. С точки зрения течения жидкости равенство нулю циркуляции означает, что в потоке нет замкнутых струек жидкости, т. е. нет водоворотов.

Изучение безвихревого поля значительно отличается тем, что это поле вполне определяется заданием *одной* скалярной функции — его потенциалом.

Трубочатые поля. Векторное поле $\vec{a} = \vec{a}(M)$, в каждой точке которого

$$\text{div } \vec{a} = 0, \quad (29)$$

называется *трубочатым (соленоидальным)*, этот термин происходит от латинского слова *solen* — трубка).

Поясним смысл этого названия. Возьмем в этом поле какую-нибудь площадку σ_1 и проведем через каждую точку ее границы векторные линии (см. п. 1). Эти линии ограничивают часть пространства, называемую *векторной трубкой* (рис. 190). Жидкость при своем течении все время движется по такой трубке, не пересекая ее стенок.

Рассмотрим часть такой трубки, ограниченную площадкой σ_1 и каким-нибудь сечением σ_2 (см. рис. 190).

Согласно условию (29) формула Остроградского — Гаусса (§ 14.2, п. 3) дает

$$\int_{\sigma_1} a_n ds + \int_{\sigma_2} a_n ds + \int_{\sigma} a_n ds = 0,$$

где σ — боковая поверхность трубки, $\vec{n}$ — внешняя нормаль.

По определению векторной трубки векторы $\vec{a}$ и $\vec{n}$ на поверхности σ ортогональны, значит, их скалярное произведение равно нулю, и поэтому равен нулю интеграл по σ . Если в интеграле по поверхности σ_2 изменить направление нормали, то придем к равенству

$$\int_{\sigma_1} a_n ds = \int_{\sigma_2} a_n ds,$$

выражающему основное свойство трубчатого поля: *поток вектора трубчатого*

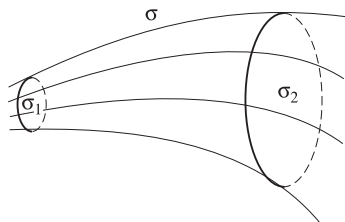


Рис. 190

поля через любое сечение векторной трубки в направлении векторных линий есть величина постоянная для данной трубки. Эта величина называется напряжением (интенсивностью) векторной трубки.

Так как для любого векторного поля $\bar{b} \operatorname{div} \operatorname{rot} \bar{b} = 0$ (см. п. 1), то:

1) поле ротора любого векторного поля трубчатое.

Справедливо и обратное предложение, которое приведем без доказательства:

2) каждое трубчатое поле является полем ротора некоторого векторного поля, т. е. если $\operatorname{div} \bar{a} = 0$, то существует такое векторное поле $\bar{b}$, что $\bar{a} = \operatorname{rot} \bar{b}$.

Вектор $\bar{b}$, ротор которого равен вектору данного трубчатого поля, называется векторным потенциалом этого трубчатого поля.

Глава 15. ЭЛЕМЕНТЫ ТЕОРИИ ФУНКЦИЙ КОМПЛЕКСНОЙ ПЕРЕМЕННОЙ

§ 15.1. Функции комплексной переменной

1. Основные определения. Внутренность круга радиуса ρ с центром в данной точке z_0 , т. е. совокупность точек удовлетворяющих неравенству $|z - z_0| < \rho$, называется *окрестностью точки* z_0 .

Число z_0 называется *пределом* последовательности комплексных чисел $\{z_n\}$ (обозначение $z_0 = \lim_{n \rightarrow \infty} z_n$), если для любого положительного числа ϵ можно подобрать такое число $N = N(\epsilon)$, что при $n > N$ выполняется неравенство $|z_0 - z_n| < \epsilon$, или $\lim_{n \rightarrow \infty} |z_0 - z_n| = 0$. Иначе говоря, как бы ни была мала окрестность точки z_0 , вне этой окрестности может оставаться лишь конечное число точек, тогда как все точки, начиная с номера $n > N$, попадут в окрестность точки z_0 .

Пусть $z_n = x_n + iy_n$, а $z_0 = x_0 + iy_0$. Так как $|z_n - z_0| = \sqrt{|x_n - x_0|^2 + |y_n - y_0|^2}$, то существование $\lim_{n \rightarrow \infty} z_n = z_0$ равносильно существованию двух пределов:

$$\lim_{n \rightarrow \infty} x_n = x_0 \quad \text{и} \quad \lim_{n \rightarrow \infty} y_n = y_0,$$

и обратно, из существования последних двух пределов следует, что

$$\lim_{n \rightarrow \infty} z_n = z_0.$$

Последовательность $\{z_n\}$, имеющая конечный предел, называется *сходящейся* последовательностью.

Если последовательность $\{z_n\}$ такова, что как бы ни было велико число M , найдется такое N , что для всех $n > N$ будет $|z_n| > M$, то говорят, что эта последовательность сходится к бесконечности. Записывают этот факт следующим образом:

$$\lim_{n \rightarrow \infty} z_n = \infty.$$

В теории функций комплексного переменного считают, что символу ∞ отвечает одна точка в комплексной плоскости; обозначают ее $z = \infty$ и называют бесконечно удаленной точкой.

Окрестностью бесконечно удаленной точки называется внешность любого круга.

При таком определении окрестности запись $\lim_{n \rightarrow \infty} z_n = \infty$ означает то же, что и в случае $\lim_{n \rightarrow \infty} z_n = z_0$. А именно, если $\lim_{n \rightarrow \infty} z_n = \infty$, то, какова бы ни была окрестность бесконечно удаленной точки, все точки последовательности $\{z_n\}$, за исключением конечного числа, находятся в этой окрестности.

Комплексная плоскость, дополненная бесконечно удаленной точкой, называется *расширенной* комплексной плоскостью.

Определение функции комплексного переменного аналогично определению функции действительного переменного.

Говорят, что на множестве E точек комплексной плоскости задана функция $w = f(z)$, если указан закон, по которому каждой точке z из множества E ставится в соответствие одно или несколько значений w . В первом случае функция называется однозначной, во втором — многозначной.

Например, функции $w = |z|$, $w = z^2$ — однозначные, функция $w = \sqrt[n]{z}$ ($n = 2, 3, \dots$) — многозначная (для каждого z ($n = 2, 3, \dots$) значений w).

Пусть $z = x + iy$, а $w = u + iv$, тогда задание функции $w = f(z)$ равносильно заданию двух функций $u = u(x; y)$ и $v = v(x, y)$ двух действительных переменных. Например, $w = z^2 = (x + y)^2 = x^2 - y^2 + 2xyi$, т. е. задание функции $w = z^2$ равносильно заданию функций $u(x; y) = x^2 - y^2$ и $v(x; y) = 2xy$.

2. Геометрический смысл функции комплексного переменного. Если значения переменной z изображать на одной комплексной плоскости, а значения w на другой, то функция $w = f(z)$ устанавливает соответствие между точками плоскости z и точками плоскости w . Мы говорим, что функция $w = f(z)$ отображает точки плоскости z на соответствующие точки плоскости w . Точки плоскости w называются образами соответствующих точек плоскости z , а точки плоскости z — прообразами.

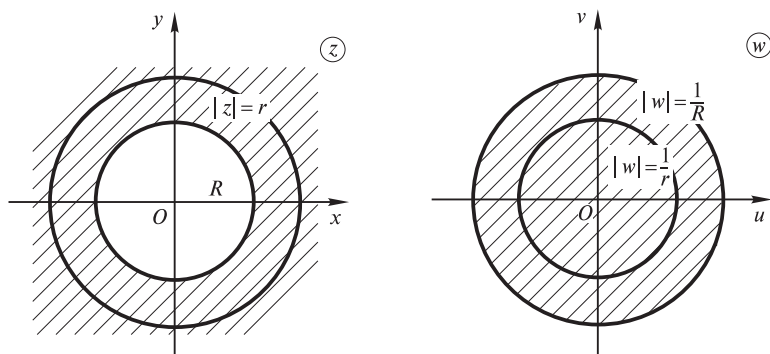


Рис. 191

П р и м е р. На какое множество отобразит функция $w = \frac{1}{z}$ внешность круга $|z| \geq R$ ($R < 1$),

Каждая окружность $|z| = r$ ($r > R$) отобразится в окружность $|w| = \frac{1}{r}$, причем граничная окружность $|z| = R$ отобразится в окружность $|w| = \frac{1}{R}$, т. е. внешность круга $|z| \geq R$ отобразится во внутренность круга $|w| = \frac{1}{R}$ (рис. 191).

3. Непрерывность функции комплексного переменного. Число w_0 называется *пределом однозначной функции* $w = f(z)$ при z стремящемся к z_0 , если для любого $\epsilon > 0$ найдется $\delta = \delta(\epsilon) > 0$ такое, что для всех значений z , удовлетворяющих неравенству $0 < |z - z_0| < \delta$, выполняется неравенство $|f(z) - w_0| < \epsilon$. Обозначается предел так же, как и для функции действительного переменного $\lim_{z \rightarrow z_0} f(z) = w_0$. Если z стремится к бесконечности, то определение предела функции следующее.

Число w_0 называется *пределом функции* $f(z)$ при z стремящемся к бесконечности, если для любого $\epsilon > 0$ найдется такое положительное число $N = N(\epsilon)$, что как только $|z| > N$, то $|f(z) - w_0| < \epsilon$.

Заметим, что теорема 4 о пределах и ее следствия 2, 3 из § 7.5 легко распространяются и на случай комплексного переменного.

Однозначная функция $f(z)$ называется *непрерывной* в точке z_0 , если она определена в точке z_0 и некоторой ее окрестности и если $\lim_{z \rightarrow z_0} f(z) = f(z_0)$.

Функция, непрерывная в каждой точке области D , называется *непрерывной* в этой области.

П р и м е р. Функции z , $\operatorname{Re} z$, $\operatorname{Im} z$, $|z|$ непрерывны во всей комплексной плоскости.

Так как определение непрерывности функции комплексного переменного с формальной стороны аналогично соответствующему определению для функции действительного переменного, то доказательства теорем об операциях над непрерывными

функциями остаются теми же в комплексной области, что и в действительном анализе. Так, *сумма, разность и произведение двух функций $f(z)$ и $\varphi(z)$, непрерывных в точке z_0 (в области G), есть функция, непрерывная в той же точке (в той же области)*; также *частное таких функций есть функция, непрерывная в точке z_0 (в области G), если $\varphi(z_0) \neq 0$ (если $\varphi(z) \neq 0$ в области G)*. Пользуясь этим предложением, мы можем, например, заключить, что любая целая рациональная функция $w = a_0 z^n + a_1 z^{n-1} + \dots + a_n$ непрерывна во всей плоскости комплексного переменного z ; далее, всякая рациональная функция $w = \frac{a_0 z^n + a_1 z^{n-1} + \dots + a_n}{b_0 z^m + b_1 z^{m-1} + \dots + a_m}$ непрерывна в каждой точке плоскости комплексного переменного z , за исключением тех значений z , при которых знаменатель равен нулю.

§ 15.2. Дифференцирование функций комплексной переменной

1. Определение аналитической функции. Пусть $w = f(z)$ однозначная функция комплексного переменного. Придадим независимому переменному $z = x + iy$ приращение $\Delta z = \Delta x + i\Delta y$, и вычислим приращение функции $\Delta w = f(z + \Delta z) - f(z)$. Мы говорим, что функция $f(z)$ дифференцируема в точке z , если отношение $\frac{\Delta w}{\Delta z}$ стремится к определенному ко-

нечному пределу, когда Δz любым образом стремится к нулю. Предел этого отношения мы называем производной функции $f(z)$ в точке z и обозначаем $f'(z)$, т. е.

$$\lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} = f'(z).$$

Требование независимости значения предела от закона стремления Δz к нулю накладывает на функцию $w = f(z)$ комплексного переменного более сильные ограничения, чем аналогичное требование для функции $y = f(x)$ действительного переменного, для которой для существования производной предел отношения $\frac{\Delta y}{\Delta x}$ должен быть одинаков только при стремлении Δx к нулю справа и слева.

О п р е д е л е н и е. Однозначная функция называется *аналитической* в области G , если она дифференцируема в каждой точке этой области.

Таким образом, функция может быть аналитической только в некоторой области; однако и о каждой отдельной точке такой области говорят, что в ней эта функция аналитическая. При этом важно заметить, что функция, аналитическая в точке, по определению должна быть аналитической в некоторой окрестности этой точки.

Так как определение производной для функции комплексного переменного с формальной стороны аналогично соответствующему

определению для функции действительного переменного, то все известные из дифференциального исчисления правила дифференцирования могут быть легко перенесены и на функции комплексной переменной. Отсюда, в частности, следует, что целая рациональная функция есть аналитическая во всей плоскости; функция рациональная есть аналитическая во всей плоскости комплексного переменного, кроме точек, в которых ее знаменатель обращается в нуль.

2. Условно Коши — Римана. *Т е о р е м а 1. Пусть функция $w = u + iv$ дифференцируема в точке z , тогда ее действительная и мнимая части удовлетворяют в ней условиям Коши — Римана:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

Д о к а з а т е л ь с т в о. Так как функция $w = u + iv$ дифференцируема в точке z , то $\lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z}$ существует и не зависит от способа стремления приращения аргумента Δz к нулю. Будем считать, что $\Delta y = 0$ и $\Delta z = \Delta x$ (точка $z + \Delta z$ приближается к точке z по прямой, параллельной действительной оси). Тогда

$$f'(z) = \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta x \rightarrow 0} \frac{\Delta u + i\Delta v}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} + i \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}.$$

С другой стороны, если считать, что $\Delta x = 0$ и $\Delta z = i\Delta y$, то

$$f'(z) = \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z} = \lim_{\Delta y \rightarrow 0} \frac{\Delta u + i\Delta v}{i\Delta y} = \frac{1}{i} \lim_{\Delta y \rightarrow 0} \frac{\Delta u}{\Delta y} + \lim_{\Delta y \rightarrow 0} \frac{\Delta v}{\Delta y} = -\frac{\partial u}{\partial y} i + \frac{\partial v}{\partial y}.$$

Итак, имеем

$$\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} i.$$

Приравнивая в последнем равенстве действительные части мнимые части, получим искомые условия Коши — Римана.

Условия Коши — Римана являются необходимыми, но не достаточными для дифференцируемости функции комплексного переменного. Чтобы функция $w = f(z) = u(x; y) + iv(x, y)$ была дифференцируемой в точке z , надо дополнительно потребовать, чтобы функции $u(x; y)$ и $v(x; y)$ были дифференцируемыми, как функции двух переменных, т. е. справедлива следующая теорема.

Т е о р е м а 2. Пусть функции $u = u(x; y)$ и $v = v(x; y)$, дифференцируемые в точке $M(x; y)$ как функции двух переменных x и y , удовлетворяют условиям Коши — Римана. Тогда функция $w = u + iv$ является дифференцируемой в точке $z = x + iy$.

3. Сопряженные гармонические функции. Рассмотрим функцию $w = u + iv$ аналитическую в некоторой области D . Ее действительная и мнимая части удовлетворяют условиям Коши — Римана:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad (1)$$

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \quad (2)$$

Дифференцируя равенство (1) по x , а равенство (2) по y и складывая полученные соотношения

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y} \text{ и } \frac{\partial^2 v}{\partial x \partial y} = -\frac{\partial^2 u}{\partial y^2},$$

найдем, что

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (3)$$

Аналогично, дифференцируя равенство (1) по y , а (2) по x и складывая результаты, найдем

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

Уравнение (3) называется *уравнением Лапласа*, а функции, удовлетворяющие уравнению (3), называются *гармоническими*.

Таким образом, мы показали, что функции $u(x; y)$ и $v(x; y)$ являются гармоническими функциями.

Две гармонические функции u и v , которые в комбинации $u + iv$ дают аналитическую функцию, называются *сопряженными гармоническими функциями*.

З а м е ч а н и е. Функция $u + iv$, скомбинированная из двух произвольных гармонических функций, может не быть аналитической.

Если известна одна гармоническая функция, например, $u(x; y)$, то вторую гармоническую $v(x; y)$, сопряженную ей, можно всегда найти, используя условия Коши — Римана.

П р и м е р. Пусть $u(x; y) = x^2 - y^2$. Легко проверить, что эта функция гармоническая. Найдем сопряженную ей $v(x; y)$. Имеем:

$$\frac{\partial u}{\partial x} = 2x; \quad \frac{\partial u}{\partial y} = -2y.$$

$$v(x; y) = \int 2x dy + C(x) = 2xy + C(x),$$

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}; \quad \frac{\partial v}{\partial x} = 2y + C'(x),$$

$$2y + C'(x) = 2y, \text{ т. е. } C'(x) = 0, \text{ откуда } C(x) = \text{const}$$

и $v(x; y) = 2xy + C$, где C — произвольная постоянная.

§ 15.3. Интегралы по комплексному переменному

1. Понятие интеграла по комплексному переменному. Интеграл от непрерывной на гладкой кривой (контуре) γ с началом в точке $A = z_0$ и концом в точке $B = z_n$ (рис. 192) функции $f(z)$ ($z = x + iy$) определяется как предел интегральной суммы

$$\int_{\gamma} f(z) dz = \lim_{\max|\Delta z_k| \rightarrow 0} \sum_{k=1}^n f(z_k) \Delta z_k, \quad (1)$$

где $\Delta z_k = z_k - z_{k-1}$.

Полагая $f(z) = u(x, y) + iv(x, y)$, $z_k = x_k + iy_k$, $\Delta z_k = \Delta x_k + i\Delta y_k$, получим

$$\begin{aligned} f(z_k) \Delta z_k &= (u(x_k, y_k) + iv(x_k, y_k))(\Delta x_k + i\Delta y_k) = \\ &= u(x_k, y_k) \Delta x_k - v(x_k, y_k) \Delta y_k + i(u(x_k, y_k) \Delta y_k + v(x_k, y_k) \Delta x_k) \end{aligned}$$

и, значит, равенство (1) примет вид

$$\int_{\gamma} f(z) dz = \int_{\gamma} u dx - v dy + i \int_{\gamma} v dx + u dy, \quad (2)$$

причем криволинейные интегралы второго рода справа существуют как от непрерывных функций по гладкой кривой (§ 12.3, п. 3), значит, существует и интеграл по комплексному переменному слева. Таким образом, интеграл по комплексному переменному сводится к изученным ранее (§ 12.3) криволинейным интегралам второго рода. Отсюда получаем следующие свойства интеграла по комплексному переменному:

1. $\int_{\gamma} f(z) dz = - \int_{\gamma^-} f(z) dz.$
2. $\int_{\gamma} cf(z) dz = c \int_{\gamma} f(z) dz.$
3. $\int_{\gamma} (f(z) \pm g(z)) dz = \int_{\gamma} f(z) dz \pm \int_{\gamma} g(z) dz.$
4. $\int_{\gamma} f(z) dz = \int_{\gamma_1} f(z) dz + \int_{\gamma_2} f(z) dz.$

(В записанных выше выражениях, c — комплексная постоянная, γ_1 и γ_2 — две кривые, составляющие кривую γ ; γ^- — кривая, совпадающая с кривой γ по форме и положению, но с противоположным направлением обхода.)

5. Если функция $f(z)$ ограничена на контуре γ : $|f(z)| \leq M$ (M — положительная постоянная), то

$$\left| \int_{\gamma} f(z) dz \right| \leq Ml$$

(l — длина контура γ).

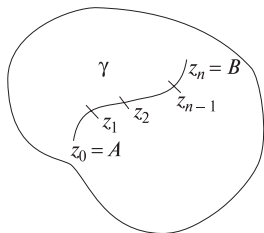


Рис. 192

Заметим, что запоминать формулу (2) нет необходимости. Надо только найти действительную и мнимую части подинтегрального выражения $f(z)dz$, заменив dz на $dx + i dy$.

Пример 1. Вычислим $\int_{\gamma} (\bar{z} - \bar{a}) dz$, где γ — окружность $|z - a| = R$, проходимая в положительном направлении.

Пусть $z = x + iy$, $a = a_1 + ia_2$.

Параметрические уравнения данной окружности имеют вид:

$$x = a_1 + R \cos t, y = a_2 + R \sin t \quad (0 \leq t \leq 2\pi).$$

Поэтому

$$\begin{aligned} \int_{\gamma} (\bar{z} - \bar{a}) dz &= \int_0^{2\pi} (R \cos t - i R \sin t) (d(a_1 + R \cos t) + i d(a_2 + R \sin t)) = \\ &= \int_0^{2\pi} (R \cos t - i R \sin t) (-R \sin t + i R \cos t) dt = i \int_0^{2\pi} (R \cos t - i R \sin t) (R \cos t + i R \sin t) dt = \\ &= i R^2 \int_0^{2\pi} (\cos^2 t + \sin^2 t) dt = i R^2 \int_0^{2\pi} dt = 2\pi R^2 i. \end{aligned}$$

Пример 2. Вычислим

$$\int_{\gamma} \frac{dz}{z - a},$$

где γ — окружность $|z - a| = R$, проходимая в положительном направлении.

Имеем (см. пример 1)

$$\int_{\gamma} \frac{dz}{z - a} = \int_{\gamma} \frac{\bar{z} - \bar{a}}{|z - a|^2} dz = \frac{1}{R^2} \int_{\gamma} (\bar{z} - \bar{a}) dz = \frac{1}{R^2} \cdot 2\pi R^2 i = 2\pi i.$$

2. Теорема Коши. Теорема 1 (основная теорема Коши). Если функция $f(z) = u + iv$ является аналитической в односвязной области G , то для любого замкнутого контура γ , лежащего в области G , справедливо равенство

$$\int_{\gamma} f(z) dz = 0. \quad (3)$$

Доказательство. Для упрощения доказательству потребуем дополнительно, чтобы производная $f'(z)$ была непрерывна в G (теорема верна и без этого дополнительного условия (см., например [13])). Тогда функции u и v вместе со своими частными производными первого порядка непрерывны в области G и в ней выполняются условия Коши — Римана. И, значит, (§ 12.3, п. 5, теорема 2) криволинейные интегралы в формуле (2) справа не зависят от пути интегрирования, что равносильно выполнению равенства (3) (§ 12.3, п. 5, теорема 1) для любого замкнутого контура γ , лежащего в G .

О п р е д е л е н и е. Будем говорить, что функция $f(z)$ является аналитической в замкнутой области $\bar{G}$, если она является аналитической в некоторой области G' , содержащей замкнутую область $\bar{G}$ (т. е. область G и ее границу Γ).

Из теоремы 1 следует теорема 2.

Т е о р е м а 2 (т е о р е м а К о ш и). Если функция $f(z)$ является аналитической в замкнутой области $\bar{G}$, то

$$\int_{\Gamma} f(z) dz = 0.$$

Последняя теорема распространяется и на случай многосвязной области (о многосвязной области см. §11.1, п. 5).

Пусть область G ограничена внешним контуром γ_0 и внутренними контурами $\gamma_1, \gamma_2, \dots, \gamma_n$, причем каждый из последних n контуров лежит вне остальных и все они расположены внутри γ_0 . В этом случае граница области G представляет собой сложный контур $\Gamma = \gamma_0 + \gamma_1^- + \gamma_2^- + \dots + \gamma_n^-$. Тогда при движении по сложному контуру Γ точки области G остаются слева — положительное направление обхода (см. рис. 193).

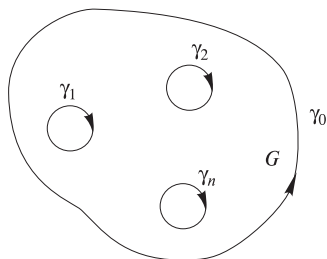


Рис. 193

Т е о р е м а 3 (т е о р е м а К о ш и). Если функция $f(z)$ является аналитической в замкнутой области $\bar{G}$, то

$$\int_{\Gamma} f(z) dz = 0,$$

или

$$\int_{\gamma_0} f(z) dz = \int_{\gamma_1} f(z) dz + \int_{\gamma_2} f(z) dz + \dots + \int_{\gamma_n} f(z) dz.$$

3. Вычисление интеграла $\oint_{\gamma} \frac{dz}{(z-a)^n}$. Вычислим интеграл $\oint_{\gamma} \frac{dz}{(z-a)^n}$, где

γ — замкнутый контур, ограничивающий односвязную область G . Рассмотрим различные случаи:

1. n — целое отрицательное число. Обозначим $m = -n$. Тогда m — целое положительное число и, значит, по теореме 3 $\oint_{\gamma} \frac{dz}{(z-a)^n} = \oint_{\gamma} (z-a)^m dz = 0$,

т. к. $(z-a)^m$ — аналитическая функция во всей комплексной плоскости, в частности в $\bar{G}$.

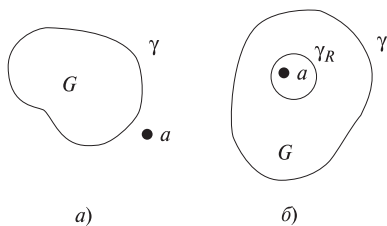


Рис. 194

2. n — целое положительное число и точка $z = a$ лежит вне области G (рис. 194, а). Тогда $\int_{\gamma} \frac{dz}{(z-a)^n} = 0$, т. к. функция $\frac{1}{(z-a)^n}$ — аналитическая в замкнутой области $\bar{G}$.

3. n — целое положительное число и точка $z = a$ лежит внутри области, ограниченной контуром γ (рис. 194, б). Опишем окружность γ_R радиуса R : $|z - a| = R$, целиком лежащую в области G . Обозначим через G' двусвязную область, ограниченную замкнутыми контурами: γ и γ_R . В области $\bar{G}'$ функция $\frac{1}{(z-a)^n}$ — аналитическая и по предыдущей теореме.

$$\int_{\gamma} \frac{dz}{(z-a)^n} = \int_{\gamma_R} \frac{dz}{(z-a)^n}. \quad (4)$$

Рассмотрим отдельно случай, когда $n \neq 1$ (случай $n = 1$ был рассмотрен ранее: п. 1, пример 2). Вычислим интеграл, стоящий в правой части равенства (4). Имеем

$$\begin{aligned} \int_{\gamma} \frac{dz}{(z-a)^n} &= \int_{\gamma_R} \frac{dz}{(z-a)^n} = \frac{1}{R^{2n}} \int (\bar{z} - \bar{a})^n dz = \frac{i}{R^{n-1}} \int_0^{2\pi} (\cos \varphi - i \sin \varphi)^{n-1} d\varphi = \\ &= \frac{i}{R^{n-1}} \left(\int_0^{2\pi} \cos(1-n)\varphi d\varphi + i \int_0^{2\pi} \sin(1-n)\varphi d\varphi \right) = 0 \end{aligned}$$

(см. § 10.4, п. 1).

4. Формула Коши. Пусть G — односвязная область, ограниченная кусочно-гладкой линией γ , $f(z)$ — функция, аналитическая в замкнутой области $\bar{G}$. Тогда имеет место следующая формула Коши

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta) d\zeta}{\zeta - z},$$

где z — любая точка внутри γ , а интегрирование по контуру γ совершается в положительном направлении.

Эта формула распространяется и на случай многосвязной области G , границей которой является сложный контур Γ , состоящий из конечного числа кусочно-гладких замкнутых линий.

Если $f(z)$ — аналитическая функция в замкнутой области $\bar{G}$, то

$$f(z) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(\zeta) d\zeta}{\zeta - z},$$

где z — любая точка области G , а интегрирование выполняется вдоль сложного контура Γ в положительном направлении.

5. Интеграл Коши и интеграл типа Коши.

Выражение

$$\frac{1}{2\pi i} \int_{\Gamma} \frac{f(\zeta) d\zeta}{\zeta - z},$$

где $f(z)$ — функция, аналитическая в замкнутой области $\bar{G}$, границей которой служит контур Γ , называется *интегралом Коши*.

Интеграл Коши изображает данную функцию $f(z)$, аналитическую в $\bar{G}$, во всякой точке, внутренней к контуру Γ .

Во всякой точке z , лежащей вне замкнутой области $\bar{G}$, интеграл Коши равен нулю, ибо функция $\frac{f(\zeta) d\zeta}{\zeta - z}$ для всех точек z , не лежащих в области $\bar{G}$, — аналитическая в $\bar{G}$.

Пусть L — произвольная кусочно-гладкая линия (не обязательно замкнутая), $\varphi(z)$ — заданная на L непрерывная функция.

Выражение

$$\frac{1}{2\pi i} \int_L \frac{\varphi(\zeta) d\zeta}{\zeta - z}$$

имеет определенное значение для любой точки z , не лежащей на L , и, значит, определяет однозначную функцию $F(z)$ во всех точках z , не принадлежащих L .

Это выражение называется *интегралом типа Коши*.

Т е о р е м а. *Функция $F(z)$, определенная интегралом типа Коши*

$$F(z) = \frac{1}{2\pi i} \int_L \frac{\varphi(\zeta) d\zeta}{\zeta - z},$$

есть аналитическая во всякой односвязной области G , не содержащей точек линии L , и для ее производной имеет место формула

$$F'(z) = \frac{1}{2\pi i} \int_L \frac{\varphi(\zeta) d\zeta}{(\zeta - z)^2}.$$

Доказывается, что это дифференцирование можно повторить n раз, поэтому функция $F(z)$, определяемая интегралом типа Коши, имеет в каждой точке z , не лежащей на L , производные всех порядков, причем

$$F^{(n)}(z) = \frac{n!}{2\pi i} \int_L \frac{\varphi(\zeta) d\zeta}{(\zeta - z)^{n+1}}.$$

6. Существование производных любого порядка для аналитической функции. Известно, что в случае функции действительной переменной из факта существования первой производной ничего нельзя сказать о

существовании производных высших порядков. Для функций комплексной переменной справедлива следующая теорема.

Т е о р е м а. *Если однозначная функция $f(z)$ имеет всюду в области G первую производную, то она имеет в этой области производные любого порядка.*

Д о к а з а т е л ь с т в о. Пусть z — произвольная точка области G , γ — кусочно-гладкий замкнутый контур, окружающий точку z и лежащий вместе со всеми своими внутренними точками в области G . По формуле Коши

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta) d\zeta}{\zeta - z}. \quad (5)$$

Но на основании п. 5 функция $f(z)$, изображаемая интегралом Коши, дифференцируема в точке z любое число раз, причем

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_{\gamma} \frac{f(\zeta) d\zeta}{(\zeta - z)^{n+1}} \quad (n = 1, 2, 3, \dots), \quad (6)$$

что и доказывает теорему.

7. Принцип максимума. Теоремы Лиувилля и Морера. Замечательным свойством аналитической функции является принцип максимума ее модуля.

Модуль функции $f(z)$, аналитической в области G и непрерывной в $\bar{G}$, не может достигать своего наибольшего значения ни в какой точке области G , кроме случая, когда $f(z) = \text{const}$.

Для аналитических функций справедливы следующие теоремы.

Т е о р е м а Л и у в и л л я (Жозеф Лиувиль (1809—1882) — французский математик). *Если функция $f(z)$ аналитична во всей плоскости и ограничена в ней по модулю, то она постоянна.*

Т е о р е м а М о р е р а (Гиацинте Морера (1856—1909) — итальянский математик). *Если функция $f(z)$ непрерывна в односвязной области G и интеграл от этой функции, взятый по любому замкнутому контуру γ , целиком лежащему в области G , равен нулю ($\int_{\gamma} f(z) dz = 0$), то функция $f(z)$ аналитична в этой области.*

§ 15.4. Равномерно сходящиеся ряды функций комплексной переменной

1. Числовые ряды. Рассмотрим числовой ряд с комплексными членами, т. е. ряд

$$\sum_{n=1}^{\infty} u_n, \quad (1)$$

где $u_n = \alpha_n + i\beta_n$ ($n = 1, 2, 3, \dots$).

Для этого ряда, как и для ряда с действительными членами § 10.1, (п. 1), вводится понятие n -й частичной суммы

$$S_n = u_1 + u_2 + \dots + u_n.$$

Ряд (1) называется *сходящимся*, если существует предел $\lim_{n \rightarrow \infty} S_n = S$, при этом число S называется *суммой* ряда (1). Здесь пользуемся понятием *предела последовательности комплексных чисел* и понимаем его в следующем смысле: число S называется пределом последовательности $\{S_n\}$, если для любого $\varepsilon > 0$ существует такой номер N , зависящий от ε , что для всех $n > N$ выполняется неравенство

$$|S_n - S| < \varepsilon, \quad (2)$$

или, что то же, $\lim_{n \rightarrow \infty} |S_n - S| = 0$.

Так же, как в случае рядов с действительными членами (§ 10.1, п. 2), устанавливается необходимый признак сходимости ряда (1): общий член u_n сходящегося ряда (1) стремится к нулю при неограниченном возрастании n , т. е. $\lim_{n \rightarrow \infty} u_n = 0$.

П р и м е р. Для ряда

$$1 + q + q^2 + \dots + q^{n-1} + \dots \quad (3)$$

(q — комплексное число, отличное от единицы) частичная сумма

$$S_n = \frac{1 - q^n}{1 - q}, \quad (4)$$

так как $(1 - q)(1 + q + q^2 + \dots + q^{n-1}) = 1 - q^n$.

Из равенства (4) имеем:

$$\left| S_n - \frac{1}{1 - q} \right| = \frac{|q|^n}{|1 - q|}.$$

Отсюда при $|q| < 1$

$$\lim_{n \rightarrow \infty} \left| S_n - \frac{1}{1 - q} \right| = 0.$$

Значит, при $|q| < 1$ ряд (3) сходится и его сумма S равна $\frac{1}{1 - q}$.

При $|q| \geq 1$ ряд (3) расходится, так как в этом случае его общий член q^{n-1} не стремится к нулю при неограниченном возрастании n .

Следующая теорема позволяет свести изучение рядов (1) с комплексными членами к исследованию таких рядов с действительными членами

$$\sum_{n=1}^{\infty} \alpha_n \quad \text{и} \quad \sum_{n=1}^{\infty} \beta_n. \quad (5)$$

Т е о р е м а 1. *Ряд (1) сходится тогда и только тогда, когда сходятся ряды (5).*

Д о к а з а т е л ь с т в о. Пусть ряд (1) сходится. Представим его частичную сумму S_n и сумму S в алгебраической форме: $S_n = a_n + ib_n$, $S = a + ib$, где a_n и b_n — частичные суммы рядов (5). Тогда условие (2) примет вид:

$$|S_n - S| = |a_n - a + i(b_n - b)| = \sqrt{(a_n - a)^2 + (b_n - b)^2} < \varepsilon \quad (n > N). \quad (6)$$

Из неравенства (6) следуют неравенства

$$|a_n - a| < \varepsilon, |b_n - b| < \varepsilon \quad (n > N), \quad (7)$$

и поэтому сходятся ряды (5) соответственно к a и b . Если же ряды (5) сходятся соответственно к a и b , то выполняются неравенства (7), а вместе с ними и неравенство $|S_n - S| < \varepsilon\sqrt{2}$, показывающее сходимость ряда (1).

Т е о р е м а 2. *Если сходится ряд*

$$\sum_{n=1}^{\infty} |u_n|, \quad (8)$$

то сходится и ряд (1).

Д о к а з а т е л ь с т в о. Имеем:

$$|u_n| = |\alpha_n + i\beta_n| = \sqrt{\alpha_n^2 + \beta_n^2} \geq |\alpha_n|, \quad |u_n| \geq |\beta_n|.$$

На основании признака сравнения рядов (§ 10.1 п. 3, теорема 2) заключаем, что сходятся ряды $\sum_{n=1}^{\infty} |\alpha_n|$ и $\sum_{n=1}^{\infty} |\beta_n|$, значит, и ряды (5). Следовательно, по теореме 1 сходится ряд (1).

Как и в случае рядов с действительными членами, ряд (1) называется *абсолютно сходящимся*, если ряд (8) сходится. Если же ряд (1) сходится, а ряд (8) расходится, то ряд (1) называется *условно сходящимся*.

На абсолютно сходящиеся ряды с комплексными членами распространяются имеющие место для абсолютно сходящихся рядов с действительными членами переместительное свойство и свойство перемножения рядов (§ 10.1, п. 5).

2. Функциональные ряды. Пусть в области G определена бесконечная последовательность однозначных функций комплексной переменной $\{f_n(z)\}$. Выражение вида

$$\sum_{n=1}^{\infty} f_n(z) \quad (9)$$

будем называть *функциональным рядом*. При фиксированном значении z_0 из области G ряд (9) превращается в числовой ряд.

Функциональный ряд (9) называется сходящимся в области G , если при любом z из области G соответствующий ему числовой ряд сходится. Если ряд (9) сходится в области G , то в этой области можно определить однозначную функцию $f(z)$, значение которой в каждой точке области G равно сумме соответствующего числового ряда. Эта функция называется *суммой* ряда (9) в области G . В силу данных определений в этом случае для любой фиксированной точки z из области G и любого заданного положительного числа ε можно указать такой номер $N = N(\varepsilon, z)$, что

$$|f(z) - S_n(z)| < \varepsilon \quad (10)$$

при $n > N$, где $S_n(z) = f_1(z) + f_2(z) + \dots + f_n(z)$ — частичная сумма ряда (9).

Если для любого положительного числа ε можно указать такой номер $N = N(\varepsilon)$, что при $n > N$ неравенство (10) выполняется сразу для всех точек z области G , то ряд (9) называется *равномерно* сходящимся в области G .

Обозначив $r_n(z) = \sum_{k=n+1}^{\infty} f_k(z)$, условие равномерной сходимости ряда

(9) можем записать в виде $|r_n(z)| < \varepsilon$ в G при $n > N(\varepsilon)$.

Как и в случае рядов функций действительного переменного (§ 10.2, п. 2), устанавливается достаточный признак равномерной сходимости (признак Вейерштрасса).

Т е о р е м а 1 (признак Вейерштрасса). Если всюду в области G члены функционального ряда (9) удовлетворяют условию

$$|f_n(z)| \leq a_n \quad (n=1, 2, 3, \dots),$$

где a_n — постоянные положительные числа, причем числовой ряд

$$a_1 + a_2 + \dots + a_n + \dots$$

сходится, то данный ряд сходится равномерно (и притом абсолютно) в области G .

Укажем и следующие теоремы.

Т е о р е м а 2. Если функции $f_n(z)$ ($n=1, 2, \dots$) непрерывны в области G , а ряд (9) сходится в этой области равномерно к функции $f(z)$, то $f(z)$ также непрерывна в области G .

Т е о р е м а 3. Если ряд (9) непрерывных функций $f_n(z)$ ($n=1, 2, \dots$) сходится равномерно в области G к функции $f(z)$, то интеграл от этой функции по любой кусочно-гладкой кривой Γ , целиком лежащей в области G , можно вычислить путем почленного интегрирования ряда (9), т. е.

$$\int_{\Gamma} f(\zeta) d\zeta = \sum_{n=1}^{\infty} \int_{\Gamma} f_n(\zeta) d\zeta.$$

Т е о р е м а 4 (теорема Вейерштрасса). Пусть функции $f_n(z)$ ($n = 1, 2, \dots$) являются аналитическими в области G , а ряд (9) сходится равномерно в любой замкнутой подобласти $\bar{G}'$ области G к функции $f(z)$. Тогда

1) $f(z)$ является аналитической функцией в области G .

$$2) f^{(k)}(z) = \sum_{n=1}^{\infty} f_n^{(k)}(z) \text{ в } G.$$

3) Ряд $\sum_{n=1}^{\infty} f_n^{(k)}(z)$ сходится равномерно в любой замкнутой подобласти $\bar{G}'$ области G .

3. Степенные ряды.

О п р е д е л е н и е. Ряд вида

$$c_0 + c_1 z + c_2 z^2 + \dots + c_n z^n + \dots, \quad (11)$$

где $c_0, c_1, c_2, \dots, c_n, \dots$ — постоянные, вообще говоря, комплексные числа, а z — комплексная переменная, называется степенным рядом с коэффициентами $c_0, c_1, c_2, \dots, c_n, \dots$.

Т е о р е м а 1 (теорема Абеля). Если степенной ряд (11) сходится при $z = z_0 \neq 0$, то он сходится абсолютно при всяком z , для которого $|z| < |z_0|$; если же при $z = z_0$ степенной ряд расходится, то он расходится и при любом значении z , удовлетворяющем условию $|z| > |z_0|$.

Доказательство точно такое же, как и доказательство теоремы Абеля для случая действительной переменной.

Рассмотрим случай, когда ряд (11) для одних значений z , отличных от нуля, сходится и для других — расходится (ряд *третьего класса*). Из теоремы Абеля следует, что существует такое положительное число R (оно называется *радиусом сходимости* ряда (11)), что ряд (11) сходится абсолютно при $|z| < R$ и расходится при $|z| > R$. Круг $|z| < R$ называется *кругом сходимости* степенного ряда (11). Если ряд (11) сходится при всех значениях z (ряд *второго класса*), то принято говорить, что его радиус сходимости $R = +\infty$. При $z = 0$ сходятся все степенные ряды рассматриваемого вида (11). Если ряд (11) сходится только при $z = 0$ (ряд *первого класса*), то полагают, что его радиус сходимости $R = 0$.

П р и м е р. Ряды

$$1 + \frac{z}{1!} + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots, \quad (12)$$

$$\frac{z}{1!} - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots + (-1)^n \frac{z^{2n+1}}{(2n+1)!} + \dots \quad (13)$$

$$1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots + (-1)^n \frac{z^{2n}}{(2n)!} + \dots \quad (14)$$

сходятся абсолютно при любом значении z . Действительно, так как соответствующие степенные ряды в действительной области (см. § 10.3,

п. 5) сходятся на всей числовой оси, то в силу теоремы Абеля ряды (12) — (14) будут сходиться абсолютно для любого комплексного z . Следовательно, для рядов (12) — (14) $R = +\infty$.

Формулы для определения радиуса сходимости степенного ряда (11) те же, что и для степенных рядов действительного переменного

$$R = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| \quad \text{и} \quad R = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{|c_n|}}$$

(при условии, что пределы существуют).

Эти формулы вытекают из признаков Даламбера и Коши.

Как и в случае действительного переменного (§ 10.3, п. 2), устанавливается следующая теорема:

Т е о р е м а 2. *Если радиус сходимости ряда (11) равен R , то для любого положительного числа $\rho < R$ степенной ряд (11) сходится равномерно в замкнутом круге $|z| \leq \rho$.*

Отсюда и с учетом теорем 3 и 4 из пункта 2 получаем:

С л е д с т в и е. *Сумма степенного ряда (11) есть функция, аналитическая внутри круга сходимости; ряд (11) можно любое число раз почленно интегрировать и дифференцировать внутри круга сходимости, причем получающиеся при этом степенные ряды имеют тот же радиус сходимости, что и исходный ряд.*

Ряд

$$c_0 + c_1(z-a) + c_2(z-a)^2 + c_3(z-a)^3 + \dots + c_n(z-a)^n + \dots, \quad (15)$$

где a — любое комплексное число, также называется степенным рядом. Этот ряд заменой переменной $z-a = \zeta$ сводится к ряду (11), причем точке $\zeta = 0$ соответствует точка $z = a$. Следовательно, если областью сходимости ряда (11) является круг $|z| < R$, то областью сходимости ряда (15) является круг $|z-a| < R$ с центром в точке $z = a$.

§ 15.5. Элементарные функции комплексной переменной

1. Показательная и тригонометрические функции комплексной переменной. Функция e^z комплексной переменной z определяется как сумма ряда

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots, \quad (1)$$

и, значит, эта функция в силу следствия предыдущего пункта является аналитической на всей комплексной плоскости. Для любых комплексных чисел z_1 и z_2 имеет место формула

$$e^{z_1} e^{z_2} = e^{z_1 + z_2}. \quad (2)$$

В самом деле, так как ряд (1) абсолютно сходится при всех значениях z , то можно применить свойство об умножении абсолютно сходящихся рядов (см. § 10.1, п. 5; § 15.4, п. 1). Получаем, что

$$\begin{aligned} e^z e^w &= \left(1 + z + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots \right) \cdot \left(1 + w + \frac{w^2}{2!} + \dots + \frac{w^n}{n!} + \dots \right) = \\ &= 1 + (z + w) + \frac{(z + w)^2}{2!} + \dots + \frac{(z + w)^n}{n!} + \dots = e^{z+w}. \end{aligned}$$

Из доказанной формулы (2) следует, что $e^z e^{-z} = e^{z-z} = e^0 = 1$, и потому $e^{-z} = \frac{1}{e^z}$. Отсюда получаем формулу деления показательной функции

$$\frac{e^z}{e^t} = e^z \cdot e^{-t} = e^{z-t},$$

т. е.

$$\frac{e^z}{e^t} = e^{z-t}.$$

Функции $\sin z$ и $\cos z$ комплексной переменной z определяются как суммы степенных рядов

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots + (-1)^n \frac{z^{2n+1}}{(2n+1)!} + \dots, \quad (3)$$

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots + (-1)^n \frac{z^{2n}}{(2n)!} + \dots, \quad (4)$$

которые, как и функция e^z , также являются аналитическими во всей комплексной плоскости.

Очевидно,

$$\sin(-z) = -\sin z, \quad \cos(-z) = \cos z. \quad (5)$$

На основе формул (1), (3) и (4), учитывая, что в абсолютно сходящемся ряде имеет место переместительное свойство (§ 10.1, п. 5; § 15.4, п. 1), имеем:

$$\begin{aligned} e^{iz} &= 1 + iz + \frac{i^2 z^2}{2!} + \frac{i^3 z^3}{3!} + \frac{i^4 z^4}{4!} + \frac{i^5 z^5}{5!} + \frac{i^6 z^6}{6!} + \frac{i^7 z^7}{7!} + \dots = \\ &= \left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots \right) + i \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots \right) = \cos z + i \sin z. \end{aligned}$$

Итак,

$$e^{iz} = \cos z + i \sin z. \quad (6)$$

Отсюда с учетом равенств (5)

$$e^{-iz} = \cos z - i \sin z. \quad (7)$$

Складывая равенства (6) и (7) и деля на 2, получаем, что

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}. \quad (8)$$

Аналогично

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}. \quad (9)$$

Формулы (6), (8), (9) называются *формулами Эйлера*. Пользуясь формулами (2) и (6), имеем

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z (\cos 2\pi + i \sin 2\pi) = e^z,$$

т. е. *показательная функция e^z имеет период $2\pi i$* .

Наконец, пользуясь формулой (6), из тригонометрической формы записи комплексного числа $z = r(\cos \varphi + i \sin \varphi)$ (§ 7.7, п. 3) получаем так называемую *показательную форму* записи комплексного числа $z = re^{i\varphi}$.

Отметим еще, что показательная функция e^z не обращается в нуль нигде на плоскости. В самом деле, полагая $z = x + yi$, имеем: $e^z = e^x \cdot e^{yi} = e^x (\cos y + i \sin y)$, откуда следует, что модуль функции e^z равен e^x . При любом действительном значении x выражение e^x не равно нулю, а следовательно, $|e^z|$ не может сделаться нулем.

Известные из тригонометрии формулы сложения и вычитания синуса и косинуса

$$\cos(z \pm t) = \cos z \cos t \mp \sin z \sin t, \quad (10)$$

$$\sin(z \pm t) = \sin z \cos t \pm \cos z \sin t$$

распространяются и на комплексную область.

В самом деле, по формуле (2) имеем: $e^{i(z+t)} = e^{iz} \cdot e^{it}$, откуда в силу формулы (6) получим:

$$\cos(z+t) + i \sin(z+t) = (\cos z + i \sin z)(\cos t + i \sin t). \quad (11)$$

Изменяя в тождестве (11) знаки у z и t , найдем:

$$\cos(z+t) - i \sin(z+t) = (\cos z - i \sin z)(\cos t - i \sin t). \quad (12)$$

Раскрывая скобки в равенствах (11) и (12) и складывая их между собой, получим:

$$\cos(z+t) = \cos z \cos t - \sin z \sin t;$$

вычитанием же равенств (11) и (12) найдем:

$$\sin(z+t) = \sin z \cos t + \cos z \sin t.$$

Итак, мы вывели формулы сложения для косинуса и синуса. Формулы вычитания получатся из формул сложения заменой t на $-t$. Наконец, полагая в формулах сложения $t = 2\pi$, получим:

$$\cos(z + 2\pi) = \cos z \cos 2\pi - \sin z \sin 2\pi = \cos z,$$

$$\sin(z + 2\pi) = \sin z \cos 2\pi + \cos z \sin 2\pi = \sin z,$$

т. е. число 2π есть период синуса и косинуса и в комплексной области; полагая же в формуле косинуса разности (10) $t = z$, найдем:

$$\cos 0 = \cos^2 z + \sin^2 z, \text{ или } \sin^2 z + \cos^2 z = 1.$$

Таким образом, мы видим, что все формулы тригонометрии остаются в силе и для комплексной области.

Определим, наконец, все те точки плоскости, в которых $\sin z$ и $\cos z$ обращаются в нуль. В силу формулы (9) равенство $\sin z = 0$ равносильно уравнению $e^{iz} = e^{-iz}$, или $e^{2iz} = 1$. Полагая $z = \alpha + \beta i$, находим:

$$e^{i2(\alpha + \beta i)} = 1, \text{ или } e^{-2\beta} \cdot e^{2i\alpha} = 1 \quad (13)$$

В уравнении (13) справа стоит единица, а слева — комплексное число, модуль которого равен $e^{-2\beta}$, аргумент же 2α . Следовательно, имеем: $e^{-2\beta} = 1$, $2\alpha = 2\pi k$, где k — целое число, откуда находим: $\beta = 0$, $\alpha = \pi k$. Итак, нули $\sin z$ будут: $z = \alpha + \beta i = \pi k$, где k — любое целое число. Аналогично покажем, что все нули $\cos z$ будут: $\frac{\pi}{2} + \pi k$, где k — любое целое число.

Заметим, что для комплексных значений аргумента уже нельзя более утверждать, что $|\sin z| \leq 1$ и $|\cos z| \leq 1$. В самом деле, например,

$$\cos i = \frac{e^{i^2} + e^{-i^2}}{2} = \frac{e^{-1} + e}{2} \approx 1,54308.$$

Согласно определению гиперболические синус и косинус будут:

$$\operatorname{sh} z = \frac{e^z - e^{-z}}{2} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots, \quad (14)$$

$$\operatorname{ch} z = \frac{e^z + e^{-z}}{2} = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \dots \quad (15)$$

Отсюда следует, что они являются аналитическими во всей комплексной плоскости.

Сравнивая формулы (14) и (15) соответственно с формулами (9) и (8), получим:

$$\operatorname{sh} z = -i \sin iz,$$

$$\operatorname{ch} z = \cos iz.$$

Наконец, отметим еще формулы дифференцирования рассмотренных выше функций:

$$\begin{aligned}\frac{de^z}{dz} &= 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots = e^z, \\ \frac{d \sin z}{dz} &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots = \cos z, \\ \frac{d \cos z}{dz} &= -z + \frac{z^3}{3!} - \frac{z^5}{5!} + \dots = -\sin z, \\ \frac{d \operatorname{sh} z}{dz} &= 1 + \frac{z^2}{2!} + \frac{z^4}{4!} - \dots = \operatorname{ch} z, \\ \frac{d \operatorname{ch} z}{dz} &= z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots = \operatorname{sh} z.\end{aligned}$$

2. Логарифмическая функция комплексной переменной. Эта функция определяется как функция, обратная показательной: число w называется логарифмом числа $z \neq 0$ (по основанию e), если

$$e^w = z \quad (16)$$

(обозначение: $w = \ln z$).

Если записать z в показательной форме: $z = r e^{i\varphi}$, то равенство (16) запишется в виде

$$r e^{i\varphi} = e^{u+iv} \quad (w = u + iv),$$

или

$$r e^{i\varphi} = e^u e^{iv}.$$

Отсюда вытекают два равенства:

$$e^u = r, \quad v = \varphi + 2k\pi \quad (k — \text{любое целое число}).$$

Из первого равенства находим $u = \ln r$, где $\ln r$ понимается в обычном смысле. Следовательно,

$$w = \ln z = \ln r + i(\varphi + 2\pi k) = \ln |z| + i \arg z + 2k\pi i$$

($k = 0, \pm 1, \pm 2, \dots$).

Видно, что $\ln z$ есть многозначная функция от z независимо от того, будет ли z действительным или комплексным.

Например, с точки зрения функций комплексной переменной $\ln 1$ равен одному из чисел $2k\pi i$ ($k = 0, \pm 1, \pm 2, \dots$). В действительном анализе для выражения $\ln 1$ выбирают среди этих чисел единственное число 0.

П р и м е р ы: 1. $\ln(-1) = i\pi + 2k\pi i = i\pi(1 + 2k)$

$$(k = 0, \pm 1, \pm 2, \dots);$$

$$2. \ln(1-i) = \ln \sqrt{2} - i \frac{\pi}{4} + 2k\pi i = \frac{1}{2} \ln 2 - i \frac{\pi}{4} (1-8k) \quad (k = 0, \pm 1, \pm 2, \dots).$$

§ 15.6. Ряд Тейлора

1. Разложение аналитической функции в степенной ряд. Как уже отмечалось (§15.4, п. 3, следствие), степенной ряд внутри своего круга сходимости определяет аналитическую функцию — сумму этого ряда. Справедливо, однако, и обратное утверждение.

Т е о р е м а. *Функция $f(z)$, аналитическая внутри круга $|z - z_0| < R$ может быть представлена в этом круге степенным рядом*

$$\sum_{k=0}^{\infty} c_k (z - z_0)^k, \quad (1)$$

причем коэффициенты этого ряда однозначно определяются по формулам:

$$c_k = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta) d\zeta}{(\zeta - z_0)^{k+1}} \quad (k = 0, 1, 2, \dots),$$

где γ — какая-нибудь окружность с центром z_0 , лежащая внутри данного круга, или с учетом формул (5) и (6) из § 15.3 — по формулам

$$c_0 = f(z_0), \quad c_k = \frac{f^{(k)}(z_0)}{k!} \quad (k = 1, 2, \dots).$$

Ряд (1) называется рядом Тейлора для функции $f(z)$.

2. Понятие голоморфной функции и его эквивалентность с понятием аналитической функции.

О п р е д е л е н и е 1. Говорят, что функция $f(z)$ *голоморфна* в точке a , если она в некоторой окрестности этой точки разлагается в степенной ряд относительно $z - a$. Свойство голоморфности функции в точке a эквивалентно ее аналитичности в этой точке. Если $f(z)$ голоморфна в точке a , то существует круг с центром в точке a некоторого радиуса ρ , внутри которого $f(z)$ разлагается в степенной ряд относительно $z - a$ и, значит, как сумма степенного ряда, будет аналитической внутри этого круга, а следовательно, и в точке a . Если же $f(z)$ аналитична в точке a , то существует круг с центром в этой точке некоторого радиуса ρ , внутри которого $f(z)$ будет аналитической функцией, и, значит, согласно приведенной выше теореме, голоморфной в точке a .

О п р е д е л е н и е 2. Функция, голоморфная в каждой точке области G , называется голоморфной в этой области.

Сказать, что функция $f(z)$ есть голоморфная в некоторой области G , равносильно утверждению, что эта функция есть аналитическая в той же области G .

3. Нули функции. Теорема единственности. Пусть $f(z)$ является аналитической функцией в области G . Точка $z_0 \in G$ называется *нулем* $f(z)$,

если $f(z_0) = 0$. В этом случае разложение функции $f(z)$ в окрестности точки z_0 в степенной ряд имеет вид

$$f(z) = c_1(z - z_0) + c_2(z - z_0)^2 + \dots + c_k(z - z_0)^k + \dots,$$

так как $c_0 = f(z_0) = 0$.

Если не только $c_0 = 0$, но и $c_1 = c_2 = \dots = c_{k-1} = 0$, а $c_k \neq 0$, то точка z_0 называется *нулем k -го порядка функции $f(z)$* . В этом случае разложение функции $f(z)$ в окрестности точки z_0 имеет вид

$$f(z) = c_k(z - z_0)^k + c_{k+1}(z - z_0)^{k+1} + \dots$$

(Это возможно, когда $f(z_0) = f'(z_0) = \dots = f^{(k-1)}(z_0) = 0$, $af^{(k)}(z_0) \neq 0$.) Если $k = 1$, то нуль называют *простым*.

Пример. Нули $\sin z$ простые, так как $\cos \pi k \neq 0$. Аналогично простые нули и для $\cos z$.

Теорема 1. Пусть функция $f(z)$ аналитична в области G и обращается в нуль в различных точках $z_1, z_2, \dots, z_n, \dots$, принадлежащих этой области. Тогда, если существует предел $\lim_{n \rightarrow \infty} z_n = a$ и точка a принадлежит области G , то $f(z) \equiv 0$ в области G .

Из этой теоремы непосредственно следует, что всякая, не равная тождественно нулю, аналитическая в области G функция $f(z)$ имеет в любой замкнутой подобласти G' области G лишь конечное число нулей; бесконечное число нулей аналитическая функция может иметь только в открытой области или в неограниченной области.

Определение. Функция $f(z)$, аналитическая во всей комплексной плоскости ($z \neq \infty$), называется *целой функцией*.

Из предыдущих рассуждений следует, что целая функция в любой ограниченной области имеет лишь конечное число нулей. Следовательно, все нули целой функции можно перенумеровать в каком-либо порядке, например в порядке возрастания их модулей. Целые функции играют важную роль как в теории функций комплексной переменной, так и ее приложениях.

Из предыдущей теоремы следует следующая теорема:

Теорема 2 (теорема единственности). Если в области G существует такая последовательность различных точек $z_1, z_2, \dots, z_n, \dots$, сходящаяся в точке a , принадлежащей области G , что значения двух аналитических в области G функций $f(z)$ и $\varphi(z)$ совпадают во всех точках указанной последовательности, то $f(z) \equiv \varphi(z)$ в области G .

Из этой теоремы следует, что если две аналитические в области G функции $f_1(z)$ и $f_2(z)$ совпадают на некоторой кривой L , принадлежащей области G , то они совпадают и во всей области G .

4. Аналитическое продолжение. Из теоремы единственности также можно сделать вывод, что если функция $f_1(z)$ аналитична в области G_1 , а функция $f_2(z)$ аналитична в области G_2 , причем области G_1 и G_2 имеют общую часть — некоторую область G , в которой $f_1(z)$ и $f_2(z)$ совпадают, то существует единственная аналитическая функция $F(z)$ такая, что

$$F(z) = \begin{cases} f_1(z) & \text{для } z \text{ из } G_1, \\ f_2(z) & \text{для } z \text{ из } G_2. \end{cases}$$

В этом случае функция $F(z)$ называется *аналитическим продолжением* функции $f_1(z)$ (соответственно $f_2(z)$) на область $G_1 + G_2$. При этом функцию $f_2(z)$ также называют аналитическим продолжением функции $f_1(z)$ на область G_2 , а функцию $f_1(z)$ — аналитическим продолжением функции $f_2(z)$ на область G_1 .

П р и м е р ы. Показательная функция e^x и тригонометрические функции $\sin x$ и $\cos x$ при действительном x разлагаются в ряды

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \quad (\S 10.3, \text{ п. 5}).$$

Заменяя в этих рядах формально x через комплексное переменное z , мы получим степенные ряды

$$f_1(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad f_2(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}, \quad f_3(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!},$$

сходящиеся во всей плоскости (§ 15.5, п. 4). Эти ряды изображают функции, аналитические во всей комплексной плоскости. Так как функции $f_1(z)$, $f_2(z)$ и $f_3(z)$ при $z = x$ совпадают соответственно с e^x , $\sin x$ и $\cos x$, то они являются аналитическими продолжениями последних функций в комплексную область.

§ 15.7. Ряд Лорана

О п р е д е л е н и е. Ряд вида

$$\sum_{k=-\infty}^{\infty} c_k (z - z_0)^k, \quad (1)$$

где k принимает все целые (положительные, отрицательные и нулевое) значения, называется *рядом Лорана*.

Этот ряд понимается как сумма двух рядов:

$$\sum_{k=0}^{\infty} c_k (z - z_0)^k \quad (2)$$

и

$$\sum_{k=1}^{\infty} c_{-k} (z - z_0)^{-k}. \quad (3)$$

Первая часть ряда Лорана — ряд, расположенный по неотрицательным степеням двучлена $(z - z_0)$ и рассмотренный выше (§15.4, п. 3), называется *правильной* частью, а вторая часть — ряд, расположенный по отрицательным степеням двучлена $(z - z_0)$, — *главной* частью ряда Лорана.

Областью сходимости ряда Лорана является общая часть областей сходимости его правильной и главной частей. Областью сходимости правильной части (см. § 15.4, п. 3) является круг с центром в точке z_0 и некоторого радиуса R_1 (значение R_1 может, в частности, равняться нулю или бесконечности). Внутри этого круга правильная часть ряда Лорана сходится к некоторой аналитической функции $f_1(z)$ (§ 15.4, п. 3) — сумме этой части ряда Лорана. Для определения области сходимости главной части ряда Лорана введем обозначение $\frac{1}{z - z_0} = \zeta$. Тогда получим ряд

$$\sum_{k=1}^{\infty} c_{-k} \zeta^k.$$

Это обычный степенной ряд, сходящийся внутри своего круга сходимости (его радиус обозначим через $\frac{1}{R_2}$) к некоторой аналитической функции $\varphi(\zeta)$.

Возвращаясь к переменной z и полагая $\varphi\left(\frac{1}{z - z_0}\right) = f_2(z)$, получим

$$f_2(z) = \sum_{k=1}^{\infty} \frac{c_{-k}}{(z - z_0)^k},$$

т. е. ряд, сходящийся к аналитической функции $f_2(z)$ вне круга радиуса R_2 с центром в точке z_0 (так же, как и R_1 , значение R_2 может, в частности, равняться нулю или бесконечности).

Итак, каждый из степенных рядов (2) и (3) сходится в своей области сходимости к соответствующей аналитической функции. Если $R_2 < R_1$ (рис. 195), то существует общая область сходимости этих рядов — круговое кольцо $R_2 < |z - z_0| < R_1$, в котором ряд (1) сходится к аналитической функции $f(z) = f_1(z) + f_2(z)$.

Если $R_2 > R_1$, то ряды (2) и (3) общей области сходимости не имеют. Тем самым ряд Лорана нигде не сходится.

Т е о р е м а. Функция $f(z)$, аналитическая в круговом кольце $R_2 < |z - z_0| < R_1$, однозначно представляется в этом кольце рядом Лорана.

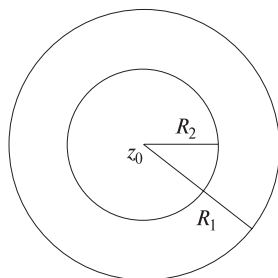


Рис. 195

Коэффициенты ряда Лорана вычисляются по формулам, подобным формулам для нахождения коэффициентов ряда Тейлора,

$$c_k = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta) d\zeta}{(\zeta - z_0)^{k+1}}, \quad (k = 0, \pm 1, \pm 2, \dots), \quad (4)$$

где γ — какая-нибудь окружность с центром z_0 , лежащая внутри данного кольца.

П р и м е р. Разложить в ряд Лорана функцию $f(z) = \frac{1}{(z-1)(z-2)}$ в кольце $1 < |z| < 2$.

Имеем:

$$f(z) = \frac{1}{z-2} - \frac{1}{z-1} = -\frac{1}{2} \cdot \frac{1}{1 - \frac{z}{2}} - \frac{1}{z} \cdot \frac{1}{1 - \frac{1}{z}} = -\sum_{n=0}^{\infty} \frac{z^n}{2^{n+1}} - \sum_{n=1}^{\infty} \frac{1}{z^n}.$$

§ 15.8. Изолированные особые точки аналитической функции

1. Три типа особых точек. Точки, в которых нарушается аналитичность функции, называются *особыми*. Если в достаточной близости к особой точке a нет других особых точек, то особая точка a называется *изолированной* особой точкой. Если a есть изолированная особая точка функции $f(z)$, то в достаточно малом круге с выколотым центром a функция $f(z)$ будет аналитической и, следовательно, разлагается в ряд Лорана

$$f(z) = \sum_{k=-\infty}^{+\infty} c_k (z-a)^k. \quad (1)$$

Возможны три различных случая:

- 1) в разложении (1) нет членов с отрицательными степенями $z-a$;
- 2) в разложении (1) есть лишь конечное число членов с отрицательными степенями $z-a$;
- 3) в разложении (1) есть бесконечное число членов с отрицательными степенями $z-a$.

В этих случаях особая точка a называется соответственно: 1) *устраняемой*; 2) *полусом*; 3) *существенно особой*.

Если a — устранимая особая точка, то в окрестности этой точки

$$f(z) = \sum_{k=0}^{+\infty} c_k (z-a)^k;$$

следовательно, после надлежащего доопределения функции в точке a ($f(a) = c_0$) функция $f(z)$ становится аналитической в точке a и «особенность устраняется». В достаточно малой окрестности устранимой особой точки функция $f(z)$ ограничена. Справедливо и обратное

утверждение: если $f(z)$ ограничена в некоторой окрестности изолированной особой точки a , $|f(z)| \leq M$, то эта точка есть устранимая особая.

Если a — полюс, то в окрестности a имеем:

$$f(z) = \frac{c_{-n}}{(z-a)^n} + \dots + \frac{c_{-1}}{z-a} + c_0 + c_1(z-a) + c_2(z-a)^2 + \dots,$$

$c_{-n} \neq 0$ (в этом случае a — полюс n -го порядка; если $n = 1$, полюс a — простой), откуда

$$f(z) = \frac{\varphi(z)}{(z-a)^n},$$

где

$$\varphi(z) = c_{-n} + c_{-n+1}(z-a) + \dots + c_0(z-a)^n + \dots$$

есть аналитическая функция в окрестности точки a , причем $\varphi(a) = c_{-n} \neq 0$.

Справедливо и обратное предложение: если $f(z) = \frac{\varphi(z)}{(z-a)^n}$, где $\varphi(z)$ ана-

литична в окрестности a и $\varphi(a) \neq 0$, то a есть полюс n -го порядка для $f(z)$.

Из такого выражения для $f(z)$ следует, что при $z \rightarrow a$ $|f(z)|$ неограниченно возрастает. Легко видеть, что если a есть нуль порядка n для $f(z)$, то a будет полюсом порядка n для $\frac{1}{f(z)}$, так как из равенства $f(z) = (z-a)^n \varphi(z)$ (где $\varphi(z)$ аналитическая в окрестности a , отличная от нуля в a) следует:

$$\frac{1}{f(z)} = \frac{1}{(z-a)^n} \cdot \frac{1}{\varphi(z)}$$

($\frac{1}{\varphi(z)}$ аналитическая в окрестности a и отлична от нуля в a).

Если a — существенно особая точка, то в любой окрестности a значения функции $f(z)$ как угодно близко подходят к любому комплексному числу (теорема Сохоцкого; Юлиан Васильевич Сохоцкий (1842—1927) — русский математик).

Если в области G функция $f(z)$ может иметь в качестве особых точек только полюсы, то $f(z)$ называется *мероморфной* в области G .

П р и м е р ы. 1. Функции $\frac{1}{\sin z}$ и $\frac{1}{\cos z}$ имеют полюсы соответственно в нулях $\sin z$ и $\cos z$, причем полюса простые.

2. Функция $e^{\frac{1}{z}}$ имеет при $z = 0$ существенно особую точку, так как разложение в ряд Лорана в окрестности этой точки будет $e^{\frac{1}{z}} = \sum_{k=0}^{\infty} \frac{1}{k! z^k}$.

2. Случай бесконечно удаленной точки. Пусть теперь $f(z)$ аналитична в окрестности бесконечно удаленной точки. Тогда вне некоторого круга с центром в нулевой точке она изобразится рядом Лорана

$$f(z) = \sum_{k=-\infty}^{+\infty} c_k z^k. \quad (2)$$

Возможны три различных случая:

- 1) в разложении (2) нет членов с положительными степенями z ;
- 2) в разложении (2) есть лишь конечное число членов с положительными степенями z ;
- 3) в разложении (2) есть бесконечное число членов с положительными степенями z .

В этих случаях бесконечно удаленная точка называется соответственно: 1) *устранимой особой точкой*; 2) *полюсом*; 3) *существенно особой точкой*. Подстановка $z = \frac{1}{\zeta}$ приводит изучение функции $f(z)$ в окрестности бесконечно удаленной точки к изучению функции $f\left(\frac{1}{\zeta}\right)$ в окрестности нулевой точки. Поэтому все заключения о поведении функции в окрестности конечной изолированной особой точки немедленно переносятся на случай бесконечно удаленной точки.

П р и м е р ы. 1. Функция $\frac{1}{1-z}$ имеет в бесконечно удаленной точке устранимую особую точку, так как при $|z| > 1$ имеем: $\frac{1}{1-z} = -\frac{1}{z} \cdot \frac{1}{1-\frac{1}{z}} = -\sum_{k=1}^{\infty} \frac{1}{z^k}$.

2. Функция z^3 имеет бесконечно удаленную точку полюсом порядка 3.

3. Для функции e^z бесконечно удаленная точка существенно особая, так как в окрестности бесконечно удаленной точки $e^z = -\sum_{k=0}^{\infty} \frac{z^k}{k!}$. Такой же тип бесконечно удаленной точки и для функций $\sin z$ и $\cos z$.

§ 15.9. Вычеты

1. Определение вычета. Пусть a — конечная изолированная особая точка аналитической функции $f(z)$; тогда в окрестности точки a эта функция изобразится рядом Лорана

$$f(z) = \sum_{k=-\infty}^{+\infty} c_k (z-a)^k. \quad (1)$$

О п р е д е л е н и е. Коэффициент при первой отрицательной степени разложения Лорана (1), т. е. число c_{-1} называется *вычетом* функции $f(z)$ относительно особой точки a , что кратко можно обозначить знаком $\operatorname{Res}_a f(z)$.

Из формулы (4) из § 15.7 при $k = -1$ найдем:

$$c_{-1} = \frac{1}{2\pi i} \int_{\gamma} f(z) dz, \quad (2)$$

где γ — достаточно малая окружность с центром a .

2. Основная теорема о вычетах. *Т е о р е м а. Если функция аналитична внутри замкнутого контура Γ и на нем, за исключением конечного числа точек внутри Γ , то интеграл $\frac{1}{2\pi i} \int_{\Gamma} f(z) dz$ равен сумме вычетов относительно особых точек, лежащих внутри Γ .*

Д о к а з а т е л ь с т в о. Пусть $a_1, a_2, \dots, a_k$ (рис. 196) — особые точки $f(z)$, лежащие внутри Γ и $\alpha_1, \alpha_2, \dots, \alpha_k$ — вычеты $f(z)$ относительно них. Пусть $\gamma_1, \gamma_2, \dots, \gamma_k$ — окружности вокруг этих точек, лежащие внутри Γ и вне друг друга. Тогда по теореме Коши для сложного контура и в силу формулы (2) получим

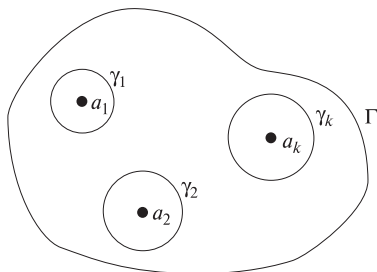


Рис. 196

$$\begin{aligned} \frac{1}{2\pi i} \int_{\Gamma} f(z) dz &= \frac{1}{2\pi i} \int_{\gamma_1} f(z) dz + \frac{1}{2\pi i} \int_{\gamma_2} f(z) dz + \dots + \frac{1}{2\pi i} \int_{\gamma_k} f(z) dz = \\ &= \alpha_1 + \alpha_2 + \dots + \alpha_k, \end{aligned}$$

что и требовалось доказать.

3. Вычисление вычета функции относительно полюса. 1) Пусть точка a — простой полюс для функции $f(z)$. Тогда в окрестности точки a имеем

$$f(z) = \frac{c_{-1}}{z-a} + c_0 + c_1(z-a) + c_2(z-a)^2 + \dots,$$

откуда

$$(z-a)f(z) = c_{-1} + c_0(z-a) + c_1(z-a)^2 + \dots$$

и потому

$$c_{-1} = \lim_{z \rightarrow a} (z-a)f(z). \quad (3)$$

Пусть теперь $f(z) = \frac{\varphi(z)}{\psi(z)}$, где $\varphi(z), \psi(z)$ аналитичны в окрестности a

и точка a есть простой нуль для $\psi(z)$. Тогда a будет простым полюсом для $f(z)$ (если $\varphi(a) \neq 0$). По формуле (3) имеем

$$c_{-1} = \lim_{z \rightarrow a} (z-a)f(z) = \lim_{z \rightarrow a} \frac{\varphi(z)}{\frac{\psi(z)}{z-a}} = \lim_{z \rightarrow a} \frac{\varphi(z)}{\frac{\psi(z) - \psi(a)}{z-a}},$$

так как согласно условию имеем $\psi(a) = 0$. Заметив, что

$$\lim_{z \rightarrow a} \varphi(z) = \lim_{z \rightarrow a} \varphi(a), \lim_{z \rightarrow a} \frac{\psi(z) - \psi(a)}{z - a} = \psi'(a) \neq 0,$$

окончательно получаем

$$c_{-1} = \frac{\varphi(a)}{\psi'(a)}. \quad (4)$$

Пр и м е р 1. $\operatorname{Res}_{\pi} \operatorname{ctg} z = \operatorname{Res}_{\pi} \frac{\cos z}{\sin z} = \frac{\cos \pi}{\cos \pi} = 1.$

2) Пусть a есть полюс порядка m для $f(z)$. Тогда в окрестности точки a имеем

$$f(z) = \frac{c_{-m}}{(z-a)^m} + \dots + \frac{c_{-1}}{z-a} + c_0 + c_1(z-a) + c_2(z-a)^2 + \dots,$$

откуда

$$(z-a)^m f(z) = c_{-m} + c_{-m+1}(z-a) + \dots + c_{-1}(z-a)^{m-1} + \dots \quad (5)$$

Продифференцировав равенство (5) $m-1$ раз, получим в правой части обыкновенный степенной ряд, свободный член которого будет $c_{-1}(m-1)!$; и, значит,

$$c_{-1} = \frac{1}{(m-1)!} \lim_{z \rightarrow a} \frac{d^{m-1}}{dz^{m-1}} ((z-a)^m f(z)). \quad (6)$$

Пр и м е р 2. $\operatorname{Res}_{-2} \frac{e^z}{(z+2)^3} = \frac{1}{2!} \lim_{z \rightarrow -2} \frac{d^2 e^z}{dz^2} = \frac{1}{2} \lim_{z \rightarrow -2} e^z = \frac{1}{2e^2}.$

4. Вычисление несобственных интегралов от функций действительного переменного. Теорию вычетов можно применять для вычисления некоторых типов несобственных интегралов. Приведем по этому поводу две теоремы, доказательства которых можно найти в более подробных курсах.

Т е о р е м а 1. Если функция $f(z)$ удовлетворяет трем требованиям: 1) $f(z)$ имеет точку $z = \infty$ своим нулем порядка не ниже второго, т. е. ряд Лорана для этой функции в окрестности точки $z = \infty$ имеет вид

$$f(z) = \frac{c_2}{z^2} + \frac{c_3}{z^3} + \dots$$

(случай $c_2 = 0$ не исключается), 2) $f(z)$ аналитична на действительной оси, 3) $f(z)$ аналитична в верхней полуплоскости $\operatorname{Im} z > 0$, исключая конечное число особых точек $a_1, a_2, \dots, a_n$, то

$$\int_{-\infty}^{+\infty} f(z) dz = 2\pi i \sum_{k=1}^n \operatorname{Res}_{a_k} f(z). \quad (7)$$

Пример 1. Вычислить интеграл $I = \int_{-\infty}^{+\infty} \frac{(x-1)dx}{(x^2+1)^2}$. Функция

$$f(z) = \frac{z-1}{(z^2+1)^2}$$

имеет точку $z = \infty$ нулем третьего порядка, а точки $z_1 = i$ и $z_2 = -i$ — полюсами второго порядка; других особых точек эта функция не имеет. В верхней полуплоскости расположен единственный полюс $z_1 = i$ функции $f(z)$. По формуле (6) находим

$$\operatorname{Res}_i f(z) = \lim_{z \rightarrow i} \frac{d}{dz} ((z-i)^2 f(z)) = \lim_{z \rightarrow i} \frac{d}{dz} \left(\frac{z-1}{(z+i)^2} \right) = \lim_{z \rightarrow i} \frac{2+i-z}{(z+i)^3} = \frac{i}{4},$$

так что в силу формулы (7)

$$I = 2\pi i \operatorname{Res}_i f(z) = -\frac{\pi}{2}.$$

Несобственный интеграл $\int_{-\infty}^{+\infty} f(x)dx$ иногда может быть вычислен и в том случае, когда точка $z = \infty$ не является нулем порядка не ниже второго для $f(z)$. Имеет место следующая теорема.

Т е о р е м а 2. Если функция $f(z)$ удовлетворяет трем требованиям: 1) $f(z) = e^{i\alpha z} g(z)$, где $\alpha > 0$ и $g(z) \rightarrow 0$ при $z \rightarrow \infty$ и при условии $\operatorname{Im} z \geq 0$ (т. е. $z \rightarrow \infty$, оставаясь в верхней полуплоскости или на действительной оси), 2) $f(z)$ аналитична на действительной оси, 3) $f(z)$ аналитична в верхней полуплоскости, исключая конечное число особых точек $a_1, a_2, \dots, a_n$, то:

$$\int_{-\infty}^{+\infty} f(x)dx = 2\pi i \sum_{k=1}^n \operatorname{Res}_{a_k} f(z). \quad (8)$$

Пример 2. Вычислить $I = \int_{-\infty}^{+\infty} \frac{x \sin x dx}{x^2 + 4}$. Рассмотрим функцию $f(z) = \frac{ze^{iz}}{z^2 + 4}$; она

удовлетворяет всем условиям теоремы 2 ($\alpha = 1$, $g(z) = \frac{z}{z^2 + 4}$). Особыми точками $f(z)$ будут

два простых полюса $z_1 = 2i$ и $z_2 = -2i$, причем в верхней полуплоскости лежит первый из них. По формуле (4)

$$\operatorname{Res}_{z_1} f(z) = \frac{z_1 e^{iz_1}}{2z_1} = \frac{1}{2e^2}$$

и, следовательно, по формуле (8), применяя формулу Эйлера,

$$\int_{-\infty}^{+\infty} \frac{xe^{ix} dx}{x^2 + 4} = \int_{-\infty}^{+\infty} \frac{x \cos x dx}{x^2 + 4} + i \int_{-\infty}^{+\infty} \frac{x \sin x dx}{x^2 + 4} = 2\pi i \cdot \frac{1}{2e^2} = \frac{\pi i}{e^2}.$$

Отделяя действительности мнимые части; находим

$$\int_{-\infty}^{+\infty} \frac{x \cos x dx}{x^2 + 4} = 0, \quad I = \int_{-\infty}^{+\infty} \frac{x \sin x dx}{x^2 + 4} = \frac{\pi}{e^2}$$

(первый из этих результатов очевиден, так как здесь подынтегральная функция нечетная).

Глава 16. ИНТЕГРАЛ ФУРЬЕ. ДЕЛЬТА-ФУНКЦИЯ

§16.1. Интеграл Фурье

1. Комплексная форма ряда Фурье. Как известно (см. § 10.4, п. 2), если периодическая функция $f(x)$ с периодом 2π является суммой равномерно сходящегося на отрезке $[-\pi, \pi]$ тригонометрического ряда

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_n \cos nx + b_n \sin nx, \quad (1)$$

то коэффициенты этого ряда определяются по формулам

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx, \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \quad (n = 1, 2, \dots). \end{aligned} \quad (2)$$

Далее по формулам Эйлера (см. §15.5, п. 1)

$$\begin{aligned} \cos nx &= \frac{e^{inx} + e^{-inx}}{2}, \\ \sin nx &= \frac{e^{inx} - e^{-inx}}{2i} = -i \frac{e^{inx} - e^{-inx}}{2}. \end{aligned}$$

Подставляя эти выражения в ряд (1), получаем:

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \frac{e^{inx} + e^{-inx}}{2} - ib_n \frac{e^{inx} - e^{-inx}}{2} \right) = \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(\frac{a_n - ib_n}{2} e^{inx} + \frac{a_n + ib_n}{2} e^{-inx} \right). \end{aligned}$$

Введем обозначения

$$\frac{a_0}{2} = c_0, \quad \frac{a_n - ib_n}{2} = c_n, \quad \frac{a_n + ib_n}{2} = c_{-n}. \quad (3)$$

Тогда

$$f(x) = c_0 + \sum_{n=1}^{\infty} (c_n e^{inx} + c_{-n} e^{-inx}),$$

или, более компактно,

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

Это и есть *комплексная форма, ряда Фурье*.

Выразим коэффициенты c_n и c_{-n} через интегралы. Пользуясь формулами (2), можем формулы (3) переписать так:

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)(\cos nx - i \sin nx) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)e^{-inx} dx, \quad (n = 0, 1, 2, \dots),$$

$$c_{-n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)e^{inx} dx, \quad (n = 1, 2, \dots).$$

2. Интегральная формула Фурье. Интеграл Фурье.

Пусть теперь функция $f(x)$, определенная на промежутке $(-\infty; +\infty)$, удовлетворяет следующим условиям:

1) На любом конечном отрезке $[-l; l]$ функция $f(x)$ удовлетворяет условиям Дирихле.

2) Существует несобственный интеграл $\int_{-\infty}^{+\infty} |f(x)| dt = Q$ (это свойство функции $f(x)$ называется ее *абсолютной интегрируемостью*).

Учитывая формулы для коэффициентов Фурье (см. § 10.4, п. 6, формулы (9), (10)):

$$a_n = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{n\pi x}{l} dx, \quad (n = 0, 1, 2, \dots),$$

$$b_n = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{n\pi x}{l} dx, \quad (n = 1, 2, \dots),$$

при $n = 1, 2, 3, \dots$ имеем:

$$a_n \cos \frac{n\pi x}{l} + b_n \sin \frac{n\pi x}{l} = \frac{1}{l} \int_{-l}^l f(t) \cos \frac{n\pi(t-x)}{l} dt.$$

Тогда

$$f(x) = \frac{1}{l} \int_{-l}^l f(t) dt + \frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \frac{n\pi(t-x)}{l} dt. \quad (4)$$

Исследуем вопрос о том, какой вид примет разложение (4) при $l \rightarrow +\infty$. Введем следующие обозначения:

$$\lambda_1 = \frac{\pi}{l}, \quad \lambda_2 = \frac{2\pi}{l}, \quad \dots, \quad \lambda_k = \frac{k\pi}{l}, \quad \dots; \quad \Delta\lambda_k = \lambda_{k+1} - \lambda_k = \frac{\pi}{l}.$$

Подставляя это в равенство (1), получаем:

$$f(x) = \frac{1}{2l} \int_{-l}^l f(t) dt + \frac{1}{\pi} \sum_{n=1}^{\infty} \left(\int_{-l}^l f(t) \cos \lambda_n(t-x) dt \right) \Delta\lambda_n. \quad (5)$$

При $l \rightarrow +\infty$ первый член в правой части равенства (5) стремится к нулю. Действительно, так как

$$\left| \frac{1}{2l} \int_{-l}^l f(t) dt \right| \leq \frac{1}{2l} \int_{-l}^l |f(t)| dt < \frac{1}{2l} \int_{-\infty}^{+\infty} |f(t)| dt = \frac{Q}{2l}$$

(эти неравенства следуют из геометрических соображений), то $\frac{1}{2l} \int_{-l}^l f(t) dt \rightarrow 0$ при $l \rightarrow +\infty$.

При любом фиксированном l выражение $\int_{-l}^l f(t) \cos \lambda_n(t-x) dt$ есть функ-

ция от переменной λ_n , которая принимает равноотстоящие значения в промежутке $(0; +\infty)$. Без доказательства укажем, что при $l \rightarrow +\infty$ формула (5) примет вид

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(t) \cos \lambda(t-x) dt. \quad (6)$$

Формула (6) называется *интегральной формулой Фурье*, а интеграл в правой части этой формулы — *интегралом Фурье* для функции $f(x)$. Представление функции $f(x)$ в виде правой части формулы (6) обычно называется *разложением этой функции в интеграл Фурье*.

В силу формулы (6)

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(t) (\cos \lambda t \cos \lambda x + \sin \lambda t \sin \lambda x) dt,$$

и, следовательно, формуле (6) можно придать такой вид:

$$f(x) = \int_0^{+\infty} (a(\lambda) \cos \lambda x + b(\lambda) \sin \lambda x) d\lambda,$$

где положено

$$a(\lambda) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \cos \lambda t dt, \quad b(\lambda) = \frac{1}{\pi} \int_{-\infty}^{+\infty} f(t) \sin \lambda t dt,$$

3. Комплексная форма интегральной формулы Фурье. В интеграле Фурье (формула (6)) внутренний интеграл — четная функция от λ . Поэтому формулу (6) можно переписать в виде

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(t) \cos \lambda(t-x) dt. \quad (6')$$

Далее согласно формуле Эйлера (§ 15.5, п. 1) имеем:

$$\cos \lambda(t-x) = \frac{1}{2} (e^{i\lambda(t-x)} + e^{-i\lambda(t-x)}).$$

Подставляя это в формулу (6'), получим:

$$f(x) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(t) e^{i\lambda(t-x)} dt + \frac{1}{4\pi} \int_{-\infty}^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(t) e^{-i\lambda(t-x)} dt. \quad (7)$$

Здесь, как нетрудно убедиться подстановкой $z = -\lambda$, интегралы, стоящие в правой части, равны друг другу. Поэтому

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(t) e^{-i\lambda(t-x)} dt. \quad (7')$$

Правая часть в формуле (7') называется *интегралом Фурье в комплексной форме* для функции $f(x)$.

4. Преобразование Фурье и его обращение. Перепишем формулу (7') в виде

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\lambda x} d\lambda \int_{-\infty}^{+\infty} f(t) e^{-i\lambda t} dt \quad (8)$$

и положим:

$$\tilde{f}(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-i\lambda t} dt. \quad (9)$$

Тогда вместо формулы (8) получим:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{f}(\lambda) e^{i\lambda x} d\lambda. \quad (10)$$

Операция, которая производится над функцией f в формуле (9), называется *преобразованием Фурье*. Операция, которая производится над функцией $\tilde{f}$ в формуле (10), называется *обратным преобразованием Фурье* (эта формула называется еще *формулой обращения*), при этом функцию f часто называют *прообразом*, а функцию $\tilde{f}$ — *образом*. Последовательное применение этих двух операций, как видно, возвращает нас к исходной функции f .

5. Спектральная функция. Как уже отмечалось, интегралы, стоящие в правой части равенства (7), равны друг другу. Поэтому формула (7) может быть переписана и в таком виде:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\lambda \int_{-\infty}^{+\infty} f(t) e^{i\lambda(x-t)} dt. \quad (11)$$

Введем теперь обозначение

$$A(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(t) e^{-i\lambda t} dt. \quad (11')$$

Эта функция от λ играет важную роль в электротехнике и носит название *спектральной функции* для $f(x)$.

$$\text{В силу (11) и (11')} \quad f(x) = \int_{-\infty}^{+\infty} A(\lambda) e^{i\lambda x} d\lambda.$$

6. Свойства преобразования Фурье. Преобразование Фурье обладает рядом полезных свойств, некоторые из которых здесь перечислим.

$$1. \widetilde{f_1 + f_2} = \widetilde{f_1} + \widetilde{f_2}, \quad \widetilde{\alpha f} = \alpha \widetilde{f} \quad (\alpha = \text{const}). \quad (12)$$

Это свойство вытекает из формулы (9).

2. Если прообраз сдвинуть на постоянную β , то его образ умножится на $e^{-i\beta\lambda}$.

Д о к а з а т е л ь с т в о. В самом деле, полагая $t - \beta = s$, имеем:

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t - \beta) e^{-i\lambda t} dt &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(s) e^{-i\lambda(\beta + s)} ds = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(s) e^{-i\lambda s} e^{-i\lambda\beta} ds = e^{-i\lambda\beta} \widetilde{f}(\lambda). \end{aligned}$$

Аналогично если образ сдвинуть на β , то прообраз умножится на $e^{i\beta x}$.

3. Если функция $f(x)$ преобразуется в $\widetilde{f}(\lambda)$ то $f(\alpha x)$ ($\alpha = \text{const} > 0$) преобразуется в $\frac{1}{\alpha} \widetilde{f}\left(\frac{\lambda}{\alpha}\right)$.

Действительно, полагая $\alpha t = s$, имеем:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(\alpha t) e^{-i\lambda t} dt = \frac{1}{\alpha} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(s) e^{-i\frac{\lambda}{\alpha}s} ds = \frac{1}{\alpha} \widetilde{f}\left(\frac{\lambda}{\alpha}\right).$$

4. Если от функции f взять первую производную, то образ умножится на $i\lambda$ (предполагается, что наряду с абсолютной интегрируемостью f на всей числовой оси на ней абсолютно интегрируема и f' , причем f' непрерывна).

В самом деле,

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f'(t) e^{-i\lambda t} dt = \frac{1}{\sqrt{2\pi}} f(t) e^{-i\lambda t} \Big|_{-\infty}^{+\infty} + \frac{i\lambda}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-i\lambda t} dt = i\lambda \widetilde{f}.$$

Мы произвели интегрирование по частям и учли, что $\lim_{t \rightarrow \pm\infty} f(t) = 0$ (см. [9], т. II).

П р и м е ч а н и е. Свойство 4 распространяется и на случай производной k -го порядка от функции f . В этом случае образ этой функции умножится на $(i\lambda)^k$.

Доказательство проводится методом математической индукции.

5. Пусть функция f зависит не только от x , но и от некоторого параметра t . Тогда $\tilde{f}$ также зависит от этого параметра ($\tilde{f} = \tilde{f}(\lambda, t)$), и в силу формул (12) запишем:
$$\frac{\tilde{f}(\lambda, t + \Delta t) - \tilde{f}(\lambda, t)}{\Delta t} = \frac{\tilde{f}(\lambda, t + \Delta t) - \tilde{f}(\lambda, t)}{\Delta t}.$$

Ходя здесь к пределу при $\Delta t \rightarrow 0$, получаем, что $\frac{\partial \tilde{f}}{\partial t} = \frac{\partial \tilde{f}}{\partial t}$ (предполагается, что эти производные существуют).

7. Свертка и преобразование Фурье. Пусть $\tilde{f}_1(\lambda)$ и $\tilde{f}_2(\lambda)$ — преобразования Фурье функций $f_1(t)$ и $f_2(t)$. Тогда

$$\begin{aligned}\tilde{f}_1(\lambda)\tilde{f}_2(\lambda) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f_1(t) e^{-i\lambda t} dt \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f_2(\tau) e^{-i\lambda \tau} d\tau = \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f_1(t) dt \int_{-\infty}^{+\infty} f_2(\tau) e^{-i(t+\tau)\lambda} d\tau.\end{aligned}$$

Сделаем замену переменной $t + \tau = x$, так что $\tau = x - t$. Будем иметь:

$$\tilde{f}_1(\lambda)\tilde{f}_2(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f_1(t) dt \int_{-\infty}^{+\infty} f_2(x-t) e^{-i\lambda x} dx,$$

откуда, меняя порядок интегрирования (предполагается, что это возможно), получаем:

$$\tilde{f}_1(\lambda)\tilde{f}_2(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\lambda x} dx \int_{-\infty}^{+\infty} f_1(t) f_2(x-t) dt. \quad (13)$$

Функция

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f_1(t) f_2(x-t) dt$$

называется *сверткой* функций f_1 и f_2 .

Формула (13) может быть теперь записана так:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi(x) e^{-i\lambda x} dx = \tilde{f}_1(\lambda)\tilde{f}_2(\lambda),$$

т. е. преобразование Фурье свертки функций f_1 и f_2 равно произведению преобразований Фурье свертываемых функций.

§ 16.2. Дельта-функция

1. Определение дельта-функции. Эта функция (она обозначается $\delta(x)$), широко применяемая в математике и ее приложениях, была введена английским физиком Полем Дираком (1902— 1984). Чтобы приближен-

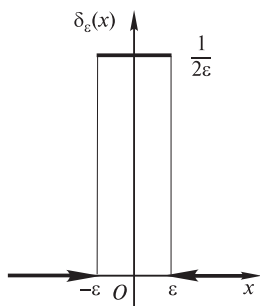


Рис. 197

но представить себе дельта-функцию, рассмотрим ступенчатую функцию (рис. 197).

$$\delta_{\epsilon}(x) = \begin{cases} \frac{1}{2\epsilon}, & |x| \leq \epsilon, \\ 0, & |x| > \epsilon. \end{cases}$$

Отсюда

$$\int_{-\infty}^{+\infty} \delta_{\epsilon}(x) dx = 1.$$

Дельта-функция получается в пределе при $\epsilon \rightarrow 0$. Формальное определение дельта-функции (которая не является функцией в обычном смысле этого термина) таково:

$$\delta = 0, \quad x \neq 0, \quad (1)$$

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1. \quad (2)$$

Это так называемая *обобщенная* функция. Теория обобщенных функций была построена советским математиком С. Л. Соболевым (1908—1989) и французским математиком Лораном Шварцем (р. 1915 г.).

Дельта-функцию можно рассматривать, например, как плотность единичной массы, сосредоточенной в точке $x = 0$ (это лишь одна из известных ее физических интерпретаций). В самом деле, обозначим эту плотность через $\rho(x)$. Тогда $\rho(x) = 0$ при $x \neq 0$, т. е. вся масса сосредоточена в точке $x = 0$, и $\int_{-\infty}^{+\infty} \rho(x) dx = 1$, ибо масса равна единице, так что $\rho(x) = \delta(x)$.

Пусть $f(x)$ — непрерывная на всей числовой оси функция. Из равенств (1) и (2) следует основное соотношение

$$\int_{-\infty}^{+\infty} \delta(x) f(x) dx = f(0). \quad (3)$$

Действительно, в силу (1)

$$\int_{-\infty}^{+\infty} \delta(x) f(x) dx = \int_{-\epsilon}^{\epsilon} \delta(x) f(x) dx, \quad (4)$$

где ϵ — малая положительная величина. В последнем интеграле промежуток интегрирования мал (его длина равна 2ϵ), поэтому на нем $f(x) \approx f(0)$, следовательно,

$$\int_{-\epsilon}^{\epsilon} \delta(x) f(x) dx = \int_{-\epsilon}^{\epsilon} \delta(x) f(0) dx = f(0) \int_{-\epsilon}^{\epsilon} \delta(x) dx. \quad (5)$$

Но с учетом равенств (1) и (2) $\int_{-\varepsilon}^{\varepsilon} \delta(x) dx = \int_{-\infty}^{+\infty} \delta(x) dx = 1$. Отсюда и из соотношений (5) и (4) получаем соотношение (3).

2. Представление дельта-функции через интеграл Фурье. Воспользуемся формулой (9) (из § 16.1) для преобразования Фурье. Пусть $f(x) = \delta(x)$. Тогда в силу формулы (3) $\tilde{f}(\lambda) \equiv \frac{1}{\sqrt{2\pi}}$. Подставляя, это выражение в формулу обращения, получаем:

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\lambda x} d\lambda.$$

3. Единичная функция Хевисайда. При помощи дельта-функции легко записываются некоторые другие функции, имеющие большое значение. Важным примером может служить *единичная функция Хевисайда* (Оливер Хевисайд (1850—1925) — английский физик):

$$e(x) = \begin{cases} 1, & \text{если } x \geq 0, \\ 0, & \text{если } x < 0. \end{cases}$$

График ее изображен на рисунке 198. Она получается при внезапном подключении какого-либо постоянного воздействия, например напряжения в электрическую цепь.

Так как $\delta(x) = 0$ при всех $x \neq 0$, то с учетом равенства (2) имеем ([9], т. II):

$$\int_{-\infty}^x \delta(t) dt = \begin{cases} 1, & \text{если } x \geq 0, \\ 0, & \text{если } x < 0. \end{cases}$$

т. е. этот интеграл равняется единичной функции Хевисайда. Итак,

$$e(x) = \int_{-\infty}^x \delta(t) dt. \quad (6)$$

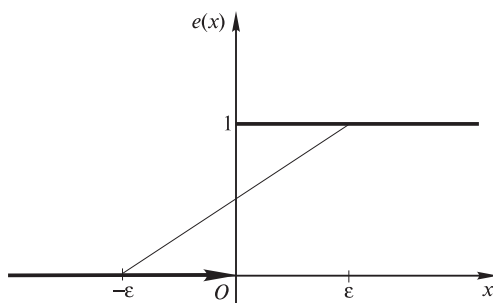


Рис. 198

Если продифференцировать равенство (6), то получится, что

$$e'(x) = \delta(x). \quad (7)$$

Это равенство понимается в обобщенном смысле. Например, можно заменить на рисунке 198 вертикальный отрезок наклонным, соединяющим точки $(-\epsilon; 0)$ и $(\epsilon; 1)$ (тонкая линия). Тогда разрывная функция заменится на непрерывную, производная которой имеет график, изображенный на рисунке 197. Если теперь перейти к пределу при $\epsilon \rightarrow 0$, то получим соотношение (7).

Таким образом, дельта-функция получается при дифференцировании разрывной функции $e(x)$.

4. Производная дельта-функции. Дельта-функцию можно не только интегрировать, но и дифференцировать; интегралы с участием $\delta'(x)$ вычисляются с помощью интегрирования по частям:

$$\int_{-\infty}^{+\infty} f(x) \delta'(x) dx = f(x) \delta(x) \Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f'(x) \delta(x) dx = -f'(0). \quad (8)$$

Здесь функция $f(x)$ непрерывна вместе со своей производной. Формулу (8) можно принять за определение производной дельта-функции.

Аналогично можно рассматривать и производные высших порядков.

Глава 17. ОСНОВНЫЕ УРАВНЕНИЯ И ОСНОВНЫЕ ЗАДАЧИ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

§ 17.1. Основные понятия о дифференциальных уравнениях с частными производными второго порядка

Дифференциальным уравнением с частными производными называется равенство, связывающее неизвестную функцию от нескольких независимых переменных, эти переменные и частные производные от неизвестной функции по независимым переменным.

Порядком дифференциального уравнения с частными производными называется порядок старшей частной производной, входящей в это уравнение.

Если $u = u(x, y)$, то в общем случае дифференциальное уравнение с частными производными второго порядка имеет вид:

$$F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial y^2}, \frac{\partial^2 u}{\partial x \partial y}\right) = 0, \quad (1)$$

где F — известная функция.

Дифференциальное уравнение с частными производными называется *линейным*, если оно линейно относительно неизвестной функции и всех ее частных производных.

Для физических приложений особый интерес представляют линейные дифференциальные уравнения с частными производными второго порядка.

Для случая двух независимых переменных приведем важнейшие типы таких уравнений.

1. Волновое уравнение

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}. \quad (2)$$

Это уравнение встречается при изучении ряда колебательных процессов (поперечные колебания упругой струны, колебание газа в трубке и др.).

2. Уравнение теплопроводности (уравнение Фурье)

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad (3)$$

описывающее процессы распространения тепла и т. д.

3. Уравнение Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0. \quad (4)$$

К исследованию этого уравнения приводит рассмотрение задач об электрических и магнитных полях, о стационарном тепловом состоянии, задач гидродинамики, диффузии и т. д. (см. [17]).

Для решения этих уравнений в различных условиях были созданы специальные приемы (так называемые «методы математической физики»).

Решением уравнения (1) называется всякая функция $u = \varphi(x, y)$, обращающая уравнение (1) в тождество.

Как известно (см. § 13.1, 13.3), общие решения обыкновенных дифференциальных уравнений содержат произвольные постоянные. Для дифференциальных уравнений с частными производными их общие решения включают произвольные функции.

§ 17.2. Вывод уравнения колебаний струны

В математической физике под *струной* понимают однородную гибкую, упругую нить (т. е. с постоянной линейной плотностью).

Пусть струна длиной l в начальный момент направлена по отрезку оси Ox от 0 до l . Предположим, что концы струны закреплены в точках $x=0$ и $x=l$. Если струну отклонить от ее первоначального положения, а потом предоставить самой себе, то точки струны будут совершать движения — говорят, что струна начнет колебаться.

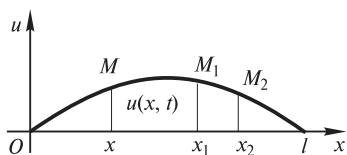


Рис. 199

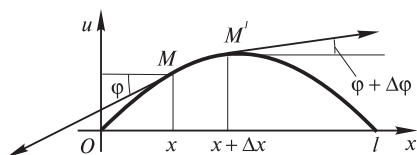


Рис. 200

Будем рассматривать малые отклонения точек струны от начального положения. В силу этого можно предполагать, что движение точек струны происходит перпендикулярно оси Ox и в одной плоскости. При этом предположении процесс колебания струны описывается одной функцией $u(x, t)$, которая дает величину перемещения точки струны с абсциссой x в момент времени t (рис. 199).

Так как мы рассматриваем малые отклонения струны в плоскости xOu , то будем предполагать, что длина элемента струны $l_{MM'}$ равна ее проекции на ось Ox , т. е. длина дуги M_1M_2 равна $x_2 - x_1$. Также будем считать, что натяжение во всех точках струны одинаковое по величине, обозначим его через T .

Рассмотрим элемент струны MM' (рис. 200).

На концах этого элемента, по касательным к струне, действуют силы T . Пусть касательные образуют с осью Ox углы φ и $\varphi + \Delta\varphi$. Тогда проекция на ось Ou сил, действующих на элемент MM' , будет равна $T \sin(\varphi + \Delta\varphi) - T \sin \varphi$. Так как угол φ мал, то можно считать $\operatorname{tg} \varphi \approx \sin \varphi$, и мы будем иметь:

$$\begin{aligned} T \sin(\varphi + \Delta\varphi) - T \sin \varphi &\approx T \operatorname{tg}(\varphi + \Delta\varphi) - T \operatorname{tg} \varphi = T \left(\frac{\partial u(x + \Delta x, t)}{\partial x} - \frac{\partial u(x, t)}{\partial x} \right) = \\ &= T \frac{\partial^2 u(x + \Theta \Delta x, t)}{\partial x^2} \Delta x \approx T \frac{\partial^2 u(x, t)}{\partial x^2} \Delta x, \end{aligned}$$

$0 < \Theta < 1$ (здесь мы применили формулу Лагранжа (§ 8.6, п. 3) к разности, стоящей в скобках).

Чтобы получить уравнение движения, надо внешние силы, приложенные к элементу MM' , приравнять силе инерции. Пусть ρ — линейная плотность струны. Тогда масса этого элемента струны будет $\rho \Delta x$, а его ускорение равно $\frac{\partial^2 u}{\partial t^2}$. Следовательно, по второму закону Ньютона $\rho \Delta x \frac{\partial^2 u}{\partial t^2} = T \frac{\partial^2 u}{\partial x^2} \Delta x$.

Сокращая на Δx и обозначая $\frac{T}{\rho} = a^2$, получаем: $\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$.

Это уравнение называется *уравнением свободных колебаний струны или одномерным однородным волновым уравнением*.

§ 17.3. Вывод акустического уравнения

Уравнение свободных колебаний струны

$$u''_{x^2} - a^2 u''_{x^2} = 0 \quad (1)$$

описывает ряд других одномерных волновых процессов (см. [17]). В частности, оно описывает (при некоторых упрощающих предположениях) продольные колебания газа (акустические колебания) в узкой трубке.

Чтобы это показать, выведем сначала *трехмерное уравнение акустики* (уравнение распространения акустических волн в пространстве). Пусть $\rho = \rho(M, t)$ — плотность газа в точке $M(x; y; z)$ в момент t . Если $\bar{a} = \bar{a}(M, t)$ означает поле скоростей частиц газа, то, как известно (см. [3]), имеет место *уравнение неразрывности* (предполагается отсутствие источников (стоков))

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \bar{a}) = 0,$$

записываемое в силу свойства 3 из примечания 2 к п. 4 § 14.2 также в виде

$$\frac{\partial \rho}{\partial t} + \rho \operatorname{div} \bar{a} + (\operatorname{grad} \rho, \bar{a}) = 0, \quad (2)$$

и *уравнение движения Эйлера* (предполагается отсутствие внешних сил)

$$\frac{\partial \bar{a}}{\partial t} + (\bar{a}, \nabla) \bar{a} = -\frac{1}{\rho} \operatorname{grad} p, \quad (3)$$

где $p = p(M, t)$ — давление газа.

Рассмотрим адиабатические движения газа, для которых справедливо соотношение

$$p = p_0 \left(\frac{\rho}{\rho_0} \right)^\gamma \quad (4)$$

(уравнение состояния), где $\gamma = \frac{c_p}{c_v}$ (c_p, c_v — удельные теплоемкости со-

ответственно при постоянном давлении и объеме), p_0 и ρ_0 — начальные значения давления и плотности. Уравнения (2)–(4) называются *уравнениями газовой динамики*.

Введем в рассмотрение величину $u = u(M, t)$, полагая

$$\rho = \rho_0(1 + u), \quad (5)$$

называемую *уплотнением* газа.

Для вывода уравнения акустики следует исключить из уравнений (2) и (3) поле $\bar{a}$ и получить, пользуясь уравнением (4), уравнение, содержащее только неизвестную функцию ρ , т. е. u .

Рассматриваемые колебания предполагаются настолько малыми, что уравнения (2), (3) и (4) можно *линеаризировать*, т. е. оставить в них (за малостью остальных членов) только линейные члены. Так, мы полагаем, что

$$\rho \operatorname{div} \bar{a} = \rho_0 \operatorname{div} \bar{a} \text{ (величинами } u \frac{\partial a_x}{\partial x}, u \frac{\partial a_y}{\partial y}, u \frac{\partial a_z}{\partial z} \text{ пренебрегаем),}$$

$$(\operatorname{grad} p, \bar{a}) = 0 \text{ (величинами } \frac{\partial u}{\partial x} a_x, \frac{\partial u}{\partial y} a_y, \frac{\partial u}{\partial z} a_z \text{ пренебрегаем),}$$

$$(\bar{a}, \nabla) \bar{a} = 0 \text{ (величинами } a_x \frac{\partial a_x}{\partial x} \text{ и т. д. пренебрегаем),}$$

$$\frac{1}{\rho} \operatorname{grad} p = \frac{1}{\rho_0} \operatorname{grad} p \text{ (так как } \frac{1}{\rho} = \frac{1}{\rho_0(1+u)} = \frac{1}{\rho_0}(1-u+u^2-\dots), \text{ а вели-}$$

чинами $u \frac{\partial p}{\partial x}$ и т. д. мы пренебрегаем) и, наконец, что $p = p_0(1+u)^\gamma = p_0(1 +$

$+\gamma u)$ (величинами u^2 и т. д. пренебрегаем). Тогда уравнения (2) и (3)

$$\text{примут вид: } \frac{\partial \rho}{\partial t} + \rho_0 \operatorname{div} \bar{a} = 0 \text{ и } \frac{\partial \bar{a}}{\partial t} = -\frac{1}{\rho_0} \operatorname{grad} p = -\frac{p_0 \gamma}{\rho_0} \operatorname{grad} u.$$

Дифференцируя первое из них по t

$$\frac{\partial^2 \rho}{\partial t^2} + \rho_0 \operatorname{div} \frac{\partial \bar{a}}{\partial t} = 0$$

(здесь использована независимость смешанных производных от порядка дифференцирования) и подставляя вместо $\frac{\partial \bar{a}}{\partial t}$ его выражение из вто-

рого, получим: $\frac{\partial^2 \rho}{\partial t^2} - p_0 \gamma \operatorname{div} \operatorname{grad} u = 0$. Но $\frac{\partial^2 \rho}{\partial t^2} = \rho_0 \frac{\partial u}{\partial t^2}$ (это следует из

(5)), а $\operatorname{div} \operatorname{grad} u = \Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$ (см. § 14.3, п. 1, (7)). Поэтому

$$\frac{\partial^2 u}{\partial t^2} - a^2 \Delta u = 0 \left(a^2 = \frac{p_0 \gamma}{\rho_0} \right). \quad (6)$$

Уравнение (6) и является *уравнением акустики*. Оно играет фундаментальную роль в теоретической физике. Коэффициент a , как и в случае уравнения колебаний струны, означает скорость распространения волн.

Что касается колебаний газа в узкой трубке, занимающей отрезок $[0; l]$ оси Ox , то в этом случае уплотнение u можно считать функцией только от x и t : $u = u(x, t)$ — и уравнение (6) сведется к уравнению (1).

§ 17.4. Вывод уравнения теплопроводности

Пусть на оси Ox лежит однородный стержень, теплоизолированный от внешней среды (нет обмена с внешней средой) и неравномерно нагретый. При этом в стержне будет происходить теплообмен между более нагретыми частями и частями, менее нагретыми. Если, например, в начальный момент времени температура точек стержня с увеличением x уменьшается, то с течением времени будет происходить поток тепла в положительном направлении оси Ox . Обозначим через $u(x, t)$ температуру в сечении стержня с абсциссой x в момент t . При этом частная производная $u'_t(x, t)$ равна скорости повышения температуры в момент времени t в данном сечении стержня, а частная производная $u'_x(x, t)$ — скорость изменения температуры в зависимости от точки x стержня. Эта величина характеризует перепад температур. Если $u'_x(x, t) < 0$, то более правые точки стержня имеют более низкую температуру. По законам физики поток тепла через сечение стержня пропорционален площади S сечения стержня, перепаду температур и промежутку времени. Иными словами, за малый промежуток времени Δt (температурные условия в этом промежутке сохраняются) через сечение стержня с абсциссой x пройдет поток тепла

$$Q(x) = -kS \frac{\partial u(x, t)}{\partial x} \Delta t. \quad (1)$$

Знак «минус» оправдан тем, что при движении тепла в положительном направлении оси Ox производная $u'_x(x, t)$, как уже отмечалось, отрицательна, а количество тепла Q , прошедшее через сечение в положительном направлении, положительно. Число k называется *коэффициентом теплопроводности* стержня. Величина k зависит от материала, из которого сделан стержень.

Проследим за изменением количества тепла и температуры в той части стержня, которая заключена между сечениями с абсциссами x и $x + \Delta x$. Количество тепла, входящее через сечение с абсциссой x за время Δt , определяется по формуле (1). Согласно этой же формуле количество тепла, выходящее через сечение с абсциссой $x + \Delta x$ за это же время Δt , равно:

$$Q(x + \Delta x) = -kS \frac{\partial u(x + \Delta x, t)}{\partial x} \Delta t.$$

Взяв разность величин входящего и выходящего тепловых потоков, получим количество тепла ΔQ , сообщенное выбранному участку стержня за время Δt :

$$\Delta Q = kS \left(\frac{\partial u(x + \Delta x, t)}{\partial x} - \frac{\partial u(x, t)}{\partial x} \right) \Delta t.$$

Отсюда, применяя формулу Лагранжа (§ 8.6, п. 3), получим:

$$\Delta Q = kS \frac{\partial^2 u(x + \Theta_1 \Delta x, t)}{\partial x^2} \Delta t \Delta x, \quad (2)$$

где $0 < \Theta_1 < 1$.

Будем считать выделенный участок стержня (между сечениями с абсциссами x и $x + \Delta x$) столь малым, что в каждый данный момент температуру всех его сечений можно считать одной и той же.

Тогда, с другой стороны, сообщенное количество тепла выбранному участку стержня за время Δt равно:

$$\Delta Q = c\rho S(u(x, t + \Delta t) - u(x, t))\Delta x,$$

где c — удельная теплоемкость вещества стержня, ρ — плотность стержня.

Отсюда, вновь применяя формулу Лагранжа, будем иметь:

$$\Delta Q = c\rho S \frac{\partial u(x, t + \Theta_2 \Delta t)}{\partial t} \Delta x \Delta t, \quad (3)$$

где $0 < \Theta_2 < 1$.

Из формул (2), (3) получаем: $c\rho \frac{\partial u(x, t + \Theta_2 \Delta t)}{\partial t} = k \frac{\partial^2 u(x + \Theta_1 \Delta x, t)}{\partial x^2}$.

Переходя здесь к пределу при $\Delta x \rightarrow 0$ и $\Delta t \rightarrow 0$, получим:

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad (4)$$

где положено $a^2 = \frac{k}{c\rho}$.

Уравнение (4) называют *уравнением теплопроводности стержня*.

§ 17.5. Классификация задач математической физики

Как уже отмечалось (см. § 17.1), дифференциальные уравнения с частными производными имеют бесчисленное множество решений. Поэтому в том случае, когда физическая задача приводится к уравнению с частными производными, для однозначной характеристики процесса необходимо к уравнению присоединить некоторые дополнительные условия.

Для задач, приводящих к уравнениям (2), (3) из § 17.1, дополнительные условия разделяются на *начальные* и *краевые*.

Начальные условия состоят в задании в один какой-нибудь момент времени, с которого начинается изучение данного физического явления (обычно при $t = 0$), значений искомой функции и ее производной (в случае уравнения (2)) или только значений самой функции (в случае уравнения (3)).

Краевые (или граничные) условия для этих задач заключаются в том, что указываются значения неизвестной функции $u(x, t)$ на концах интервала изменения координаты x .

Если процесс протекает в бесконечном интервале изменения координаты x (бесконечная струна, бесконечный стержень — § 17.6), то краевые условия отпадают и получается *задача* только с начальными условиями, или, как ее часто называют, *задача Коши*.

Если ставится задача для конечного промежутка, то должны быть заданы и начальные и краевые условия. Тогда говорят о *смешанной задаче* (см. § 17.7).

Например, рассмотрим простейшую задачу о поперечных колебаниях струны, закрепленной на концах. В этой задаче $u(x, t)$ даст отклонение струны от оси Ox . Если концы струны $0 \leq x \leq l$ закреплены, то должны выполняться краевые условия

$$u|_x=0=0, u|_{x=l}=0, \quad (1)$$

или, более кратко,

$$u(0, t) = 0, u(l, t) = 0.$$

Краевые условия (1), где правые части — нули, называются *однородными*.

Однако заданием только этих краевых условий закон колебания не определяется однозначно. Он будет зависеть еще и от начальной формы струны, и от начальной скорости струны в каждой точке. Итак, кроме краевых условий, должны быть заданы следующие начальные условия: $u(x, 0) = f(x)$, $u'_t(x, 0) = F(x)$. Здесь $f(x)$ и $F(x)$ — известные, заранее заданные функции: первая из них в качестве графика имеет форму струны в начальный момент, вторая указывает, какова в начальный момент скорость каждой точки струны.

Из физических соображений ясно, что, придав некоторое начальное положение и некоторую начальную скорость каждой точке струны, а затем отпустив струну, мы вызовем вполне определенное движение точек струны. Следовательно, задание краевых и начальных условий однозначно определяет некоторый закон колебания струны. Этот закон будет найден позднее (см. § 17.7).

Переходя к уравнению (4) из § 17.1, заметим, что в это уравнение время t не входит и обе независимые переменные являются координатами точки. Для задач, приводящих к уравнению (4), ставятся только краевые условия, т. е. указывается поведение неизвестной функции на контуре области (см. задачу Дирихле — § 17.8).

Кроме того, следует иметь в виду, что каждое из выведенных уравнений носит идеализированный характер, т. е. отражает лишь наиболее существенные черты соответствующего физического процесса.

Данные, входящие как в уравнение, так и в дополнительные условия в физической задаче, определяются из экспериментальных данных и могут считаться известными лишь приближенно. Поэтому мы должны быть уверенными в том, что решение задачи при приближенных исходных данных будет близко к тем решениям, которые получались бы при точных исходных данных. Таким образом, важно, чтобы малые изменения данных задачи вызывали лишь малые изменения в ее решении, т. е., как говорят, чтобы решение было устойчиво относительно исходных данных.

Задача математической физики считается поставленной правильно (*корректно*), если решение задачи, удовлетворяющее всем ее условиям, существует, единственно и устойчиво.

Все задачи, которые мы будем рассматривать, принадлежат к типу задач, имеющих единственное решение, устойчивое относительно исходных данных, т. е. задачи поставлены корректно.

§ 17.6. Задача Коши

1. Колебания бесконечной струны. Метод Даламбера. Рассмотрим задачу о колебаниях бесконечной струны. Если струна очень длинная, то на колебания, возникающие где-то в ее середине, концы струны будут оказывать малое влияние. Поэтому, рассматривая свободные колебания неограниченной струны, мы должны решить уравнение

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

только при начальных условиях

$$u(x, 0) = f(x), \quad (2)$$

$$u'_t(x, 0) = F(x), \quad (3)$$

причем ввиду неограниченности струны функции $f(x)$ и $F(x)$ заданы на всей числовой оси.

Такая задача, как уже отмечалось (§ 17.5), называется задачей с начальными условиями или задачей Коши. Метод решения ее, который мы сейчас изложим, называется *методом Даламбера* или *методом бегущих волн*.

Прежде всего заметим, что если $\varphi(p)$ и $\psi(q)$ (p, q — произвольные независимые переменные) — любые дважды дифференцируемые функции, то функция

$$u(x, t) = \varphi(x - at) + \psi(x + at) \quad (4)$$

является решением уравнения (1).

Действительно, последовательно дифференцируя, находим:

$$\begin{aligned}u'_x &= \varphi'(x - at) + \psi'(x + at), \\u''_{x^2} &= \varphi''(x - at) + \psi''(x + at), \\u'_t &= -a\varphi'(x - at) + a\psi'(x + at), \\u''_{t^2} &= a^2\varphi''(x - at) + a^2\psi''(x + at),\end{aligned}\tag{5}$$

и, следовательно, $u''_{t^2} = a^2 u''_{x^2}$.

Решение (4), зависящее от двух произвольных функций, называется *решением Даламбера*. Для определения этих функций воспользуемся начальными условиями (2), (3). Полагая в формулах (4), (5) $t = 0$, получим:

$$\varphi(x) + \psi(x) = f(x), \quad -a\varphi'(x) + a\psi'(x) = F(x).\tag{6}$$

Интегрируя последнее равенство на отрезке $[0; x]$, получим:

$$-a[\varphi(x) - \varphi(0)] + a[\psi(x) - \psi(0)] = \int_0^x F(y)dy,$$

которое приведем к виду

$$-\varphi(x) + \psi(x) = \frac{1}{a} \int_0^x F(y)dy + C,\tag{7}$$

где $C = -\varphi(0) + \psi(0)$. Из равенств (6) и (7) находим искомые функции $\varphi(x)$ и $\psi(x)$:

$$\varphi(x) = \frac{1}{2}f(x) - \frac{1}{2a} \int_0^x F(y)dy - \frac{C}{2},\tag{8}$$

$$\psi(x) = \frac{1}{2}f(x) + \frac{1}{2a} \int_0^x F(y)dy + \frac{C}{2}.\tag{9}$$

Заменяя в формулах (8) и (9) x соответственно на $x - at$ и $x + at$ и подставляя полученные выражения в формулу (4), найдем решение задачи Коши:

$$u(x, t) = \frac{f(x - at) + f(x + at)}{2} + \frac{1}{2a} \int_{x-at}^{x+at} F(y)dy.\tag{10}$$

Формула (10) называется *формулой Даламбера*.

2. Теплопроводность в бесконечном стержне. В случае бесконечного стержня краевые условия отсутствуют и задача Коши сводится к отысканию решения уравнения

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2},\tag{11}$$

определенного для всех x , $-\infty < x < +\infty$ и $t > 0$, и удовлетворяющего начальному условию

$$u(x, 0) = f(x). \quad (12)$$

Для решения этой задачи совершим преобразование Фурье по x (см. § 16.1, пп. 4, 6) обеих частей уравнения (11), считая t параметром.

В силу свойств 1, 5 и 4 из п. 6 (§ 16.1) уравнение (11) при этом перейдет в

$$\frac{\partial \tilde{u}}{\partial t} = a^2 (\hat{\lambda})^2 \tilde{u} = -a^2 \lambda^2 \tilde{u}. \quad (13)$$

При любом фиксированном λ уравнение (13) — обыкновенное линейное однородное дифференциальное уравнение с постоянными коэффициентами. Решая его, получим:

$$\tilde{u} = C(\lambda) e^{-a^2 \lambda^2 t}, \quad (14)$$

где постоянная C может зависеть от λ .

Для определения $C(\lambda)$ совершим преобразование Фурье над обеими частями начального условия (12), что дает

$$\tilde{u}(\lambda, 0) = \tilde{f}(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\lambda x} dx.$$

Но в силу равенства (14) $\tilde{u}(\lambda, 0) = C(\lambda)$. Поэтому

$$C(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-i\lambda x} dx.$$

Подставляя это выражение в равенство (14) и изменяя обозначение переменной интегрирования, получаем:

$$\tilde{u}(\lambda, t) = \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(s) e^{-i\lambda s} ds \right) e^{-a^2 \lambda^2 t}.$$

Совершая обратное преобразование, получаем искомое решение

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\lambda x} d\lambda \int_{-\infty}^{+\infty} f(s) e^{-i\lambda s - a^2 \lambda^2 t} ds.$$

Упростим ответ. Для этого внесем $e^{i\lambda x}$ под знак внутреннего интеграла и поменяем порядок интегрирования. Будем иметь:

$$u(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(s) ds \int_{-\infty}^{+\infty} e^{-a^2 \lambda^2 t + i(x-s)\lambda} d\lambda. \quad (15)$$

Внутренний интеграл в равенстве (15) равен (см. [2]):

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-a^2 \lambda^2 t + i(x-s)\lambda} d\lambda = \frac{1}{2\sqrt{\pi a^2 t}} e^{-\frac{(x-s)^2}{4a^2 t}}.$$

Отсюда и из (15) находим:

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}} \int_{-\infty}^{+\infty} f(s) e^{-\frac{(x-s)^2}{4a^2 t}} ds.$$

Последняя формула была получена в 1823 г. Пуассоном.

§ 17.7. Смешанная задача для одномерного однородного волнового уравнения и ее решение методом Фурье

1. Решение смешанной задачи методом Фурье. Рассмотрим указанную в § 17.5 задачу о свободных колебаниях струны, закрепленной на обоих концах. Как там отмечалось, эта задача сводится к решению уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

при краевых условиях

$$u(0, t) = 0, u(l, t) = 0 \quad (2)$$

и начальных условиях

$$u(x, 0) = f(x), u'_t(x, 0) = F(x). \quad (3)$$

Уравнение (1) будем решать *методом Фурье* — разделения переменных. Суть его заключается в том, что мы отыскиваем частные решения уравнения (1), удовлетворяющие пока только краевым условиям (2), в виде произведения двух функций, каждая из которых зависит только от одной переменной:

$$u(x, t) = X(x)T(t). \quad (4)$$

При этом мы ищем нетривиальные решения, т. е. тождественно не равные нулю. Подставляя функцию (4) в уравнение (1), получаем:

$$X(x)T''(t) = a^2 X''(x)T(t),$$

или

$$\frac{T''(t)}{a^2 T(t)} = \frac{X''(x)}{X(x)}.$$

Мы получили равенство, в котором левая часть не зависит от x , а правая — от t . Следовательно, обе части равенства не зависят ни от x , ни от t , т. е. являются постоянными. Обозначим эту постоянную символом $-\lambda^2$ (поскольку отношение $\frac{u''_{tt}}{u} = \frac{T''}{T}$ ускорения к отклонению является отрицательным, так как вектор ускорения направлен здесь к положению равновесия):

$$\frac{T''(t)}{a^2 T(t)} = \frac{X''(x)}{X(x)} = -\lambda^2.$$

Отсюда

$$T'''(t) + \lambda^2 a^2 T(t) = 0,$$

$$X''(x) + \lambda^2 X(x) = 0.$$

Из этих уравнений находим (см. § 13.4, п. 2):

$$T(t) = A \cos \lambda at + B \sin \lambda at,$$

$$X(x) = C \cos \lambda x + D \sin \lambda x.$$

Таким образом,

$$u(x, t) = (A \cos \lambda at + B \sin \lambda at)(C \cos \lambda x + D \sin \lambda x).$$

Полагая здесь $x = 0$, получим:

$$u(0, t) = (A \cos \lambda at + B \sin \lambda at) C \equiv 0.$$

Значит, $C = 0$. Далее, поскольку

$$u(l, t) = (A \cos \lambda at + B \sin \lambda at) D \sin \lambda l \equiv 0,$$

то λ следует выбрать так, чтобы

$$\sin \lambda l = 0, \lambda l = k\pi, \lambda = \frac{k\pi}{l}.$$

Эти значения λ называются *собственными значениями* для данной краевой задачи, а соответствующие им функции

$$X(x) = D \sin \frac{k\pi}{l} x$$

собственными функциями.

Итак, частное решение уравнения (1), удовлетворяющее краевым условиям, найдено:

$$u_k(x, t) = \left(a_k \cos \frac{k\pi a}{l} t + b_k \sin \frac{k\pi a}{l} t \right) \sin \frac{k\pi}{l} x. \quad (5)$$

Здесь числа $a_k = AD$, $b_k = BD$ произвольны.

Заметим теперь, что уравнение (1) линейное и однородное и потому сумма его решений также будет решением. Так что функция

$$u(x, t) = \sum_{k=1}^{\infty} u_k(x, t) = \sum_{k=1}^{\infty} \left(a_k \cos \frac{k\pi at}{l} + b_k \sin \frac{k\pi at}{l} \right) \sin \frac{k\pi}{l} x \quad (6)$$

является решением уравнения (1). Ясно, что эта функция удовлетворяет краевым условиям (2), так как им удовлетворяет каждая из функций $u_k(x, t)$.

Подберем теперь числа a_k и b_k так, чтобы удовлетворить и начальным условиям (3) уравнения (1). Имеем:

$$u(x, 0) = f(x) = \sum_{k=1}^{\infty} a_k \sin \frac{k\pi}{l} x.$$

Это разложение функции $f(x)$ в ряд Фурье по синусам. Следовательно (§ 10.4),

$$a_k = \frac{2}{l} \int_0^l f(x) \sin \frac{k\pi x}{l} dx, \quad k = 1, 2, \dots \quad (7)$$

Далее, так как

$$u'_t(x, 0) = F(x) = \sum_{k=1}^{\infty} \frac{k\pi a}{l} b_k \sin \frac{k\pi}{l} x,$$

то

$$\frac{k\pi a}{l} b_k = \frac{2}{l} \int_0^l F(x) \sin \frac{k\pi x}{l} dx,$$

откуда

$$b_k = \frac{2}{k\pi a} \int_0^l F(x) \sin \frac{k\pi x}{l} dx, \quad k = 1, 2, \dots \quad (8)$$

Таким образом, найдено решение уравнения свободного колебания струны, удовлетворяющего указанным краевым и начальным условиям:

$$u(x, t) = \sum_{k=1}^{\infty} \left(a_k \cos \frac{k\pi a}{l} t + b_k \sin \frac{k\pi a}{l} t \right) \sin \frac{k\pi}{l} x,$$

где a_k и b_k находятся из равенств (7) и (8).

2. Стоячие волны. Представим решение $u_k(x, t)$, определенное формулой (5), в другом виде, обозначив

$$a_k = N_k \sin \varphi_k, \quad b_k = N_k \cos \varphi_k,$$

откуда

$$N_k = \sqrt{a_k^2 + b_k^2}, \quad \operatorname{tg} \varphi_k = \frac{a_k}{b_k}.$$

Тогда получим:

$$u_k(x, t) = N_k \left(\sin \varphi_k \cos \frac{\pi k a t}{l} + \cos \varphi_k \sin \frac{\pi k a t}{l} \right) \sin \frac{\pi k x}{l},$$

или

$$u_k(x, t) = N_k \sin \frac{\pi k x}{l} \sin \left(\frac{\pi k a t}{l} + \varphi_k \right). \quad (8')$$

Из полученной формулы (8') видно, что все точки струны совершают гармонические колебания с одной и той же частотой $\omega_k = \frac{k\pi a}{l}$ (собст-



Рис. 201

венной частотой струны) и фазой φ_k . Амплитуда колебания зависит от абсциссы x точки струны и равна $N_k \sin \frac{k\pi x}{l}$. Движение струны такого типа называется *стоячей волной*.

Точки, для которых $x = \frac{l}{k}, \frac{2l}{k}, \dots, \frac{(k-1)l}{k}$, остаются на месте — это *узлы*. Тем самым струна, колеблющаяся по закону (8'), разбивается на k участков, середины которых — *пучности* — колеблются с наибольшей амплитудой. На рисунке 201 изображены последовательные положения струны, колеблющейся по закону (8') для $k = 1, 2, 3$.

Колебания струны воспринимаются нами обычно по звуку, издаваемому струной. Не останавливаясь на процессе распространения колебаний в воздухе и восприятия звуковых колебаний нашим ухом, можно сказать, что звук струны является наложением *простых тонов*, соответствующих стоячим волнам, на которые разлагается колебание (8'). Это разложение звука на простые тоны не является операцией только математического характера. Выделение простых тонов можно произвести экспериментально при помощи резонаторов.

Высота тона зависит от частоты колебаний, соответствующих этому тону. Самый низкий тон, который может создавать струна, определяется составляющей $u_1(x, t)$ с частотой $\omega_1 = \frac{a\pi}{l} = \frac{\pi}{l} \sqrt{\frac{T}{\rho}}$ и называется *ос-*

новым тоном струны. Остальные тона, соответствующие частотам, кратным ω_1 , называются *обертонами*. Решение (6) складывается из основного тона и обертонов. Их суммарное действие приводит к созданию *тембра звука*, издаваемого закрепленной струной.

Наконец, заметим, что так как решение (6) есть сумма или наложение бесконечного числа стоячих волн, то изложенный метод Фурье часто называют также *методом стоячих волн*.

3. Единственность решения смешанной задачи. Допустим, что существуют две функции u_1 и u_2 , удовлетворяющие уравнению (1), краевым условиям (2) и начальным условиям (3). Тогда их разность $u = u_1 - u_2$ должна удовлетворять тому же уравнению (1), краевым условиям (2) и нулевым начальным условиям:

$$u(x, 0) = 0, \quad u'_t(x, 0) = 0. \quad (9)$$

Мы предположим пока, что начальные условия для функции u могут быть произвольными, ненулевыми, а именно (3). Тогда, умножая уравнение (1) на u'_t и интегрируя его по x на отрезке $[0; l]$, получим:

$$\int_0^l u'_t u''_{t^2} dx - a^2 \int_0^l u'_t u''_{x^2} dx = 0. \quad (10)$$

Так как

$$u'_t u''_{t^2} = \frac{1}{2} \frac{\partial}{\partial t} (u'_t)^2,$$

то

$$\int_0^l u'_t u''_{t^2} dx = \frac{1}{2} \frac{\partial}{\partial t} \int_0^l u'^2_t dx \quad (11)$$

(предполагая, что здесь дифференцирование по t можно вносить под знак интеграла). Далее, интегрируя по частям, найдем, что

$$\int_0^l u'_t u''_{x^2} dx = u'_t u'_x \Big|_0^l - \int_0^l u'_x u''_{tx} dx.$$

Но при краевых условиях (2)

$$u'_t u'_x \Big|_0^l = 0$$

(заметим, что если $u(0, t) = 0$, то и $u'_t(0, t) = 0$; аналогично если $u(l, t) = 0$, то и $u'_t(l, t) = 0$). Следовательно,

$$\int_0^l u'_t u''_{x^2} dx = - \int_0^l u'_x u''_{tx} dx = - \frac{1}{2} \frac{\partial}{\partial t} \int_0^l u'^2_x dx \quad (12)$$

(предполагаем, что здесь дифференцирование по t можно вносить под знак интеграла). С учетом равенств (11) и (12) равенство (10) может быть переписано в виде

$$\frac{1}{2} \frac{\partial}{\partial t} \int_0^l (u'^2_t + a^2 u'^2_x) dx = 0,$$

откуда следует, что

$$\int_0^l (u'^2_t + a^2 u'^2_x) dx = C = \text{const}. \quad (13)$$

Это равенство, называемое *интегралом энергии*, представляет собой, как это сейчас будет ниже показано, *математическое выражение закона сохранения энергии для свободных колебаний любой физической природы при нулевых краевых условиях*, т. е. при отсутствии притока энергии колебаний. Заметим, что неоднородность в краевых условиях означает наличие постоянно действующих факторов, подводящих или рассеивающих энергию. Неоднородность же в начальных условиях (напри-

мер, начальные отклонения и начальные скорости точек струны) означает лишь, что в начальный момент процесс обладал некоторым запасом энергии, который, как увидим ниже, он сохраняет в течение всего колебания.

Умножив обе части равенства (13) на $\frac{1}{2}\rho$ (ρ — линейная плотность струны), перепишем его в виде

$$\int_0^l \left(\frac{1}{2} \rho u_t'^2 + \frac{1}{2} T u_x'^2 \right) dx = E, \quad (14)$$

где $T = a^2 \rho$ и постоянная $E = \frac{1}{2} \rho C$. Заметим теперь, что

$$\int_0^l \frac{1}{2} \rho u_t'^2 dx = K(t) \quad (15)$$

является *кинетической энергией* струны в момент t , а

$$\int_0^l \frac{1}{2} T u_x'^2 dx = U(t) \quad (16)$$

ее *потенциальной энергией* в момент t . Первое следует из того, что кинетическая энергия элемента dx равна $\frac{1}{2} m v^2$, где $m = \rho dx$. а $v = u_t'(x, t)$.

Второе усматривается следующим образом. Потенциальная энергия струны есть работа силы натяжения, т. е. для элемента струны dx

$$T(ds - dx) = T(\sqrt{1 + u_x'^2} - 1)dx.$$

Но

$$\sqrt{1 + u_x'^2} - 1 = \frac{1}{2} u_x'^2 \quad (\S 17.3).$$

Таким образом, потенциальная энергия элемента струны dx в любой момент t равна $\frac{1}{2} T u_x'^2 dx$.

Итак, в силу формул (14), (15), (16) в любой момент t колебания должно иметь место равенство

$$K(t) + U(t) = E, \quad (17)$$

где E — полная энергия струны.

Положим в формулах (15) и (16) $t = 0$ и учтем, что $u_t'(x, 0) = F(x)$ и $u(x, 0) = f(x)$, и, значит, $u_x'(x, 0) = f'(x)$. Тогда в силу равенства (17) будем иметь:

$$E = K(0) + U(0) = \frac{1}{2} \rho \int_0^l F^2(x) dx + \frac{1}{2} T \int_0^l f'^2(x) dx,$$

т. е. что в любой момент t колебания выполняется равенство

$$\frac{1}{2}\rho \int_0^l u_t'^2 dx + \frac{1}{2}T \int_0^l u_x'^2 dx = \frac{1}{2}\rho \int_0^l F^2(x) dx + \frac{1}{2}T \int_0^l f'^2(x) dx. \quad (18)$$

Это равенство и показывает, что *полная энергия струны остается в течение всего колебания равной ее начальной энергии.*

Возвращаясь к доказательству единственности решения смешанной задачи, вспомним, что разность $u = u_1 - u_2$ удовлетворяет нулевым начальным условиям (9), так что $F(x) \equiv 0$ и $f(x) \equiv 0$. Тогда из формулы (18) найдем, что для этой разности для всех $t \geq 0$

$$\frac{1}{2}\rho \int_0^l u_t'^2 dx + \frac{1}{2}T \int_0^l u_x'^2 dx = 0.$$

Отсюда вытекает, что

$$\frac{1}{2}\rho u_t'^2(x, t) + \frac{1}{2}T u_x'^2(x, t) \equiv 0,$$

и, значит,

$$u_t'(x, t) \equiv 0 \text{ и } u_x'(x, t) \equiv 0.$$

а это означает, что $u(x, t) = C = \text{const}$. Но так как $u(x, 0) = 0$, то $C = 0$, и, следовательно, $u(x, t) \equiv 0$, т. е. $u_1 \equiv u_2$.

§ 17.8. Задача Дирихле для круга

Пусть в плоскости xOy имеется круг радиуса R с центром в начале координат и на его окружности задана некоторая функция $f(\varphi)$, где φ — полярный угол. Требуется найти функцию $u(r, \varphi)$, непрерывную в этом круге, включая его границу, удовлетворяющую внутри его уравнению Лапласа:

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} = 0, \quad (1)$$

и на границе данного круга принимающую заданные значения

$$u|_{r=R} = f(\varphi). \quad (2)$$

Перепишем уравнение (1) в виде

$$r^2 \frac{\partial^2 u}{\partial r^2} + r \frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial \varphi^2} = 0. \quad (3)$$

Будем искать решение методом разделения переменных, полагая

$$u = \Phi(\varphi) R(r). \quad (4)$$

Подставляя это выражение в уравнение (3), получим:

$$r^2\Phi(\varphi)R''(r) + r\Phi(\varphi)R'(r) + \Phi''(\varphi)R(r) = 0$$

или

$$\frac{\Phi''(\varphi)}{\Phi(\varphi)} = -\frac{r^2 R''(r) + rR'(r)}{R(r)}. \quad (5)$$

Так как левая часть равенства (5) не зависит от r , а правая — от φ , то они равны постоянному числу. Обозначая эту постоянную через $-k^2$ (ниже объясняется, почему мы не могли приравнять $+k^2$), получаем два уравнения:

$$\Phi''(\varphi) + k^2\Phi(\varphi) = 0, \quad (6)$$

$$r^2 R''(r) + rR'(r) - k^2 R(r) = 0. \quad (7)$$

Общее решение уравнения (6) будет

$$\Phi = A\cos k\varphi + B\sin k\varphi. \quad (8)$$

Решение уравнения (7) будем искать в форме $R(r) = r^m$. Подставляя $R(r) = r^m$ в уравнение (7), получим:

$$r^2 m(m-1)r^{m-2} + rmr^{m-1} - k^2 r^m = 0$$

или

$$m^2 - k^2 = 0.$$

Итак, имеются два частных линейно независимых решения уравнения (7): r^k и r^{-k} . Следовательно (§ 13.4, п. 1, теорема 1), общее решение уравнения (7) будет

$$R(r) = Cr^k + Dr^{-k}. \quad (9)$$

Выражения (8) и (9) подставляем в равенство (4):

$$u_k = (A_k \cos k\varphi + B_k \sin k\varphi)(C_k r^k + D_k r^{-k}). \quad (10)$$

Функция (10) будет решением уравнения (3) при любом значении k , отличном от нуля. Если $k = 0$, то уравнения (6) и (7) принимают вид:

$$\Phi''(\varphi) = 0, \quad rR''(r) + R'(r) = 0.$$

Их общие решения соответственно есть

$$\begin{aligned} \Phi &= A_0 + B_0\varphi, \\ R(r) &= C_0 + D_0 \ln r, \end{aligned}$$

и, следовательно,

$$u_0 = (A_0 + B_0\varphi)(C_0 + D_0 \ln r). \quad (11)$$

Это решение уравнения (3) должно быть периодической функцией от φ , так как при одном и том же значении r при φ и $\varphi + 2\pi$ мы должны иметь одно и то же значение решения, потому что рассматривается одна и та же точка круга. Поэтому в формуле (11) должно $B_0 = 0$. Далее мы ищем решение, непрерывное в круге. Следовательно, в равенствах (10) и (11) должно соответственно быть $D_k = 0$ и $D_0 = 0$.

Таким образом, правая часть равенства (11) обращается в произведение $A_0 C_0$, которое обозначим через $\frac{A_0}{2}$, т. е. $u_0 = \frac{A_0}{2}$.

Будем составлять решение нашей задачи в виде суммы решений вида (10), так как сумма решений есть решение. Сумма должна быть периодической функцией от φ . Это будет так, если каждое слагаемое будет периодической функцией от φ . Для этого k должно принимать целые значения. (Заметим, что если бы мы приравнивали части равенства (5) числу $+k^2$, то не получили бы периодического решения, так как тогда в выражение Φ не вошли бы $\cos k\varphi$ и $\sin k\varphi$; $R(r)$ не влияет на периодичность, ибо от φ не зависит.) Мы можем ограничиться только положительными значениями $k = 1, 2, \dots$, так как в силу произвольности постоянных A, B , С отрицательные значения k новых частных решений не дают. Итак,

$$u(r, \varphi) = \frac{A_0}{2} + \sum_{n=1}^{\infty} (A_n \cos n\varphi + B_n \sin n\varphi) r^n \quad (12)$$

(постоянная C_n включена в A_n и B_n). Подберем теперь произвольные постоянные A_n и B_n так, чтобы удовлетворялось краевое условие (2). Подставляя в равенство (12) $r = R$, получаем:

$$f(\varphi) = \frac{A_0}{2} + \sum_{n=1}^{\infty} (A_n \cos n\varphi + B_n \sin n\varphi) R^n.$$

Это — разложение функции $f(\varphi)$ в ряд Фурье. Следовательно (§ 10.4, п. 2),

$$A_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) dt, \quad A_n = \frac{1}{\pi R^n} \int_{-\pi}^{\pi} f(t) \cos ntdt, \quad B_n = \frac{1}{\pi R^n} \int_{-\pi}^{\pi} f(t) \sin ntdt. \quad (13)$$

Таким образом, ряд (12) с коэффициентами, определенными по формулам (13), будет решением нашей задачи. Преобразуем формулу (12). Подставляя вместо A_0, A_n и B_n их выражения по формулам (13) и производя тригонометрические преобразования, получим:

$$u(r, \varphi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left[1 + 2 \sum_{n=1}^{\infty} \left(\frac{r}{R} \right)^n \cos n(t - \varphi) \right] dt.$$

Но

$$1 + 2 \sum_{n=1}^{\infty} \left(\frac{r}{R} \right)^n \cos n(t - \varphi) = 1 + \sum_{n=1}^{\infty} \left(\frac{r}{R} \right)^n [e^{in(t-\varphi)} + e^{-in(t-\varphi)}] =$$

$$= \frac{1 - \left(\frac{r}{R} \right)^2}{1 - 2 \frac{r}{R} \cos(t - \varphi) + \left(\frac{r}{R} \right)^2} = \frac{R^2 - r^2}{R^2 - 2Rr \cos(t - \varphi) + r^2}.$$

Поэтому

$$u(r, \varphi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{R^2 - r^2}{R^2 - 2Rr \cos(t - \varphi) + r^2} dt. \quad (14)$$

Путем анализа этой формулы доказывается, что если функция $f(\varphi)$ непрерывная, то функция $u(r, \varphi)$, определенная интегралом (14), удовлетворяет уравнению (3) и при $r \rightarrow R - 0$ будет $u(r, \varphi) \rightarrow f(\varphi)$, т. е. $u(r, \varphi)$ является решением поставленной задачи Дирихле для круга.

Интеграл (14), дающий решение задачи Дирихле для круга, называется *интегралом Пуассона*, а подынтегральное выражение

$$\Phi(r, t, R, \varphi) = \frac{R^2 - r^2}{R^2 + r^2 - 2Rr \cos(t - \varphi)}$$

называется *ядром Пуассона*.

Отметим следующие свойства ядра Пуассона:

1. $\Phi(r, t, R, \varphi) > 0$.
2. $\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{R^2 - r^2}{R^2 - 2Rr \cos(t - \varphi) + r^2} dt = 1$.

Глава 18. ТЕОРИЯ ВЕРОЯТНОСТЕЙ

§ 18.1. Основные понятия. Определение вероятности

1. Понятие о случайном событии. Опыт, эксперимент, наблюдение явления называются *испытанием*. Испытаниями, например, являются: бросание монеты, выстрел из винтовки, бросание игральной кости (кубика с нанесенным на каждую грань числом очков — от одного до шести).

Результат (исход) испытания называется *событием*. Событиями являются: выпадение герба или выпадение цифры, попадание в цель или промах, появление того или иного числа очков на брошенной игральной кости.

Для обозначения событий используются большие буквы латинского алфавита: A , B , C и т. д.

О п р е д е л е н и е 1. Два события называются *совместимыми*, если появление одного из них не исключает появления другого в одном и том же испытании.

П р и м е р 1. Испытание: однократное бросание игральной кости. Событие A — появление четырех очков, событие B — появление четного числа очков. События A и B совместимы.

О п р е д е л е н и е 2. Два события называются *несовместимыми*, если появление одного из них исключает появление другого в одном и том же испытании.

П р и м е р 2. Испытание: однократное бросание монеты. Событие A — выпадение герба, событие B — выпадение цифры. Эти события несовместимы, так как появление одного из них исключает появление другого.

Несовместимость более чем двух событий означает их попарную несовместимость.

П р и м е р 3. Испытание: однократное бросание игральной кости. Пусть события $A_1, A_2, A_3, A_4, A_5, A_6$ — соответственно выпадение одного очка, двух, трех и т. д. Эти события являются несовместимыми.

О п р е д е л е н и е 3. Два события A и B называются *противоположными*, если в данном испытании они несовместимы и одно из них обязательно происходит.

Событие, противоположное событию A , обозначают через $\bar{A}$.

П р и м е р 4. Испытание однократное: бросание монеты. Событие A — выпадение герба, событие B — выпадение цифры. Эти события противоположны, так как исходами бросания могут быть лишь они и появление одного из них исключает появление другого, т. е. $A = \bar{B}$ или $\bar{A} = B$.

О п р е д е л е н и е 4. Событие называется *достоверным*, если в данном испытании оно является единственно возможным его исходом, и *невозможным*, если в данном испытании оно заведомо не может произойти.

П р и м е р 5. Испытание: извлечение шара из урны, в которой все шары белые. Событие A — вынут белый шар — достоверное событие; событие B — вынут черный шар — невозможное событие.

Заметим, что достоверное и невозможное события в данном испытании являются противоположными.

О п р е д е л е н и е 5. Событие A называется *случайным*, если оно объективно может наступить или не наступить в данном испытании.

П р и м е р 6. Событие A_6 — выпадение шести очков при бросании игральной кости — случайное. Оно может наступить, но оно может и не наступить в данном испытании.

2. Алгебра событий.

О п р е д е л е н и е 1. *Суммой событий A и B* называется событие $C = A + B$, состоящее в наступлении по крайней мере одного из событий A или B .

П р и м е р. Испытание: стрельба двух стрелков (каждый делает по выстрелу). Событие A — попадание в мишень первым стрелком, событие B — попадание в мишень вторым стрелком. Суммой событий A и B будет событие $C = A + B$, состоящее в попадании в мишень по крайней мере одним стрелком.

Аналогично суммой конечного числа событий $A_1, A_2, \dots, A_k$ называется событие $A = A_1 + A_2 + \dots + A_k$, состоящее в наступлении хотя бы одного из событий A_i ($i = 1, 2, \dots, k$).

Из определения 1 непосредственно следует, что $A + B = B + A$. Справедливо также и сочетательное свойство. Однако $A + A = A$ (а не $2A$, как в алгебре).

О п р е д е л е н и е 2. *Произведением событий A и B* называется событие $C = AB$, состоящее в том, что в результате испытания произошли и событие A , и событие B .

Аналогично произведением конечного числа событий $A_1, A_2, \dots, A_k$ называется событие $A = A_1 A_2 \dots A_k$, состоящее в том, что в результате испытания произошли все указанные события.

В условиях предыдущего примера произведением событий A и B будет событие $C = AB$, состоящее в попадании в мишень двух стрелков.

Из определения 2 непосредственно следует, что $AB = BA$. Справедливы также сочетательный и дистрибутивный законы. Однако $AA = A$ (а не A^2).

3. Классическое определение вероятности. Можно ли как-то измерить возможность появления некоторого случайного события? Другими словами, можно ли охарактеризовать эту возможность некоторым числом?

Всякое испытание влечет за собой некоторую совокупность исходов — результатов испытания, т. е. событий. Во многих случаях возможно перечислить все события, которые могут быть исходами данного испытания.

О п р е д е л е н и е 1. Говорят, что совокупность событий образует *полную группу событий* для данного испытания, если его результатом обязательно становится хотя бы одно из них.

Приведем примеры полных групп событий: выпадение герба и выпадение цифры при одном бросании монеты; попадание в цель и промах при одном выстреле; выпадение одного, двух, трех, четырех, пяти, шести очков при одном бросании игральной кости.

Рассмотрим полную группу попарно несовместимых событий $U_1, U_2, \dots, U_n$, связанную с некоторым испытанием. Предположим, что в этом испытании осуществление каждого из событий $U_1, U_2, \dots, U_n$ равновозможно, т. е. условия испытания не создают преимуществ в появлении какого-либо события перед другими возможными.

О п р е д е л е н и е 2. События $U_1, U_2, \dots, U_n$, образующие полную группу попарно несовместимых и равновозможных событий, будем называть *элементарными событиями*.

П р и м е р 1. Вернемся к опыту с подбрасыванием игральной кости. Пусть U_i — событие, состоящее в том, что кость выпала гранью с цифрой i . Как уже отмечалось, события $U_1, U_2, \dots, U_6$ образуют полную группу попарно несовместимых событий. Так как кость предполагается однородной и симметричной, то события $U_1, U_2, \dots, U_6$ являются и равновозможными, т. е. являются элементарными.

О п р е д е л е н и е 3. Событие A называется *благоприятствующим* событию B , если наступление события A влечет за собой наступление события B .

П р и м е р 2. Пусть при бросании игральной кости события U_2, U_4, U_6 — появление соответственно двух, четырех, шести очков и A — событие, состоящее в появлении четного очка; события U_2, U_4, U_6 благоприятствуют событию A .

О п р е д е л е н и е 4 (классическое определение вероятности). *Вероятностью $P(A)$ события A называется отношение $\frac{m}{n}$ числа элементарных событий, благоприятствующих событию A , к числу всех элементарных событий, т. е. $P(A) = \frac{m}{n}$.*

П р и м е р 3. Вычислим вероятность выпадения герба при одном бросании монеты. Очевидно, событие A — выпадение герба и событие B — выпадение цифры — образуют полную группу несовместимых и равновозможных событий для данного испытания. Значит, здесь $n = 2$. Событию A благоприятствует лишь одно событие — само A , т. е. здесь $m = 1$. Поэтому $P(A) = \frac{1}{2}$.

П р и м е р 4. Очевидно, что в опыте с игральной костью (пример 1) $P(U_i) = \frac{1}{6}, i = 1, \dots, 6$.

Пример 5. Найти вероятность того, что при бросании игральной кости выпадет число очков, делящееся на 2 (событие A).

Число элементарных событий здесь 6. Число благоприятствующих элементарных событий 3 (выпадение 2, 4 или 6). Поэтому $P(A) = \frac{3}{6} = \frac{1}{2}$.

Из приведенного классического определения вероятности вытекают следующие ее свойства:

1. *Вероятность достоверного события равна единице.*

Действительно, достоверному событию должны благоприятствовать все n элементарных событий, т. е. $m = n$, и, следовательно,

$$P(A) = \frac{m}{n} = \frac{n}{n} = 1.$$

2. *Вероятность невозможного события равна нулю.*

В самом деле, невозможному событию не может благоприятствовать ни одно из элементарных событий, т. е. $m = 0$, откуда

$$P(A) = \frac{m}{n} = \frac{0}{n} = 0.$$

3. *Вероятность случайного события есть положительное число, заключенное между нулем и единицей.*

Действительно, случайному событию благоприятствует лишь часть из общего числа n элементарных событий. Поэтому в этом случае $0 < m < n$ и, значит, $0 < \frac{m}{n} < 1$. Следовательно, $0 < P(A) < 1$.

Итак, вероятность любого события удовлетворяет двойному неравенству $0 \leq P(A) \leq 1$.

4. Относительная частота. Статистическое определение вероятности.

Классическое определение вероятности не является пригодным для изучения произвольных случайных событий. Так, оно неприемлемо, если результаты испытания не равновозможны. Например, при бросании неправильной игральной кости выпадение ее различных граней не равновозможно.

В таких случаях используется так называемое статистическое определение вероятности.

Пусть произведено n испытаний, при этом некоторое событие A наступило m раз.

О п р е д е л е н и е 1. Число m называется *абсолютной частотой* (или просто *частотой*) события A , а отношение $P^*(A) = \frac{m}{n}$ называется *относительной частотой* события A .

Пример 1. При транспортировке из 10 000 арбузов испортилось 26. Здесь $m=26$ — абсолютная частота испорченных арбузов, а $P^*(A) = \frac{26}{10000} = 0,0026$ — относительная.

Результаты многочисленных опытов и наблюдений помогают заключить: при проведении серий из n испытаний, когда число n сравнительно мало, относительная частота $P^*(A)$ принимает значения, которые могут довольно сильно отличаться друг от друга. Но с увеличением n — числа испытаний в сериях — относительная частота

$$P^*(A) = \frac{m}{n}$$

приближается к некоторому числу $P(A)$, стабилизируясь возле него и принимая все более устойчивые значения.

Пример 2. Было проведено 10 серий бросаний монеты, по 1000 бросаний в каждой. Относительные частоты выпадения герба оказались равными 0,501; 0,485; 0,509; 0,536; 0,485; 0,488; 0,500; 0,497; 0,494; 0,484. (Данные взяты из книги [18].)

Эти частоты группируются около числа 0,5.

Определение 2 (статистическое определение вероятности). Вероятностью события A в данном испытании называется число $P(A)$, около которого группируются значения относительной частоты при больших n .

В примере 2 вероятность в статистическом смысле равна 0,5.

Таким образом, относительная частота события приближенно совпадает с его вероятностью в статистическом смысле, если число испытаний достаточно велико (имеется огромный опытный материал по проверке последнего утверждения).

С этой точки зрения величина $m = nP(A)$ представляет собой среднее значение числа появления события A при n испытаниях.

Статистическое определение вероятности, использующее статистическую обработку данных, находит широкое применение.

При широких предположениях доказывается, что вероятности события в классическом и статистическом смыслах совпадают между собой.

§ 18.2. Свойства вероятности

1. Теорема сложения вероятностей несовместимых событий.

Теорема. Вероятность суммы двух несовместимых событий A и B равна сумме вероятностей этих событий:

$$P(A + B) = P(A) + P(B). \quad (1)$$

Доказательство. Используем классическое определение вероятности. Предположим, что в данном испытании число всех элементарных событий равно n , событию A благоприятствуют k элементарных событий, событию B — l элементарных событий. Так как A и B — несовместимые события, то ни одно из элементарных событий $U_1, U_2, \dots, U_n$ не может одновременно благоприятствовать и событию A , и событию B . Следовательно, событию $A + B$ будет благоприятствовать $k + l$ элементарных событий. По определению вероятности имеем:

$$P(A) = \frac{k}{n}, \quad P(B) = \frac{l}{n}, \quad P(A + B) = \frac{k + l}{n},$$

откуда и следует утверждение теоремы.

Совершенно так же теорема формулируется и доказывается для любого конечного числа попарно несовместимых событий.

С л е д с т в и е. Сумма вероятностей противоположных событий A и $\bar{A}$ равна единице:

$$P(A) + P(\bar{A}) = 1. \quad (2)$$

Так как события A и $\bar{A}$ несовместимы, то по доказанной теореме $P(A) + P(\bar{A}) = P(A + \bar{A})$. Событие $A + \bar{A}$ есть достоверное событие (ибо одно из событий A или $\bar{A}$ произойдет). Поэтому $P(A + \bar{A}) = 1$, что приводит к искомому соотношению (2).

П р и м е р. В урне 10 шаров: 3 красных, 5 синих, 2 белых. Какова вероятность вынуть цветной шар, если вынимается один шар? Вероятность вынуть красный шар $P(A) = \frac{3}{10}$, синий $P(B) = \frac{5}{10}$. Так как события A и B несовместимы, то по доказанной выше теореме $P(A + B) = P(A) + P(B) = \frac{3}{10} + \frac{5}{10} = 0,8$.

2. Теорема умножения вероятностей.

О п р е д е л е н и е 1. Два события A и B называются *независимыми*, если вероятность появления каждого из них не зависит от того, появилось другое событие или нет. В противном случае события A и B называют *зависимыми*. (Несколько событий $A_1, \dots, A_k$ называются *независимыми в совокупности*, если вероятность появления любого из них не зависит от того, произошли ли какие-либо другие рассматриваемые события или нет.)

П р и м е р 1. Пусть в урне находятся 2 белых и 2 черных шара. Пусть событие A — вынут белый шар. Очевидно, $P(A) = \frac{1}{2}$. После первого испытания вынутый шар кладется обратно в урну, шары перемешиваются и снова вынимается шар. Событие B — во втором

испытании вынут белый шар — также имеет вероятность $P(B) = \frac{1}{2}$, т. е. события A и B независимые.

Предположим теперь, что вынутый шар в первом испытании не кладется обратно в урну. Тогда если произошло A , т. е. в первом испытании вынут белый шар, то вероятность события B уменьшается $\left(P(B) = \frac{1}{3}\right)$; если в первом испытании был вынут черный шар, то вероятность события B увеличивается $\left(P(B) = \frac{2}{3}\right)$. Здесь вероятность события B зависит от появления или не появления события A , т. е. события A и B зависимые.

О п р е д е л е н и е 2. Пусть A и B — зависимые события. *Условной вероятностью $P_A(B)$ события B называется вероятность события B , найденная в предположении, что событие A уже наступило.* Так, в примере 1 $P_A(B) = \frac{1}{3}$.

Заметим, что если события A и B независимые, то

$$P_A(B) = P(B).$$

Т е о р е м а 1. *Вероятность произведения двух зависимых событий A и B равна произведению вероятности одного из них на условную вероятность другого, найденную в предположении, что первое событие уже наступило:*

$$P(AB) = P(A)P_A(B). \quad (3)$$

Д о к а з а т е л ь с т в о. Пусть из всего числа n элементарных событий k благоприятствуют событию A , и пусть из этих k событий l благоприятствуют событию B , а значит, и событию AB . Тогда

$$P(AB) = \frac{l}{n} = \frac{k}{n} \cdot \frac{l}{k} = P(A)P_A(B),$$

что и доказывает искомое равенство (3).

З а м е ч а н и е. Применив формулу (3) к событию BA , получим:

$$P(BA) = P(B)P_B(A).$$

Так как $AB = BA$ (см. § 18.1, п. 2), то получаем, что

$$P(A)P_A(B) = P(B)P_B(A). \quad (4)$$

П р и м е р 2. В условиях примера 1 берем тот случай, когда вынутый шар в первом испытании не кладется обратно в урну. Поставим следующий вопрос: какова вероятность вынуть первый и второй раз белые шары? По формуле (3) имеем:

$$P(AB) = \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}.$$

Т е о р е м а 2. Вероятность произведения двух независимых событий A и B равно произведению вероятностей этих событий:

$$P(AB) = P(A)P(B). \quad (5)$$

Действительно, если A и B — независимые события, то $P_A(B) = P(B)$ и формула (3) превращается в формулу (5).

Отметим, что в случае независимости событий эта теорема распространяется на любое конечное их число.

П р и м е р 3. Найти вероятность одновременного поражения цели двумя орудиями при одном залпе (из обоих орудий), если вероятность поражения цели первым орудием (событие A) равна 0,8, а вторым (событие B) — 0,7.

События A и B независимы, поэтому по теореме 2 искомая вероятность

$$P(AB) = 0,7 \cdot 0,8 = 0,56.$$

3. Теорема сложения вероятностей совместимых событий.

Т е о р е м а. Вероятность суммы двух совместимых событий A и B равна сумме вероятностей этих событий без вероятности их произведения:

$$P(A + B) = P(A) + P(B) - P(AB). \quad (6)$$

Д о к а з а т е л ь с т в о. Пусть из всего числа n элементарных событий k благоприятствуют событию A , l — событию B и m — одновременно событиям A и B . Отсюда событию $A + B$ благоприятствуют $k + l - m$ элементарных событий. Тогда

$$P(A + B) = \frac{k + l - m}{n} = \frac{k}{n} + \frac{l}{n} - \frac{m}{n} = P(A) + P(B) - P(AB).$$

З а м е ч а н и е. Если события A и B несовместимы, то их произведение AB есть невозможное событие, и, следовательно, $P(AB) = 0$, т. е. формула (1) является частным случаем формулы (6).

П р и м е р. Вероятности попадания в цель при стрельбе первого и второго орудий соответственно равны $P(A) = 0,7$ и $P(B) = 0,8$. Найти вероятность попадания при одном залпе (из обоих орудий) хотя бы одним из орудий.

Очевидно, события A и B совместимы и независимы. Поэтому

$$P(A + B) = P(A) + P(B) - P(AB) = 0,7 + 0,8 - 0,7 \cdot 0,8 = 0,94.$$

4. Формула полной вероятности. Пусть событие A может произойти лишь при условии появления одного из n попарно несовместимых событий $B_1, B_2, \dots, B_n$, образующих полную группу событий. События $B_1, B_2, \dots, B_n$ будем называть *гипотезами* для события A . Тогда

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A). \quad (7)$$

Эта формула полной вероятности.

В самом деле, событие A может наступить только при условии наступления одного из событий $B_1, B_2, \dots, B_n$, т. е.

$$A = B_1 A + B_2 A + \dots + B_n A,$$

причем ввиду несовместимости событий $B_1, B_2, \dots, B_n$ события $B_1 A, B_2 A, \dots, B_n A$ также несовместимы. Поэтому на основании теорем сложения и умножения вероятностей имеем:

$$\begin{aligned} P(A) &= P(B_1 A) + P(B_2 A) + \dots + P(B_n A) = \\ &= P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + \dots + P(B_n)P_{B_n}(A). \end{aligned}$$

Пример. Для приема зачета преподаватель заготовил 50 задач: 20 задач по дифференциальному исчислению, 30 по интегральному исчислению. Для сдачи зачета студент должен решить первую же доставшуюся наугад задачу. Какова вероятность для студента сдать зачет, если он умеет решить 18 задач по дифференциальному исчислению и 15 задач по интегральному исчислению?

Вероятность получить задачу по дифференциальному исчислению (событие B_1) равна $P(B_1) = 0,4$, по интегральному исчислению (событие B_2) равна $P(B_2) = 0,6$. Если событие A означает, что задача решена, то $P_{B_1}(A) = 0,9$, $P_{B_2}(A) = 0,5$. Теперь по формуле (7) имеем: $P(A) = 0,4 \cdot 0,9 + 0,6 \cdot 0,5 = 0,66$.

5. Формула Байеса. Пусть в условиях рассуждения, относящегося к формуле полной вероятности, произведено одно испытание, в результате которого произошло событие A . Спрашивается: как изменились (в связи с тем, что событие A уже произошло) вероятности гипотез, т. е. величины $P(B_k)$, $k = 1, \dots, n$?

Найдем условную вероятность $P_A(B_k)$.

По теореме умножения вероятностей и формуле (4) (см. п. 2) имеем:

$$P(AB_k) = P(A)P_A(B_k) = P(B_k)P_{B_k}(A).$$

Отсюда

$$P_A(B_k) = \frac{P(B_k)P_{B_k}(A)}{P(A)}.$$

Наконец, используя формулу полной вероятности, находим:

$$P_A(B_k) = \frac{P(B_k)P_{B_k}(A)}{\sum_{j=1}^n P(B_j)P_{B_j}(A)} \quad (k = 1, 2, \dots, n). \quad (8)$$

Формулы (8) называются *формулами Байеса* (Томас Бейес, или Байес (1702—1761), — английский математик).

Пример. Партия деталей изготовлена тремя рабочими, причем первый рабочий изготовил 25% всех деталей, второй — 35%, третий — 40%. В продукции первого рабочего

брак составляет 5%, в продукции второго — 4% и в продукции третьего — 2%. Случайно выбранная для контроля деталь оказалась бракованной. Какова вероятность того, что она изготовлена вторым рабочим?

Введем обозначения для событий: A — выбранная для контроля деталь оказалась бракованной; B_1, B_2, B_3 — эта деталь изготовлена соответственно первым, вторым и третьим рабочим.

Имеем:

$$P(B_1) = 0,25; P(B_2) = 0,35; P(B_3) = 0,40; P_{B_1}(A) = 0,05; P_{B_2}(A) = 0,04; P_{B_3}(A) = 0,02.$$

$$P_A(B_2) = \frac{0,35 \cdot 0,04}{0,25 \cdot 0,05 + 0,35 \cdot 0,04 + 0,40 \cdot 0,02} = \frac{28}{69}.$$

§ 18.3. Основные формулы комбинаторики

Комбинаторикой называется раздел математики, изучающий вопрос о том, сколько комбинаций определенного типа можно составить из данных предметов (элементов).

Как при решении задач с использованием классического определения вероятности, так и в дальнейшем нам понадобятся некоторые формулы комбинаторики. Приведем наиболее употребительные из них.

О п р е д е л е н и е 1. *Размещениями* из n различных элементов по m элементов ($m \leq n$) называются комбинации, составленные из данных n элементов по m элементов, которые отличаются либо самими элементами, либо порядком элементов.

Например, из трех элементов a, b, c можно составить по два элемента следующие размещения: ab, ac, bc, ba, ca, cb .

Число различных размещений из n элементов по m элементов определяется с помощью формулы $A_n^m = n(n-1)(n-2) \dots (n-m+1)$.

О п р е д е л е н и е 2. *Перестановками* из n различных элементов называются размещения из этих n элементов по n .

Как видно из определений 1 и 2, перестановки можно считать частным случаем размещений при $m = n$. Следовательно, число всех перестановок из n элементов вычисляется по формуле

$$P_n = n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1 = n!$$

О п р е д е л е н и е 3. *Сочетаниями* из n различных элементов по m элементов называются комбинации, составленные из данных n элементов по m элементов, которые отличаются хотя бы одним элементом.

Отметим разницу между сочетаниями и размещениями: в первых не учитывается порядок элементов.

Число сочетаний из n элементов по m элементов вычисляется по формуле

$$C_n^m = \frac{n!}{m!(n-m)!}.$$

Приведем, наконец, один из примеров применения формул комбинаторики к нахождению вероятности события.

П р и м е р. Набирая номер телефона, абонент забыл две последние цифры и, помня лишь, что эти цифры различны, набрал их наудачу. Какова вероятность того, что номер набран правильно?

Две последние цифры можно набрать A_{10}^2 способами, а благоприятствовать событию M (цифры набраны правильно) будет только один способ. Поэтому

$$P(M) = \frac{1}{A_{10}^2} = \frac{1}{10 \cdot 9} = \frac{1}{90}.$$

§ 18.4. Дискретные случайные величины

1. Понятие случайные величины.

О п р е д е л е н и е 1. *Случайной величиной* называется переменная величина, которая в зависимости от исхода испытания случайно принимает одно значение из множества возможных значений.

П р и м е р ы. 1) Число очков, выпавших при однократном бросании игральной кости, есть случайная величина, она может принять одно из значений: 1, 2, 3, 4, 5, 6.

2) Прирост веса домашнего животного за месяц есть случайная величина, которая может принять значение из некоторого числового промежутка.

3) Число родившихся мальчиков среди пяти новорожденных есть случайная величина, которая может принять значения 0, 1, 2, 3, 4, 5.

4) Расстояние между эпицентром взрыва бомбы и целью, на которую она сброшена, есть случайная величина, которая может принять любое неотрицательное значение.

Случайные величины будем обозначать прописными буквами X, Y, Z , а их возможные значения — соответствующими строчными буквами x, y, z . Например, если случайная величина X имеет три возможных значения, то они будут обозначены так: x_1, x_2, x_3 .

О п р е д е л е н и е 2. Случайная величина, принимающая различные значения, которые можно записать в виде конечной или бесконечной последовательности, называется *дискретной случайной величиной*.

Случайные величины из примеров 1) и 3) дискретные.

О п р е д е л е н и е 3. Случайная величина, которая может принимать все значения из некоторого числового промежутка, называется *непрерывной случайной величиной*.

Случайные величины из примеров 2) и 4) являются непрерывными.

О п р е д е л е н и е 4. Под *суммой (произведением)* случайных величин X и Y понимают случайную величину $Z = X + Y$ ($Z = XY$), возможные значения которой состоят из сумм (произведения) каждого возможного значения величины X и каждого возможного значения величины Y .

2. Законы распределения дискретных случайных величин. Рассмотрим дискретную случайную величину X с конечным множеством возможных значений. Величина X считается заданной, если перечислены все ее возможные значения, а также вероятности, с которыми величина X может принять эти значения. Указанный перечень возможных значений и их вероятностей называется *законом распределения* дискретной случайной величины. Закон распределения дискретной случайной величины может быть задан с помощью таблицы:

x	x_1	x_2	x_3	$\dots$	x_{n-1}	x_n
p	p_1	p_2	p_3	$\dots$	p_{n-1}	p_n

В верхней строке выписываются все возможные значения $x_1, x_2, \dots, x_n$ величины X , в нижней строке выписываются вероятности $p_1, p_2, \dots, p_n$ значений $x_1, x_2, \dots, x_n$. Читается таблица следующим образом: случайная величина X может принять значение x_i с вероятностью p_i ($i = 1, 2, \dots, n$).

Так как в результате испытания величина X всегда примет одно из значений $x_1, x_2, \dots, x_n$, то $p_1 + p_2 + \dots + p_n = 1$.

П р и м е р. В денежной лотерее разыгрывается 1 выигрыш в 100 000 р., 10 выигрышей по 10 000 р. и 100 выигрышей по 100 р. при общем числе билетов 10 000. Найти закон распределения случайного выигрыша X для владельца одного лотерейного билета.

Здесь возможные значения для X есть: $x_1 = 0, x_2 = 100, x_3 = 10\,000, x_4 = 100\,000$. Вероятность их будет: $p_2 = 0,01, p_3 = 0,001, p_4 = 0,0001, p_1 = 1 - 0,01 - 0,001 - 0,0001 = 0,9889$.

Следовательно, закон распределения выигрыша X может быть задан таблицей:

x	0	100	10 000	100 000
p	0,9889	0,01	0,001	0,0001

§ 18.5. Математическое ожидание дискретной случайной величины

1. Понятие математического ожидания. Закон распределения полностью задает дискретную случайную величину. Однако часто встречаются случаи, когда закон распределения случайной величины неизвестен. В таких случаях случайную величину изучают по ее числовым характеристикам. Одной из таких характеристик является математическое ожидание.

Пусть некоторая дискретная случайная величина X с конечным числом своих значений задана законом распределения с помощью таблицы, приведенной в п. 2 § 18.4.

О п р е д е л е н и е. *Математическим ожиданием $M(X)$ дискретной случайной величины X называется сумма произведений всех возможных значений величины X на соответствующие вероятности:*

$$M(X) = x_1 p_1 + x_2 p_2 + \dots + x_n p_n. \quad (1)$$

П р и м е р. Найти математическое ожидание выигрыша X в примере из § 18.4 (п. 2). Используя полученную там таблицу, имеем:

$$M(X) = 0 \cdot 0,9889 + 100 \cdot 0,01 + 10\,000 \cdot 0,001 + 100\,000 \cdot 0,0001 = 21 \text{ р.}$$

Очевидно, $M(X) = 21$ р. есть справедливая цена одного лотерейного билета.

Т е о р е м а. *Математическое ожидание дискретной случайной величины X приближенно равно среднему арифметическому всех ее значений (при достаточно большом числе испытаний).*

Д о к а з а т е л ь с т в о. Предположим, что произведено n испытаний, в которых дискретная случайная величина X приняла значения $x_1, \dots, x_k$ соответственно $m_1, \dots, m_k$ раз, так, что $m_1 + \dots + m_k = n$. Тогда среднее арифметическое всех значений, принятых величиной X , выразится равенством

$$x_{\text{cp}} = \frac{x_1 m_1 + x_2 m_2 + \dots + x_k m_k}{n}, \text{ или } x_{\text{cp}} = x_1 \frac{m_1}{n} + x_2 \frac{m_2}{n} + \dots + x_k \frac{m_k}{n}.$$

Так как коэффициент $\frac{m_i}{n}$ является относительной частотой события «величина X приняла значение x_i » ($i = 1, 2, \dots, k$), то

$$x_{\text{cp}} = x_1 p_1 + x_2 p_2 + \dots + x_k p_k.$$

Из статистического определения вероятности следует, что при достаточно большом числе испытаний $p_i^* \approx p_i$ ($i = 1, 2, \dots, k$). Поэтому

$$x_{\text{cp}} \approx x_1 p_1 + x_2 p_2 + \dots + x_k p_k \text{ или } x_{\text{cp}} \approx M(X).$$

П р и м е ч а н и е. В связи с только что установленной теоремой математическое ожидание случайной величины называют также ее *средним значением*.

2. Свойства математического ожидания дискретной случайной величины. 1. *Математическое ожидание постоянной величины C равно этой величине.*

Постоянную C можно рассматривать как дискретную случайную величину, принимающую лишь одно значение C с вероятностью $p = 1$. Поэтому $M(C) = C \cdot 1 = C$.

2. *Постоянный множитель можно выносить за знак математического ожидания, т. е. $M(CX) = CM(X)$.*

Используя соотношение (1), имеем:

$$\begin{aligned} M(CX) &= Cx_1p_1 + Cx_2p_2 + \dots + Cx_np_n = \\ &= C(x_1p_1 + x_2p_2 + \dots + x_np_n) = CM(X). \end{aligned}$$

В дальнейшем часто ради краткости вместо слов «математическое ожидание» будем писать *МО*.

Следующие два свойства (3—4) примем без доказательства (доказательства см. в [4]).

3. *МО суммы двух случайных величин X и Y равно сумме их МО:*

$$M(X + Y) = M(X) + M(Y).$$

О п р е д е л е н и е. Случайные величины X и Y называются *независимыми*, если закон распределения каждой из них не зависит от того, какое возможное значение приняла другая величина.

Примером двух независимых случайных величин могут служить суммы выигрышей по каждому из двух билетов по двум различным денежно-вещевым лотереям. Здесь ставший известным размер выигрыша по билету одной лотереи не влияет на ожидаемый размер выигрыша и соответствующую ему вероятность по билету другой лотереи. Несколько случайных величин называются независимыми, если закон распределения любой из них не зависит от того, какие возможные значения приняли остальные случайные величины.

4. *МО произведения двух независимых случайных величин равно произведению их математических ожиданий: $M(XY) = M(X)M(Y)$.*

Следствием свойств 2 и 3 является свойство 5.

5. *МО разности двух случайных величин X и Y равно разности их математических ожиданий: $M(X - Y) = M(X) - M(Y)$.*

П р и м е ч а н и е 1. Свойства 3 и 4 имеют место и для любого конечного числа случайных величин.

П р и м е ч а н и е 2. Если множество возможных значений дискретной случайной величины X бесконечно, то математическое ожидание $M(X)$ определяется суммой числового ряда $M(X) = \sum_{k=1}^{\infty} x_k p_k$ при условии, что этот ряд абсолютно сходится (в противном случае говорят, что математическое ожидание $M(X)$ не существует). Перечисленные свойства *МО* остаются в силе (см. [5] и для таких случайных величин).

П р и м е р. Найти математическое ожидание случайной величины $Z = X + 2Y$, если известны математические ожидания случайных величин X и Y : $M(X) = 5$, $M(Y) = 3$. Используя свойства 3 и 2 математического ожидания, получим:

$$M(Z) = M(X + 2Y) = M(X) + M(2Y) = M(X) + 2M(Y) = 5 + 2 \cdot 3 = 11.$$

§ 18.6. Дисперсия дискретной случайной величины

1. Понятие дисперсии. Математическое ожидание не дает полной характеристики закона распределения случайной величины. Покажем это на примере. Пусть заданы две дискретные случайные величины X и Y своими законами распределения:

X	-2	0	2
P	0,4	0,2	0,4

Y	-100	0	100
P	0,3	0,4	0,3

Несмотря на то что MO величин X и Y одинаковы, возможные значения величин X и Y «разбросаны» или «рассеяны» около своих MO по-разному: возможные значения величины X расположены гораздо ближе к своему MO , чем значения величины Y .

Из сказанного вытекает необходимость введения новой числовой характеристики случайной величины, по которой можно судить о «рассеянии» возможных значений этой случайной величины.

Пусть задана дискретная случайная величина X с помощью таблицы (см. § 18.4, п. 2).

О п р е д е л е н и е 1. *Отклонением случайной величины X от ее математического ожидания (или просто отклонением случайной величины X) называется случайная величина $X - M(X)$.*

Вычислим теперь MO отклонения $X - M(X)$. Пользуясь свойствами 5 и 1 (§ 18.5, п. 2), получим: $M[X - M(X)] = M(X) - M(X) = 0$.

Следовательно, справедлива следующая теорема.

Т е о р е м а. *Математическое ожидание отклонения $X - M(X)$ равно нулю: $M[X - M(X)] = 0$.*

Из теоремы видно, что с помощью отклонения $X - M(X)$ не удастся определить среднее отклонение возможных значений величины X от ее MO , т. е. степень рассеяния величины X . Однако можно освободиться от этого недостатка, если рассматривать квадрат отклонения случайной величины X .

Запишем закон распределения случайной величины $(X - M(X))^2$:

$(X - M(X))^2$	$(x_1 - M(X))^2$	$(x_2 - M(X))^2$	...	$(x_n - M(X))^2$
P	P_1	P_2	...	P_n

О п р е д е л е н и е 2. *Дисперсией $D(X)$ дискретной случайной величины X называется MO квадрата отклонения случайной величины X от ее MO : $D(X) = M[(X - M(X))^2]$.*

Из закона распределения величины $[X - M(X)]^2$ следует, что

$$D(X) = [x_1 - M(X)]^2 p_1 + [x_2 - M(X)]^2 p_2 + \dots + [x_n - M(X)]^2 p_n.$$

2. Свойства дисперсии дискретной случайной величины.

1. Дисперсия дискретной случайной величины X равна разности между $МО$ квадрата величины X и квадратом ее $МО$:

$$D(X) = M(X^2) - M^2(X).$$

Действительно, используя свойства $МО$, имеем:

$$\begin{aligned} D(X) &= M[(X - M(X))^2] = M[X^2 - 2XM(X) + M^2(X)] = \\ &= M(X^2) - 2M(X) \cdot M(X) + M^2(X) = \\ &= M(X^2) - 2M^2(X) + M^2(X) = M(X^2) - M^2(X). \end{aligned}$$

С помощью этого свойства и свойства $МО$ устанавливаются следующие свойства.

2. Дисперсия постоянной величины C равна нулю.

3. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат: $D(CX) = C^2 D(X)$.

4. Дисперсия суммы двух независимых случайных величин равна сумме дисперсий этих величин: $D(X + Y) = D(X) + D(Y)$.

Методом математической индукции это свойство распространяется и на случай любого конечного числа слагаемых.

Следствием свойств 3 и 4 является свойство 5.

5. Дисперсия разности двух независимых случайных величин X и Y равна сумме их дисперсий: $D(X - Y) = D(X) + D(Y)$.

П р и м е р. Дисперсия случайной величины X равна 3. Найти дисперсию следующих величин: а) $-3X$; б) $4X + 3$.

Согласно свойствам 2, 3 и 4 дисперсии имеем:

а) $D(-3X) = 9D(X) = 9 \cdot 3 = 27$;

б) $D(4X + 3) = D(4X) + D(3) = 16D(X) + 0 = 16 \cdot 3 = 48$.

3. Среднее квадратическое отклонение.

О п р е д е л е н и е. Средним квадратическим отклонением $\sigma(X)$ случайной величины X называется корень квадратный из ее дисперсии: $\sigma(X) = \sqrt{D(X)}$.

Введение среднего квадратического отклонения объясняется тем, что дисперсия измеряется в квадратных единицах относительно размерности самой случайной величины. Например, если возможные значения некоторой случайной величины измеряются в метрах, то ее дисперсия измеряется в квадратных метрах. В тех случаях, когда нужно

иметь числовую характеристику рассеяния возможных значений в той же размерности, что и сама случайная величина, и используется среднее квадратическое отклонение.

П р и м е р. Случайная величина X — число очков, выпавших при однократном бросании игральной кости. Определить $\sigma(X)$.

Имеем:

$$M(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3,5;$$

$$D(X) = (1 - 3,5)^2 \cdot \frac{1}{6} + (2 - 3,5)^2 \cdot \frac{1}{6} + (3 - 3,5)^2 \cdot \frac{1}{6} +$$

$$+ (4 - 3,5)^2 \cdot \frac{1}{6} + (5 - 3,5)^2 \cdot \frac{1}{6} + (6 - 3,5)^2 \cdot \frac{1}{6} = \frac{35}{12};$$

$$\sigma(X) = \sqrt{\frac{35}{12}} \approx 1,71.$$

4. Нормированные случайные величины.

О п р е д е л е н и е. Случайная величина Y называется *нормированной*, если ее математическое ожидание равно 0, а дисперсия 1:

$$M(Y) = 0, D(Y) = 1.$$

От любой случайной величины X можно перейти к нормированной случайной величине Y с помощью линейного преобразования:

$$Y = \frac{X - m}{\sigma},$$

где m — математическое ожидание величины X , а σ — ее среднее квадратическое отклонение.

В самом деле, в силу свойств математического ожидания (§ 18.5, п. 2) и дисперсии (§ 18.6, п. 2) имеем:

$$M(Y) = \frac{1}{\sigma} M(X - m) = \frac{1}{\sigma} (M(X) - m) = \frac{1}{\sigma} (m - m) = 0,$$

$$D(Y) = \frac{1}{\sigma^2} D(X - m) = \frac{1}{\sigma^2} (D(X) + D(m)) = \frac{1}{\sigma^2} \cdot \sigma^2 = 1.$$

§ 18.7. Основные законы распределения дискретных случайных величин

1. Биномиальное распределение. Пусть производится n испытаний, причем вероятность появления события A в каждом испытании равна p и не зависит от исхода других испытаний (независимые испытания). Такая последовательность испытаний называется *схемой Бернулли*. Так как вероятность наступления события A в одном испытании равна p , то вероятность его ненаступления равна $q = 1 - p$.

В условиях схемы Бернулли найдем вероятность того, что при n испытаниях событие A наступит m раз ($m \leq n$).

Предположим, что событие A наступило в первых m испытаниях m раз и не наступило во всех последующих испытаниях. Это сложное событие можно записать в виде произведения:

$$\underbrace{AA \dots A}_{m \text{ раз}} \quad \underbrace{\bar{A} \bar{A} \dots \bar{A}}_{n-m \text{ раз}}$$

Общее число сложных событий, в которых событие A наступает m раз, равно числу сочетаний из n элементов по m элементов. При этом вероятность каждого сложного события равна $p^m q^{n-m}$. Так как эти сложные события являются несовместимыми, то вероятность их суммы равна сумме их вероятностей (§ 18.2, п. 1). Итак, если $P_n(m)$ есть вероятность появления события A в n испытаниях m раз, то

$$P_n(m) = C_n^m p^m q^{n-m} \quad (1)$$

или

$$P_n(m) = \frac{n!}{m!(n-m)!} p^m q^{n-m}. \quad (1')$$

Формула (1) ((1')) называется *формулой Бернулли*.

П р и м е р. Пусть всхожесть семян данного растения составляет 90%. Найти вероятность того, что из четырех посеянных семян взойдут: а) три; б) не менее трех.

а) В данном случае $n=4$, $m=3$, $p=0,9$, $q=1-p=0,1$. Применяя формулу Бернулли (1'), получим: $P_4(3) = \frac{4!}{3!1!} (0,9)^3 \cdot 0,1 = 0,2916$.

б) Искомое событие A состоит в том, что из четырех семян взойдут или три, или четыре. По теореме сложения вероятностей $P(A) = P_4(3) + P_4(4)$. Но $P_4(4) = (0,9)^4 = 0,6561$. Поэтому $P(A) = 0,2916 + 0,6561 = 0,9477$.

Снова рассмотрим n независимых испытаний, в каждом из которых наступает событие A с вероятностью p . Обозначим через X случайную величину, равную числу появлений события A в n испытаниях.

Понятно, что событие A может вообще не наступить, наступить один раз, два раза и т. д. и, наконец, наступить n раз. Следовательно, возможными значениями величины X будут числа $0, 1, 2, \dots, n-1, n$. По формуле Бернулли (1) можно найти вероятности этих значений:

$$\begin{aligned} P_n(0) &= q^n, \\ P_n(1) &= C_n^1 q^{n-1} p, \\ &\dots \dots \dots \\ P_n(n) &= p^n. \end{aligned}$$

Запишем полученные данные в виде таблицы распределения:

X	0	1	...	m	...	n
P	q^n	$C_n^1 p q^{n-1}$	...	$C_n^m p^m q^{n-m}$	...	p^n

Построенный закон распределения дискретной случайной величины X называется *законом биномиального распределения*.

Найдем $M(X)$. Очевидно, что X_i — число появлений события A в каждом испытании — представляет собой случайную величину со следующим распределением:

X_i	0	1
P_i	q	p

Поэтому $M(X_i) = 0 \cdot q + 1 \cdot p = p$. Но так как $X = X_1 + X_2 + \dots + X_n$, то $M(X) = np$.

Найдем далее $D(X)$ и $\sigma(X)$. Так как величина X_i^2 имеем распределение

X_i^2	0^2	1^2
P_i	q	p

то $M(X_i^2) = 0^2 \cdot q + 1^2 \cdot p = p$. Поэтому

$$D(X_i) = M(X_i^2) - M^2(X_i) = p - p^2 = p(1 - p) = pq.$$

Наконец, в силу независимости величин $X_1, X_2, \dots, X_n$,

$$D(X) = D(X_1) + D(X_2) + \dots + D(X_n) = npq.$$

Отсюда

$$\sigma(X) = \sqrt{npq}.$$

2. Распределение Пуассона. Пусть производится серия n независимых испытаний ($n = 1, 2, 3, \dots$), причем вероятность появления данного события A в этой серии $P(A) = p_n > 0$ зависит от ее номера n и стремится к нулю при $n \rightarrow \infty$ (последовательность «редких событий»). Предположим, что для каждой серии среднее значение числа появлений события A постоянно, т. е. $np_n = \mu = \text{const}$. Отсюда $p_n = \frac{\mu}{n}$.

Исходя из формулы Бернулли (1), для вероятности появления события A в n -й серии ровно m раз имеем выражение

$$P_n(m) = C_n^m p_n^m (1 - p_n)^{n-m} = C_n^m \left(\frac{\mu}{n}\right)^m \left(1 - \frac{\mu}{n}\right)^{n-m}.$$

Пусть m фиксировано. Тогда

$$\begin{aligned} \lim_{n \rightarrow \infty} C_n^m \left(\frac{\mu}{n}\right)^m &= \lim_{n \rightarrow \infty} \frac{n(n-1)(n-2) \dots (n-(m-1))}{m! n^m} \mu^m = \\ &= \frac{\mu^m}{m!} \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{m-1}{n}\right) = \frac{\mu^m}{m!}; \\ \lim_{n \rightarrow \infty} \left(1 - \frac{\mu}{n}\right)^{n-m} &= e^{-\mu} \end{aligned}$$

(здесь использован второй замечательный предел; § 7.3, п. 2). Поэтому

$$\lim_{n \rightarrow \infty} P_n(m) = \frac{\mu^m}{m!} e^{-\mu}.$$

Если n велико, то в силу определения предела (§ 7.3, п.1) вероятность $P_n(m)$ сколь угодно мало отличается от $\frac{\mu^m}{m!} e^{-\mu}$. Отсюда при больших n для искомой вероятности $P_n(m)$ имеем приближенную формулу Пуассона (для простоты знак приближенного равенства опущен)

$$P_n(m) = \frac{\mu^m}{m!} e^{-\mu}, \text{ где } \mu = np_n.$$

П р и м е р. Завод отправил на базу 500 доброкачественных изделий. Вероятность того, что в пути изделие повредится, равна 0,002. Найти вероятность того, что на базу придут 3 негодных изделия.

По условию $n=500$, $p_n=0,002$, $m=3$. Поэтому $\mu = 500 \cdot 0,002 = 1$ и искомая вероятность

$$P_{500}(3) = \frac{1^3}{3!} e^{-1} \approx 0,06.$$

О п р е д е л е н и е. Говорят, что случайная величина X распределена по закону Пуассона, если эта величина задана таблицей:

X	0	1	2	3	...
P	$e^{-\mu}$	$\mu e^{-\mu}$	$\frac{\mu^2}{2!} e^{-\mu}$	$\frac{\mu^3}{3!} e^{-\mu}$	...

где μ — фиксированное положительное число (разным значениям μ отвечают разные распределения Пуассона).

Найдем математическое ожидание дискретной величины X , распределенной по закону Пуассона. Согласно определению математического ожидания (§ 18.5, п. 2, примечание 2) имеем:

$$M(X) = \sum_{k=0}^{\infty} k \frac{\mu^k}{k!} e^{-\mu} = \mu e^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!} = \mu e^{-\mu} e^{\mu} = \mu.$$

Найдем далее $D(X)$. Сначала найдем $M(X^2)$:

$$\begin{aligned} M(X^2) &= \sum_{k=0}^{\infty} k^2 \frac{\mu^k}{k!} e^{-\mu} = \mu e^{-\mu} \sum_{k=1}^{\infty} k \frac{\mu^{k-1}}{(k-1)!} = \\ &= \mu e^{-\mu} \left(\mu \sum_{k=2}^{\infty} \frac{\mu^{k-2}}{(k-2)!} + e^{\mu} \right) = \mu e^{-\mu} (\mu e^{\mu} + e^{\mu}) = \mu^2 + \mu. \end{aligned}$$

Теперь по известной формуле (§ 18.6, п. 2)

$$D(X) = M(X^2) - M^2(X) = \mu^2 + \mu - \mu^2 = \mu.$$

3. Локальная и интегральная предельные теоремы Лапласа. Если число испытаний n велико, то вычисления по формуле Бернулли становятся затруднительными. Лаплас получил важную приближенную формулу для вероятности $P_n(m)$ появления события A ровно m раз, если n достаточно большое число. Им же получена приближенная формула и для суммы вида $\sum_{m=k}^l P_n(m)$.

Локальная предельная теорема Лапласа (доказательство см. в [5]). Пусть $p = P(A)$ — вероятность события A , причем $0 < p < 1$. Тогда вероятность того, что в условиях схемы Бернулли событие A при n испытаниях появится ровно m раз, выражается приближенной формулой Лапласа

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \varphi\left(\frac{m - np}{\sqrt{npq}}\right), \quad (2)$$

где $q = 1 - p$, $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Для функции $\varphi(x)$ имеется таблица (см. приложение 3) ее значений для положительных значений x (функция $\varphi(x)$ четная).

Пример 1. Вероятность поражения цели стрелком при одиночном выстреле равна $p = 0,2$. Какова вероятность того, что при 100 выстрелах цель будет поражена ровно 20 раз?

Здесь $p = 0,2$, $q = 0,8$, $n = 100$ и $m = 20$. Отсюда $\sqrt{npq} = \sqrt{100 \cdot 0,2 \cdot 0,8} = 4$, и, следовательно, $t = \frac{m - np}{\sqrt{npq}} = \frac{20 - 100 \cdot 0,2}{4} = 0$. Учитывая, что $\varphi(0) = \frac{1}{\sqrt{2\pi}} \approx 0,40$, из формулы (2) по-

лучаем: $P_{100}(20) \approx 0,40 \cdot \frac{1}{4} = 0,10$ (для получения приближенного равенства $\frac{1}{\sqrt{2\pi}} \approx 0,40$ можно

использовать калькулятор).

Перейдем к интегральной предельной теореме Лапласа. Поставим следующий вопрос: какова вероятность того, что в условиях схемы Бернулли событие A , имеющее вероятность $P(A) = p$ ($0 < p < 1$), при n испытаниях (как и прежде, число испытаний велико) появится не менее k раз и не более l раз? Эту искомую вероятность обозначим через $P_n(k, l)$.

На основании теоремы сложения вероятностей для несовместимых событий (§ 18.2, п. 1) получим:

$$P_n(k, l) = \sum_{m=k}^l P_n(m). \quad (3)$$

Установим приближенную формулу Лапласа для подсчета суммы (3) при больших m и n . Используя локальную теорему Лапласа, с учетом (3) приближенно будем иметь:

$$P_n(k, l) \approx \sum_{m=k}^l \frac{1}{\sqrt{npq}} \varphi(x_m).$$

где

$$x_m = \frac{m - np}{\sqrt{npq}}, \quad m = k, k+1, \dots, l \quad (4)$$

и

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Далее в силу (4) имеем:

$$\Delta x_m = x_{m+1} - x_m = \frac{(m+1) - np}{\sqrt{npq}} - \frac{m - np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}}, \quad (5)$$

и потому

$$P_n(k, l) \approx \sum_{m=k}^l \varphi(x_m) \Delta x_m.$$

Здесь сумма справа является интегральной суммой для функции $\varphi(x)$ на отрезке $x_k \leq x \leq x_l$, причем, как следует из (5), при $n \rightarrow \infty$ $\Delta x_m \rightarrow 0$. Следовательно, при $n \rightarrow \infty$ предел указанной интегральной суммы есть определенный интеграл $\int_{x_k}^{x_l} \varphi(x) dx$.

Поэтому

$$P_n(k, l) \approx \int_{x_k}^{x_l} \varphi(x) dx = \frac{1}{\sqrt{2\pi}} \int_{x_k}^{x_l} e^{-\frac{x^2}{2}} dx, \quad (6)$$

где

$$x_k = \frac{k - np}{\sqrt{npq}}, \quad x_l = \frac{l - np}{\sqrt{npq}}. \quad (7)$$

Это составляет содержание *интегральной предельной теоремы Лапласа*.
Введем функцию

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt, \quad (8)$$

называемую *функцией Лапласа* или *интегралом вероятностей*. Очевидно, $\Phi(x)$ есть первообразная для функции $\phi(x)$. Так как $\phi(x) > 0$ в $(-\infty; +\infty)$, то $\Phi(x)$ — возрастающая функция в этом интервале (см. § 8.7, п. 1, теорема 2). На основании формулы Ньютона — Лейбница (§ 9.7, п. 2) из формулы (6) будем иметь:

$$P_n(k, l) \approx \Phi(x_l) - \Phi(x_k). \quad (9)$$

Это *интегральная формула Лапласа*.

Как известно (см. § 9.5, примечание), интеграл $\frac{1}{\sqrt{2\pi}} \int e^{-\frac{x^2}{2}} dx$ не берется в элементарных функциях. Поэтому для функции (8) составлена таблица (см. приложение 4) ее значений для положительных значений x , так как $\Phi(0) = 0$ и функция $\Phi(x)$ нечетная:

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int_0^{-x} e^{-\frac{t^2}{2}} dt = -\frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz = -\Phi(x) \quad (t = -z, dt = -dz).$$

Пример 2. Вероятность того, что изделие не прошло проверку ОТК, равно 0,2. Найти вероятность того, что среди 400 случайно отобранных изделий окажутся непроверенными от 70 до 100 изделий.

Здесь $n = 400$, $k = 70$, $l = 100$, $p = 0,2$, $q = 0,8$. Поэтому в силу равенств (7) $x_k = -1,25$, $x_l = 2,5$ и согласно формуле (9)

$$P_{400}(70, 100) \approx \Phi(2,5) - \Phi(-1,25) = \Phi(2,5) + \Phi(1,25) \approx 0,4938 + 0,3944 = 0,8882.$$

§ 18.8. Непрерывные случайные величины

1. Интегральная функция распределения. Для непрерывной случайной величины в отличие от дискретной нельзя построить таблицу распределения. Поэтому непрерывные случайные величины изучаются другим способом, который мы сейчас рассмотрим.

Пусть X — непрерывная случайная величина с возможными значениями из некоторого интервала $(a; b)$ и x — действительное число. Под выражением $X < x$ понимается событие «случайная величина X приняла

значение, меньшее x ». Вероятность этого события $P(X < x)$ есть некоторая функция переменной x :

$$F(x) = P(X < x).$$

О п р е д е л е н и е. *Интегральной функцией распределения* или *кратко функцией распределения* непрерывной случайной величины X называется функция $F(x)$, равная вероятности того, что X приняла значение, меньшее x :

$$F(x) = P(X < x). \quad (1)$$

Отметим, что интегральная функция распределения совершенно так же определяется и для дискретных случайных величин.

Укажем свойства, которыми обладает функция $F(x)$.

$$1. 0 \leq F(x) \leq 1.$$

Это свойство следует из того, что $F(x)$ есть вероятность.

2. $F(x)$ — неубывающая функция, т. е. если $x_1 < x_2$, то

$$F(x_1) \leq F(x_2).$$

Д о к а з а т е л ь с т в о 1. Предположим, что $x_1 < x_2$. Событие « X примет значение, меньшее x_2 » можно представить в виде суммы двух несовместимых событий: « X примет значение, меньшее x_1 » и « X примет значение, удовлетворяющее неравенствам $x_1 \leq X < x_2$ ». Обозначим вероятности последних двух событий соответственно через $P(X < x_1)$ и $P(x_1 \leq X < x_2)$. По теореме о вероятности суммы двух несовместимых событий имеем:

$$P(X < x_2) = P(X < x_1) + P(x_1 \leq X < x_2),$$

откуда с учетом (1)

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1). \quad (2)$$

Так как вероятность любого события есть число неотрицательное, то $P(x_1 \leq X < x_2) \geq 0$, и, значит, $F(x_2) \geq F(x_1)$.

Формула (2) утверждает свойство 3.

3. Вероятность попадания случайной величины X в полуинтервал $[a; b)$ равна разности между значениями интегральной функции распределения в правом и левом концах интервала $(a; b)$:

$$P(a \leq X < b) = F(b) - F(a). \quad (3)$$

П р и м е р. Случайная величина X задана интегральной функцией распределения:

$$F(x) = \begin{cases} 0 & \text{при } x \leq -1, \\ \frac{x}{4} + \frac{1}{4} & \text{при } -1 < x \leq 3, \\ 1 & \text{при } x > 3. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, принадлежащее полуинтервалу $[0; 2)$. Так как на полуинтервале $[0; 2)$ $F(x) = \frac{x}{4} + \frac{1}{4}$, то $P(0 \leq X < 2) = F(2) - F(0) = \frac{1}{2} + \frac{1}{4} - \frac{1}{4} = \frac{1}{2}$.

В дальнейшем случайную величину X будем называть непрерывной, если непрерывна ее интегральная функция распределения $F(x)$.

4. Вероятность того, что непрерывная случайная величина X примет какое-либо заранее заданное значение, равна нулю:

$$P(X = x_1) = 0. \quad (4)$$

Д о к а з а т е л ь с т в о 2. Положив в формуле (2) $x_2 = x_1 + \Delta x$, будем иметь:

$$P(x_1 \leq X < x_1 + \Delta x) = F(x_1 + \Delta x) - F(x_1). \quad (5)$$

Так как $F(x)$ — непрерывная функция, то, перейдя в (5) к пределу при $\Delta x > 0$, получим искомое равенство (4).

Из свойства 4 следует свойство 5.

5. Вероятности попадания непрерывной случайной величины в интервал, сегмент и полуинтервал с одними и теми же концами одинаковы:

$$P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b). \quad (6)$$

6. Если возможные значения случайной величины X принадлежат интервалу $(a; b)$, то: 1) $F(x) = 0$ при $x \leq a$; 2) $F(x) = 1$ при $x \geq b$.

Д о к а з а т е л ь с т в о. 1) Пусть $x_1 \leq a$. Тогда событие $X < x_1$ невозможно и, следовательно, вероятность его равна нулю. 2) Пусть $x_2 \geq b$. Тогда событие $X < x_2$ достоверно и, следовательно, вероятность его равна 1.

С л е д с т в и е. Если возможные значения непрерывной случайной величины расположены на всей числовой оси, то справедливы следующие предельные соотношения:

$$F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1.$$

2. Дифференциальная функция распределения. Дифференциальной функцией распределения непрерывной случайной величины X (или ее плотностью вероятности) называется функция $f(x)$, равная производной интегральной функции распределения: $f(x) = F'(x)$.

Так как $F(x)$ — неубывающая функция (п. 1), то $f(x) \geq 0$ (см. § 8.7, п. 1).

Т е о р е м а. Вероятность попадания непрерывной случайной величины X в интервал $(a; b)$ равна определенному интегралу от ее плотности вероятности, взятому в пределах от a до b :

$$P(a < X < b) = \int_a^b f(x) dx. \quad (7)$$

Д о к а з а т е л ь с т в о. Так как $F(x)$ является первообразной для $f(x)$, то на основании формулы Ньютона — Лейбница (§ 9.7, п. 2) имеем:

$$\int_a^b f(x)dx = F(b) - F(a). \quad (8)$$

Теперь с учетом соотношений (3), (6), (8) получим искомое равенство.

Из равенства (7) следует, что геометрически (см. § 9.6, п. 2) вероятность $P(a < X < b)$ представляет собой площадь криволинейной трапеции, ограниченной графиком плотности вероятности $y = f(x)$ и отрезками прямых $y = 0$, $x = a$ и $x = b$.

С л е д с т в и е. В частности, если $f(x)$ — четная функция и концы интервала симметричны относительно начала координат, то

$$P(-a < X < a) = P(|X| < a) = 2 \int_0^a f(x)dx. \quad (9)$$

Действительно, в этом случае аналогично случаю коэффициентов Фурье для четной функции (§ 10.4, п. 4) $\int_{-a}^a f(x)dx = 2 \int_0^a f(x)dx$.

П р и м е р. Задана плотность вероятности случайной величины X :

$$f(x) = \begin{cases} 0, & x < 0, \\ 2x, & 0 \leq x \leq 1, \\ 0, & x > 1. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значение, принадлежащее интервалу $(0,5; 1)$.

Согласно формуле (7) искомая вероятность

$$P(0,5 < X < 1) = 2 \int_{0,5}^1 x dx = x^2 \Big|_{0,5}^1 = 0,75.$$

Заменяя в формуле (8) a на $-\infty$ и b на x , получим:

$$F(x) - F(-\infty) = \int_{-\infty}^x f(x)dx,$$

откуда в силу найденного выше следствия (п. 1)

$$F(x) = \int_{-\infty}^x f(x)dx. \quad (10)$$

Формула (10) дает возможность отыскать интегральную функцию распределения $F(x)$ по ее плотности вероятности.

Отметим, что из формулы (10) и из только что отмеченного следствия вытекает, что

$$\int_{-\infty}^{+\infty} f(x) dx = 1. \quad (11)$$

3. Математическое ожидание и дисперсия непрерывной случайной величины.

О п р е д е л е н и е 1. Математическим ожиданием ($МО$) непрерывной случайной величины X с плотностью вероятности $f(x)$ называется величина несобственного интеграла (если он сходится):

$$M(X) = \int_{-\infty}^{+\infty} xf(x) dx. \quad (12)$$

О п р е д е л е н и е 2. Дисперсией непрерывной случайной величины X , математическое ожидание которой $M(X) = a$ и функция $f(x)$ является ее плотностью вероятности, называется величина несобственного интеграла (если он сходится):

$$D(X) = \int_{-\infty}^{+\infty} (x - a)^2 f(x) dx. \quad (13)$$

Можно показать (см. [5]), что $МО$ и дисперсия непрерывной случайной величины имеют те же свойства, что и $МО$ и дисперсия дискретной случайной величины.

Для непрерывной случайной величины X среднее квадратическое отклонение $\sigma(X)$ определяется, как и для дискретной величины, формулой $\sigma(X) = \sqrt{D(X)}$.

П р и м е р. Случайная величина X задана плотностью вероятности:

$$f(x) = \begin{cases} 0, & x < 0, \\ \frac{x}{2}, & 0 \leq x \leq 2, \\ 0, & x > 2. \end{cases}$$

Определить математическое ожидание, дисперсию и среднее квадратическое отклонение величины X .

Согласно формулам (12) и (13) имеем:

$$\begin{aligned} M(X) &= \int_{-\infty}^{+\infty} xf(x) dx = \frac{1}{2} \int_0^2 x^2 dx = \frac{x^3}{6} \Big|_0^2 = \frac{4}{3}, \\ D(X) &= \int_{-\infty}^{+\infty} \left(x - \frac{4}{3}\right)^2 f(x) dx = \int_0^2 \left(x - \frac{4}{3}\right)^2 \frac{1}{2} x dx = \\ &= \frac{1}{2} \int_0^2 \left(x^3 - \frac{8}{3}x^2 + \frac{16}{9}x\right) dx = \frac{2}{9} \end{aligned}$$

и, наконец,

$$\sigma(X) = \sqrt{D(X)} = \frac{\sqrt{2}}{3} \approx 0,47$$

(для получения последнего приближенного равенства можно использовать калькулятор).

4. Равномерное распределение. Распределение вероятностей непрерывной случайной величины X , принимающей все свои значения из отрезка $[a; b]$, называется *равномерным*, если ее плотность вероятности на этом отрезке постоянна, а вне его равна нулю, т. е.

$$f(x) = \begin{cases} 0 & \text{при } x < a, \\ c & \text{при } a \leq x \leq b, \\ 0 & \text{при } x > b. \end{cases}$$

Отсюда

$$\int_{-\infty}^{+\infty} f(x) dx = \int_a^b c dx = c(b-a). \quad (14)$$

Но, как известно (см. п. 2),

$$\int_{-\infty}^{+\infty} f(x) dx = 1. \quad (15)$$

Из сравнения равенств (14), (15) получаем: $c = \frac{1}{b-a}$. График функции $f(x)$ изображен на рисунке 202. Определим интегральную функцию распределения $F(x) = \int_{-\infty}^x f(x) dx$.

Если $x < a$, то $f(x) = 0$, и, следовательно $F(x) = 0$. Если $a \leq x \leq b$, то $f(x) = \frac{1}{b-a}$, и, следовательно, $F(x) = \int_a^x \frac{1}{b-a} dx = \frac{x-a}{b-a}$.

Если $b < x$, то $f(x) = 0$, и, следовательно,

$$F(x) = \int_{-\infty}^x f(x) dx = \int_a^b \frac{1}{b-a} dx = 1.$$

График функции $F(x)$ показан на рисунке 203.

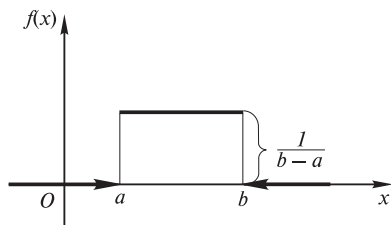


Рис. 202

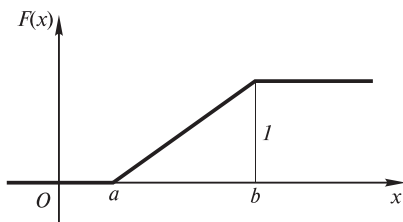


Рис. 203

Наконец

$$M(X) = \int_{-\infty}^{+\infty} f(x)xdx = \int_a^b \frac{1}{b-a} xdx = \frac{1}{2}(a+b),$$

$$D(X) = \int_{-\infty}^{+\infty} (x - M(X))^2 f(x)dx = \int_a^b \frac{1}{b-a} \left(x - \frac{1}{2}(a+b) \right)^2 dx = \frac{(b-a)^2}{12}.$$

П р и м е р. На отрезке $[a; b]$ наугад указывают точку. Какова вероятность того, что эта точка окажется в левой половине отрезка?

Обозначим через X случайную величину, равную координате выбранной точки. X распределена равномерно (в этом и состоит точный смысл слов «наугад указывают точку»), а так как середина отрезка $[a; b]$ имеет координату $\frac{a+b}{2}$, то искомая вероятность равна (см. п. 2):

$$P\left(a < X < \frac{a+b}{2}\right) = \int_a^{\frac{a+b}{2}} f(x)dx = \int_a^{\frac{a+b}{2}} \frac{1}{b-a} dx = \frac{1}{2}.$$

Впрочем, этот результат был ясен с самого начала.

5. Нормальный закон распределения. Закон распределения вероятностей непрерывной случайной величины X называется *нормальным* или *законом Гаусса*, если ее плотность вероятности есть

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad (16)$$

где σ и a — постоянные, причем $\sigma > 0$.

Убедимся, что функция (16) удовлетворяет условию (11) из § 18.8 (п. 2). Действительно, перейдя в интеграле

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx \quad (17)$$

к новой переменной

$$t = \frac{x-a}{\sigma\sqrt{2}}, \quad (18)$$

получим интеграл

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \int_0^{+\infty} e^{-t^2} dt.$$

Но $\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$ (§ 12.1, п. 6). Следовательно,

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt = 1. \quad (19)$$

Значит, интеграл (17) тоже равен единице.

Покажем, что $M(X) = a, \sigma = \sqrt{D(X)}$, или $\sigma^2 = D(X)$.

Согласно формуле (12) получаем:

$$M(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

Введя новую переменную t по формуле (18), с учетом равенства (19) получим:

$$\begin{aligned} M(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (a + t\sigma\sqrt{2}) e^{-t^2} \sigma\sqrt{2} dt = \\ &= \frac{a}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2} dt + \frac{\sigma\sqrt{2}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} t e^{-t^2} dt = a - \frac{\sigma}{\sqrt{2\pi}} e^{-t^2} \Big|_{-\infty}^{+\infty} = a. \end{aligned}$$

Аналогично устанавливается, что

$$D(X) = \sigma^2.$$

Отсюда

$$\sigma = \sqrt{D(X)}.$$

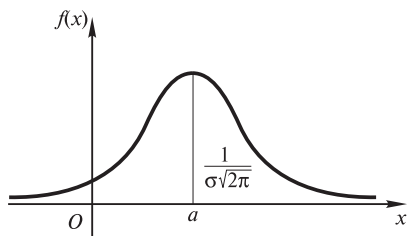


Рис. 204

В § 8.8 (пример 2) был построен график функции $y = e^{-x^2}$ (кривая Гаусса). С учетом графика этой функции график функции (16) будет иметь вид, изображенный на рисунке 204.

Нормальное распределение с параметрами $a=0$ и $\sigma=1$ называется *нормированным*. Плотность вероятности в случае такого распределения будет

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Пусть случайная величина X распределена по нормальному закону. Тогда вероятность того, что X примет значение, принадлежащее интервалу $(\alpha; \beta)$, согласно теореме из п. 2, будет

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}} dx.$$

Сделаем в этом интеграле замену переменной, полагая $\frac{x-a}{\sigma} = t$. Тогда: $x = a + \sigma t, dx = \sigma dt$ и

$$P(\alpha < X < \beta) = \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt,$$

откуда с учетом того, что функция $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ есть первообразная для $\varphi(x)$ (§ 18.7, п. 3), и формулы Ньютона — Лейбница (§ 9.7, п. 2) будем иметь:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - \alpha}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \quad (20)$$

Пример 1. Пусть случайная величина X распределена по нормальному закону с параметрами $a = 30$ и $\sigma = 10$. Найти вероятность того, что X примет значение, принадлежащее интервалу $(10; 50)$.

Пользуясь формулой (20), получим:

$$P(10 < X < 50) = \Phi\left(\frac{50 - 30}{10}\right) - \Phi\left(\frac{10 - 30}{10}\right) = \Phi(2) - \Phi(-2) = 2\Phi(2).$$

По таблице приложения 4 находим $\Phi(2) = 0,4772$. Отсюда искомая вероятность

$$P(10 < X < 50) = 2 \cdot 0,4772 = 0,9544.$$

Часто требуется вычислить вероятность того, что отклонение нормально распределенной случайной величины X от ее математического ожидания по абсолютной величине меньше заданного положительного числа δ , т. е. найти $P(|X - a| < \delta)$. Используя формулу (20) и нечетность функции $\Phi(x)$, имеем:

$$P(|X - a| < \delta) = P(a - \delta < X < a + \delta) = \Phi\left(\frac{\delta}{\sigma}\right) - \Phi\left(-\frac{\delta}{\sigma}\right) = 2\Phi\left(\frac{\delta}{\sigma}\right),$$

т. е.

$$P(|X - a| < \delta) = 2\Phi\left(\frac{\delta}{\sigma}\right). \quad (21)$$

Пример 2. Пусть случайная величина X распределена по нормальному закону с параметрами $a = 20$ и $\sigma = 10$. Найти $P(|X - 20| < 3)$.

Используя формулу (21), имеем:

$$P(|X - 20| < 3) = 2\Phi\left(\frac{3}{10}\right).$$

По таблице приложения 4 находим: $\Phi(0,3) = 0,1179$.

Поэтому $P(|X - 20| < 3) = 0,2358$.

Нормальное распределение вероятностей имеет в теории вероятностей большое значение. Нормальному закону подчиняется вероятность при стрельбе по цели, в измерениях и т. п. В частности, оказывается, что закон распределения суммы достаточно большого числа независимых случайных величин, влияние каждой из которых на всю сумму ничтожно мало, близок к нормальному распределению. Этот факт, назы-

ваемый *центральной предельной теоремой*, был доказан выдающимся русским математиком А.М. Ляпуновым (1857 — 1918).

§ 18.9. Закон больших чисел

1. Неравенство Чебышева (П.Л. Чебышев (1821 — 1894) — выдающийся русский математик).

Л е м м а. Пусть X — случайная величина, принимающая только неотрицательные значения. Тогда

$$P(X \geq 1) \leq M(X). \quad (1)$$

Д о к а з а т е л ь с т в о. Для простоты докажем это утверждение для дискретной случайной величины X , принимающей значения $x_1, x_2, \dots, x_n$, при условии $x_i \geq 0$. По теореме сложения вероятностей для несовместимых событий (§ 18.2, п. 1) имеем:

$$P(X \geq 1) = \sum_{x_i \geq 1} P(X = x_i),$$

где суммирование распространено на все значения x_i , большие единицы или равные ей. Но для $x_i \geq 1$, очевидно,

$$P(X = x_i) \leq x_i P(X = x_i).$$

Поэтому

$$P(X \geq 1) = \sum_{x_i \geq 1} P(X = x_i) \leq \sum_{x_i \geq 1} x_i P(X = x_i). \quad (2)$$

Добавим к правой части неравенства (2) сумму $\sum_{x_i < 1} x_i P(X = x_i)$, где $x_i < 1$.

Эта сумма неотрицательна, так как $x_i \geq 0$ по условию, а вероятность $P(X = x_i) \geq 0$. Поэтому

$$\sum_{x_i \geq 1} x_i P(X = x_i) \leq \sum_{x_i \geq 1} x_i P(X = x_i) + \sum_{x_i < 1} x_i P(X = x_i) = \sum_{i=1}^n x_i P(X = x_i). \quad (3)$$

Последняя сумма распространена на все значения x_i , принимаемые случайной величиной X . Следовательно (см. § 18.5, п. 1),

$$\sum_{i=1}^n x_i P(X = x_i) = M(X).$$

Отсюда, сопоставляя соотношения (2) и (3), имеем искомое неравенство (1).

Т е о р е м а. Для любой случайной величины X при каждом положительном числе ε имеет место неравенство

$$P(|X - M(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}. \quad (4)$$

Неравенство (4) называется *неравенством Чебышева*.

Доказательство. Так как событие $|X - M(X)| \geq \varepsilon$ равносильно событию

$$\frac{(X - M(X))^2}{\varepsilon^2} \geq 1,$$

то

$$P(|X - M(X)| \geq \varepsilon) = P\left(\frac{(X - M(X))^2}{\varepsilon^2} \geq 1\right).$$

Случайная величина $\frac{(X - M(X))^2}{\varepsilon^2}$ неотрицательна, и, значит, согласно лемме, свойству 2 математического ожидания (§ 18.5, п. 2) и определению дисперсии (§ 18.6, п. 1)

$$P\left(\frac{(X - M(X))^2}{\varepsilon^2} \geq 1\right) \leq M\left(\frac{(X - M(X))^2}{\varepsilon^2}\right) = \frac{1}{\varepsilon^2} M((X - M(X))^2) = \frac{D(X)}{\varepsilon^2}.$$

Поэтому

$$P(|X - M(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}.$$

Пример. Пусть случайная величина X имеет $D(X) = 0,001$. Какова вероятность того, что X отличается от $M(X)$ более чем на 0,1?

По неравенству Чебышева

$$P(|X - M(X)| \geq 0,1) \leq \frac{D(X)}{0,1^2} = \frac{0,001}{0,01} = 0,1.$$

Примечание. Отметим другую форму неравенства Чебышева. Так как событие, выражаемое неравенством $|X - M(X)| < \varepsilon$, противоположно событию, выражаемому неравенством $|X - M(X)| \geq \varepsilon$, то (§ 18.2, п. 1 следствие)

$$P(|X - M(X)| < \varepsilon) + P(|X - M(X)| \geq \varepsilon) = 1.$$

Отсюда с учетом неравенства (4) получаем такую форму неравенства Чебышева:

$$P(|X - M(X)| < \varepsilon) \geq 1 - \frac{D(X)}{\varepsilon^2}. \quad (5)$$

2. Закон больших чисел Чебышева. Докажем закон больших чисел в широкой и удобной для практики форме, полученной П.Л. Чебышевым.

Теорема (Теорема Чебышева: закон больших чисел). Если дисперсии независимых случайных величин $X_1, X_2, \dots, X_n$ ограничены одной и той же постоянной C : $D(X_i) \leq C$ ($i = 1, 2, \dots, n$), то, каково бы ни было $\varepsilon > 0$, вероятность выполнения неравенства $|\bar{X} - M(\bar{X})| < \varepsilon$, где $\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$, будет сколь угодно

близка к единице, если число случайных величин n достаточно велико, т. е.

$$\lim_{n \rightarrow \infty} P(|\bar{X} - M(\bar{X})| < \varepsilon) = 1. \quad (6)$$

Д о к а з а т е л ь с т в о. Применяя неравенство Чебышева в форме (5) к величине $\bar{X}$, имеем:

$$P(|\bar{X} - M(\bar{X})| < \varepsilon) \geq 1 - \frac{D(\bar{X})}{\varepsilon^2}. \quad (7)$$

Пользуясь свойствами дисперсии (§ 18.6, п. 2) и условием теоремы, получим:

$$D(\bar{X}) = \frac{1}{n^2}(D(X_1) + \dots + D(X_n)) \leq \frac{nC}{n^2} = \frac{C}{n}.$$

Отсюда с учетом неравенства (7) и того, что вероятность любого события не превосходит единицы (§ 18.1, п. 3), получим:

$$1 \geq P(|\bar{X} - M(\bar{X})| < \varepsilon) \geq 1 - \frac{C}{\varepsilon^2 n}. \quad (8)$$

Наконец, переходя в неравенстве (8) к пределу при $n \rightarrow \infty$, приходим к искомому соотношению (6).

Ч а с т н ы й с л у ч а й т е о р е м ы Ч е б ы ш е в а. Если все X_k имеют одинаковые математические ожидания $M(X_1) = \dots = M(X_n) = a$ и $D(X_k) \leq C$ ($k = 1, 2, \dots, n$), то

$$\lim_{n \rightarrow \infty} P(|\bar{X} - a| < \varepsilon) = 1. \quad (9)$$

Действительно, в условиях рассматриваемого частного случая равенство (6) имеет вид (9).

Сущность теоремы Чебышева состоит в следующем. Несмотря на то, что каждая из независимых случайных величин X_k может принять значение, далекое от математического ожидания $M(X_k)$, среднее арифметическое $\bar{X}$ достаточно большого числа случайных величин с большой вероятностью весьма близко к среднему арифметическому их математических ожиданий.

Теорема Чебышева имеет большое практическое значение. Пусть, например, измеряется некоторая физическая величина. Обычно принимают в качестве искомого значения измеряемой величины среднее арифметическое результатов нескольких измерений. Можно ли считать такой подход верным? Теорема Чебышева (ее частный случай) отвечает на этот вопрос положительно.

На теореме Чебышева основан широко применяемый в статистике выборочный метод, согласно которому по сравнительно небольшой случайной выборке выносят суждение, касающееся всей совокупности исследуемых объектов.

Из теоремы Чебышева (частный случай) следует следующая теорема, называемая теоремой Бернулли, являющаяся простейшей формой закона больших чисел:

Т е о р е м а Б е р н у л л и. Пусть m — число наступлений события A в n независимых испытаниях и p есть вероятность наступления события A в каждом из испытаний. Тогда, каково бы ни было положительное число ε ,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1. \quad (10)$$

Д о к а з а т е л ь с т в о. Обозначим через X_k случайную величину, равную числу наступлений события A в k -м испытании, где $k = 1, 2, \dots, n$. Тогда имеем (§ 18.7, п. 1):

$$m = X_1 + X_2 + \dots + X_n, \\ M(X_k) = p, \quad D(X_k) = pq \leq \frac{1}{4},$$

и все условия частного случая теоремы Чебышева выполнены. Равенство (9) превращается в (10).

Практический смысл теоремы Бернулли следующий: при постоянстве вероятности случайного события A во всех испытаниях, при неограниченном возрастании числа испытаний можно с вероятностью, как угодно близкой к единице (т. е. как угодно близкой к достоверности), утверждать, что наблюдаемая относительная частота случайного события будет как угодно мало отклоняться от его вероятности.

§ 18.10. Использование теории вероятностей при обработке экспериментальных данных

1. Закон распределения ошибок. Ошибкой или погрешностью измерения называется разность $x - a$ между результатом измерения x и истинным значением измеряемой величины a .

Ошибки измерений в основном можно подразделить на три группы: 1) грубые ошибки; 2) систематические ошибки; 3) случайные ошибки.

Грубые ошибки возникают от невнимательности при чтении показаний прибора, неправильной записи показаний, неправильном использовании прибора. Эти ошибки могут быть исключены соблюдением правил измерения.

Систематические ошибки происходят, например, от несовершенства приборов, от личных качеств наблюдателя и могут быть устранены соответствующими поправками.

Будем считать, что результаты измерения не содержат грубых и систематических ошибок.

Однако и после устранения этих ошибок результаты измерения все еще будут содержать неустраняемые, неизбежные ошибки, которые получили название *случайных ошибок измерения*. Эти ошибки вызываются многочисленными трудно уловимыми причинами, каждая из которых приводит лишь к незначительному колебанию результатов измерения (например, при взвешивании на аналитических весах к таким причинам относятся незначительные колебания температуры и влажности воздуха, колебания стола, попадание соринок на взвешиваемый предмет и т. д.). Ошибку измерения можно считать суммой большого числа независимых случайных величин, которая по центральной предельной теореме должна быть распределена нормально. При этом принимают еще, что *МО* случайных ошибок равно нулю.

Итак, принимается положение: при отсутствии грубых и систематических ошибок ошибка измерения есть случайная величина (обозначим ее через T), распределенная нормально, причем ее *МО* равно нулю, т. е. плотность вероятности величины T равна

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}},$$

где σ — среднее квадратическое отклонение величины T , характеризующее разброс результатов измерения вокруг измеряемой величины.

В силу предыдущего результат измерения есть также случайная величина (обозначим ее через X), связанная с T зависимостью $X = a + T$. Отсюда в силу известных свойств *МО* и дисперсии (§ 18.5, 18.6) имеем:

$$M(X) = a, \sigma(X) = \sigma(T) = \sigma.$$

Таким образом, предполагая измерение свободным от грубых и систематических ошибок, можно считать, что возможный результат измерения есть случайная величина X , математическое ожидание которой равно истинному значению a измеряемой величины. Так как величина T подчиняется нормальному закону распределения, то возможный результат измерения $X = a + T$ как случайная величина, линейно зависящая от T , также подчиняется (см. [14]) нормальному закону распределения. В этом заключается *основной закон ошибок*.

Заметим, что случайная ошибка измерения, как и результаты измерения, всегда выражается в некоторых целых единицах, связанных с шагом шкалы измерительного прибора; в теории удобнее считать случайную ошибку непрерывной случайной величиной, что упрощает расчеты.

2. Оценка истинного значения измеряемой величины. Пусть для определения неизвестной физической постоянной a производится n независимых равноточных измерений (т. е. измерений, проводимых в

одинаковых условиях; эти условия считаются выполненными, если измерения проводятся одним прибором), причем считается, что грубые и систематические ошибки отсутствуют. Возможный результат каждого из n измерений есть случайная величина, которую обозначим через X_i (i — номер измерения). Так как каждое измерение не зависит от результатов других измерений, то имеем n случайных независимых величин $X_1, X_2, \dots, X_n$. Обозначим через $x_1, x_2, \dots, x_n$ фактически полученные результаты n измерений величины a . Таким образом, x_i есть одно из возможных значений X_i .

На основании закона больших чисел Чебышева (§ 18.9, п. 2) можно утверждать, что с практической достоверностью для достаточно большого числа n измерений средняя арифметическая результатов измерений отличается от истинного значения физической постоянной сколь угодно мало, т. е. с вероятностью, сколь угодно близкой к единице, имеет место приближенное равенство

$$a \approx \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Оценим точность этого приближенного равенства. Для этого прежде всего заметим, что в силу основного закона ошибок (п. 1) каждый возможный результат измерения X_i есть случайная величина, подчиняющаяся нормальному закону распределения вероятностей с одним и тем же математическим ожиданием, равным истинному значению a измеряемой величины: $M(X_i) = a$ ($i = 1, 2, \dots, n$). Так как все измерения равноточные, то дисперсии всех случайных величин должны быть одинаковыми, т. е. $D(X_i) = \sigma^2$.

При измерении возможны две ситуации:

а) известно σ (это характеристика прибора и комплекса условий, при которых производятся наблюдения), требуется по результатам измерений оценить a ;

б) σ не известно, требуется по результатам измерений оценить a и σ .

Эти очень важные задачи будут обсуждаться в § 19.3.

§ 18.11. Двумерные случайные величины

В различных задачах практики встречаются случайные величины, возможные значения которых определяются не одним числом, а несколькими. Так, при вытачивании на станке цилиндрического бруска его размеры (диаметр основания и высота) являются случайными величинами. Таким образом, здесь мы имеем дело с совокупностью двух случайных величин. Можно привести примеры, в которых рассматривается совокупность трех и более случайных величин.

Двумерная дискретная случайная величина (X, Y) задается таблицей значений ее составляющих X и Y и вероятностей. Общий вид такой таблицы:

$X \backslash Y$	y_1	y_2	...	y_m
x_1	p_{11}	p_{12}	...	p_{1m}
x_2	p_{21}	p_{22}	...	p_{2m}
...	...	...	...	...
x_n	p_{n1}	p_{n2}	...	p_{nm}

Здесь, например, p_{12} есть вероятность того, что двумерная величина примет значение (x_1, y_2) .

Указанной таблицей задан закон распределения дискретной двумерной случайной величины.

Непрерывная двумерная случайная величина может аналогично непрерывной одномерной величине определяться интегральной функцией распределения $F(x, y)$ и дифференциальной функцией $f(x, y)$ (плотностью вероятности двумерной случайной величины).

Аналогично одномерному случаю вводятся понятия математического ожидания $M(X, Y)$, дисперсии $D(X, Y)$ и среднего квадратического отклонения $\sigma(X, Y)$.

Глава 19. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Мы приступаем к изучению элементов математической статистики, в которой разрабатываются научно обоснованные методы сбора статистических данных и их обработки.

§ 19.1. Выборочный метод

1. Генеральная совокупность и выборка. Пусть требуется изучить множество однородных объектов (это множество называется *статистической совокупностью*) относительно некоторого качественного или коли-

качественного признака, характеризующего эти объекты. Например, если имеется партия деталей, то качественным признаком может служить стандартность детали, а количественным — контролируемый размер детали.

Лучше всего произвести сплошное обследование, т. е. изучить каждый объект. Однако в большинстве случаев по разным причинам это сделать невозможно. Препятствовать сплошному обследованию может большое число объектов, недоступность их. Если, например, нужно знать среднюю глубину воронки при взрыве снаряда из опытной партии, то, производя сплошное обследование, мы уничтожим всю партию.

Если сплошное обследование невозможно, то из всей совокупности выбирают для изучения часть объектов.

Статистическая совокупность, из которой отбирают часть объектов, называется *генеральной совокупностью*. Множество объектов, случайно отобранных из генеральной совокупности, называется *выборкой*.

Число объектов генеральной совокупности и выборки называется соответственно *объемом* генеральной совокупности и *объемом* выборки.

П р и м е р. Плоды дерева (200 штук) обследуют на наличие специфического для данного сорта вкуса. Для этого отбирают 10 шт. Здесь 200 — объем генеральной совокупности, а 10 — объем выборки.

Если выборку отбирают по одному объекту, который обследуют и снова возвращают в генеральную совокупность, то выборка называется *повторной*. Если объекты выборки уже не возвращаются в генеральную совокупность, то выборка называется *бесповторной*. На практике чаще используется бесповторная выборка. Если объем выборки составляет небольшую долю объема генеральной совокупности, то разница между повторной и бесповторной выборками незначительна.

Свойства объектов выборки должны правильно отражать свойства объектов генеральной совокупности, или, как говорят, выборка должна быть *репрезентативной* (представительной). Считается, что выборка репрезентативна, если все объекты генеральной совокупности имеют одинаковую вероятность попасть в выборку, т. е. выбор производится случайно. Например, для того, чтобы оценить будущий урожай, можно сделать выборку из генеральной совокупности еще не созревших плодов и исследовать их характеристики (массу, качество и пр.). Если вся выборка будет сделана с одного дерева, то она не будет репрезентативной. Репрезентативная выборка должна состоять из случайно выбранных плодов со случайно выбранных деревьев.

2. Статистическое распределение выборки. Полигон. Гистограмма.

Пусть из генеральной совокупности извлечена выборка, причем x_1 наблюдалось n_1 раз, $x_2 - n_2$ раз, $x_k - n_k$ раз и $n_1 + n_2 + \dots + n_k = n$ — объем выборки. Наблюдаемые значения $x_1, x_2, \dots, x_k$ называются *вариантами*, а последовательность вариантов, записанная в возрастающем порядке, — *вариационным рядом*. Числа наблюдений $n_1, n_2, \dots, n_k$ называют *частотами*, а их отношения к объему выборки $\frac{n_1}{n} = p_1^*, \frac{n_2}{n} = p_2^*, \dots, \frac{n_k}{n} = p_k^*$ — *относительными частотами*. Отметим, что сумма относительных частот равна единице:

$$p_1^* + p_2^* + \dots + p_k^* = \frac{n_1}{n} + \frac{n_2}{n} + \dots + \frac{n_k}{n} = \frac{n}{n} = 1.$$

Статистическим распределением выборки называют перечень вариантов и соответствующих им частот или относительных частот. Статистическое распределение можно задать также в виде последовательности интервалов и соответствующих им частот (непрерывное распределение). В качестве частоты, соответствующей интервалу, принимают сумму частот вариантов, попавших в этот интервал.

Заметим, что в теории вероятностей под распределением понимают соответствие между возможными значениями случайной величины и ее вероятностями, а в математической статистики — соответствие между наблюдаемыми вариантами и их частотами или относительными частотами.

П р и м е р 1. Перейти от частот к относительным частотам в следующем распределении выборки объема $n = 20$:

Варианта x_i	2	6	12
Частота n_i	3	10	7

Найдем относительные частоты:

$$p_1^* = \frac{3}{20} = 0,15; \quad p_2^* = \frac{10}{20} = 0,50; \quad p_3^* = \frac{7}{20} = 0,35.$$

Поэтому получаем следующее распределение:

Варианта x_i	2	6	12
Относительная частота p_i^*	0,15	0,50	0,35

Для графического изображения статистического распределения используются *полигоны* и *гистограммы*.

Для построения полигона на оси Ox откладывают значения вариант x_i , на оси Oy — значения частот n_i (относительных частот p_i^*).

Пример 2. На рисунке 205 изображен полигон следующего распределения:

Варианта x_i	1	2	3	5
Относительная частота p_i^*	0,4	0,2	0,3	0,1

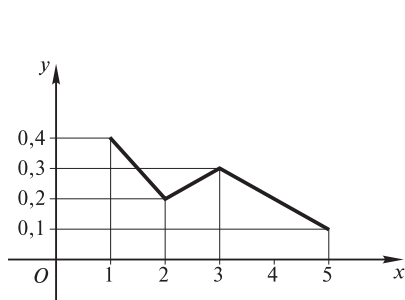


Рис. 205

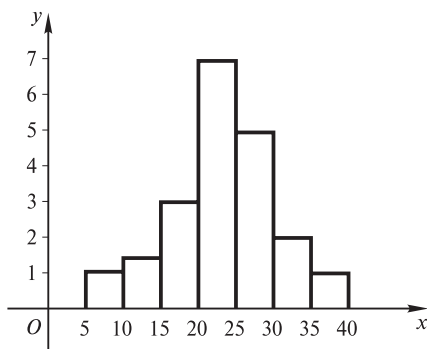


Рис. 206

Полигоном обычно пользуются в случае небольшого количества вариантов. В случае большого количества вариантов и в случае непрерывного распределения признака чаще строят гистограммы. Для этого интервал, в котором заключены все наблюдаемые значения признака, разбивают на несколько частичных интервалов длиной h и находят для каждого частичного интервала n_i — сумму частот вариантов, попавших в i интервал. Затем на этих интервалах, как на основаниях, строят прямоугольники с высотами $\frac{n_i}{h}$ (или $\frac{n_i}{nh}$, где n — объем выборки). Площадь i частичного прямоугольника равна $\frac{hn_i}{h} = n_i$ (или $\frac{hn_i}{nh} = \frac{n_i}{n} = p_i^*$). Следовательно, площадь гистограммы равна сумме всех частот (или относительных частот), т. е. объему выборки (или единице).

Пример 3. На рисунке 206 изображена гистограмма непрерывного распределения объема $n = 100$, приведенного в следующей таблице:

Частичный интервал h	Сумма частот вариант частичного интервала n_i	$\frac{n_i}{h}$
5—10	4	0,8
10—15	6	1,2
15—20	16	3,2
20—25	36	7,2
25—30	24	4,8
30—35	10	2,0
35—40	4	0,8

§ 19.2. Оценки параметров генеральной совокупности по ее выборке

1. Выборка как набор случайных величин. Пусть имеется некоторая генеральная совокупность, каждый объект которой наделен количественным признаком X . При случайном извлечении объекта из генеральной совокупности становится известным значение x признака X этого объекта. Таким образом, мы можем рассматривать извлечение объекта из генеральной совокупности как испытание, X — как случайную величину, а x — как одно из возможных значений X .

Допустим, что из теоретических соображений удалось установить, к какому типу распределений относится признак X . Естественно, возникает задача оценки (приближенного нахождения) параметров, которыми определяется это распределение. Например, если известно, что изучаемый признак распределен в генеральной совокупности нормально, то необходимо оценить, т. е. приблизительно найти математическое ожидание и среднее квадратическое отклонение, так как эти два параметра полностью определяют нормальное распределение.

Обычно в распоряжении исследователя имеются лишь данные выборки генеральной совокупности, например, значения количественного признака $x_1, x_2, \dots, x_n$, полученные в результате n наблюдений (здесь и далее наблюдения предполагаются независимыми). Через эти данные и выражают оцениваемый параметр.

Опытные значения признака X можно рассматривать и как значения разных случайных величин $X_1, X_2, \dots, X_n$ с тем же распределением, что и X , и, следовательно, с теми же числовыми характеристиками, которые имеет X . Значит, $M(X) = M(X)$ и $D(X) = D(X)$. Величины $X_1, X_2, \dots, X_n$ можно считать независимыми в силу независимости наблюдений. Значения $x_1, x_2, \dots, x_n$ в этом случае называются *реализациями* случайных величин $X_1, X_2, \dots, X_n$. Отсюда и из предыдущего следует, что найти оценку неизвестного параметра — это значит найти функцию от наблюдаемых случайных величин $X_1, X_2, \dots, X_n$, которая и дает приближенное значение оцениваемого параметра.

2. Генеральная и выборочная средние. Методы их расчета. Пусть изучается дискретная генеральная совокупность объема N относительно количественного признака X .

О п р е д е л е н и е 1. Генеральной средней $\bar{x}_r$ (или a) называется среднее арифметическое значение признака генеральной совокупности.

Если все значения $x_1, x_2, \dots, x_N$ признака генеральной совокупности объема N различны, то

$$\bar{x}_r = \frac{1}{N}(x_1 + x_2 + \dots + x_N).$$

Если же значения признака $x_1, x_2, \dots, x_k$ имеют соответственно частоты $N_1, N_2, \dots, N_k$, причем $N_1 + N_2 + \dots + N_k = N$, то

$$\bar{x}_r = \frac{1}{N}(x_1 N_1 + x_2 N_2 + \dots + x_k N_k)$$

или

$$\bar{x}_r = \frac{1}{N} \sum_{i=1}^k x_i N_i. \quad (1)$$

Как уже отмечалось (п. 1), извлечение объекта из генеральной совокупности есть наблюдение случайной величины X .

Пусть все значения $x_1, x_2, \dots, x_N$ различны. Так как каждый объект может быть извлечен с одной и той же вероятностью $\frac{1}{N}$, то

$$M(X) = x_1 \cdot \frac{1}{N} + x_2 \cdot \frac{1}{N} + \dots + x_N \cdot \frac{1}{N} = \bar{x}_r,$$

т. е.

$$M(X) = \bar{x}_r. \quad (2)$$

Такой же итог следует, если значения $x_1, x_2, \dots, x_k$ имеют соответственно частоты $N_1, N_2, \dots, N_k$.

В случае непрерывного распределения признака X по определению полагают $\bar{x}_r = M(X)$.

Пусть для изучения генеральной совокупности относительно количественного признака X произведена выборка объема n .

О п р е д е л е н и е 2. Выборочной средней $\bar{x}_b$ называется среднее арифметическое значений признака выборочной совокупности.

Если все значения $x_1, x_2, \dots, x_n$ признака выборки объема n различны, то

$$\bar{x}_b = \frac{1}{n}(x_1 + x_2 + \dots + x_n). \quad (3)$$

Если же значения признака $x_1, x_2, \dots, x_k$ имеют соответственно частоты $n_1, n_2, \dots, n_k$, причем $n_1 + n_2 + \dots + n_k = n$, то

$$\bar{x}_b = \frac{1}{n}(x_1 n_1 + x_2 n_2 + \dots + x_k n_k)$$

или

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^k x_i n_i. \quad (4)$$

Пример 1. Выборочным путем были получены следующие данные о массе 20 морских свинок при рождении (в г): 30, 30, 25, 32, 30, 25, 33, 32, 29, 28, 27, 36, 31, 34, 30, 23, 28, 31, 36, 30. Найти выборочную среднюю $\bar{x}_B$.

Согласно формуле (4) имеем:

$$\bar{x}_B = \frac{30 \cdot 5 + 25 \cdot 2 + 32 \cdot 2 + 33 + 29 + 28 \cdot 2 + 27 + 36 \cdot 2 + 31 \cdot 2 + 34 + 23}{20} = 30.$$

Итак, $\bar{x}_B = 30$ г.

Ниже, не уменьшая общности рассуждений, будем считать значения $x_1, x_2, \dots, x_n$ признака различными.

Разумеется, выборочная средняя для различных выборок того же объема n из той же генеральной совокупности будет получаться, вообще говоря, различной. И это не удивительно — ведь извлечение i -го по счету объекта есть наблюдение случайной величины X_i , а их среднее арифметическое

$$\bar{X} = \frac{1}{n} (X_1 + X_2 + \dots + X_n)$$

есть тоже случайная величина.

Таким образом, всевозможные, которые могут получиться выборочные средние, есть возможные значения случайной величины $\bar{X}$, которая называется выборочной средней случайной величиной.

Найдем $M(\bar{X})$, пользуясь тем, что $M(X_i) = M(X)$ (см. п. 1).

С учетом свойств MO (§ 18.5, 18.8) получаем:

$$\begin{aligned} M(\bar{X}) &= M \left[\frac{1}{n} (X_1 + X_2 + \dots + X_n) \right] = \frac{1}{n} [M(X_1) + M(X_2) + \dots + M(X_n)] = \\ &= \frac{1}{n} [M(X) + M(X) + \dots + M(X)] = \frac{1}{n} \cdot na = a. \end{aligned}$$

Итак, $M(\bar{X})$ (MO выборочной средней) совпадает с a (генеральной средней).

Теперь найдем $D(\bar{X})$. Так как $D(X_i) = D(X)$ (п. 1) и $X_1, X_2, \dots, X_n$ независимы, то согласно свойствам дисперсии (§ 18.6, 18.8) получаем:

$$\begin{aligned} D(\bar{X}) &= D \left[\frac{1}{n} (X_1 + X_2 + \dots + X_n) \right] = \frac{1}{n^2} [D(X_1) + D(X_2) + \dots + D(X_n)] = \\ &= \frac{1}{n^2} [D(X) + D(X) + \dots + D(X)] = \frac{1}{n^2} \cdot nD(X) = \frac{D(X)}{n}, \end{aligned}$$

т. е.

$$D(\bar{X}) = \frac{D(X)}{n}. \quad (5)$$

Наконец, отметим, что если варианты x_i — большие числа, то для облегчения вычисления выборочной средней применяют следующий прием. Пусть C — константа.

Так как

$$\sum_{i=1}^n x_i = \sum_{i=1}^n (x_i - C) + nC,$$

то формула (3) преобразуется к виду:

$$\bar{x}_B = C + \frac{1}{n} \sum_{i=1}^n (x_i - C). \quad (6)$$

Константу C (так называемый *ложный* нуль) берут такой, чтобы, во-первых, разности $x_i - C$ были небольшими и, во-вторых, число C было по возможности «круглым».

Пример 2. Имеется выборка:

$$x_1 = 71,88; x_2 = 71,93; x_3 = 72,05; x_4 = 72,07;$$

$$x_5 = 71,90; x_6 = 72,02; x_7 = 71,93; x_8 = 71,77;$$

$$x_9 = 72,71; x_{10} = 71,96.$$

Берем $C = 72,00$ и вычисляем разности: $\alpha_i = x_i - C$.

$$\alpha_1 = -0,12; \alpha_2 = -0,07; \alpha_3 = 0,05; \alpha_4 = 0,07;$$

$$\alpha_5 = -0,10; \alpha_6 = 0,02; \alpha_7 = -0,07; \alpha_8 = -0,23;$$

$$\alpha_9 = 0,11; \alpha_{10} = -0,04.$$

Их сумма: $\alpha_1 + \alpha_2 + \dots + \alpha_{10} = -0,38$; их среднее арифметическое: $\frac{1}{10}(\alpha_1 + \alpha_2 + \dots + \alpha_{10}) = -0,038 \approx -0,04$. Выборочная средняя:

$$\bar{x}_B \approx 72,00 - 0,04 = 71,96.$$

3. Генеральная и выборочная дисперсии. Для того чтобы охарактеризовать рассеяние значений количественного признака X генеральной совокупности вокруг своего среднего значения, вводят следующую характеристику — генеральную дисперсию.

Определение 1. *Генеральной дисперсией* D_r называется среднее арифметическое квадратов отклонений значений признака X генеральной совокупности от генеральной средней $\bar{x}_a$.

Если все значения $x_1, x_2, \dots, x_N$ признака генеральной совокупности объема N различны, то

$$D_r = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}_r)^2.$$

Если же значения признака $x_1, x_2, \dots, x_k$ имеют соответственно частоты $N_1, N_2, \dots, N_k$, причем $N_1 + N_2 + \dots + N_k = N$, то

$$D_r = \frac{1}{N} \sum_{i=1}^k (x_i - \bar{x}_r)^2 N_i. \quad (7)$$

П р и м е р 1. Генеральная совокупность задана таблицей распределения:

x_i	2	4	5	6
N_i	8	9	10	3

Найти генеральную дисперсию. Согласно формулам (1) и (7) имеем:

$$\bar{x}_r = \frac{2 \cdot 8 + 4 \cdot 9 + 5 \cdot 10 + 6 \cdot 3}{8 + 9 + 10 + 3} = \frac{120}{30} = 4,$$

$$D_r = \frac{(2-4)^2 \cdot 8 + (4-4)^2 \cdot 9 + (5-4)^2 \cdot 10 + (6-4)^2 \cdot 3}{30} = \frac{54}{30} = 1,8.$$

Генеральным соединим квадратическим отклонением (стандартом) называется $\sigma_r = \sqrt{D_r}$.

Пусть все значения $x_1, x_2, \dots, x_N$ различны.

Найдем дисперсию признака X , рассматриваемого как случайную величину:

$$D(X) = M[(X - M(X))^2].$$

Так как $M(X) = \bar{x}_r$ и $P\{X = x_i\} = \frac{1}{N}$ (см. п. 2), то

$$D(X) = (x_1 - \bar{x}_r)^2 \cdot \frac{1}{N} + (x_2 - \bar{x}_r)^2 \cdot \frac{1}{N} + \dots + (x_N - \bar{x}_r)^2 \cdot \frac{1}{N} = D_r,$$

т. е.

$$D(X) = D_r.$$

Таким образом, дисперсия $D(X)$ равна D_r .

Такой же итог следует, если значения $x_1, x_2, \dots, x_k$ имеют соответственно частоты $N_1, N_2, \dots, N_k$.

В случае непрерывного распределения признака X по определению полагают:

$$D_r = D(X). \quad (8)$$

С учетом формулы (8) формула (5) (п. 2) переписывается в виде:

$$D(\bar{X}) = \frac{D_r}{n},$$

откуда $\sqrt{D(\bar{X})} = \frac{\sqrt{D_r}}{\sqrt{n}}$ или $\sigma(\bar{X}) = \frac{\sigma_r}{\sqrt{n}}$. Величина $\sigma(\bar{X})$ называется *средней квадратической ошибкой*.

Для того чтобы охарактеризовать рассеяние наблюдаемых значений количественного признака выборки вокруг своего среднего значения $\bar{x}_B$ вводят нижеследующую характеристику.

О п р е д е л е н и е 2. *Выборочной дисперсией* D_B называется среднее арифметическое квадратов отклонений наблюдаемых значений признака X от выборочной средней $\bar{x}_B$.

Если все значения $x_1, x_2, \dots, x_n$ признака выборки объема n различны, то

$$D_B = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_B)^2. \quad (9)$$

Если же значения признака $x_1, x_2, \dots, x_k$ имеют соответственно частоты $n_1, n_2, \dots, n_k$, причем $n_1 + n_2 + \dots + n_k = n$, то

$$D_B = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x}_B)^2 n_i. \quad (10)$$

П р и м е р 2. Выборочная совокупность задана таблицей распределения:

x_i	1	2	3	4
n_i	20	15	10	5

Найти выборочную дисперсию. Согласно формулам (4) и (10) имеем:

$$\begin{aligned} \bar{x}_B &= \frac{1 \cdot 20 + 2 \cdot 15 + 3 \cdot 10 + 4 \cdot 5}{20 + 15 + 10 + 5} = \frac{100}{50} = 2, \\ D_B &= \frac{(1-2)^2 \cdot 20 + (2-2)^2 \cdot 15 + (3-2)^2 \cdot 10 + (4-2)^2 \cdot 5}{50} = \frac{50}{50} = 1. \end{aligned}$$

Выборочным средним квадратическим отклонением (стандартом) называется квадратный корень из выборочной дисперсии:

$$\sigma_B = \sqrt{D_B}.$$

В условиях примера 2 получаем, что $\sigma_B = \sqrt{D_B} = \sqrt{1} = 1$.

Ниже, не уменьшая общности рассуждений, будем считать значения $x_1, x_2, \dots, x_n$ признака различными.

Выборочную дисперсию, рассматриваемую нами как случайную величину, будем обозначать $\tilde{S}^2$:

$$\tilde{S}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Справедлива следующая теорема.

Т е о р е м а. *МО выборочной дисперсии равно $\frac{n-1}{n} D_r$, т. е.*

$$M(\tilde{S}^2) = \frac{n-1}{n} D_r.$$

Доказательство. С учетом свойств MO (гл. 18) получаем:

$$M(\tilde{S}^2) = M\left[\frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2\right] = \frac{1}{n}\sum_{i=1}^n M[(X_i - \bar{X})^2].$$

Вычислим одно слагаемое $M[(X_i - \bar{X})^2]$. Имеем:

$$M[(X_i - \bar{X})^2] = M(X_i^2 - 2X_i\bar{X} + \bar{X}^2) = M(X_i^2) - 2M(X_i\bar{X}) + M(\bar{X}^2).$$

Вычислим по отдельности эти MO . Согласно свойству 1 дисперсии (гл. 18) и формулам (2), (8) имеем:

$$M(X_i^2) = M(X^2) = D(X) + M^2(X) = D_r + a^2.$$

Далее, с учетом свойства 4 MO

$$\begin{aligned} M(X_i\bar{X}) &= M\left[X_i \cdot \frac{1}{n}(X_1 + X_2 + \dots + X_n)\right] = \\ &= \frac{1}{n}[M(X_iX_1) + M(X_iX_2) + \dots + M(X_iX_n)]. \end{aligned}$$

То слагаемое этой суммы, у которого второй индекс равен i , т.е. $M(X_iX_i)$, равно $M(X_i^2) = D_r + a^2$. У всех остальных слагаемых $M(X_iX_j)$ индексы разные. Поэтому в силу независимости X_i и X_j (см. гл. 18)

$$M(X_iX_j) = M(X_i)M(X_j) = M(X)M(X) = M^2(X) = a^2.$$

Так как имеем $n - 1$ таких слагаемых, то

$$M(X_i\bar{X}) = \frac{1}{n}[D_r + a^2 + (n-1)a^2] = a^2 + \frac{D_r}{n}.$$

В силу свойства 1 дисперсии

$$M(\bar{X}^2) = D(\bar{X}) + M^2(\bar{X}).$$

Нами уже найдены (п. 2 и п. 3)

$$M(\bar{X}) = M(X) = a, D(\bar{X}) = \frac{D_r}{n}.$$

Поэтому

$$M(\bar{X}^2) = \frac{D_r}{n} + a^2.$$

Таким образом,

$$M[(X_i - \bar{X})^2] = D_r + a^2 - 2\left(a^2 + \frac{D_r}{n}\right) + \frac{D_r}{n} + a^2 = \frac{n-1}{n}D_r$$

и не зависит от индекса суммирования. Поэтому

$$M(\tilde{S}^2) = \frac{1}{n} \cdot n \frac{n-1}{n} D_r = \frac{n-1}{n} D_r.$$

Теорема доказана.

В заключение настоящего пункта отметим, что если варианты x_i — большие числа, то для облегчения вычисления выборочной дисперсии D_B формулу (9) преобразуют к следующему виду:

$$D_B = \frac{1}{n} \sum_{i=1}^n (x_i - C)^2 - (\bar{x}_B - C)^2, \quad (11)$$

где C — ложный нуль.

Действительно, с учетом формулы (3) имеем:

$$\begin{aligned} \sum_{i=1}^n (x_i - C)^2 - n(\bar{x}_B - C)^2 &= \sum_{i=1}^n x_i^2 - 2C \sum_{i=1}^n x_i + nC^2 - n\bar{x}_B^2 + 2nC\bar{x}_B - nC^2 = \\ &= \sum_{i=1}^n x_i^2 - 2nC\bar{x}_B - n\bar{x}_B^2 + 2nC\bar{x}_B = \sum_{i=1}^n x_i^2 - n\bar{x}_B^2 = \sum_{i=1}^n x_i^2 - 2n\bar{x}_B^2 + n\bar{x}_B^2 = \\ &= \sum_{i=1}^n x_i^2 - 2\bar{x}_B \sum_{i=1}^n x_i + n\bar{x}_B^2 = \sum_{i=1}^n (x_i^2 - 2x_i\bar{x}_B + \bar{x}_B^2) = \sum_{i=1}^n (x_i - \bar{x}_B)^2, \end{aligned}$$

откуда

$$D_B = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_B)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - C)^2 - (\bar{x}_B - C)^2.$$

Пр и м е р 3. Для выборки, указанной в примере 2 из п. 2, имеем (ложный нуль остается прежним $C = 72,00$):

$$\begin{aligned} \sum_{i=1}^{10} (x_i - C)^2 &= \sum_{i=1}^{10} \alpha_i^2 = 0,0144 + 0,0049 + 0,0025 + 0,0049 + 0,0100 + 0,0004 + \\ &+ 0,0049 + 0,0529 + 0,0121 + 0,0016 = 0,1086; \\ (\bar{x}_B - C)^2 &= (-0,038)^2 \approx 0,0014. \end{aligned}$$

Наконец, согласно формуле (11)

$$D_B \approx \frac{1}{10} \cdot 0,1086 - 0,0014 = 0,0094$$

4. Оценки параметров распределения. Уже говорилось (п. 1) о том, что одной из задач статистики является оценка параметров распределения случайной величины X по данным выборки. При этом в теоретических рассуждениях считают, что генеральная совокупность бесконечна. Это делается для того, чтобы можно было переходить к пределу при $n \rightarrow \infty$, где n — объем выборки. Для оценки параметров распределения X из данных выборки составляют выражения, которые должны служить оценками неизвестных параметров. Например, $\bar{X}$ (п. 2) является оценкой генеральной средней, а $\tilde{S}^2$ (п. 3) — оценкой генеральной дисперсии D_r . Обозначим через Θ оцениваемый параметр, через $\hat{\Theta}_n$ — оценку этого параметра ($\hat{\Theta}_n$ является выражением, составленным из $X_1, X_2, \dots, X_n$ (см. п. 1)). Для того чтобы оценка $\hat{\Theta}_n$ давала хорошее приближение, она должна удовлетворять определенным требованиям. Укажем эти требования.

Несмещенной называют оценку $\hat{\Theta}_n$, *МО* которой равно оцениваемому параметру Θ , т. е. $M(\hat{\Theta}_n) = \Theta$, в противном случае оценка называется *смещенной*.

Пр и м е р 1. Оценка $\bar{X}$ является несмещенной оценкой генеральной средней a , так как $M(\bar{X}) = a$ (см. п. 2).

Пр и м е р 2. Оценка $\tilde{S}^2$ является смещенной оценкой генеральной дисперсии D_r , так как согласно установленной выше теореме (п. 3)

$$M(\tilde{S}^2) = \frac{n-1}{n} D_r \neq D_r.$$

Пр и м е р 3. Наряду с выборочной дисперсией $\tilde{S}^2$ рассматривают еще так называемую исправленную дисперсию $S^2 = \frac{n-1}{n} \tilde{S}^2$, которая является также оценкой генеральной дисперсии. Для S^2 с учетом установленной выше теоремы (п. 3) имеем:

$$M(S^2) = M\left(\frac{n}{n-1} \tilde{S}^2\right) = \frac{n}{n-1} M(\tilde{S}^2) = \frac{n}{n-1} \cdot \frac{n-1}{n} D_r = D_r.$$

Таким образом, оценка S^2 в отличие от оценки $\tilde{S}^2$ является несмещенной оценкой генеральной дисперсии. Явное выражение для S^2 имеет вид:

$$S^2 = \frac{n}{n-1} \tilde{S}^2 = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

т. е.

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2. \quad (12)$$

Естественно в качестве приближенного неизвестного параметра брать несмещенные оценки, для того чтобы не делать систематической ошибки в сторону завышения или занижения.

Состоятельной называют такую оценку $\hat{\Theta}_n$ параметра Θ , что для любого наперед заданного числа $\varepsilon > 0$ вероятность $P\{|\hat{\Theta}_n - \Theta| < \varepsilon\}$ при $n \rightarrow \infty$ стремится к 1. В таком случае говорят, что $\hat{\Theta}_n$ сходится к Θ по вероятности. Это значит, что при достаточно больших n можно с вероятностью, близкой к 1, т. е. почти наверное, утверждать, что оценка $\hat{\Theta}_n$ отличается от оцениваемого параметра Θ меньше чем на ε .

Очевидно, такому требованию должна удовлетворять всякая оценка, пригодная для практического использования.

Заметим, что несмещенная оценка $\hat{\Theta}_n$ будет состоятельной, если при $n \rightarrow \infty$ ее дисперсия стремится к нулю: $D(\hat{\Theta}_n) \rightarrow 0$. Это следует из неравенства Чебышева (5) (см. § 18.9, п.1).

Пр и м е р 4. Как установлено выше (см. п. 3), $D(\bar{X}) = \frac{D_r}{n}$. Отсюда следует, что несмещенная оценка $\bar{X}$ является и состоятельной, так как

$$\lim_{n \rightarrow \infty} D(\bar{X}) = \lim_{n \rightarrow \infty} \frac{D_r}{n} = D_r \lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Можно показать, что несмещенная оценка S^2 является также состоятельной. Поэтому в качестве оценки генеральной дисперсии принимают исправленную дисперсию. Заметим, что оценки S^2 и $\tilde{S}^2$ отличаются множителем $\frac{n}{n-1}$, который стремится к 1 при $n \rightarrow \infty$. На практике $\tilde{S}^2$ и S^2 не различают при $n > 30$.

Для оценки генерального среднего квадратического отклонения используют исправленное среднее квадратическое отклонение, которое равно квадратному корню из исправленной дисперсии:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}. \quad (13)$$

Левые части формул (12), (13), в которых случайные величины $X_1, X_2, \dots, X_n$ заменены их реализациями $x_1, x_2, \dots, x_n$ и $\bar{X}$ — выборочной средней $\bar{x}_B$, будем обозначать соответственно через s^2 и s .

Отметим, что если варианты x_i — большие числа, то для облегчения вычисления s^2 формулу для s^2 аналогично формуле (9) преобразуют к виду:

$$s^2 = \frac{n}{n-1} \left[\frac{1}{n} \sum_{i=1}^n (x_i - C)^2 - (\bar{x}_B - C)^2 \right], \quad (14)$$

где C — ложный нуль.

Оценки, обладающие свойствами несмещенности и состоятельности, при ограниченном числе опытов могут отличаться дисперсиями.

Ясно, что, чем меньше дисперсия оценки, тем меньше вероятность грубой ошибки при определении приближенного значения параметра. Поэтому необходимо, чтобы дисперсия оценки была минимальной. Оценка, обладающая таким свойством, называется *эффективной*.

Из отмеченных требований, предъявляемых к оценке, наиболее важными являются требования несмещенности и состоятельности.

П р и м е р 5. С плодового дерева случайным образом отобрано 10 плодов. Их веса $x_1, x_2, \dots, x_{10}$ (в граммах) записаны в первой колонке приведенной ниже таблицы. Обработаем статистические данные выборки. Для вычисления $\bar{x}_B$ и s по формулам (6) и (14) введем ложный нуль $C = 250$ и все необходимые при этом вычисления сведем в указанную таблицу:

i	x_i	$x_i - C$	$(x_i - C)^2$
1	225	− 25	625
2	274	24	576
3	305	55	3025
4	253	3	9
5	220	− 30	900
6	245	− 5	25
7	211	− 39	1521
8	234	− 16	256
9	230	− 20	400
10	231	− 19	261
Сумма		− 72	7598

Следовательно.

$$\bar{x}_B = 250 + \frac{1}{10} \sum_{i=1}^{10} (x_i - 250) = 250 + \frac{1}{10} (-72) = 250 - 7,2 \approx 243 \text{ (г);}$$

$$s = \sqrt{\frac{10}{9} \left[\frac{1}{10} \sum_{i=1}^{10} (x_i - 250)^2 - (250 - 7,2 - 250)^2 \right]} = \sqrt{\frac{10}{9} [759,8 - (-7,2)^2]} \approx 28 \text{ г.}$$

Отсюда: $\frac{s}{\sqrt{10}} \approx 9 \text{ г.}$

Итак, оценка генеральной средней веса плода равна 243 г со средней квадратической ошибкой 9 г.

Оценка генерального среднего квадратического отклонения веса плода равна 28 г.

П р и м е р 6. Через каждый час измерялось напряжение тока в электросети. Результаты измерений в вольтах представлены в следующей таблице:

i	1	2	3	4	5	6	7	8	9	10	11	12
x_i	222	219	224	220	218	217	221	220	215	218	223	225

i	13	14	15	16	17	18	19	20	21	22	23	24
x_i	220	226	221	216	211	219	220	221	222	218	221	219

Найти оценки для математического ожидания и дисперсии результатов измерений. Оценки для математического ожидания и дисперсии найдем по формулам (6) и (14), положив $C=220$. Все необходимые вычисления приведены в нижеследующей таблице:

i	$x_i - C$	$(x_i - C)^2$	i	$x_i - C$	$(x_i - C)^2$	i	$x_i - C$	$(x_i - C)^2$
1	2	4	9	-5	25	17	1	1
2	-1	1	10	-2	4	18	-1	1
3	4	16	11	3	9	19	0	0
4	0	0	12	5	25	20	1	1
5	-2	4	13	0	0	21	2	4
6	-3	9	14	6	36	22	-2	4
7	1	1	15	1	1	23	1	1
8	0	0	16	-4	16	24	-1	1
Сумма	1	35		4	116		1	13

Следовательно,

$$\bar{x}_n = 220 + \frac{1}{24} \sum_{i=1}^{24} (x_i - 220) = 220 + \frac{6}{24} = 220,25 \text{ В};$$

$$s^2 = \frac{24}{23} \left[\frac{1}{24} \sum_{i=1}^{24} (x_i - 220)^2 - (220,25 - 220)^2 \right] \approx 7,06 \text{ В}.$$

§ 19.3. Доверительные интервалы для параметров нормального распределения

1. Надежность. Доверительные интервалы. Пусть Θ — оцениваемый параметр, $\tilde{\Theta}_n$ — его оценка, составленная из $X_1, X_2, \dots, X_n$.

Если известно, что оценка $\tilde{\Theta}_n$ является несмещенной и состоятельной, то по данным выборки вычисляют значение $\tilde{\Theta}_n$ и считают его приближением истинного значения Θ . При этом среднее квадратическое отклонение (если его вообще вычисляют) оценивает порядок ошибки. Такие оценки называют *точечными*. Например, в предыдущем параграфе речь шла о точечных оценках генеральной средней и генеральной дисперсии. В общем случае, когда о распределении признака X ничего неизвестно, это уже немало.

Если же о распределении имеется какая-либо информация, то можно сделать больше.

В данном параграфе речь будет идти об оценке параметров μ и σ случайной величины, имеющей нормальное распределение. Это очень важный случай. Например (см. §18.10), результат измерения имеет нормальное распределение. В этом случае становится возможным применять так называемое интервальное оценивание, к изложению которого мы и переходим.

Пусть $\delta > 0$ — некоторое число. Если выполняется неравенство $|\Theta - \tilde{\Theta}_n| < \delta$, т. е. $-\delta < \Theta - \tilde{\Theta}_n < \delta$, что можно записать в виде: $\tilde{\Theta}_n - \delta < \Theta < \tilde{\Theta}_n + \delta$, то говорят, что интервал $(\tilde{\Theta}_n - \delta, \tilde{\Theta}_n + \delta)$ покрывает параметр Θ . Однако невозможно указать оценку $\tilde{\Theta}_n$, такую, чтобы событие $\{|\Theta - \tilde{\Theta}_n| < \delta\}$ было достоверным, поэтому мы будем говорить о вероятности этого события. Число δ называется *точностью* оценки $\tilde{\Theta}_n$.

О п р е д е л е н и е. *Надежностью (доверительной вероятностью) оценки $\tilde{\Theta}_n$ параметра Θ для заданного $\delta > 0$ называется вероятность γ того, что интервал $(\tilde{\Theta}_n - \delta, \tilde{\Theta}_n + \delta)$ покрывает параметр Θ , т. е.*

$$\gamma = P\{\tilde{\Theta}_n - \delta < \Theta < \tilde{\Theta}_n + \delta\} = P\{|\tilde{\Theta}_n - \Theta| < \delta\}.$$

Заметим, что после того, как по данным выборки вычислена оценка $\tilde{\Theta}_n$, событие $\{|\tilde{\Theta}_n - \Theta| < \delta\}$ становится или достоверным, или невозможным, так как интервал $(\tilde{\Theta}_n - \delta, \tilde{\Theta}_n + \delta)$ или покрывает Θ , или нет. Но дело в том, что параметр Θ нам неизвестен. Поэтому мы называем на-

дежностью γ уже вычисленной оценки $\tilde{\Theta}_n$ вероятность того, что интервал $(\tilde{\Theta}_n - \delta, \tilde{\Theta}_n + \delta)$ найденный для произвольной выборки, покрывает Θ . Если мы сделаем много выборок объема n и для каждой из них построим интервал $(\tilde{\Theta}_n - \delta, \tilde{\Theta}_n + \delta)$, то доля тех выборок, чьи интервалы покрывают Θ , равна γ .

Иными словами γ есть мера нашего доверия вычисленной оценке $\tilde{\Theta}_n$.

Ясно, что, чем меньше число δ , тем меньше надежность γ .

О п р е д е л е н и е. *Доверительным интервалом* называется найденный по данным выборки интервал $(\tilde{\Theta}_n - \delta, \tilde{\Theta}_n + \delta)$, который покрывает параметр Θ с заданной надежностью γ .

Надежность γ обычно принимают равной 0,95, или 0,99, или 0,999.

Конечно, нельзя категорически утверждать, что найденный доверительный интервал покрывает параметр Θ . Но в этом можно быть уверенным на 95% при $\gamma = 0,95$, на 99% при $\gamma = 0,99$ и т. д. Это значит, что если сделать много выборок, то для 95% из них (если, например, $\gamma = 0,95$) вычисленные доверительные интервалы действительно покрывают Θ .

2. Доверительный интервал для МО при известном σ . В некоторых случаях среднее квадратическое отклонение σ ошибки измерения (а вместе с нею и самого измерения) бывает известно. Например, если измерения производятся одним и тем же прибором при одних и тех же условиях, то σ для всех измерений одно и то же и обычно бывает известно.

Итак, пусть случайная величина X распределена нормально с параметрами a и σ , причем σ известно. Построим доверительный интервал, покрывающий неизвестный параметр a с заданной надежностью γ . Данные выборки есть реализации случайных величин $X_1, X_2, \dots, X_n$, имеющих нормальное распределение с параметрами a и σ (§19.2, п. 1). Оказывается, что и выборочная средняя случайная величина $\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$ тоже имеет нормальное распределение (это мы примем без доказательства). При этом (см. §19.2, пп. 2, 3):

$$M(\bar{X}) = a, \quad \sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}.$$

Потребуем, чтобы выполнялось соотношение $P(|\bar{X} - a| < \delta) = \gamma$, где γ — заданная надежность. Пользуясь формулой (21) (§18.8, п. 5), получим:

$$P(|\bar{X} - a| < \delta) = 2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right)$$

или

$$P(|\bar{X} - a| < \delta) = 2\Phi(t),$$

где

$$t = \frac{\delta\sqrt{n}}{\sigma}. \quad (1)$$

Найдя из равенства (1) $\delta = \frac{t\sigma}{\sqrt{n}}$, можем написать:

$$P\left(|\bar{X} - a| < \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t).$$

Так как P задана и равна γ , то окончательно имеем (для получения рабочей формулы выборочную среднюю заменяем на $\bar{x}_B$):

$$P\left(\bar{x}_B - \frac{t\sigma}{\sqrt{n}} < a < \bar{x}_B + \frac{t\sigma}{\sqrt{n}}\right) = 2\Phi(t) = \gamma.$$

Смысл полученного соотношения таков: *с надежностью γ можно утверждать, что доверительный интервал $\left(\bar{x}_B - \frac{t\sigma}{\sqrt{n}}, \bar{x}_B + \frac{t\sigma}{\sqrt{n}}\right)$ покрывает неизвестный параметр a ; точность оценки $\delta = \frac{t\sigma}{\sqrt{n}}$. Здесь число t определяется из равенства $\Phi(t) = \frac{\gamma}{2}$ (оно следует из $2\Phi(t) = \gamma$) по таблице приложения 4.*

Как уже упоминалось, надежность γ обычно принимают равной или 0,95, или 0,99, или 0,999.

П р и м е р. Признак X распределен в генеральной совокупности нормально с известным $\sigma = 0,40$. Найти по данным выборки доверительный интервал для a с надежностью $\gamma = 0,99$, если $n = 20$, $\bar{x}_B = 6,34$.

Для $\Phi(t) = \frac{\gamma}{2} = \frac{0,99}{2} = 0,495$ находим по таблице приложения 4 $t = 2,58$. Следовательно, $\delta = 2,58 \cdot \frac{0,40}{\sqrt{20}} \approx 0,23$. Концы доверительного интервала $6,34 - 0,23 = 6,11$ и $6,34 + 0,23 = 6,57$. Итак, доверительный интервал $(6,11; 6,57)$ покрывает a с надежностью 0,99.

3. Доверительный интервал для МО при неизвестном σ . Пусть случайная величина X имеет нормальное распределение с неизвестными нам параметрами a и σ . Оказывается, что случайная величина (ее возможные значения будем обозначать через t):

$$T = \frac{\bar{X} - a}{\frac{S}{\sqrt{n}}},$$

где n — объем выборки, $\bar{X}$ — выборочная средняя, S — исправленное среднее квадратическое отклонение, имеет распределение, не зависящее от a и σ . Оно называется распределением Стьюдента*.

* Стьюдент — псевдоним английского статистика Госсета.

Плотность вероятности распределения Стьюдента дается формулой

$$S(t, n) = B_n \left(1 + \frac{t^2}{n-1} \right)^{-\frac{n}{2}},$$

где коэффициент B_n зависит от объема выборки.

Потребуем, чтобы выполнялось соотношение

$$P(|T| < t_\gamma) = \gamma,$$

где γ — заданная надежность.

Так как $S(t, n)$ — четная функция от t , то, пользуясь формулой (9) (§18.8), получим:

$$P\left(\frac{|\bar{X} - a|\sqrt{n}}{S} < t_\gamma\right) = 2 \int_0^{t_\gamma} S(t, n) dt = \gamma.$$

Отсюда:

$$P\left(\bar{X} - \frac{t_\gamma S}{\sqrt{n}} < a < \bar{X} + \frac{t_\gamma S}{\sqrt{n}}\right) = \gamma.$$

Следовательно, приходим к утверждению: *с надежностью γ можно утверждать, что доверительный интервал $\left(\bar{x}_B - \frac{t_\gamma s}{\sqrt{n}}, \bar{x}_B + \frac{t_\gamma s}{\sqrt{n}}\right)$ покрывает неизвестный параметр a ; точность оценки $\delta = \frac{t_\gamma s}{\sqrt{n}}$.* Здесь случайные величины $\bar{X}$ и S заменены неслучайными величинами $\bar{x}_B$ и s , найденными по выборке.

В приложении 5 приведена таблица значений $t_\gamma = t(\gamma, n)$ для различных значений n и обычно задаваемых значений надежности.

Заметим, что при $n \geq 30$ распределение Стьюдента практически не отличается от нормированного нормального распределения (§18.8, п. 5).

Это связано с тем, что $\lim_{n \rightarrow \infty} \left(1 + \frac{t^2}{n-1} \right)^{-\frac{n}{2}} = e^{-\frac{t^2}{2}}$.

Пример. Признак X распределен в генеральной совокупности нормально. Найти доверительный интервал для $\bar{x}_r$ с надежностью $\gamma = 0,99$, если $n = 20$, $\bar{x}_B = 6,34$, $s = 0,40$. Для надежности $\gamma = 0,99$ и $n = 20$ находим по таблице приложения 5 $t_\gamma = 2,861$. Следовательно, $\delta = 2,861 \cdot \frac{0,40}{\sqrt{20}} \approx 0,26$. Концы доверительного интервала $6,34 - 0,26 = 6,08$ и $6,34 + 0,26 = 6,60$. Итак, доверительный интервал $(6,08; 6,60)$ покрывает $\bar{x}_r$ с надежностью 0,99.

4. Доверительный интервал для среднего квадратического отклонения.

Для нахождения доверительного интервала для среднего квадратиче-

ского отклонения σ будем использовать следующее предложение, устанавливаемое аналогично двум предыдущим (п. 2 и 3).

С надежностью γ можно утверждать, что доверительный интервал $(s - sq; s + sq)$ покрывает неизвестный параметр σ ; точность оценки $\delta = sq$.

В приложении 6 приведена таблица значений $q = q(\gamma, n)$ для различных значений n и обычно задаваемых значений надежности γ .

П р и м е р 1. Признак X распределен в генеральной совокупности нормально. Найти доверительный интервал для σ_r с надежностью $\gamma = 0,95$, если $n = 20$, $s = 0,40$.

Для надежности $\gamma = 0,95$ и $n = 20$ находим в таблице приложения 6 $q = 0,37$. Далее $sq = 0,40 \cdot 0,37 \approx 0,15$. Концы доверительного интервала $0,40 - 0,15 = 0,25$ и $0,40 + 0,15 = 0,55$. Итак, доверительный интервал $(0,25; 0,55)$ покрывает σ_r с надежностью 0,95.

П р и м е р 2. На ферме испытывалось влияние витаминов на прибавку в массе телят. Для этой цели было осмотрено 20 телят одного возраста. Средняя масса их оказалась равной 340 кг, а «исправленное» среднее квадратическое отклонение — 20 кг.

Определить: 1) доверительный интервал для математического ожидания a с надежностью 0,95; 2) доверительный интервал для среднего квадратического отклонения с той же надежностью.

При решении задачи исходить из предположения, что данные пробы взяты из нормальной генеральной совокупности.

Р е ш е н и е. 1) Согласно условиям задачи $x_a = 340$; $s = 20$; $\gamma = 0,95$; $n = 20$. Находим в таблице приложения 5 $t_\gamma = 2,093$. Следовательно, $\delta = 2,093 \cdot \frac{20}{\sqrt{20}} \approx 9,4$. Концы доверительного интервала $340 - 9,4 = 330,6$ и $340 + 9,4 = 349,4$. Итак, доверительный интервал $(330,6; 349,4)$ покрывает a с надежностью 0,95.

Можно считать, что в данном случае истинная масса измерена достаточно точно (отклонение порядка $\frac{9,4}{340} \approx 0,03$).

2) Для надежности $\gamma = 0,95$ и $n = 20$ находим в таблице приложения 6 $q = 0,37$. Далее $sq = 20 \cdot 0,37 = 7,4$. Концы доверительного интервала $20 - 7,4 = 12,6$ и $20 + 7,4 = 27,4$. Таким образом, $12,6 < \sigma < 27,4$, откуда можно заключить, что σ определено неудовлетворительно (отклонение порядка $\frac{sq}{s} = q \approx 0,4$ — почти половина!). Чтобы сузить доверительный интервал при той же надежности, необходимо увеличить число проб n .

П р и м е ч а н и е. Выше предполагалось, что $q < 1$. Если $q > 1$, то, учитывая, что $\sigma > 0$, получаем $0 < \sigma < s + sq$. Значения q и в этом случае определяются по таблице приложения 6.

П р и м е р 3. Признак X генеральной совокупности распределен нормально. По выборке объема $n = 10$ найдено «исправленное» среднее квадратическое отклонение $s = 0,16$. Найти доверительный интервал для σ_r надежностью 0,999.

Для надежности $\gamma = 0,999$ и $n = 10$ по таблице приложения 6 находим $q = 1,80$.

Следовательно, искомый доверительный интервал таков:

$$0 < \sigma_r < 0,16 + 0,16 \cdot 1,80$$

или

$$0 < \sigma_r < 0,448.$$

5. Оценка истинного значения измеряемой величины. Пусть производится n независимых равнооточных измерений некоторой физической величины, истинное значение a которой неизвестно. Будем рассматривать результаты отдельных измерений как случайные величины $X_1, X_2, \dots, X_n$. Эти величины независимы (измерения независимы), имеют одно и то же математическое ожидание a (истинное значение измеряемой величины), одинаковые дисперсии σ^2 (измерения равнооточны) и распределены нормально (такое допущение подтверждается опытом). Таким образом, все предположения, которые были сделаны при выводе доверительных интервалов в пунктах 2 и 3 настоящего параграфа, выполняются, следовательно, мы вправе использовать полученные в них предложения. Так как обычно σ неизвестно, следует пользоваться предложением, найденным в пункте 3 данного параграфа.

П р и м е р. По данным девяти независимых равнооточных измерений физической величины найдены среднее арифметическое результатов отдельных измерений $\bar{x}_B = 42,319$ и «исправленное» среднее квадратическое отклонение $s = 5,0$. Требуется оценить истинное значение a измеряемой величины с надежностью $\gamma = 0,99$.

Истинное значение измеряемой величины равно ее математическому ожиданию. Поэтому задача сводится к оценке математического ожидания (при неизвестном σ) при помощи доверительного интервала

$$\bar{x}_B - \frac{t_\gamma s}{\sqrt{n}} < a < \bar{x}_B + \frac{t_\gamma s}{\sqrt{n}},$$

покрывающего a с заданной надежностью $\gamma = 0,99$.

Пользуясь таблицей приложения 5 по $\gamma = 0,99$ и $n = 9$, находим $t_\gamma = 3,36$.

Найдем точность оценки:

$$\delta = \frac{t_\gamma s}{\sqrt{n}} = 3,36 \cdot \frac{5}{\sqrt{9}} = 3,36 \cdot \frac{5}{3} = 5,60.$$

Концы доверительного интервала

$$42,319 - 5,60 = 36,719$$

и

$$42,319 + 5,60 = 47,919.$$

Итак, с надежностью $\gamma = 0,99$ истинное значение измеряемой величины a заключено в доверительном интервале $36,719 < a < 47,919$.

6. Оценка точности измерений. В теории ошибок принято точность измерений (точность прибора) характеризовать с помощью среднего квадратического отклонения σ случайных ошибок измерений. Для оценки σ используют «исправленное» среднее квадратическое отклонение s . Поскольку обычно результаты измерений независимы, имеют одно и то же математическое ожидание (истинное значение измеряемой величины) и одинаковую дисперсию (в случае равноточных измерений), то утверждение, приведенное в п. 4, применимо для оценки точности измерений.

Пример. По 16 независимым равноточным измерениям найдено «исправленное» среднее квадратическое отклонение $s = 0,4$. Найти точность измерений с надежностью $\gamma = 0,99$.

Как отмечено выше, точность измерений характеризуется средним квадратическим отклонением σ случайных ошибок измерений. Поэтому задача сводится к отысканию доверительного интервала $(s - sq; s + sq)$, покрывающего σ с заданной надежностью $\gamma = 0,99$ (см. п. 4). По таблице приложения 6 по $\gamma = 0,99$ и $n = 16$ найдем $q = 0,70$. Следовательно, искомый доверительный интервал таков:

$$0,4(1 - 0,70) < \sigma < 0,4(1 + 0,70)$$

или

$$0,12 < \sigma < 0,68.$$

§ 19.4. Проверка статистических гипотез

Пусть по выборке объема n получено эмпирическое распределение:

Варианты	x_1	x_2	...	x_m
Эмпирические (наблюдаемые) частоты	n_1	n_2	...	n_m

По данным наблюдения выдвигают гипотезу о законе распределения генеральной совокупности, например, предполагают, что генеральная совокупность распределена равномерно или нормально. Такие гипотезы называются *статистическими*. Затем для тех же объектов, которые попали в выборку, вычисляют частоты, уже исходя из теоретической гипотезы. В результате получаются частоты (их называют *выравнивающими частотами*), которые, вообще говоря, отличаются от наблюдавшихся. Как определить, правильно или нет выдвинута гипотеза, т. е. случайны ли расхождения наблюдавшихся и выравнивающих частот или эти расхождения являются следствием неправильности гипотезы? Для решения этого вопроса применяют критерии согласия

эмпирических наблюдений выдвинутой гипотезе. Имеется несколько критериев согласия: χ^2 («хи квадрат») К. Пирсона, Колмогорова, Смирнова и др. Мы познакомимся с критерием согласия χ^2 («хи квадрат») Пирсона.

Предположим, что на основе приведенного выше распределения выдвинута гипотеза H : генеральная совокупность имеет нормальное распределение. Для вычисления выравнивающих частот поступают следующим образом:

1) находят значения $\bar{x}_B, \sigma_B = \sqrt{D_B}$;

2) выравнивающие частоты n'_i ищут по формуле

$$n'_i = \frac{nh}{\sigma_B} \varphi(u_i),$$

где n — сумма наблюдавшихся частот, h — разность между двумя соседними вариантами, $u_i = \frac{x_i - \bar{x}_B}{\sigma_B}$ и $\varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$.

В результате получено множество выравнивающих частот:

$$n'_1, n'_2, \dots, n'_m.$$

Обозначим через χ^2 сумму квадратов разностей между эмпирическими и выравнивающими частотами, деленных на соответствующие выравнивающие частоты:

$$\chi^2 = \sum_{i=1}^m \frac{(n_i - n'_i)^2}{n'_i}. \quad (1)$$

Для данной выборки по формуле (1) находим значение случайной величины χ^2 . Обозначим его через χ_0^2 . Затем определяется число $k = m - 3$, называемое *числом степеней свободы*, где m — число различных вариантов выборки.

Теперь проверка гипотезы H проводится так. Задаются уровнем значимости p , т. е. столь малой вероятностью p , при которой о событии $\{\chi_0^2 > \chi^2\}$, имеющем вероятность p , можно с большой уверенностью сказать, что в единичном испытании оно не произойдет. В таблице значений χ^2 по заданному уровню значимости p и числу степеней свободы k (приложение 7) находят значение $\chi^2(p; k)$. Если окажется, что $\chi_0^2 > \chi^2(p; k)$, то гипотеза H отвергается на уровне значимости p , так как произошло событие, которое не должно было произойти при верной гипотезе H ; если же $\chi_0^2 < \chi^2$, то H принимается на уровне значимости p . Обычно в качестве p берут либо 0,05, либо 0,01, либо 0,001.

* Из этой формулы видно, что чем меньше различие между эмпирическими и выравнивающими частотами, тем меньше будет χ^2 .

П р и м е р. При уровне значимости 0,05 проверить гипотезу о нормальном распределении генеральной совокупности, если известны эмпирические и теоретические частоты:

эмпирические частоты... 6 13 38 74 106 85 30 14

теоретические частоты... 3 14 42 82 99 76 37 13

Вычислим χ_0^2 , для чего составим расчетную таблицу:

i	n_i	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2$	$\frac{(n_i - n'_i)^2}{n'_i}$
1	6	3	3	9	3
2	13	14	- 1	1	0,07
3	38	42	- 4	16	0,38
4	74	82	- 8	64	0,78
5	106	99	7	49	0,49
6	85	76	9	81	1,07
7	30	37	- 7	49	1,32
8	14	13	1	1	0,08
					$\chi_0^2 = 7,19$

Найдем число степеней свободы, учитывая, что число различных вариантов $m = 8$. Имеем: $k = 8 - 3 = 5$. По уровню значимости $p = 0,05$ и числу степеней свободы $k = 5$ по таблице значений χ^2 (приложение 7) находим: $\chi^2(0,05; 5) = 11,1$. Так как $\chi_0^2 < \chi^2(0,05; 5)$, нет оснований отвергнуть гипотезу H .

§ 19.5. Линейная корреляция

1. Корреляционная зависимость. Часто приходится иметь дело с более сложной зависимостью, чем функциональная. Такова, например, связь между осадками и урожаем или связь между толщиной снегового покрова зимой и объемом стока последующего половодья. Здесь каждому значению одной величины соответствует множество возможных значений другой величины. Подобного рода зависимости относятся к *корреляционным* зависимостям.

О п р е д е л е н и е 1. Две случайные величины X и Y находятся в *корреляционной зависимости*, если каждому значению любой из этих величин соответствует определенное распределение вероятностей другой величины.

О п р е д е л е н и е 2. *Условным математическим ожиданием* (кратко УМО) дискретной случайной величины X при $Y = y$ (y — определенное

возможное значение Y) называется сумма произведений возможных значений величины X на их условные вероятности.

$$M_y(X) = \sum_{i=1}^n x_i P_y(X = x_i),$$

где $P_y(X = x_i)$ — условная вероятность равенства $X = x_i$ при условии, что $Y = y$.

Для непрерывных величин

$$M_y(X) = \int_{-\infty}^{+\infty} x \varphi_y(x) dx,$$

где $\varphi_y(x)$ — плотность вероятности случайной непрерывной величины X при условии $Y = y$.

УМО $M_y(X)$ есть функция от y : $M_y(X) = f(y)$, которую называют функцией *регрессии* величины X на величину Y .

Аналогично определяется УМО случайной величины Y и функция регрессии Y на X :

$$M_x(Y) = g(x).$$

Уравнение $x = f(y)$ ($y = g(x)$) называется *уравнением регрессии* X на Y (Y на X), а линия на плоскости, соответствующая этому уравнению, называется *линией регрессии*.

Линия регрессии Y на X (X на Y) показывает, как в среднем зависит Y от X (X от Y).

Пример 1. X и Y независимы, $M(X) = a$, $M(Y) = b$. Тогда $g(x) = M_x(Y) = M(Y) = b$; $f(y) = M_y(X) = M(X) = a$. Линии регрессии изображены на рисунке 207.

Пример 2. X и Y связаны линейной зависимостью: $Y = AX + B$, $A \neq 0$. Тогда функция регрессии Y на X будет иметь вид:

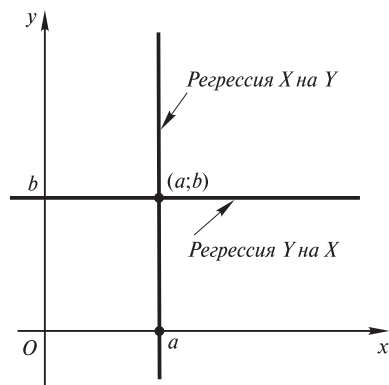


Рис. 207

$$g(x) = M_x(Y) = M(Ax + B) = Ax + B.$$

Так как $X = \frac{1}{A}(Y - B)$, то функция регрессии X на Y имеет вид:

$$f(y) = M_y(X) = M\left[\frac{1}{A}(y - B)\right] = \frac{1}{A}(y - B).$$

Значит, линия регрессии X на Y : $x = \frac{1}{A}(y - B)$,

т. е. $y = Ax + B$. Таким образом, в случае линейной зависимости X и Y линия регрессии X на Y и Y на X совпадают, и эта линия — прямая.

2. Коэффициент корреляции. Для характеристики корреляционной зависимости между случайными величинами вводится понятие коэффициента корреляции.

О п р е д е л е н и е 1. Если X и Y — независимые случайные величины, то (см. § 18.5)

$$M(XY) = M(X)M(Y). \quad (1)$$

Если же X и Y не являются независимыми случайными величинами, то, вообще говоря, $M(XY) \neq M(X)M(Y)$.

Условились за меру связи (зависимости) двух случайных величин X и Y принять безразмерную величину r , определяемую соотношением:

$$r = \frac{M(XY) - M(X)M(Y)}{\sigma(X)\sigma(Y)} \quad (2)$$

или более кратко соотношением:

$$r = \frac{\mu}{\sigma_1 \sigma_2}, \quad (3)$$

где

$$\mu = M(XY) - M(X)M(Y), \quad \sigma_1 = \sigma(X), \quad \sigma_2 = \sigma(Y),$$

и называемую *коэффициентом корреляции*.

Легко видеть, что

$$\mu = M(XY) - M(X)M(Y) = M[(X - M(X))(Y - M(Y))].$$

О п р е д е л е н и е 2. Случайные величины X и Y называются *некоррелированными*, если $r = 0$, и *коррелированными*, если $r \neq 0$.

П р и м е р 1. Независимые случайные величины X и Y являются некоррелированными, так как в силу соотношения (1) $r = 0$.

П р и м е р 2. Пусть случайные величины X и Y связаны линейной зависимостью: $Y = AX + B$, $A \neq 0$. Найдем коэффициент корреляции. Имеем:

$$\begin{aligned} \mu &= M(XY) - M(X)M(Y) = M(AX^2 + BX) - M(X)M(AX + B) = \\ &= AM(X^2) + BM(X) - AM^2(X) - BM(X) = A(M(X^2) - M^2(X)) = A\sigma^2(X) \quad (\text{см. гл. 18}), \\ \sigma^2(Y) &= D(Y) = D(AX + B) = D(AX) = A^2 D(X) = A^2 \sigma^2(X), \end{aligned}$$

откуда:

$$\sigma(Y) = |A|\sigma(X).$$

Поэтому

$$|r| = \frac{|A|\sigma^2(X)}{|A|\sigma^2(X)} = 1.$$

Таким образом, коэффициент корреляции случайных величин, связанных линейной зависимостью, равен ± 1 (точнее, $r = 1$, если $A > 0$, и $r = -1$, если $A < 0$).

Отметим некоторые свойства коэффициента корреляции.

Из примера 1 следует:

1. Если X и Y — независимые случайные величины, то коэффициент корреляции равен нулю.

Заметим, что обратное утверждение, вообще говоря, неверно.

2. Укажем без доказательства, что $|r| \leq 1$. При этом если $|r| = 1$, то между случайными величинами X и Y имеет место функциональная, а именно линейная зависимость.

3. Как видно из формулы (2), коэффициент корреляции характеризует относительную величину отклонения математического ожидания произведения $M(XY)$ от произведения математических ожиданий $M(X)M(Y)$ величин X и Y . Так как это отклонение имеет место только для зависимых величин, то можно сказать, что коэффициент корреляции характеризует тесноту зависимости между X и Y .

3. Линейная корреляция. Этот вид корреляционной зависимости встречается довольно часто.

О п р е д е л е н и е. Корреляционная зависимость между случайными величинами X и Y называется *линейной корреляцией*, если обе функции регрессии $f(y)$ и $g(x)$ являются линейными. В этом случае обе линии регрессии являются прямыми; они называются *прямыми регрессии*.

Выведем уравнения прямой регрессии Y на X , т. е. найдем коэффициенты линейной функции $g(x) = Ax + B$.

Обозначим $M(X) = a$, $M(Y) = b$, $M[(X - a)^2] = \sigma_1^2$, $M[(Y - b)^2] = \sigma_2^2$. С использованием свойств *МО* (гл. 18) находим:

$$M(Y) = M[g(X)] = M(AX + B) = AM(X) + B,$$

т. е. $b = Aa + B$, откуда: $B = b - Aa$.

Далее с помощью тех же свойств *МО* имеем:

$$\begin{aligned} M(XY) &= M[Xg(X)] = M(AX^2 + BX) = \\ &= AM(X^2) + BM(X) = AM(X^2) + (b - Aa)a, \end{aligned}$$

откуда:

$$A = \frac{\mu}{M(X^2) - a^2}$$

или согласно свойству 1 дисперсии (гл. 18)

$$A = \frac{\mu}{\sigma_1^2}.$$

Полученный коэффициент называется *коэффициентом регрессии* Y на X и обозначается через $\rho(Y/X)$.

$$\rho(Y/X) = \frac{\mu}{\sigma_1^2}. \quad (4)$$

Таким образом, уравнение прямой регрессии Y на X имеет вид:

$$y = \rho(Y/X)(x - a) + b. \quad (5)$$

Аналогично можно получить уравнение прямой регрессии X на Y :

$$y = \rho(X/Y)(y - b) + a, \quad (6)$$

где

$$\rho(X/Y) = \frac{\mu}{\sigma_2^2} \quad (7)$$

есть коэффициент регрессии X на Y .

Уравнения прямых регрессии можно записать в более симметричном виде, если воспользоваться коэффициентом корреляции. С учетом этого коэффициента имеем:

$$\rho(Y/X) = r \frac{\sigma_2}{\sigma_1}, \quad \rho(X/Y) = r \frac{\sigma_1}{\sigma_2} \quad (8)$$

и поэтому уравнения прямых регрессии принимают вид:

$$\begin{aligned} y - b &= r \frac{\sigma_2}{\sigma_1} (x - a), \\ x - a &= r \frac{\sigma_1}{\sigma_2} (y - b). \end{aligned}$$

Из уравнений прямых регрессии видно, что обе эти прямые проходят через точку $(a; b)$; угловые коэффициенты прямых регрессии равны соответственно (обозначения углов см. рис. 208):

$$\operatorname{tg} \alpha = r \frac{\sigma_2}{\sigma_1}; \quad \operatorname{tg} \beta = \frac{1}{r} \frac{\sigma_2}{\sigma_1}.$$

Так как $|r| \leq 1$, то $|\operatorname{tg} \alpha| \leq |\operatorname{tg} \beta|$. Это означает, что прямая регрессии Y на X имеет меньший наклон к оси абсцисс, чем прямая регрессии X на Y . Чем ближе $|r|$ к 1, тем меньше угол между прямыми регрессии. Эти прямые сливаются тогда и только тогда, когда $|r| = 1$.

При $r = 0$ прямые регрессии имеют уравнения $y = b$; $x = a$.

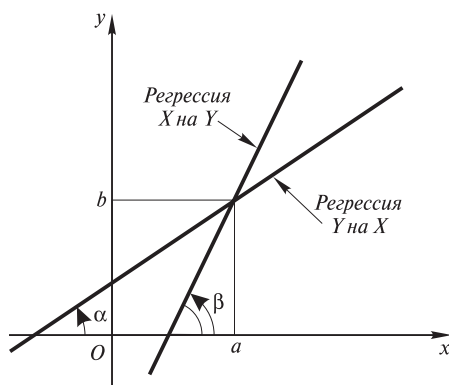


Рис. 208

В этом случае $M_x(Y) = b = M(Y)$; $M_y(X) = a = M(X)$.

Из формул (8) видно, что коэффициенты регрессии имеют тот же знак, что и коэффициент корреляции r , и связаны соотношением:

$$\rho(Y/X)\rho(X/Y) = r^2.$$

4. Расчет прямых регрессии. Пусть проведено n опытов, в результате которых получены следующие значения системы величин $(X; Y): (x_i, y_i)$, $i=1, 2, \dots, n$. За приближенные значения $M(X)$, $M(Y)$, $D(X)$ и $D(Y)$ принимают их выборочные значения

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y}_B = \frac{1}{n} \sum_{i=1}^n y_i, \quad s_1^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_B)^2, \\ s_2^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}_B)^2.$$

Оценкой для μ служит величина

$$\mu_B = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_B)(y_i - \bar{y}_B).$$

Заменяя в соотношениях (3), (4), (7) величины μ , σ_1 , σ_2 их выборочными значениями μ_B , s_1 , s_2 , получим приближенные значения коэффициента корреляции и коэффициентов регрессии

$$r \approx \frac{\mu_B}{s_1 s_2}, \quad \rho(Y/X) \approx \frac{\mu_B}{s_1^2}, \quad \rho(X/Y) \approx \frac{\mu_B}{s_2^2}$$

$(\frac{\mu_B}{s_1 s_2}, \frac{\mu_B}{s_1^2}, \frac{\mu_B}{s_2^2})$ — выборочные коэффициенты соответственно корреляции* и регрессии).

Подставляя в уравнения (5) и (6) вместо a , b , $\rho(Y/X)$ и $\rho(X/Y)$ их приближенные значения, получим выборочные уравнения прямых регрессии:

$$y - \bar{y}_B = \frac{\mu_B}{s_1^2} (x - \bar{x}_B), \\ x - \bar{x}_B = \frac{\mu_B}{s_2^2} (y - \bar{y}_B).$$

П р и м е р. Найти выборочное уравнение прямой регрессии Y на X по данным $n=10$ наблюдений. Результаты наблюдений и нужные вычисления собраны в таблице ($C=70$ и $C'=9,0$ — ложные нули).

* Выборочный коэффициент корреляции $\frac{\mu_B}{s_1 s_2}$ обозначим через r_B .

x_i	y_i	$x_i - C$	$x_i - C'$	$x_i - \bar{x}_B$	$(x_i - \bar{x}_B)^2$	$y_i - \bar{y}_B$	$(x_i - \bar{x}_B) \times (y_i - \bar{y}_B)$
71	8,6	1	-0,4	-4,5	20,25	-0,48	2,16
72	8,9	2	-0,1	-3,5	12,25	-0,18	0,63
73	8,9	3	-0,1	-2,5	6,25	-0,18	0,45
74	9,0	4	0,0	-1,5	2,25	0,08	-0,12
75	9,1	5	0,1	-0,5	0,25	0,02	-0,01
76	9,2	6	0,2	0,5	0,25	0,12	0,06
77	9,2	7	0,2	1,5	2,25	0,12	0,18
78	9,2	8	0,2	2,5	6,25	0,12	0,30
79	9,3	9	0,3	3,5	12,25	0,22	0,77
80	9,4	10	0,4	4,5	20,25	0,32	1,44
Сумма		55	0,8		82,5		5,86
		$\bar{x}_B = 75,5$	$\bar{y}_B = 9,08$		$s_1^2 = 9,17$		$\mu_B = 0,65$

Вычисляем:

$$\frac{\mu_B}{s_1^2} = \frac{0,65}{9,17} \approx 0,071.$$

Уравнение искомой прямой имеет вид:

$$y - 9,08 = 0,071(x - 75,5)$$

или

$$y = 0,071x + 3,72.$$

Глава 20. ЭЛЕМЕНТЫ ВАРИАЦИОННОГО И ОПЕРАЦИОННОГО ИСЧИСЛЕНИЙ

§ 20.1. Элементы вариационного исчисления

1. Функционал. Если M — множество функций и к каждой функции $f(x)$, принадлежащей M , относится определенное число, то говорят, что на множестве M задан функционал.

Пример 1. Пусть M — множество функций $y = y(x)$, непрерывных на отрезке $[0; 1]$. Рассмотрим определенный интеграл

$$I = I\{y\} = \int_0^1 y^2 dx. \quad (1)$$

Если вместо $y(x)$ подставлять различные конкретные функции, то будут получаться конкретные числовые значения I . Например, выбрав $y = x^2$, получим:

$$I = \int_0^1 (x^2)^2 dx = \int_0^1 x^4 dx = \frac{x^5}{5} \Big|_0^1 = \frac{1}{5};$$

выбрав $y = x^3$, получим $I = \frac{1}{7}$ и т. д.

Таким образом, формула (1) задает закон, согласно которому каждой функции $y(x)$ соответствует значение I , т. е. формула (1) определяет функционал.

Пример 2. M — множество функций $y = y(x)$, определенных на отрезке $[a, b]$ и обладающих на нем непрерывной производной. Формула

$$I = I\{y\} = \int_a^b (xy + y') dx$$

определяет функционал; длина дуги l линии $y = y(x)$, $a \leq x \leq b$, где $y(x)$ принадлежит множеству M , есть также функционал:

$$l\{y\} = \int_a^b \sqrt{1 + y'^2} dx.$$

Пример 3. Если M — множество функций $y = y(x)$, в область определения которых входит точка x_0 , то значения функций $y(x)$ в точке x_0 образуют функционал

$$I = I\{y\} = y(x_0).$$

2. Основная задача вариационного исчисления. В 1696 г. Иоганн Бернулли (1667—1748, швейцарский математик) предложил задачу о линии быстрейшего ската — брахистохроне.

В вертикальной плоскости xOy (ось Oy направлена вертикально вниз) даны две точки $O(0; 0)$ и $B(b; y_b)$ (рис. 209). Требуется соединить эти точки такой кривой $y = y(x)$, что материальной точке, скользящей без трения по этой кривой, под действием одной только силы тяжести, понадобится на путь от O до B наименьшее время.

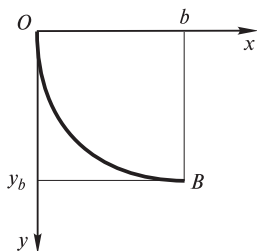


Рис. 209

Математическая формулировка задачи основывается на физическом положении, что при движении по кривой $y = y(x)$ скорость $v = \frac{ds}{dt}$ (ds — дифференциал дуги кривой $y = y(x)$) равна $\sqrt{2gy}$ (g — ускорение свободно-

го падения). Это следует из закона сохранения энергии $\frac{mv^2}{2} = mgy$.

Поэтому

$$dt = \frac{ds}{\sqrt{2gy}}, \text{ или } dt = \frac{\sqrt{1+y'^2} dx}{\sqrt{2g}\sqrt{y}},$$

откуда время, затраченное на путь от O до B будет:

$$T = \frac{1}{\sqrt{2g}} \int_a^b \frac{\sqrt{1+y'^2} dx}{\sqrt{y}}. \quad (2)$$

Этим равенством решение задачи Бернулли сведено к разысканию такой кривой, проходящей через точки O и B , или, что то же самое, такой функции $y = y(x)$, удовлетворяющей условиям $y(0) = 0$, $y(b) = y_b$, для которой значение интеграла (2) будет наименьшим. Таким образом, задача Бернулли сводится к нахождению минимума функционала (2) (это решение будет дано в п. 6).

Имеется много других задач на экстремум функционалов. Все такие задачи и составляют предмет вариационного исчисления. Таким образом, основная задача вариационного исчисления — это *задача о нахождении экстремума заданного функционала*.

3. Вариация функционала. При исследовании функционала бывает важно выяснить, как меняется его значение при малом изменении функции, от которой он зависит. Рассмотрим этот вопрос на примере функционала (1). Пусть в правую часть была сначала подставлена некоторая функция $y(x)$, а затем некоторая новая функция $y(x) + \delta y(x)$, где $\delta y(x)$ — *вариация y* — произвольная функция, принимающая малые значения. (Например, могло быть сначала $y = x^2$, а затем $y = x^2 + \alpha x^3$, где постоянная α мала.) Тогда и значение функционала немного изменится и станет равным:

$$\int_0^1 (y + \delta y)^2 dx = \int_0^1 y^2 dx + 2 \int_0^1 y \delta y dx + \int_0^1 (\delta y)^2 dx.$$

Таким образом, приращение этого значения равно:

$$\Delta I = 2 \int_0^1 y \delta y dx + \int_0^1 (\delta y)^2 dx. \quad (3)$$

Если временно зафиксировать функцию $y(x)$ и менять ее вариацию, то мы видим, что в правой части равенства (3) в зависимости от δy первый член линейный, т. е. подынтегральное выражение относительно δy линейно, тогда как второй квадратичный (в общем случае присутствуют также и члены более высоких степеней). Так как значения δy малы,

то главную роль в правой части играет линейный член, тогда как квадратичный имеет высший порядок малости. Этот линейный член в приращении функционала называется *вариацией функционала* и обозначается δI , т. е. для примера 1

$$\delta I = 2 \int_0^1 y \delta y dx. \quad (4)$$

Таким образом, с точностью до членов высшего порядка малости

$$\Delta I \approx \delta I. \quad (5)$$

В тех случаях, когда членами второго и следующих порядков малости можно пренебречь, говорят просто, что вариация функционала — это его бесконечно малое приращение, полученное за счет бесконечно малого изменения (варьирования) функции, от которой зависит значение функционала. Если вспомнить дифференциальное исчисление, то можно усмотреть полную аналогию между понятиями дифференциала функции и вариации функционала.

4. Необходимое условие экстремума функционала. Пусть функция $y(x)$ реализует экстремум функционала (1); другими словами, значение функционала I для функции $y(x)$ больше (в случае максимума) значений этого функционала для всех функций, достаточно близких к $y(x)$, или меньше (в случае минимума) значений этого функционала для всех функций, достаточно близких к $y(x)$. При этом ΔI будет отрицательное в случае максимума и положительное в случае минимума для всех указанных функций, т. е. в обоих случаях не меняет знака. Но отсюда следует, что

$$\delta I = 0. \quad (6)$$

В самом деле, из формулы (3) вытекает, что если $\delta I \neq 0$, то ΔI и δI имеют одинаковый знак; но в силу линейной зависимости δI от δy при перемене знака у δy и δI изменит знак (см. формулу (4)), что противоречит предыдущему.

Условие (6) и является необходимым условием экстремума.

5. Уравнение Эйлера. В конкретных примерах вариация функционалов вычисляется с помощью формулы Тейлора (см. § 8.9). Пусть, например, рассматривается функционал вида

$$I = I\{y\} = \int_a^b F(x, y) dx,$$

где y считается зависящим от x , $y = y(x)$. (Считаем, что подинтегральные функции в рассматриваемых здесь функционалах непрерывны и обладают непрерывными частными производными до нужного порядка.)

Тогда

$$\Delta I = \int_a^b (F(x, y + \delta y) - F(x, y)) dx. \quad (7)$$

Но так как $F(x, y + \delta y) = F(x, y) + F'_y(x, y)\delta y + \dots$ где после слагаемого $F'_y(x, y)\delta y$ идут члены высшего порядка малости, то, подставляя в (7) и отбрасывая эти члены, получим:

$$\delta I = \int_a^b F'_y(x, y)\delta y dx.$$

Для функционала

$$I = I\{y\} = \int_a^b F(x, y, y') dx \quad (8)$$

аналогично получаем:

$$\delta I = \int_a^b (F'_y(x, y, y')\delta y + F'_{y'}(x, y, y')\delta y') dx, \quad (9)$$

где $\delta y'$ можно понимать и как $\delta(y)'$, и как $(\delta y)'$, так как производная от разности двух функций равна разности их производных.

Рассмотрим функционал (8), где F — заданная функция, a, b — заданные пределы интегрирования. Пусть ищется экстремум этого функционала при заданных краевых условиях

$$y(a) = y_a, y(b) = y_b. \quad (10)$$

На основании формул (6) и (9) получаем:

$$\int_a^b (F'_y(x, y, y')\delta y + F'_{y'}(x, y, y')\delta y') dx = 0. \quad (11)$$

Это равенство должно выполняться для любой вариации δy , удовлетворяющей соотношениям

$$\delta y(a) = 0, \delta y(b) = 0, \quad (12)$$

которые нужны для того, чтобы $y + \delta y$ удовлетворяла тем же условиям (10), что и y .

Интегрируя по частям второе слагаемое в (11) и применяя соотношение (12), получим:

$$0 = \int_a^b F'_y \delta y dx + F'_{y'} \delta y \Big|_a^b - \int_a^b \frac{d}{dx} (F'_{y'}) \delta y dx = \int_a^b \left(F'_y - \frac{d}{dx} F'_{y'} \right) \delta y dx.$$

Отсюда в силу произвола в выборе δy вытекает, что выражение в скобках последнего интеграла тождественно равно нулю.

Итак, мы пришли к так называемому уравнению Эйлера

$$F'_{yy'}(x, y, y') - F''_{xy'}(x, y, y') - F''_{yy'}(x, y, y')y' - F''_{y'^2}(x, y, y')y'' = 0.$$

Видно, что это обыкновенное дифференциальное уравнение второго порядка.

6. Случаи понижения порядка уравнения Эйлера.

а) Подынтегральная функция F в (8) явно не зависит от x , т. е. $F = F(y, y')$.

В этом случае уравнение Эйлера после умножения на y' может быть переписано в виде $\frac{d}{dx}(F - y' F_{y'}) = 0$, откуда получаем дифференциальное уравнение первого порядка

$$F - y' F_{y'} = C, \quad (13)$$

где C — произвольная постоянная.

б) Подынтегральная функция F в (8) явно не зависит от y , т. е. $F_y' = 0$. Тогда $\frac{d}{dx} F_{y'} = 0$, и, следовательно, также имеем дифференциальное уравнение первого порядка $F_{y'}(x, y') = C$.

Применим формулу (13) к задаче о *брахистохроне*, которая отмечалась в п. 2. Там эта задача сведена была к нахождению наименьшего значения функционала (2).

Так как в (2) подынтегральная функция явно не зависит от x , то уравнение Эйлера может быть записано в форме (13).

Имеем:

$$\frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} - y' \frac{y'}{\sqrt{2gy} \cdot \sqrt{1 + y'^2}} = C_1.$$

Отсюда

$$y = \frac{\bar{C}_1}{1 + y'^2}, \quad (14)$$

где

$$\bar{C}_1 = \frac{1}{C_1^2 2g}.$$

Для решения уравнения (14) положим $y' = \operatorname{tg} \varphi$. Тогда

$$y = \bar{C}_1 \cos^2 \varphi, \quad (15)$$

Откуда

$$y' = -2\bar{C}_1 \cos \varphi \sin \varphi \frac{d\varphi}{dx},$$

и, значит,

$$\operatorname{tg} \varphi = -2\bar{C}_1 \sin \varphi \cos \varphi \frac{d\varphi}{dx}, \text{ или } dx = -2\bar{C}_1 \cos^2 \varphi d\varphi.$$

Отсюда и из равенства (15)

$$x = -\frac{\bar{C}_1}{2}(2\varphi + \sin 2\varphi) + C_2, \quad y = \frac{\bar{C}_1}{2}(1 + \cos 2\varphi).$$

Положив здесь $2\varphi = \pi - \theta$, $r = \frac{\bar{C}_1}{2}$, $C = C_2 - \frac{C_1\pi}{2}$, получим:

$$x = r(\theta - \sin \theta) + C, \quad y = r(1 - \cos \theta).$$

Так как $y = 0$ при $x = 0$, то $\theta = 0$ и $C = 0$. Так что

$$x = r(\theta - \sin \theta), \quad y = r(1 - \cos \theta).$$

Это — уравнение циклоиды. Число r может быть найдено из условия $y(b) = y_b$.

Итак, брахистохроной является циклоида.

§ 20.2. Элементы операционного исчисления

Операционное исчисление является одним из методов прикладного математического анализа. С его помощью удастся во многих случаях упростить решение задач, встречающихся в механике, электронике, автоматике и других областях науки и техники. Операционное исчисление составляет теоретическую основу целого ряда инженерных методов расчета и проектирования автоматических систем.

В этой главе мы рассмотрим некоторые понятия операционного исчисления и применим его к решению линейных дифференциальных уравнений.

1. Операторы и изображения. Рассмотрим функцию $f(t)$, определенную на всей числовой оси и обладающую следующими свойствами:

а) функция $f(t)$ на любом конечном интервале оси Ot или непрерывна, или имеет конечное число точки разрыва первого рода;

б) $f(t) = 0$ при $t < 0$;

в) существуют такие числа $M > 0$ и $s_0 \geq 0$, что $|f(t)| < Me^{s_0 t}$ для всех t .

Условие б) вводится в связи с тем, что во многих задачах физики и техники аргумент t рассматривается как время. Поэтому не имеет значения, как ведет себя функция $f(t)$ до некоторого начального момента времени, который всегда можно принять равным нулю.

Условие в) накладывает ограничение на характер роста функции $f(t)$ при $t \rightarrow +\infty^*$ и обеспечивает тем самым существование некоторых встречающихся в дальнейшем несобственных интегралов. Оно означает, что функция $f(t)$ при $t \rightarrow \infty$ растет не быстрее показательной функции. В частности, условию в) удовлетворяют, например, любая ограниченная функция (в этом случае, очевидно, $s_0 = 0$) и степенные функции t^k ($k > 0$). Число s_0 называется *показателем роста*.

* В дальнейшем ради краткости будем писать ∞ вместо $+\infty$.

Каждой функции $f(t)$ поставим в соответствие функцию $F(p)$ комплексного переменного $p = s + i\tau$ ($s > s_0$), полагая

$$F(p) = \int_0^{\infty} f(t) e^{-pt} dt. \quad (1)$$

Функция $F(p)$ называется *изображением* функции $f(t)$, а функция $f(t)$, удовлетворяющая указанным выше трем условиям, называется *оригиналом*. Правая часть равенства (1) называется *интегралом Лапласа* для функции $f(t)$, а определяемая им функция комплексного переменного p называется *преобразованием Лапласа* от функции $f(t)$.

Если $F(p)$ является изображением функции $f(t)$, то это записывается так: $f(t) \leftrightarrow F(p)$ или $L\{f(t)\} = F(p)$ (стрелка направлена всегда от изображения к оригиналу).

З а м е ч а н и е. Если функция $f(t)$ не удовлетворяет хотя бы одному из указанных выше трех условий, то она не является оригиналом. Так, например, функции $\lg t$ и $\frac{1}{t}$ не могут быть оригиналами, так как они имеют точки разрыва 2-го рода. Функция e^{t^2} также не может быть оригиналом, так как не удовлетворяет условию в) — она растет быстрее, чем функция $Me^{s_0 t}$ (каковы бы ни были M и s_0).

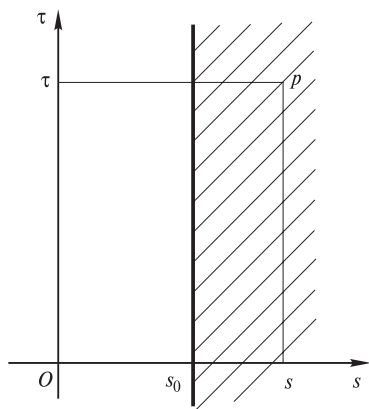


Рис. 210

2. Существование изображений.

Имеют место следующие теоремы, которые мы приведем без доказательства.

Т е о р е м а 1 (теорема существования изображения). *Для всякого оригинала $f(t)$ существует изображение $F(p)$, определенное в полуплоскости $\text{Re } p = s > s_0$, где s_0 — показатель роста оригинала (рис. 210). В каждой точке этой полуплоскости функция $F(p)$ имеет производную любого порядка. Кроме того, если $\text{Re } p = s \rightarrow \infty$, то изображение $F(p) \rightarrow 0$.*

Интеграл, стоящий к правой части равенства (1), определяется следующим образом:

$$\begin{aligned} \int_0^{\infty} f(t) e^{-pt} dt &= \int_0^{\infty} f(t) e^{-(s+i\tau)t} dt = \int_0^{\infty} f(t) e^{-st} e^{-i\tau t} dt = \int_0^{\infty} f(t) (\cos \tau t - i \sin \tau t) e^{-st} dt = \\ &= \int_0^{\infty} f(t) e^{-st} \cos \tau t dt - i \int_0^{\infty} f(t) e^{-st} \sin \tau t dt \end{aligned}$$

Т е о р е м а 2 (теорема единственности оригинала). Если $F(p)$ является изображением двух оригиналов $f_1(t)$ и $f_2(t)$, то эти оригиналы совпадают во всех точках, в которых они непрерывны.

В заключение отметим, что изображение обладает свойством линейности. Это значит, что:

1) изображение произведения оригинала на число равно произведению изображения на это число, т. е. если

$$f(t) \leftrightarrow F(p), \text{ то } cf(t) \leftrightarrow cF(p),$$

2) изображение алгебраической суммы нескольких оригиналов равно алгебраической сумме изображений этих оригиналов, т. е. если, например,

$$f_1(t) \leftrightarrow F_1(p), f_2(t) \leftrightarrow F_2(p),$$

то

$$f_1(t) \pm f_2(t) \leftrightarrow F_1(p) \pm F_2(p).$$

Эти свойства вытекают из свойства линейности интеграла. Так, например,

$$\begin{aligned} f_1(t) \pm f_2(t) &\leftrightarrow \int_0^{\infty} [f_1(t) \pm f_2(t)] e^{-pt} dt = \\ &= \int_0^{\infty} f_1(t) e^{-pt} dt \pm \int_0^{\infty} f_2(t) e^{-pt} dt = F_1(p) \pm F_2(p). \end{aligned}$$

3. Примеры вычислений изображений. В дальнейшем нам понадобится следующая лемма.

Л е м м а. Если для комплексного числа $z = x + iy$ его действительная часть $\operatorname{Re} z = x > 0$ и b — действительная переменная, то $\lim_{b \rightarrow \infty} e^{-zb} = 0$.

Д о к а з а т е л ь с т в о. По формуле Эйлера имеем

$$e^{-zb} = e^{-(x+iy)b} = e^{-xb} e^{-iyb} = e^{-xb} (\cos yb - i \sin yb).$$

Следовательно,

$$\lim_{b \rightarrow \infty} e^{-zb} = \lim_{b \rightarrow \infty} e^{-xb} (\cos yb - i \sin yb) = \lim_{b \rightarrow \infty} e^{-xb} \cos yb - i \lim_{b \rightarrow \infty} e^{-xb} \sin yb = 0,$$

так как $\lim_{b \rightarrow \infty} e^{-xb} = 0$ при $x > 0$.

А. Изображение единичной функции Хевисайда. Эта функция имеет вид (см. § 16.2, п. 3):

$$\epsilon(t) = \begin{cases} 0, & \text{если } t < 0, \\ 1, & \text{если } t \geq 0. \end{cases}$$

Используя формулу (1) находим:

$$\epsilon(t) \leftrightarrow \int_0^{\infty} 1 e^{-pt} dt = \lim_{b \rightarrow \infty} \int_0^b e^{-pt} dt = \lim_{b \rightarrow \infty} \left[-\frac{e^{-pt}}{p} \right]_0^b = -\frac{1}{p} (\lim_{b \rightarrow \infty} e^{-pb} - 1) = \frac{1}{p}$$

(так как $\operatorname{Re} p = s > s_0 \geq 0$, то по лемме $\lim_{b \rightarrow \infty} e^{-pb} = 0$. Итак,

$$e(t) \leftarrow \frac{1}{p}.$$

В. Изображение показательной функции. Найдем изображение функции

$$f(t) = \begin{cases} 0, & \text{если } t < 0, \\ e^{\alpha t}, & \text{если } t \geq 0. \end{cases}$$

где α — комплексное число. На основании формулы (1) имеем:

$$\begin{aligned} f(t) &\leftarrow \int_0^{\infty} f(t) e^{-pt} dt = \int_0^{\infty} e^{\alpha t} e^{-pt} dt = \int_0^{\infty} e^{-(p-\alpha)t} dt = \lim_{b \rightarrow \infty} \int_0^b e^{-(p-\alpha)t} dt = \\ &= \frac{\lim_{b \rightarrow \infty} [e^{-(p-\alpha)t} \Big|_0^b]}{-(p-\alpha)} = \lim_{b \rightarrow \infty} \left[\frac{-e^{-(p-\alpha)b}}{p-\alpha} + \frac{1}{p-\alpha} \right] = \frac{1}{p-\alpha}, \end{aligned}$$

если $\operatorname{Re}(p-\alpha) > 0$, или $\operatorname{Re} p > \operatorname{Re} \alpha$, так как при этом условии на основании леммы $\lim_{b \rightarrow \infty} e^{-(p-\alpha)b} = 0$. Итак, для данной функции имеем:

$$f(t) \leftarrow \frac{1}{p-\alpha}, \operatorname{Re} p > \operatorname{Re} \alpha. \quad (2)$$

В дальнейшем мы будем пользоваться более простой записью оригинала, оставляя лишь то его выражение, которое он имеет при $t \geq 0$. В частности, формулу (2) кратко запишем так:

$$e^{\alpha t} \leftarrow \frac{1}{p-\alpha}, \operatorname{Re} p > \operatorname{Re} \alpha. \quad (3)$$

Аналогично получим

$$e^{-\alpha t} \leftarrow \frac{1}{p+\alpha}, \operatorname{Re} p > \operatorname{Re}(-\alpha). \quad (4)$$

С. Изображения функций $\sin \omega t$ и $\cos \omega t$ (ω — действительное число).

По известным формулам (см. § 15.5, п.1) имеем:

$$\sin \omega t = \frac{1}{2i}(e^{i\omega t} - e^{-i\omega t}),$$

$$\cos \omega t = \frac{1}{2}(e^{i\omega t} + e^{-i\omega t}).$$

На основании формул (3) и (4) и свойства линейности изображения получим:

$$\sin \omega t = \frac{1}{2i}(e^{i\omega t} - e^{-i\omega t}) \leftarrow \frac{1}{2i} \left[\frac{1}{p-i\omega} - \frac{1}{p+i\omega} \right] = \frac{\omega}{p^2 + \omega^2}, \operatorname{Re} p > 0,$$

$$\cos \omega t = \frac{1}{2}(e^{i\omega t} + e^{-i\omega t}) \leftarrow \frac{1}{2} \left[\frac{1}{p-i\omega} + \frac{1}{p+i\omega} \right] = \frac{p}{p^2 + \omega^2}, \operatorname{Re} p > 0.$$

Итак,

$$\sin \omega t \leftrightarrow \frac{\omega}{p^2 + \omega^2}, \operatorname{Re} p > 0, \quad (5)$$

$$\cos \omega t \leftrightarrow \frac{p}{p^2 + \omega^2}, \operatorname{Re} p > 0. \quad (6)$$

Д. Изображение гиперболических функций $\operatorname{sh} \omega t$ и $\operatorname{ch} \omega t$. Аналогично находим изображения гиперболических синуса и косинуса $\operatorname{sh} \omega t$ и $\operatorname{ch} \omega t$ (ω — действительное число):

$$\operatorname{sh} \omega t = \frac{1}{2}(e^{\omega t} - e^{-\omega t}) \leftrightarrow \frac{1}{2} \left[\frac{1}{p - \omega} - \frac{1}{p + \omega} \right] = \frac{\omega}{p^2 - \omega^2}, \operatorname{Re} p > |\omega|,$$

$$\operatorname{ch} \omega t = \frac{1}{2}(e^{\omega t} + e^{-\omega t}) \leftrightarrow \frac{1}{2} \left[\frac{1}{p - \omega} + \frac{1}{p + \omega} \right] = \frac{p}{p^2 - \omega^2}, \operatorname{Re} p > |\omega|.$$

Итак,

$$\operatorname{sh} \omega t \leftrightarrow \frac{\omega}{p^2 - \omega^2}, \operatorname{Re} p > |\omega|,$$

$$\operatorname{ch} \omega t \leftrightarrow \frac{p}{p^2 - \omega^2}, \operatorname{Re} p > |\omega|.$$

Е. Изображение степенной функции t^n . По формуле (1) имеем:

$$t^n \leftrightarrow \int_0^{\infty} t^n e^{-pt} dt.$$

Выполняя n раз интегрирование по частям и замечая, что при $\operatorname{Re} p > 0$

$$\lim_{t \rightarrow \infty} t^k e^{-pt} = 0 \quad (k = 1, 2, \dots, n)$$

получим:

$$\int_0^{\infty} t^n e^{-pt} dt = \frac{n!}{p^{n+1}}, \operatorname{Re} p > 0.$$

Таким образом,

$$t^n \leftrightarrow \frac{n!}{p^{n+1}}, \operatorname{Re} p > 0. \quad (7)$$

П р и м е р 1. Найти оригинал, изображение которого имеет вид:

$$F(p) = \frac{2p+1}{p^2+4}.$$

Р е ш е н и е. Так как

$$F(p) = \frac{2p+1}{p^2+4} = 2 \frac{p}{p^2+2^2} + \frac{1}{2} \frac{2}{p^2+2^2},$$

то на основании формул (5) и (6) и свойства линейности получим:

$$f(t) = 2 \cos 2t + \frac{1}{2} \sin 2t.$$

Пример 2. Найти изображение функции $f(t) = t^3 - 5t^2 + 2t + 6$.

Решение. На основании формул (2) и (7) и свойства линейности получим:

$$f(t) \leftrightarrow \frac{3!}{p^{3+1}} - 5 \frac{2!}{p^{2+1}} + 2 \frac{1!}{p^{1+1}} + 6 \frac{1}{p} = \frac{6}{p^4} - \frac{10}{p^3} + \frac{2}{p^2} + \frac{6}{p}.$$

4. Основные теоремы операционного исчисления.

Теорема подобия. Пусть $f(t) \leftrightarrow F(p)$. Тогда $f(\alpha t) \leftrightarrow \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right)$, где $\alpha > 0$.

Доказательство. По определению изображения имеем:

$$f(\alpha t) \leftrightarrow \int_0^{\infty} f(\alpha t) e^{-pt} dt.$$

Произведем в этом интеграле замену переменной, полагая $\alpha t = y$. Тогда $dy = \alpha dt$ и, следовательно,

$$\int_0^{\infty} f(\alpha t) e^{-pt} dt = \int_0^{\infty} f(y) e^{-p \frac{y}{\alpha}} \frac{dy}{\alpha} = \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right).$$

Итак,

$$f(\alpha t) \leftrightarrow \frac{1}{\alpha} F\left(\frac{p}{\alpha}\right). \quad (8)$$

Соотношение (8) показывает, что при умножении независимой переменной оригинала на положительное число его изображение и независимая переменная изображения делятся на это число. В этом и состоит теорема подобия.

Теорема сдвига. Если $f(t) \leftrightarrow F(p)$, то $e^{-\alpha t} f(t) \leftrightarrow F(p + \alpha)$. При этом предполагается, что α — действительное число и $\operatorname{Re}(p + \alpha) > s_0$.

Доказательство. Найдем изображение функции $e^{-\alpha t} f(t)$:

$$e^{-\alpha t} f(t) \leftrightarrow \int_0^{\infty} e^{-\alpha t} f(t) e^{-pt} dt = \int_0^{\infty} f(t) e^{-(p+\alpha)t} dt = F(p + \alpha).$$

Итак, если $f(t) \leftrightarrow F(p)$, то $e^{-\alpha t} f(t) \leftrightarrow F(p + \alpha)$.

На основании этой теоремы и формул (5) и (6) находим:

$$\begin{aligned} e^{-\alpha t} \sin \omega t &\leftrightarrow \frac{\omega}{(p + \alpha)^2 + \omega^2}, \\ e^{-\alpha t} \cos \omega t &\leftrightarrow \frac{p + \alpha}{(p + \alpha)^2 + \omega^2}. \end{aligned} \quad (9)$$

Пример 1. Найти изображение функции $e^{-2t}(\sin t + 3 \cos t)$.

Решение. Используя свойства линейности и формулы (9), получим:

$$e^{-2t}(\sin t + 3 \cos t) \leftrightarrow \frac{1}{(p+2)^2 + 1} + 3 \frac{p+2}{(p+2)^2 + 1} = \frac{3p+7}{p^2 + 4p + 5}.$$

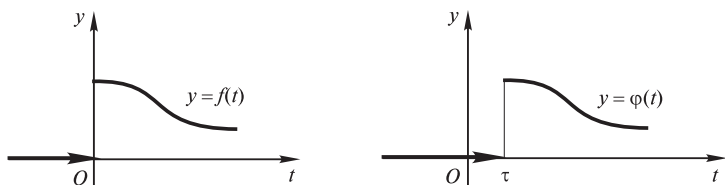


Рис. 211

Пример 2. Найти оригинал по его изображению $\frac{2p+1}{p^2+2p+2}$.

Решение. Используя свойства линейности и формулы (9), находим:

$$\frac{2p+1}{p^2+2p+2} = \frac{2(p+1)-1}{(p+1)^2+1} = \frac{2(p+1)}{(p+1)^2+1} - \frac{1}{(p+1)^2+1} \rightarrow 2e^{-t} \cos t - e^{-t} \sin t = e^{-t}(2 \cos t - \sin t).$$

Найдем изображение функции $e^{-\alpha t} t^n$.

На основании формулы (7) имеем: $f(t) = t^n \leftrightarrow F(p) = \frac{n!}{p^{n+1}}$.

Применяя теорему смещения, получим: $e^{-\alpha t} t^n \leftrightarrow \frac{n!}{(p+\alpha)^{n+1}}$.

Заменяя в этой формуле $-\alpha$ на α , найдем: $e^{\alpha t} t^n \leftrightarrow \frac{n!}{(p-\alpha)^{n+1}}$.

Пусть $f(t)$ — оригинал. Рассмотрим функцию $\varphi(t)$, определенную следующим образом:

$$\varphi(t) = \begin{cases} 0, & \text{если } t < \tau, \\ f(t-\tau), & \text{если } t \geq \tau, \end{cases}$$

где τ — положительное число.

График функции $y = \varphi(t)$ получается из графика оригинала $y = f(t)$ сдвигом последнего графика вправо вдоль оси Ot на величину τ (рис. 211).

Теорема запаздывания. Пусть $\tau > 0$ и $f(t) \leftrightarrow F(p)$. Тогда $f(t-\tau) \leftrightarrow e^{-p\tau} F(p)$.

5. Дифференцирование и интегрирование оригиналов. Пусть $F(p)$ — изображение оригинала $f(t)$. Требуется найти изображение производной $f'(t)$.

Теорема 1. Если $f(t) \leftrightarrow F(p)$ и $f'(t)$ является оригиналом, то

$$f'(t) \leftrightarrow pF(p) - f(0). \quad (10)$$

Доказательство. На основании формулы, определяющей изображение, имеем:

$$f'(t) \leftrightarrow \int_0^{\infty} f'(t) e^{-pt} dt.$$

Полученный интеграл интегрируем по частям. Полагая $u = e^{-pt}$, $dv = f'(t)dt$, откуда $du = -pe^{-pt}dt$, $v = f(t)$, находим:

$$\begin{aligned} \int_0^\infty f'(t)e^{-pt} dt &= \lim_{b \rightarrow \infty} \left[f(t)e^{-pt} \Big|_0^b + p \int_0^b f(t)e^{-pt} dt \right] = \\ &= \lim_{b \rightarrow \infty} \left[f(b)e^{-pb} - f(0) + p \int_0^b f(t)e^{-pt} dt \right] = p \int_0^\infty f(t)e^{-pt} dt - f(0), \end{aligned}$$

так как $\lim_{b \rightarrow \infty} f(b)e^{-pb} = 0$.* Но $\int_0^\infty f(t)e^{-pt} dt = F(p)$.

Следовательно,

$$f'(t) \leftrightarrow pF(p) - f(0).$$

Теорема доказана.

Применяя формулу (10) ко второй производной $f''(t)$, получим

$$f''(t) \leftrightarrow pF_1(p) - f'(0),$$

где $F_1(p)$ — изображение $f'(t)$. Но так как по формуле (10)

$$F_1(t) = pF(p) - f(0),$$

то

$$f''(t) \leftrightarrow p[pF(p) - f(0)] - f'(0) = p^2F(p) - pf(0) - f'(0). \tag{11}$$

Аналогично

$$f'''(t) \leftrightarrow p^3F(p) - p^2f(0) - pf'(0) - f''(0) \tag{12}$$

и т. д.

$$f^{(n)}(t) \leftrightarrow p^nF(p) - p^{n-1}f(0) - p^{n-2}f'(0) - \dots - pf^{(n-2)}(0) - f^{(n-1)}(0).^{**} \tag{13}$$

В частности, когда $f(0) = f'(0) = \dots = f^{(n-1)}(0) = 0$, т. е. когда все начальные значения равны нулю, получим:

$$\begin{aligned} f'(t) &\leftrightarrow pF(p), \\ f''(t) &\leftrightarrow p^2F(p), \\ &\dots\dots\dots \\ f^{(n)}(t) &\leftrightarrow p^nF(p). \end{aligned} \tag{14}$$

В самом деле, $|f(b)e^{-pb}| = |f(b)| |e^{-(s+ic)b}| = |f(b)| |e^{-sb}| e^{-rcb} = |f(b)| e^{-sb}$. Но по свойству в) оригинала $|f(b)| < Me^{s_0b}$, следовательно, $|f(b)e^{-pb}| < Me^{s_0b} e^{-sb} = Me^{-(s-s_0)b}$. Согласно теореме существования изображения $F(p)$ существует для $\text{Re } p = s > s_0$, т. е. для $s - s_0 > 0$. Поэтому $Me^{-(s-s_0)b} \rightarrow 0$ при $b \rightarrow \infty$ и, значит, $\lim_{b \rightarrow \infty} f(b)e^{-pb} = 0$.

** При выводе формул (11)—(13) предполагается, что производные $f''(t)$, $f'''(t)$, ..., $f^{(n)}(t)$ являются оригиналами.

Итак, при нулевых начальных условиях n -кратное дифференцирование оригинала сводится к умножению его изображения на p^n .

Т е о р е м а 2. Если $f(t) \leftrightarrow F(p)$, то

$$\int_0^t f(x) dx \leftrightarrow \frac{F(p)}{p}. \quad (15)$$

Д о к а з а т е л ь с т в о. Можно показать, что функция $g(t) = \int_0^t f(x) dx$ является оригиналом. Пусть ее изображение есть $\Phi(p): g(t) \leftrightarrow \Phi(p)$. Так как $g(0) = \int_0^0 f(x) dx = 0$, то по первой из формул (14) имеем:

$$g'(t) \leftrightarrow p\Phi(p).$$

Но

$$g'(t) = \frac{d}{dt} \int_0^t f(x) dx = f(t) \leftrightarrow F(p).$$

Следовательно,

$$F(p) = p\Phi(p).$$

Отсюда

$$\Phi(p) = \frac{F(p)}{p},$$

т. е.

$$\int_0^t f(x) dx \leftrightarrow \frac{F(p)}{p}.$$

Соотношение (15) показывает, что операции интегрирования оригинала соответствует алгебраическое действие над изображением, а именно его деление на p .

6. Сводная таблица оригиналов и их изображений. Приведем таблицу некоторых оригиналов и их изображений.

N°	Оригинал $f(t)$	Изображение $F(p) = \int_0^{\infty} f(t)e^{-pt} dt$	Примечание
I	$e_0(t)$	$\frac{1}{p}$	
II	$e^{\alpha t}$	$\frac{1}{p - \alpha}$	
III	$e^{-\alpha t}$	$\frac{1}{p + \alpha}$	
IV	$\sin \omega t$	$\frac{\omega}{p^2 + \omega^2}$	

V	$\cos \omega t$	$\frac{p}{p^2 + \omega^2}$	
VI	$\operatorname{sh} \omega t$	$\frac{\omega}{p^2 - \omega^2}$	
VII	$\operatorname{ch} \omega t$	$\frac{p}{p^2 - \omega^2}$	
VIII	$e^{-\alpha t} \sin \omega t$	$\frac{\omega}{(p + \alpha)^2 + \omega^2}$	
IX	$e^{-\alpha t} \cos \omega t$	$\frac{p + \alpha}{(p + \alpha)^2 + \omega^2}$	
X	t^n	$\frac{n!}{p^{n+1}}$	
XI	$e^{-\alpha t} t^n$	$\frac{n!}{(p + \alpha)^{n+1}}$	
XII	$e^{\alpha t} t^n$	$\frac{n!}{(p - \alpha)^{n+1}}$	
XIII	$f(\alpha t)$	$\frac{1}{\alpha} F\left(\frac{p}{\alpha}\right)$	теорема подобия
XIV	$e^{-\alpha t} f(t)$	$F(p + \alpha)$	теорема смещения
XV	$f'(t)$	$pF(p) - f(0) = F_1(p)$	дифференцирование оригинала
XVI	$f''(t)$	$p^2 F(p) - pf(0) - f'(0)$	дифференцирование оригинала
XVII	$\int_0^t f(x) dx$	$\frac{F(p)}{p}$	интегрирование оригинала

Пользуясь сводной таблицей, можно по приведенному в ней изображению (оригиналу) найти его оригинал (изображение).

Следует заметить, что если изображение является правильной рациональной дробью, то для нахождения соответствующего оригинала следует изображение разложить на сумму простейших дробей и применить свойство линейности.

П р и м е р. Дано изображение $F(p) = \frac{p^2 + 2p + 2}{(p - 2)^2(p + 3)}$. Найти оригинал.

Р е ш е н и е. Разлагая данную правильную дробь на простейшие дроби, получим:

$$F(p) = \frac{p^2 + 2p + 2}{(p - 2)^2(p + 3)} = \frac{4}{5(p - 2)} + \frac{2}{(p - 2)^2} + \frac{1}{5(p + 3)}.$$

На основании формул (II), (XII) и (III) таблицы находим: $\frac{1}{p - 2} \Rightarrow e^{2t}$, $\frac{1}{(p - 2)^2} \Rightarrow te^{2t}$, $\frac{1}{p + 3} \Rightarrow e^{-3t}$. Следовательно, на основании свойства линейности $F(p) \Rightarrow \frac{4}{5}e^{2t} + 2te^{2t} + \frac{1}{5}e^{-3t}$.

7. Применение операционного исчисления к решению линейных дифференциальных уравнений с постоянными коэффициентами. Для простоты ограничимся линейными дифференциальными уравнениями второго порядка с постоянными коэффициентами:

$$y''(t) + \alpha_1 y'(t) + \alpha_2 y(t) = f(t), \quad (16)$$

где α_1 и α_2 — действительные числа. Требуется найти частное решение $y(t)$ этого уравнения, удовлетворяющее начальным условиям $y(0) = y_0$, $y'(0) = y'_0$, где y_0 и y'_0 — заданные числа.

Предположим, что искомое решение $y(t)$, его производные $y'(t)$, $y''(t)$ и правая часть дифференциального уравнения $f(t)$ являются оригиналами. Обозначив $y(t) \leftrightarrow Y(p)$, $f(t) \leftrightarrow F(p)$ и используя формулы (XV), (XVI), а также начальные условия, найдем изображения $y'(t)$ и $y''(t)$:

$$y'(t) \leftrightarrow pY(p) - y_0, \quad y''(t) \leftrightarrow p^2 Y(p) - py_0 - y'_0.$$

На основании свойства линейности перейдем в уравнении (16) к изображениям:

$$y''(t) + \alpha_1 y'(t) + \alpha_2 y(t) \leftrightarrow p^2 Y(p) - py_0 - y'_0 + \\ + \alpha_1 (pY(p) - y_0) + \alpha_2 Y(p) = F(p)$$

или

$$(p^2 + \alpha_1 p + \alpha_2)Y(p) = F(p) + py_0 + y'_0 + \alpha_1 y_0. \quad (17)$$

Уравнение (17) называют *вспомогательным уравнением*, или *уравнением в изображениях*, соответствующим дифференциальному уравнению (16). Таким образом, вместо дифференциального уравнения (16) для оригинала $y(t)$ мы получили линейное алгебраическое уравнение (17) для его изображения $Y(p)$. Из уравнения (17) находим:

$$Y(p) = \frac{F(p) + py_0 + y'_0 + \alpha_1 y_0}{p^2 + \alpha_1 p + \alpha_2}. \quad (18)$$

Формула (18) дает так называемое *операторное решение* уравнения (16). Оригинал $y(t)$, для которого функция $Y(p)$, определяемая формулой (18), является изображением, и служит искомым решением дифференциального уравнения (16).

П р и м е р 1. Найти решение дифференциального уравнения

$$y''(t) - 3y' + 2y = 2e^{3t},$$

удовлетворяющее начальным условиям $y(0) = 1$, $y'(0) = 3$.

Р е ш е н и е. По формуле (II) таблицы находим изображение правой части уравнения: $2e^{3t} \leftrightarrow \frac{2}{p-3}$. Используя формулу (18) и замечая, что $\alpha_1 = -3$, $\alpha_2 = 2$, $y_0 = 1$, $y'_0 = 3$, получим изображение решения:

$$Y(p) = \frac{\frac{2}{p-3} + p \cdot 1 + 3 + (-3) \cdot 1}{p^2 - 3p + 2} = \frac{p^2 - 3p + 2}{(p-3)(p^2 - 3p + 2)} = \frac{1}{p-3}.$$

На основании формулы (II) таблицы находим оригинал, который является искомым решением уравнения: $y(t) = e^{3t}$.

П р и м е р 2. Найти решение дифференциального уравнения

$$y''(t) + 4y = \sin t,$$

удовлетворяющее начальным условиям $y(0) = 1$, $y'(0) = 1$.

Р е ш е н и е. По формуле (IV) таблицы находим изображение правой части уравнения:

$$\sin t \leftrightarrow \frac{1}{p^2 + 1}.$$

По формуле (18) находим изображение решения, принимая во внимание, что $\alpha_1 = 0$, $\alpha_2 = 4$, $y_0 = 1$, $y'_0 = 1$,

$$Y(p) = \frac{\frac{1}{p^2 + 1} + p + 1}{p^2 + 4} = \frac{1 + (p+1)(p^2 + 1)}{(p^2 + 4)(p^2 + 1)} = \frac{p^3 + p^2 + p + 2}{(p^2 + 4)(p^2 + 1)}.$$

Преобразуем полученное выражение. Разлагая $Y(p)$ на простейшие дроби, получим:

$$Y(p) = \frac{p + \frac{2}{3}}{p^2 + 4} + \frac{1}{3(p^2 + 1)}.$$

или

$$Y(p) = \frac{p}{p^2 + 4} + \frac{2}{3(p^2 + 4)} + \frac{1}{3(p^2 + 1)}.$$

Так как

$$\frac{1}{3} \cdot \frac{1}{p^2 + 1} \rightarrow \frac{1}{3} \sin t, \quad \frac{p}{p^2 + 4} \rightarrow \cos 2t, \quad \frac{1}{3} \cdot \frac{1}{p^2 + 4} \rightarrow \frac{1}{3} \sin 2t,$$

то

$$y(t) = \frac{1}{3} \sin t + \cos 2t + \frac{1}{3} \sin 2t.$$

Глава 21. ЛИНЕЙНЫЕ ПРОСТРАНСТВА

§ 21.1. Основные понятия

1. Понятие о линейном пространстве. Ситуация, когда имеется множество каких-то элементов, которые можно *складывать между собой* и *умножать на числа*, получая в результате элементы того же самого множества, встречаются в математике очень часто. В § 2.1 изучались линейные операции над векторами: сложение, умножение вектора

на число. Складывать между собой и умножать на числа можно многочлены — в результате получаются такие же многочлены. Если складываются и умножаются на числа многочлены, степени которых не превосходят данного числа n , то и полученные при этом многочлены будут степени не выше n . Складывать между собой и умножать на числа можно и произвольные функции — в результате снова получаются функции. Если функции, к которым применяются эти операции, непрерывны на каком-то отрезке $[a, b]$ (или, скажем, на всей числовой оси), то и полученные в результате функции обладают тем же свойством. Наконец, и просто числа можно складывать между собой и умножать на числа; более того, вместо одного числа можно рассматривать пары, тройки и вообще упорядоченные наборы (строки), состоящие из n чисел: $(x_1, x_2, \dots, x_n)$. Строки можно складывать $(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$ и умножать на числа: $c(x_1, x_2, \dots, x_n) = (cx_1, cx_2, \dots, cx_n)$, получая всякий раз такую же строку.

Все это дает основание для рассмотрения так называемых линейных пространств, т. е. множеств элементов, над которыми можно производить линейные операции, удовлетворяющие определенным условиям, и в результате получать элементы того же множества.

О п р е д е л е н и е 1. Множество L элементов $x, y, z, \dots$ называется *линейным пространством* (или *линейным векторным пространством*), если для любых двух его элементов (еще называемых векторами) x, y определена сумма $x + y \in L$ и для любого $x \in L$ и любого действительного числа α определено *произведение* $\alpha x \in L$, причем для любых элементов x, y и z из L и для любых действительных чисел α и β выполнены следующие условия:

- 1) $x + y = y + x$;
- 2) $(x + y) + z = x + (y + z)$;
- 3) существует такой (нулевой) элемент $0 \in L$, что $x + 0 = x$;
- 4) для каждого x существует такой элемент $-x$ (называемый *противоположным* к x), что $x + (-x) = 0$;
- 5) $1 \cdot x = x$;
- 6) $\alpha(\beta x) = (\alpha\beta)x$;
- 7) $(\alpha + \beta)x = \alpha x + \beta x$;
- 8) $\alpha(x + y) = \alpha x + \alpha y$.

О п р е д е л е н и е 2. *Линейным подпространством* L_1 линейного пространства L называется множество L_1 , содержащееся в L , и относительно линейных операций, определенных в L , само являющееся линейным пространством.

Приведем примеры линейных пространств.

1. Множество всех свободных векторов с линейными операциями, описанными в § 2.1, является линейным пространством, так как условия 1—8 выполняются. Обозначим его через B_3 . Множество всех свободных векторов, параллельных некоторой плоскости, является линейным пространством (его обозначим через B_2), причем B_2 — линейное подпространство B_3 .

2. Пусть $\{P_n(x)\}$ — множество всех алгебраических многочленов степени не выше n (n — натуральное число). Операции сложения таких многочленов и умножения их на действительные числа определим обычными правилами. Очевидно, условия 1—8 выполняются, поэтому $\{P_n(x)\}$ является линейным пространством. Заметим, что $\{P_{n-1}(x)\}$ — линейное подпространство $\{P_n(x)\}$.

З а м е ч а н и е. Множество всех многочленов степени, точно равной натуральному числу n , не является линейным пространством, так как сумма двух таких многочленов может оказаться многочленом степени ниже n , т. е. не принадлежать рассматриваемому множеству.

3. Рассмотрим множество A_n , элементами которого служат упорядоченные совокупности n произвольных действительных чисел $(x_1, x_2, \dots, x_n)$. Каждый элемент этого множества будем обозначать одним символом, например, $x, y, \dots$ и писать $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$; действительные числа $x_1, x_2, \dots, x_n$ называются *координатами* элемента x .

Линейные операции над элементами A_n определяются формулами

$$\begin{aligned}x + y &= ((x_1 + y_1), (x_2 + y_2), \dots, (x_n + y_n)), \\ \alpha x &= (\alpha x_1, \alpha x_2, \dots, \alpha x_n).\end{aligned}$$

Отметим, что элемент $0 = (0, 0, \dots, 0)$ является нулевым, элемент $-x = (-x_1, -x_2, \dots, -x_n)$ — противоположным элементу $x = (x_1, x_2, \dots, x_n)$.

Легко проверить, что все условия 1—8 будут выполнены.

Из условий 1—8 можно получить ряд утверждений, справедливых для любых линейных пространств. В качестве примера приведем без доказательства две теоремы.

Т е о р е м а 1. *В любом линейном пространстве существует единственный нулевой элемент и для каждого x существует единственный противоположный элемент.*

Т е о р е м а 2. *В любом линейном пространстве нулевой элемент 0 равен произведению произвольного элемента x на действительное число $\alpha = 0$; для каждого элемента x противоположный элемент x' равен произведению этого элемента на действительное число $\alpha = -1$, т. е. $x' = (-1)x$.*

Разностью $x - y$ векторов x и y называется вектор $z = x + (-y)$.

2. Базис и размерность линейного пространства. В § 2.1 (п. 3) было введено понятие линейной зависимости векторов, обобщением кото-

рого является понятие *линейной зависимости элементов* любого линейного пространства.

Пусть $x, y, \dots, u$ — элементы пространства L , а $\alpha, \beta, \dots, \lambda$ — действительные числа.

Тогда выражение $\alpha x + \beta y + \dots + \lambda u$ называется *линейной комбинацией* элементов $x, y, \dots, u$.

Элементы $x, y, \dots, u$ пространства L называются *линейно зависимыми*, если существуют действительные числа $\alpha, \beta, \dots, \lambda$ не все равные нулю, для которых имеет место равенство

$$\alpha x + \beta y + \dots + \lambda u = 0. \quad (1)$$

Элементы $x, y, \dots, u$ пространства L называются *линейно независимыми*, если равенство (1) имеет место только при $\alpha = 0, \beta = 0, \dots, \lambda = 0$.

Справедлива следующая теорема 1, аналогичная известному предложению (§ 2.1, п. 3).

Т е о р е м а 1. *Элементы $x, y, \dots, u$ пространства L линейно зависимы тогда и только тогда, когда, по меньшей мере один из них является линейной комбинацией остальных.*

Справедливы и следующие предложения.

Т е о р е м а 2. *Если среди элементов $x, y, \dots, u$ имеется нулевой элемент, то эти элементы линейно зависимы.*

Действительно, если, например, $x = 0$, то равенство (1) выполняется при $\alpha = 1, \beta = 0, \dots, \lambda = 0$.

Аналогично устанавливается теорема 3.

Т е о р е м а 3. *Если часть элементов $x, y, \dots, u$ линейно зависима, то все они также линейно зависимы.*

Рассмотрим вопрос о линейной зависимости элементов пространства A_n .

Докажем, что n элементов

$$\begin{aligned} e_1 &= (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \\ e_3 &= (0, 0, 1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 0, 1) \end{aligned} \quad (2)$$

являются линейно независимыми, а совокупность этих n элементов и еще одного любого элемента $x = (x_1, x_2, \dots, x_n)$ этого пространства уже образуют линейно зависимую систему.

Действительно, линейная комбинация указанных элементов $e_1, e_2, \dots, e_n$ в силу условий 1—8 линейного пространства представляет собой элемент $\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = (\alpha_1, \alpha_2, \dots, \alpha_n)$, который является нулевым лишь при $\alpha_1 = 0, \alpha_2 = 0, \dots, \alpha_n = 0$. Это и означает, что элементы $\alpha_1, \alpha_2, \dots, \alpha_n$ линейно независимы.

На основании теоремы 1 система элементов $e_1, e_2, \dots, e_n$ и $x = (x_1, x_2, \dots, x_n)$ уже будет линейно зависимой, так как элемент x

представляет собой линейную комбинацию элементов $e_1, e_2, \dots, e_n$, т. е. $x = (x_1, x_2, \dots, x_n) = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$.

О п р е д е л е н и е 1. Линейное пространство L называется n -мерным, если в нем существуют n линейно независимых элементов, а любые $(n + 1)$ элементов уже являются линейно зависимыми. Число n в этом случае называется *размерностью* пространства L . Итак, размерность пространства L — это наибольшее возможное количество линейно независимых элементов в нем.

Так, размерность B_3 равна 3, размерность B_2 равна 2, размерность A_n равна n .

О п р е д е л е н и е 2. Совокупность n линейно независимых элементов n -мерного пространства L называется его *базисом*.

Аналогично результатам п. 6 из § 2.1 устанавливается следующее предложение.

Т е о р е м а 4. *Каждый элемент x линейного n -мерного пространства L можно представить, и притом единственным способом, в виде линейной комбинации элементов базиса, т. е. $x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$ ($e_1, e_2, \dots, e_n$ — базис; $x_1, x_2, \dots, x_n$ — действительные числа, называемые *координатами* элемента x в базисе $e_1, e_2, \dots, e_n$).*

Приведем примеры базисов конкретных линейных пространств. Базис B_3 образует любая тройка некопланарных векторов (§ 2.1, п. 6), базис B_2 образуют любые два неколлинеарные вектора (§ 2.1, п. 6). Множество n элементов (2) образуют базис в линейном пространстве A_n .

Так же, как и теорема 2 из п. 6 § 2.1, устанавливается следующая теорема.

Т е о р е м а 5. *При сложении двух элементов линейного пространства L их координаты складываются, при умножении элемента на любое действительное число α все его координаты умножаются на это число.*

Рассмотрим линейное пространство $\{P_2(x)\}$, т. е. множество всех многочленов степени ≤ 2 (п. 1). Все эти многочлены представляют собой линейные комбинации степеней $x^2, x, 1$, причем сами эти степени линейно независимы, ни одна из них не является линейной комбинацией остальных. Отсюда с учетом теоремы 1 сразу следует, что рассматриваемое пространство является трехмерным. Указанные степени образуют в этом пространстве базис.

Аналогично получается, что пространство многочленов степени $\leq n$ имеет размерность $n + 1$.

§ 21.2. Евклидово пространство

О п р е д е л е н и е. Линейное пространство называется *евклидовым*, если в нем определена операция скалярного умножения: любым

двум элементам x и y сопоставлено вещественное число, обозначаемое символом (x, y) и называемое *скалярным произведением* векторов x и y , причем выполняются следующие условия, каковы бы ни были векторы x, y, z и число α :

1. $(x, y) = (y, x)$.
2. $(x, x) > 0$, если $x \neq 0$.
3. $(x + y, z) = (x, z) + (y, z)$.
4. $(\alpha x, y) = \alpha(x, y)$.

Отметим, что из условия 4 при $\alpha = 0$ следует равенство $(0, y) = 0$, т. е. скалярное произведение нулевого вектора на любой вектор равно нулю; равен нулю, в частности, скалярный квадрат нулевого вектора.

Приведем примеры евклидовых пространств.

1. В линейном пространстве B_3 скалярное произведение двух векторов $\vec{a}$ и $\vec{b}$ определим так, как в § 2.2 (п. 1); условия 1—4 для него будут выполнены (см. § 2.2, п. 1, свойства 1—4). Следовательно, линейное пространство B_3 всех свободных векторов с указанным определением скалярного произведения является евклидовым пространством. Евклидово пространство образует также совокупность векторов, параллельных некоторой плоскости (или некоторой прямой), т. е. линейные пространства B_2 и B_1 .

2. Рассмотрим n -мерное линейное пространство A_n , упорядоченных совокупностей n действительных чисел. Скалярное произведение двух его элементов $x = (x_1, x_2, \dots, x_n)$ и $y = (y_1, y_2, \dots, y_n)$ определим формулой

$$(x, y) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n \quad (1)$$

по аналогии с формулой (7) из § 2.2. Легко видеть, что условия 1—4 для него выполняются. Таким образом, рассматриваемое линейное пространство со скалярным произведением (1) является евклидовым пространством; его обозначим через E_n .

В евклидовом пространстве вводят понятие длины (нормы) элемента x по формуле $|x| = \sqrt{(x, x)}$.

Из условия 4 скалярного произведения следует, что $|x| > 0$ для $x \neq 0$; $|x| = 0$ при $x = 0$.

С помощью условий 1 и 3 получаем $|\alpha x| = \sqrt{(\alpha x, \alpha x)} = \sqrt{\alpha^2 (x, x)} = |\alpha| |x|$, т. е. $|\alpha x| = |\alpha| |x|$, откуда, в частности, следует, что $|0| = 0$. Элемент, длина которого равна единице, называется *нормированным*. Для каждого элемента $x \neq 0$ можно получить нормированный, умножив x на число $\alpha = \frac{1}{|x|}$; последнее действие называется *нормированием* элемента

(для обычных векторов употребляемый термин — «единичный вектор» — см. § 2.1).

Докажем, что

$$|(x, y)| \leq |x||y|. \quad (2)$$

Действительно,

$$0 \leq |x + \alpha y|^2 = (x + \alpha y, x + \alpha y) = (y, y)\alpha^2 + 2(x, y)\alpha + (x, x).$$

Полученную сумму справа рассматриваем как квадратичный трехчлен относительно α . Так как трехчлен сохраняет знак, то его дискриминант неположителен, т. е.

$$(x, y)^2 - (x, x)(y, y) \leq 0$$

или

$$(x, y)^2 \leq (x, x)(y, y), \quad (3)$$

откуда

$$|(x, y)| = \sqrt{(x, y)^2} \leq \sqrt{(x, x)(y, y)} = \sqrt{(x, x)}\sqrt{(y, y)} = |x||y|,$$

что и требовалось доказать.

Неравенство (2) носит название неравенства Коши — Буняковского (В. Я. Буняковский (1804—1889) — русский математик).

С использованием неравенства (2) имеем

$$\begin{aligned} |x + y|^2 &= (x + y, x + y) = (x, x) + 2(x, y) + (y, y) \leq \\ &\leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2, \end{aligned}$$

откуда

$$|x + y| \leq |x| + |y|. \quad (4)$$

Неравенство (4) называется *неравенством треугольника*.

В евклидовом пространстве B_3 с обычным определением скалярного произведения норма вектора совпадает с его длиной и

$$(\bar{a}, \bar{b})^2 \leq |\bar{a}|^2 |\bar{b}|^2, \quad |\bar{a} + \bar{b}| \leq |\bar{a}| + |\bar{b}|.$$

В евклидовом пространстве E_n со скалярным произведением (1) норма элемента $x = (x_1, x_2, \dots, x_n)$ определяется формулой

$$|x| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2},$$

а неравенства (3) и (4) принимают вид:

$$\begin{aligned} (x_1 y_1 + x_2 y_2 + \dots + x_n y_n)^2 &\leq (x_1^2 + x_2^2 + \dots + x_n^2)(y_1^2 + y_2^2 + \dots + y_n^2), \\ \sqrt{(x_1 + y_1)^2 + (x_2 + y_2)^2 + \dots + (x_n + y_n)^2} &\leq \\ &\leq \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} + \sqrt{y_1^2 + y_2^2 + \dots + y_n^2}. \end{aligned}$$

В евклидовом пространстве вводится понятие угла между ненулевыми элементами x и y по формуле

$$\cos(\widehat{x, y}) = \frac{(x, y)}{|x||y|} = \frac{(x, y)}{\sqrt{(x, x)}\sqrt{(y, y)}}. \quad (5)$$

Для векторов это определение совпадает с обычным определением угла. Отметим, что $|\cos(\widehat{x, y})| \leq 1$ в силу неравенства (2).

Если элементы x и y линейно зависимы, т. е. $y = \alpha x$, то из формулы (5) следует, что $\cos(\widehat{x}, \widehat{y}) = 1$ при $\alpha > 0$, $\cos(\widehat{x}, \widehat{y}) = -1$ при $\alpha < 0$. Обратное также верно: если $\cos(\widehat{x}, \widehat{y}) = \pm 1$, то элементы x и y линейно зависимы.

Элементы x и y называются *ортогональными*, если $(x, y) = 0$. Система элементов $e_1, e_2, \dots, e_n$ называется *ортогональной*, если они попарно ортогональны. Ортогональная система элементов всегда линейно независима. В самом деле, равенство

$$\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = 0 \quad (6)$$

выполнено лишь при $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$, так как при скалярном умножении равенства (6) на любой вектор e_k этой системы получаем $\alpha_k (e_k, e_k) = 0$, откуда $\alpha_k = 0$ ($k = 1, 2, \dots, n$).

В конечномерном евклидовом пространстве наиболее удобно пользоваться *ортогональным базисом*, т. е. базисом, который одновременно является ортогональной системой. Действительно, если $e_1, e_2, \dots, e_n$ такой базис, то любой элемент x можно представить в виде:

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n,$$

откуда (после скалярного умножения на e_k)

$$\alpha_k = \frac{(x, e_k)}{(e_k, e_k)} \quad (k = 1, 2, \dots, n). \quad (7)$$

Формула (7) упрощается, если элементы $e_1, e_2, \dots, e_n$ являются нормированными; тогда знаменатель равен единице.

Ортогональный базис из нормированных элементов называется *евклидовым базисом*. В трехмерном евклидовом пространстве B_3 таким базисом является базис $\vec{i}, \vec{j}, \vec{k}$.

§ 21.3. Линейные операторы

1. Определение и примеры.

О п р е д е л е н и е. Будем говорить, что в линейном пространстве L задан *оператор*, или *преобразование*, $\mathcal{A}$, если каждому элементу x из L по некоторому правилу поставлен в соответствие элемент $\mathcal{A}x$, принадлежащий L , т. е. $\mathcal{A}$ — это функция с областью определения L и множеством значений, принадлежащим L . Оператор $\mathcal{A}$ называется *линейным*, если для любых двух элементов x и y и для любого действительного числа λ выполняются равенства $\mathcal{A}(x + y) = \mathcal{A}x + \mathcal{A}y$, $\mathcal{A}(\lambda x) = \lambda \mathcal{A}x$.

Пусть $\mathcal{A}$ — линейный оператор в линейном пространстве L и $e_1, e_2, \dots, e_n$ — базис в L . Рассмотрим элементы $\mathcal{A}e_1, \mathcal{A}e_2, \dots, \mathcal{A}e_n$. Так как это

через $\mathcal{E}$. Таким образом, $\mathcal{E}x = x$. Поэтому $\mathcal{E}e_k = e_k$, $k = 1, 2, \dots, n$, и следовательно, матрица оператора $\mathcal{E}$ имеет вид

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

т. е. является единичной матрицей E .

П р и м е р 4. Обозначим через $\mathcal{O}$ так называемый нулевой оператор, определяемый равенством $\mathcal{O}x = 0$ для всех x из L . Матрица этого оператора состоит из одних нулей.

2. Действия над линейными операторами. Пусть $\mathcal{A}$ и $\mathcal{B}$ — два произвольных линейных оператора в линейном пространстве L , λ — произвольное действительное число и x — любой элемент из L .

О п р е д е л е н и е 1. Суммой линейных операторов $\mathcal{A}$ и $\mathcal{B}$ называется оператор $\mathcal{C}$, определяемый равенством $\mathcal{C}x = \mathcal{A}x + \mathcal{B}x$. Обозначение, $\mathcal{C} = \mathcal{A} + \mathcal{B}$

О п р е д е л е н и е 2. Произведением линейного оператора $\mathcal{A}$ на число λ называется оператор $\mathcal{C}$, определяемый равенством $\mathcal{C}x = \lambda(\mathcal{A}x)$.

Обозначение: $\mathcal{C} = \lambda \mathcal{A}$.

О п р е д е л е н и е 3. Произведением линейного оператора $\mathcal{A}$ на линейный оператор $\mathcal{B}$ называется оператор $\mathcal{C}$ определяемый равенством $\mathcal{C}x = \mathcal{A}(\mathcal{B}x)$.

Обозначение: $\mathcal{C} = \mathcal{A}\mathcal{B}$.

Операторы $\mathcal{A} + \mathcal{B}$ и $\mathcal{A}\mathcal{B}$ являются линейными. Докажем это, например, для оператора $\mathcal{A} + \mathcal{B}$. Действительно, для любых элементов x и y и любого действительного числа λ имеем:

$$\begin{aligned} (\mathcal{A} + \mathcal{B})(x + y) &= \mathcal{A}(x + y) + \mathcal{B}(x + y) = \mathcal{A}x + \mathcal{A}y + \mathcal{B}x + \mathcal{B}y = (\mathcal{A} + \mathcal{B})x + (\mathcal{A} + \mathcal{B})y, \\ (\mathcal{A} + \mathcal{B})(\lambda x) &= \mathcal{A}(\lambda x) + \mathcal{B}(\lambda x) = \lambda(\mathcal{A}x) + \lambda(\mathcal{B}x) = \\ &= \lambda(\mathcal{A}x + \mathcal{B}x) = \lambda(\mathcal{A} + \mathcal{B})x. \end{aligned}$$

Т е о р е м а. Пусть $\mathcal{A}$ и $\mathcal{B}$ — два линейных оператора в линейном пространстве L ; λ — произвольное действительное число; $e_1, e_2, \dots, e_n$ — базис в пространстве L ; $A = (a_{ik})$, $B = (b_{ik})$ — матрицы линейных операторов $\mathcal{A}$, $\mathcal{B}$ в заданном базисе. Тогда в этом базисе: 1) матрица линейного оператора $\mathcal{A} + \mathcal{B}$ есть $A + B$; 2) матрица линейного оператора $\lambda \mathcal{A}$ есть λA ; 3) матрица линейного оператора $\mathcal{A}\mathcal{B}$ есть AB .

Д о к а з а т е л ь с т в о. Ограничимся доказательством утверждения 1). Имеем:

$$(\mathcal{A} + \mathcal{B})e_k = \mathcal{A}e_k + \mathcal{B}e_k = \sum_{i=1}^n a_{ik} e_i + \sum_{i=1}^n b_{ik} e_i = \sum_{i=1}^n (a_{ik} + b_{ik}) e_i, \quad (1)$$

$k = 1, 2, \dots, n$. Здесь использованы определения суммы линейных операторов и матрицы линейного оператора. Равенство (1) означает, что на

пересечении i строки и k столбца матрицы линейного оператора $\mathcal{A} + \mathcal{B}$ в базисе $e_1, e_2, \dots, e_n$ находится число $a_{ik} + b_{ik}$, т. е. эта матрица имеет вид $A + B$.

Из этой теоремы и известных свойств матриц (§ 3.1) вытекают, в частности, следующие свойства сложения и умножения операторов:

1. $\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$,
2. $(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C})$,
3. $\mathcal{A} + \mathcal{O} = \mathcal{A}$,
4. $\mathcal{A}(\mathcal{B}\mathcal{C}) = (\mathcal{A}\mathcal{B})\mathcal{C}$,
5. $0 \cdot \mathcal{A} = \mathcal{O}$,
6. $\mathcal{A}\mathcal{E} = \mathcal{E}\mathcal{A} = \mathcal{A}$,

произведение же $\mathcal{A}\mathcal{B}$, вообще говоря, отличается от произведения $\mathcal{B}\mathcal{A}$, так как произведение матриц, вообще говоря, не подчиняется переместительному закону (§ 3.1, п. 5).

Обратным к линейному оператору $\mathcal{A}$ называется оператор $\mathcal{A}^{-1}$ такой, что $\mathcal{A}\mathcal{A}^{-1} = \mathcal{A}^{-1}\mathcal{A} = \mathcal{E}$, где $\mathcal{E}$ — тождественный оператор. Оператор $\mathcal{A}$ имеет обратный (и в том случае $\mathcal{A}$ называется *невырожденным*) в том и только в том случае, когда его матрица A невырождена, при этом матрица обратного оператора $\mathcal{A}^{-1}$ равна обратной матрице A^{-1} .

Очевидно, рассмотренные выше в примерах 1 и 3 операторы — невырожденные, а в примерах 2—4 — вырожденные.

3. Собственные векторы и собственные значения линейного оператора.

О п р е д е л е н и е 1. *Собственным вектором* линейного оператора $\mathcal{A}$, действующего в линейном пространстве L , называется ненулевой вектор x из L такой, что для некоторого числа λ выполняется соотношение

$$\mathcal{A}x = \lambda x. \quad (2)$$

Число λ в этом случае называется *собственным значением (числом)* линейного оператора $\mathcal{A}$. Говорят, что собственный вектор x принадлежит собственному значению λ .

Из (2) следует, что при действии линейного оператора собственный вектор просто умножается на некоторое число. Рассмотрим матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

соответствующую линейному оператору $\mathcal{A}$ в базисе $e_1, e_2, \dots, e_n$ и век-

тор-столбец $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, составленный из координат собственного вектора

Всякое решение λ уравнения (6) обладает тем свойством, что соответствующая ему система уравнений (5) имеет ненулевое решение $x_1, x_2, \dots, x_n$. При этом вектор-столбец $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ является собственным векто-

ром матрицы A , а вектор $x_1 e_1 + x_2 e_2 + \dots + x_n e_n$ — собственным вектором линейного оператора $\mathcal{A}$.

Заметим, что корни характеристического уравнения матрицы называются также собственными значениями, собственными числами или корнями матрицы.

Справедлива (см. [1]) следующая теорема.

Т е о р е м а. *Характеристический многочлен линейного оператора не зависит от выбора базиса.*

П р и м е р 1. Для тождественного оператора все ненулевые векторы линейного пространства L являются, очевидно, собственными (с собственным значением, равным единице).

П р и м е р 2. Для нулевого оператора все ненулевые векторы пространства являются собственными (с собственным значением, равным нулю).

П р и м е р 3. Найти собственные числа и собственные векторы линейного оператора с матрицей $A = \begin{pmatrix} 5 & 8 \\ 4 & 9 \end{pmatrix}$.

Характеристическое уравнение имеет вид $\begin{vmatrix} 5-\lambda & 8 \\ 4 & 9-\lambda \end{vmatrix} = 0$, или $\lambda^2 - 14\lambda + 13 = 0$; характеристические числа $\lambda_1 = 1, \lambda_2 = 13$.

Для определения координат собственных векторов получаем две системы линейных уравнений:

$$\begin{cases} (5-\lambda_1)x_1 + 4x_2 = 0, \\ 8x_1 + (9-\lambda_1)x_2 = 0, \end{cases} \quad \begin{cases} (5-\lambda_2)x_1 + 4x_2 = 0, \\ 8x_1 + (9-\lambda_2)x_2 = 0. \end{cases}$$

Так как $\lambda_1 = 1$, то первую систему можно записать следующим образом:

$$\begin{cases} 4x_1 + 4x_2 = 0, \\ 8x_1 + 8x_2 = 0, \end{cases}$$

т. е. $x_1 + x_2 = 0$ или $x_2 = -x_1$. Следовательно, решение этой системы имеет вид $x_1 = c_1, x_2 = -c_1$, где c_1 — произвольная величина. Таким образом, собственному числу $\lambda = 1$ соответствует семейство собственных векторов $x = c_1 e_2 - c_1 e_1$, т. е. $x = c_1(e_2 - e_1)$.

Значение $\lambda_2 = 13$ приводит к системе уравнений $\begin{cases} -8x_1 + 4x_2 = 0, \\ 8x_1 - 4x_2 = 0. \end{cases}$ т. е. $x_2 = 2x_1$. Полагая

$x_1 = c_2$, получаем $x_2 = 2c_2$. Следовательно, собственному числу $\lambda = 13$ соответствует семейство собственных векторов $x = c_2(e_1 + 2e_2)$.

Придавая в равенствах $x = c_1(e_2 - e_1), x = c_2(e_1 + 2e_2)$ величинам c_1 и c_2 всевозможные числовые значения, будем получать всевозможные векторы линейного оператора $\mathcal{A}$.

Глава 22. НЕКОТОРЫЕ ЧИСЛЕННЫЕ МЕТОДЫ

§ 22.1. Численное интегрирование

Один из численных методов (метод Гаусса) уже был рассмотрен ранее (см. § 3.4, пп. 6 и 9). Ниже рассмотрим еще два других метода: численное интегрирование и численное решение уравнений.

Пусть речь идет о вычислении интеграла

$$\int_a^b f(x)dx, \quad (1)$$

где $f(x)$ — функция, непрерывная на сегменте $[a; b]$. Формула Ньютона — Лейбница не всегда позволяет вычислить интеграл (1), так как далеко не всегда мы знаем первообразную для подынтегральной функции. Здесь на помощь приходит приближенное вычисление определенных интегралов. Кстати, на практике часто и не требуется знать точное значение данного интеграла. Рассмотрим три способа приближенного вычисления определенных интегралов: метод прямоугольников, метод трапеций и метод параболических трапеций, или метод Симпсона (Томас Симпсон (1710—1761) — английский математик).

Если $f(x) \geq 0$ на сегменте $[a; b]$, то, как известно (см. § 9.6), интеграл (1) представляет собой площадь криволинейной трапеции $aABb$, ограниченной сверху графиком функции $y = f(x)$ снизу отрезком $[a; b]$ оси Ox , с боков отрезками прямых $x = a$, $x = b$. Составим интегральную сумму, соответствующую делению сегмента $[a; b]$ на n равных частей точками $a = x_0 < x_1 < x_2 < \dots < x_{k-1} < x_k < \dots < x_n = b$ и выбору точек $\tau_k = x_{k-1}$, ($k = 1, 2, \dots, n$): $\sum_{k=1}^n f(x_{k-1})\Delta x_k$. Но $\Delta x_k = \frac{b-a}{n}$. Значит,

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

Сумма $\sum_{k=1}^n f(x_{k-1})\Delta x_k$ есть площадь ступенчатой фигуры, изображенной на рисунке 212. Поэтому площадь указанной криволинейной трапеции приближенно равна площади этой ступенчатой фигуры, т. е.

$$\int_a^b f(x)dx \approx \frac{b-a}{n} \sum_{k=1}^n f(x_{k-1}). \quad (2)$$

Если $\tau_k = x_k$ ($k = 1, 2, \dots, n$), то аналогично получим приближенную формулу

$$\int_a^b f(x)dx \approx \frac{b-a}{n} \sum_{k=1}^n f(x_k). \quad (3)$$

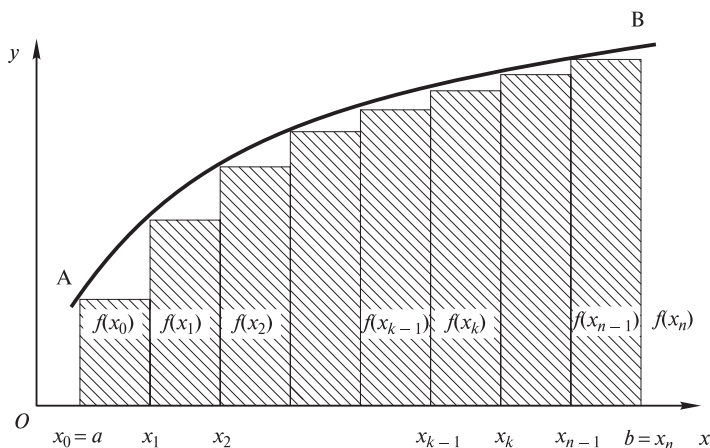


Рис. 212

Формулы (2) и (3) называются *формулами прямоугольников*. Если функция $f(x)$ возрастает на $[a; b]$, то формула (2) дает значение интеграла (1) по недостатку, а формула (3) — по избытку. Поэтому для повышения точности естественно взять среднее арифметическое этих формул:

$$\int_a^b f(x) dx \approx \frac{b-a}{n} \left(\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-1} f(x_k) \right). \quad (4)$$

Эту формулу называют *формулой трапеций*, так как ее геометрический смысл связан с заменой площади каждой прямоугольной полоски, на которые разбивается криволинейная трапеция, на площадь прямолинейной трапеции.

П р и м е р. Вычислим значение интеграла

$$\int_0^2 e^{-x^2} dx: \quad (5)$$

а) по формулам прямоугольников (2) и (3) при $n=10$ и $n=20$; б) по формуле трапеций при $n=10$ и $n=20$.

При рассмотрении этого примера будем использовать, например, микрокалькулятор «Электроника БЗ-36».

Вычислим сначала приближенное значение интеграла (5) по формуле прямоугольников (2) при $n=10$. В этом случае $x_0=0, x_1=0,2, x_2=0,4, x_3=0,6, x_4=0,8, x_5=1, x_6=1,2, x_7=1,4, x_8=1,6, x_9=1,8$ и потому

$$\int_0^2 e^{-x^2} dx \approx \frac{2-0}{10} (e^{-0^2} + e^{-0,2^2} + \dots + e^{-1,8^2}).$$

Значение подынтегральной функции в точке x , вычисляется по программе $xx = /- / Fe^x F\Pi^x$. Следовательно, программа вычисления интеграла имеет вид:

```

0x = /- / Fe^x F\Pi^+
0.2x = /- / Fe^x F\Pi^+
. . . . .
1.8x = /- / Fe^x F\Pi^+
2 ÷ 10x F\Pi =

```

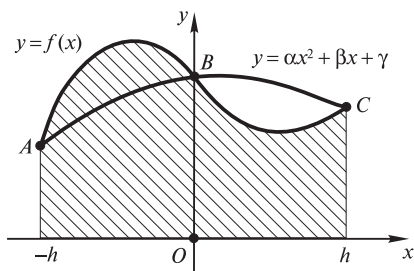


Рис. 213

О т в е т: 0,98.

Вычисление по формуле (3) при $n = 10$ производится по аналогичной программе с той лишь разницей, что начальным значением аргумента является 0,2, а конечным значением — число 2. Получаем ответ: 0,78. Среднее арифметическое этих ответов 0,88 дает приближенное значение интеграла по формуле трапеций.

Аналогично производятся подсчеты при $n = 20$. Ответ по формуле трапеций: 0,882.

Метод Симпсона дает более точную приближенную формулу вычисления интеграла (1).

Предположим сначала, что пределы интегрирования имеют вид $a = -h$, $b = h$, т. е. расположены симметрично относительно точки O . Если h мало, то кривую $y = f(x)$ приближенно можно заменить параболой

$$y = \alpha x^2 + \beta x + \gamma \equiv F(x), \quad (6)$$

проходящей через точки $A(-h; f(-h))$, $B(0; f(0))$ и $C(h; f(h))$ (рис. 213). Тогда

$$\int_{-h}^h f(x) dx$$

приближенно будет равен интегралу

$$\int_{-h}^h F(x) dx = \left[\frac{\alpha x^3}{3} + \frac{\beta x^2}{2} + \gamma x \right]_{-h}^h = \frac{2}{3} \alpha h^3 + 2\gamma h = \frac{2}{3} h(\alpha h^2 + 3\gamma).$$

Полагая в выражении (6) последовательно $x = -h, 0, h$, получаем:

$$F(-h) = \alpha h^2 - \beta h + \gamma, \quad F(0) = \gamma, \quad F(h) = \alpha h^2 + \beta h + \gamma.$$

Отсюда

$$F(-h) + 4F(0) + F(h) = 2(\alpha h^2 + 3\gamma).$$

Поэтому

$$\int_{-h}^h F(x) dx \approx \frac{h}{3} (F(-h) + 4F(0) + F(h)).$$

Но

$$F(-h) = f(-h), \quad F(0) = f(0), \quad F(h) = f(h).$$

И значит,

$$\int_{-h}^h f(x)dx \approx \frac{h}{3}(f(-h) + 4f(0) + f(h))$$

(формула Симпсона).

Случай общего расположения пределов интегрирования сводится к рассмотренному, если перенести начало координат в точку $x = \frac{a+b}{2}$ оси абсцисс. Тогда будем иметь

$$\int_a^b f(x)dx \approx \frac{b-a}{6}(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)). \quad (7)$$

Для увеличения точности вычисления интеграла (1) разобьем отрезок $[a; b]$ точками $a = x_0, x_1, \dots, x_{2n-1}, x_{2n} = b$ на $2n$ равных частей ($f(x_k) = y_k$). Применяя к отрезку $[x_{2k-2}; x_{2k}]$ ($k = 1, 2, \dots, n$) формулу (7), получим:

$$\int_{x_{2k-2}}^{x_{2k}} f(x)dx \approx \frac{b-a}{6n}(y_{2k-2} + 4y_{2k-1} + y_{2k}).$$

Отсюда с использованием свойства 4 (см. § 9.7, п. 1) получим:

$$\begin{aligned} \int_a^b f(x)dx &\approx \frac{b-a}{6n} \sum_{k=1}^n (y_{2k-2} + 4y_{2k-1} + y_{2k}) = \\ &= \frac{b-a}{6n} ((y_0 + y_{2n}) + 4(y_1 + y_3 + \dots + y_{2n-1}) + 2(y_2 + y_4 + \dots + y_{2n-2})). \end{aligned} \quad (8)$$

Заметим, что все четыре формулы (2), (3), (4), (8) тем точнее, чем больше n .

Наконец, отметим, что каждый из изложенных методов содержит четкий алгоритм их нахождения. Это позволяет широко применять эти методы для вычислений на ЭВМ.

§ 22.2. Численное решение уравнений

1. Постановка вопроса. Число x_0 называется *корнем* функции $f(x)$, если равно нулю значение этой функции при $x = x_0$. Это число x_0 вместе с тем называется *корнем уравнения*

$$f(x) = 0. \quad (1)$$

Такие уравнения могут быть алгебраическими, если функция $f(x)$ алгебраическая или трансцендентными в противном случае.

Из школьной алгебры известно, как решить уравнение (1), если $f(x)$ линейная или квадратичная функция. Существуют методы решения алгебраических уравнений третьей и четвертой степени. Формулы, выражающие корни этих уравнений через их коэффициенты, являются довольно громоздкими, поэтому ими редко пользуются. Что касается алгебраических уравнений степени n ($n \geq 5$), то в общем случае для них нет таких формул: эти уравнения в радикалах не решаются.

С другой стороны, при применении точных методов полученными результатами во многих случаях можно воспользоваться лишь приближенно, например результатом $x = \sqrt{3}$ и т. п. По указанным причинам широко пользуются приближенными методами решения уравнений (1).

Прежде можно начать с примерного, хотя и довольно грубого построения графика функции $f(x)$. Если при этом обнаружится, что на каком-нибудь отрезке $[a; b]$ эта функция непрерывна и принимает на его концах значения противоположных знаков, то в силу теоремы 2 из п. 2 § 7.6 функция $f(x)$ должна иметь на интервале $(a; b)$ хотя бы один нуль, т. е. уравнение (1) имеет там по крайней мере один корень. Если к тому же функция $f(x)$ на этом интервале монотонна, то такой корень здесь только один, т. е. этот корень отделен от остальных. Если обозначить этот, пока неизвестный, корень через α , то можно ручаться, что $a < \alpha < b$. Для дальнейшего уточнения значения α применяются различные методы (см. ниже).

2. Метод проб. Пусть для определенности $f(a) < 0, f(b) > 0$. Тогда берут произвольное значение c между a и b и вычисляют $f(c)$, причем здесь существен только знак $f(c)$. Допустим, что $f(c) > 0$. Значит, $a < \alpha < c$. В этом случае берут какое-нибудь значение d между a и c и вычисляют $f(d)$; если $f(d) < 0$, то $d < \alpha < c$, и т. д. При этом значения $c, d, \dots$ берутся более или менее произвольными, удобными для вычисления. Заметим, что если $|f(a)|$ значительно меньше, чем $f(b)$, то довольно вероятно, что α окажется ближе к a , чем к b , и поэтому c следует взять поближе к a и т. п.

3. Метод хорд. Метод хорд заключается в том, что в качестве c берется не произвольное значение между a и b , а абсцисса x_1 точки пересечения оси Ox с хордой графика, проведенной через точки $M(a; f(a))$ и $N(b; f(b))$ (рис. 214). Другими словами, приближенно принимаем дугу графика за отрезок прямой, т. е. производим линейную интерполяцию, что является достаточно обоснованным, если интервал $(a; b)$ не слишком велик. Для нахождения x_1 напишем уравнение хорды MN (см. § 1.5, п. 2)

$$\frac{x-a}{b-a} = \frac{y-f(a)}{f(b)-f(a)}.$$

Полагая здесь $y = 0$, найдем, что

$$x_1 = a - \frac{(b-a)f(a)}{f(b)-f(a)}. \quad (2)$$

Если необходимо, это построение можно повторить. Применив к отрезку $[x_1; b]$ (рис. 214) формулу (2), получим для искомого корня второе приближение

$$x_2 = x_1 - \frac{(b-x_1)f(x_1)}{f(b)-f(x_1)}$$

и т. д.

$$x_n = x_{n-1} - \frac{(b-x_{n-1})f(x_{n-1})}{f(b)-f(x_{n-1})}. \quad (3)$$

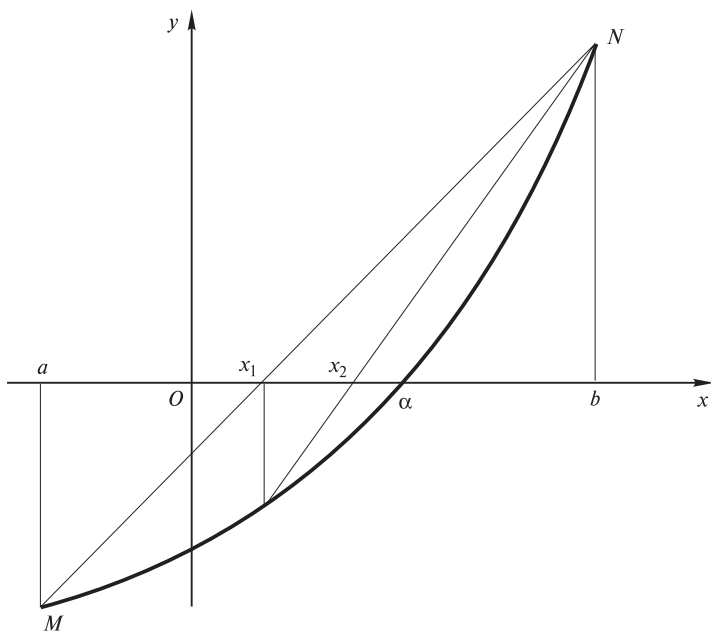


Рис. 214

Из геометрических соображений следует, что последовательность чисел $x_1, x_2, \dots, x_n, \dots$ сходится к корню α , т. е.

$$\lim_{n \rightarrow \infty} x_n = \alpha.$$

Пользуясь формулой (3), вычисления проводят до тех пор, пока в двух последовательных приближениях не станут совпадать десятичные знаки с учетом заданной точности.

П р и м е р. Функция $f(x) = x^3 - 2x^2 - 4x - 7$ на концах отрезка $[3; 4]$ имеет противоположные знаки, на нем непрерывна и производная $f'(x) = 3x^2 - 6x - 4$ сохраняет в интервале $(3; 4)$ знак плюс. Значит, внутри интервала $(3; 4)$ лежит один корень уравнения

$$x^3 - 2x^2 - 4x - 7 = 0.$$

Найдем его с точностью до 0,01. Формула (2) дает:

$$x_1 = 3 - \frac{(4-3)(-10)}{9-(-10)} = 3 + \frac{10}{19} \approx 3,53.$$

Теперь вычисляем $f(3,53) \approx -2,05$. Из двух отрезков $[3; 3,53]$, $[3,53; 4]$ выбираем второй, ибо на его концах знаки $f(x)$ противоположны. Находим второе приближение:

$$x_2 = 3,53 - \frac{(4-3,53)f(3,53)}{f(4) - f(3,53)} \approx 3,62.$$

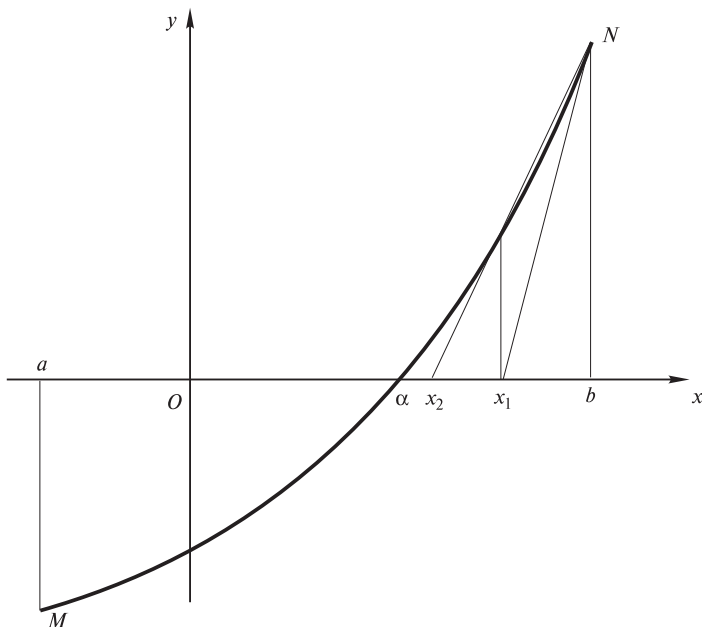


Рис. 215

Значение $f(3,62) \approx -0,24$ отрицательно, поэтому берем отрезок $[3,62; 4]$. Находим:

$$x_3 = 3,62 - \frac{(4 - 3,62)f(3,62)}{f(4) - f(3,62)} \approx 3,63$$

и $f(3,63) \approx 0,04$. По ходу выкладок надо ожидать, что x_4 будет отличаться от x_3 менее чем на 0,01 и, значит, x_3 дает искомое приближение.

4. Метод касательных. В методе касательных (он же называется методом Ньютона) за s берется абсцисса x_1 точки пересечения оси Ox с касательной, проведенной к графику в одном из концов рассматриваемой дуги (в верхнем, если дуга MN обращена вогнутостью кверху, и в нижнем, если — книзу). Уравнение касательной, изображенной на рисунке 215, имеет вид (см. § 8.2, п. 2):

$$y - f(b) = f'(b)(x - b),$$

откуда, полагая $y = 0$, находим:

$$x_1 = b - \frac{f(b)}{f'(b)}, \quad (4)$$

если $f(b) \neq 0$. И здесь построение можно повторить (см. рис. 215):

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \quad (5)$$

и т. д.

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}.$$

Как и в методе хорд, последовательность $\{x_n\}$ сходится к корню α . Степень приближения можно определить так же, как и в способе хорд.

П р и м е р. Вычислить с точностью до 0,01 корень уравнения $x^3 - 2x^2 - 4x - 7 = 0$, содержащийся в интервале (3; 4).

Имеем $f'(x) = 3x^2 - 4x - 4$, $f''(x) = 6x - 4$. Обе производные сохраняют в интервале (3; 4) знак плюс. Поэтому используем формулу (4). Первое приближение:

$$x_1 = 4 - \frac{f(4)}{f'(4)} = 4 - \frac{9}{28} \approx 3,68.$$

Далее находим $f(3,68) \approx 1,03$, $f'(3,68) \approx 21,9$ и по формуле (5):

$$x_2 = 3,68 - \frac{1,03}{21,9} \approx 3,68 - 0,047 = 3,633.$$

Последующие приближения будут все меньше и меньше, причем по ходу выкладок можно предвидеть, что дальнейшие уточнения корня не повлияют на цифру сотен. Поэтому с точностью до 0,01 искомое приближение есть 3,63.

5. Метод итераций. Методы, описанные выше (проб, хорд, касательных), принадлежат к числу *итерационных* методов (иначе говоря, методов последовательных приближений), в которых некоторый единообразный процесс последовательно повторяется («итерируется», от латинского «итерация» — повторение), в результате чего получаются все более точные приближенные решения. Это единообразие имеет многочисленные удобства, в частности, в применении быстродействующих вычислительных машин.

В общем виде в применении к уравнению (1) метод итераций выглядит так: уравнение (1) приводится к виду $x = \varphi(x)$, где $\varphi(x)$ удовлетворяет в $(a; b)$ условию $|\varphi'(x)| < 1$, если такое приведение возможно. Вычислительная формула метода имеет вид

$$x_{n+1} = \varphi(x_n) \quad (n = 0, 1, 2, \dots),$$

где x_0 — любое число из промежутка $(a; b)$, в котором содержится корень c . Можно доказать, что $\lim_{n \rightarrow \infty} x_n = c$.

П р и м е р. Уравнение $x^3 - 6x + 2 = 0$ можно записать в виде $x = \frac{1}{6}(x^3 + 2) \equiv \varphi(x)$. В интервале (0; 1) имеем $0 < \varphi'(x) = \frac{1}{2}x^2 < 1$.

Вычисления ведем по формуле

$$x_{n+1} = \frac{1}{6}(x_n^3 + 2).$$

Положив $x_0 = 0$, получим $x_1 = \frac{1}{3}$, затем $x_2 = \frac{55}{162}$ и т. д.

Приложения

Приложение 1

Основные сведения из школьной математики

Числовые неравенства

Если $a > b$, то $b < a$.

Если $a > b$ и $b > c$, то $a > c$.

Если $a > b$, то $a + c > b + c$.

Если $a > b$ и $c > 0$, то $ac > bc$.

Если $a > b$ и $c < 0$, то $ac < bc$.

Если $a > b$ и $c > d$, то $a + c > b + d$.

Если $a > 0$, $b > 0$, $c > 0$, $d > 0$, причем $a > b$ и $c > d$, то $ac > bd$.

Если $a > b > 0$ и n — натуральное число, то $a^n > b^n$.

Разложение на множители

$$a^2 - b^2 = (a - b)(a + b),$$

$$a^3 - b^3 = (a - b)(a^2 + ab + b^2),$$

$$a^3 + b^3 = (a + b)(a^2 - ab + b^2),$$

$$a^2 + 2ab + b^2 = (a + b)^2,$$

$$a^2 - 2ab + b^2 = (a - b)^2,$$

$$a^3 + 3a^2b + 3ab^2 + b^3 = (a + b)^3,$$

$$a^3 - 3a^2b + 3ab^2 - b^3 = (a - b)^3,$$

$$ax^2 + bx + c = a(x - x_1)(x - x_2), \text{ где } x_1 \text{ и } x_2 \text{ — корни уравнения } ax^2 + bx + c = 0,$$

$$x^n - a^n = (x - a)(x^{n-1} + ax^{n-2} + a^2x^{n-3} + \dots + a^{n-2}x + a^{n-1}).$$

Квадратное уравнение $ax^2 + bx + c = 0$

$$x_{1,2} = \frac{-b \pm \sqrt{D}}{2a} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \text{ (формула корней квадратного уравнения).}$$

$$x_{1,2} = \frac{-k \pm \sqrt{k^2 - ac}}{a}, \text{ где } k = \frac{b}{2} \text{ (формула корней квадратного уравнения в случае, если } b \text{ —}$$

четное число).

Теорема Виета. Если x_1 и x_2 — корни приведенного квадратного уравнения $x^2 + px + q = 0$, то

$$x_1 + x_2 = -p,$$

$$x_1 x_2 = q.$$

Арифметическая прогрессия

$a_{n+1} = a_n + d$ (определение арифметической прогрессии),

$a_n = a_1 + d(n-1)$ (формула n -го члена),

$a_n = \frac{a_{n-1} + a_{n+1}}{2}$ (характеристическое свойство),

$S_n = \frac{a_1 + a_n}{2} \cdot n = \frac{2a_1 + d(n-1)}{2} \cdot n$ (формула суммы n первых членов).

Геометрическая прогрессия

$b_{n+1} = b_n q$, $b_1 \neq 0$, $q \neq 0$ (определение геометрической прогрессии),

$b_n = b_1 q^{n-1}$ (формула n -го члена),

$b_n^2 = b_{n-1} b_{n+1}$ (характеристическое свойство),

$S_n = \frac{b_n q - b_1}{q - 1} = \frac{b_1 (q^n - 1)}{q - 1}$ (формула суммы n первых членов),

$S = \frac{b_1}{1 - q}$ (формула суммы бесконечной геометрической прогрессии при $|q| < 1$).

Тригонометрия

1. Свойства тригонометрических функций:

$$\sin(-x) = -\sin x,$$

$$\cos(-x) = \cos x,$$

$$\operatorname{tg}(-x) = -\operatorname{tg} x,$$

$$\operatorname{ctg}(-x) = -\operatorname{ctg} x,$$

$$\sin(x + 2\pi k) = \sin x,$$

$$\cos(x + 2\pi k) = \cos x,$$

$$\operatorname{tg}(x + \pi k) = \operatorname{tg} x,$$

$$\operatorname{ctg}(x + \pi k) = \operatorname{ctg} x,$$

где k — любое целое число.

2. Таблица значений тригонометрических функций некоторых углов:

Функция	Аргумент α						
	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$
$\sin \alpha$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	0	-1
$\cos \alpha$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	-1	0
$\operatorname{tg} \alpha$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	—	0	—
$\operatorname{ctg} \alpha$	—	$\sqrt{3}$	1	$\frac{\sqrt{3}}{3}$	0	—	0

П р и м е ч а н и е. Связь между градусной и радианной мерами измерения угла:

$$1^\circ = \frac{\pi}{180} \text{ рад.}$$

3. Формулы, связывающие тригонометрические функции одного и того же аргумента:

$$\begin{aligned}\sin^2 \alpha + \cos^2 \alpha &= 1; \\ \operatorname{tg} \alpha &= \frac{\sin \alpha}{\cos \alpha}; \operatorname{ctg} \alpha = \frac{\cos \alpha}{\sin \alpha}; \\ 1 + \operatorname{tg}^2 \alpha &= \frac{1}{\cos^2 \alpha}; 1 + \operatorname{ctg}^2 \alpha = \frac{1}{\sin^2 \alpha}.\end{aligned}$$

4. Формулы двойного угла:

$$\begin{aligned}\sin 2\alpha &= 2 \sin \alpha \cos \alpha; \\ \cos 2\alpha &= \cos^2 \alpha - \sin^2 \alpha; \\ \operatorname{tg} 2\alpha &= \frac{2 \operatorname{tg} \alpha}{1 - \operatorname{tg}^2 \alpha}.\end{aligned}$$

5. Формулы понижения степени:

$$\begin{aligned}\sin^2 \alpha &= \frac{1 - \cos 2\alpha}{2}; \\ \cos^2 \alpha &= \frac{1 + \cos 2\alpha}{2}.\end{aligned}$$

6. Формулы сложения и вычитания аргументов:

$$\begin{aligned}\sin(\alpha \pm \beta) &= \sin \alpha \cos \beta \pm \cos \alpha \sin \beta; \\ \cos(\alpha \pm \beta) &= \cos \alpha \cos \beta \mp \sin \alpha \sin \beta; \\ \operatorname{tg}(\alpha \pm \beta) &= \frac{\operatorname{tg} \alpha \pm \operatorname{tg} \beta}{1 \mp \operatorname{tg} \alpha \operatorname{tg} \beta}.\end{aligned}$$

7. Формулы сложения и вычитания тригонометрических функций:

$$\begin{aligned}\sin \alpha + \sin \beta &= 2 \sin \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}; \\ \sin \alpha - \sin \beta &= 2 \sin \frac{\alpha - \beta}{2} \cos \frac{\alpha + \beta}{2}; \\ \cos \alpha + \cos \beta &= 2 \cos \frac{\alpha + \beta}{2} \cos \frac{\alpha - \beta}{2}; \\ \cos \alpha - \cos \beta &= -2 \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}; \\ \operatorname{tg} \alpha \pm \operatorname{tg} \beta &= \frac{\sin(\alpha \pm \beta)}{\cos \alpha \cos \beta}.\end{aligned}$$

8. Формулы преобразования произведения тригонометрических функций в сумму и разность:

$$\begin{aligned}\sin \alpha \sin \beta &= \frac{1}{2} (\cos(\alpha - \beta) - \cos(\alpha + \beta)), \\ \cos \alpha \cos \beta &= \frac{1}{2} (\cos(\alpha - \beta) + \cos(\alpha + \beta)), \\ \sin \alpha \cos \beta &= \frac{1}{2} (\sin(\alpha - \beta) + \sin(\alpha + \beta)).\end{aligned}$$

9. Знаки тригонометрических функций по четвертям:

Функция	Четверть			
	I	II	III	IV
sin	+	+	−	−
cos	+	−	−	+
tg	+	−	+	−
ctg	+	−	+	−

10. Формулы приведения:

Функция	Аргумент t						
	$\frac{\pi}{2} - \alpha$	$\frac{\pi}{2} + \alpha$	$\pi - \alpha$	$\pi + \alpha$	$\frac{3\pi}{2} - \alpha$	$\frac{3\pi}{2} + \alpha$	$2\pi - \alpha$
$\sin t$	$\cos \alpha$	$\cos \alpha$	$\sin \alpha$	$-\sin \alpha$	$-\cos \alpha$	$-\cos \alpha$	$-\sin \alpha$
$\cos t$	$\sin \alpha$	$-\sin \alpha$	$-\cos \alpha$	$-\cos \alpha$	$-\sin \alpha$	$\sin \alpha$	$\cos \alpha$
$\operatorname{tg} t$	$\operatorname{ctg} \alpha$	$-\operatorname{ctg} \alpha$	$-\operatorname{tg} \alpha$	$\operatorname{tg} \alpha$	$\operatorname{ctg} \alpha$	$-\operatorname{ctg} \alpha$	$-\operatorname{tg} \alpha$
$\operatorname{ctg} t$	$\operatorname{tg} \alpha$	$-\operatorname{tg} \alpha$	$-\operatorname{ctg} \alpha$	$\operatorname{ctg} \alpha$	$\operatorname{tg} \alpha$	$-\operatorname{tg} \alpha$	$-\operatorname{ctg} \alpha$

11. Решение простейших тригонометрических уравнений:

$$\sin x = a, |a| \leq 1, x = (-1)^n \arcsin a + \pi n,$$

$$\cos x = a, |a| \leq 1, x = \pm \arccos a + 2\pi n,$$

$$\operatorname{tg} x = a, x = \operatorname{arctg} a + \pi n,$$

$$\operatorname{ctg} x = a, x = \operatorname{arcctg} a + \pi n, n - \text{целое число}.$$

12. Обратные тригонометрические функции:

$$-\frac{\pi}{2} \leq \arcsin x \leq \frac{\pi}{2}, 0 \leq \arccos x \leq \pi,$$

$$-\frac{\pi}{2} < \operatorname{arctg} x < \frac{\pi}{2}, 0 < \operatorname{arcctg} x < \pi.$$

$$\arcsin(-x) = -\arcsin x, \arccos(-x) = \pi - \arccos x,$$

$$\operatorname{arctg}(-x) = -\operatorname{arctg} x, \operatorname{arcctg}(-x) = \pi - \operatorname{arcctg} x.$$

13. Теорема синусов:

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma},$$

14. Теорема косинусов:

$$a^2 = b^2 + c^2 - 2bc \cos \alpha.$$

15. Площадь треугольника:

$$S = \frac{1}{2}ah_a, S = \frac{1}{2}ab \sin \gamma,$$

$$S = \sqrt{p(p-a)(p-b)(p-c)} — \text{формула Герона.}$$

Приложение 2

Таблица интегралов

(a, b, m, n — постоянные)

$$1. \int (ax + b)^n dx = \frac{(ax + b)^{n+1}}{a(n+1)} + C, \quad n \neq -1.$$

$$2. \int \frac{dx}{ax + b} = \frac{1}{a} \ln|ax + b| + C.$$

$$3. \int \frac{xdx}{ax + b} = \frac{1}{a^2} (ax - b \ln|ax + b|) + C.$$

$$4. \int \frac{x^2 dx}{ax + b} = \frac{x^2}{2a} - \frac{bx}{a^2} + \frac{b^2}{a^3} \ln|ax + b| + C.$$

$$5. \int \frac{dx}{x(ax + b)} = -\frac{1}{b} \ln \left| \frac{ax + b}{x} \right| + C.$$

$$6. \int \frac{dx}{x^2(ax + b)} = -\frac{1}{bx} + \frac{a}{b^2} \ln \left| \frac{ax + b}{x} \right| + C.$$

$$7. \int \frac{xdx}{(ax + b)^2} = \frac{1}{a^2} \left(\ln|ax + b| + \frac{b}{ax + b} \right) + C.$$

$$8. \int \frac{x^2 dx}{(ax + b)^2} = \frac{1}{a^3} \left(ax - 2b \ln|ax + b| - \frac{b^2}{ax + b} \right) + C.$$

$$9. \int \frac{dx}{x(ax + b)^2} = \frac{1}{b(ax + b)} - \frac{1}{b^2} \ln \left| \frac{ax + b}{x} \right| + C.$$

$$10. \int \frac{dx}{a^2 x^2 + b^2} = \frac{1}{ab} \operatorname{arctg} \frac{a}{b} x + C.$$

$$11. \int \frac{dx}{a^2 x^2 - b^2} = \frac{1}{2ab} \ln \left| \frac{ax - b}{ax + b} \right| + C.$$

$$12. \int \frac{xdx}{a^2 x^2 \pm b^2} = \frac{1}{2a^2} \ln|a^2 x^2 \pm b^2| + C.$$

$$13. \int \frac{x^2 dx}{a^2 x^2 + b^2} = \frac{1}{a^2} x - \frac{b}{a^3} \operatorname{arctg} \frac{a}{b} x + C.$$

$$14. \int \frac{x^2 dx}{a^2 x^2 - b^2} = \frac{1}{a^2} x + \frac{b}{2a^3} \ln \left| \frac{ax - b}{ax + b} \right| + C.$$

$$15. \int \frac{dx}{x(a^2 x^2 + b^2)} = \frac{1}{2b^2} \ln \frac{x^2}{a^2 x^2 + b^2} + C.$$

$$16. \int \frac{dx}{x(a^2x^2 - b^2)} = \frac{1}{2b^2} \ln \left| \frac{a^2x^2 - b^2}{x^2} \right| + C.$$

$$17. \int \frac{dx}{x^2(a^2x^2 + b^2)} = -\frac{1}{b^2x} - \frac{a}{b^3} \operatorname{arctg} \frac{a}{b}x + C.$$

$$18. \int \frac{dx}{x^2(a^2x^2 - b^2)} = \frac{1}{b^2x} + \frac{a}{2b^3} \ln \left| \frac{ax - b}{ax + b} \right| + C.$$

$$19. \int x\sqrt{ax + b} dx = \frac{2}{15a^2} (3ax - 2b)\sqrt{(ax + b)^3} + C.$$

$$20. \int x^2\sqrt{ax + b} dx = \frac{2}{105a^3} (15a^2x^2 - 12abx + 8b^2)\sqrt{(ax + b)^3} + C.$$

$$21. \int \frac{xdx}{\sqrt{ax + b}} = \frac{2(ax - 2b)}{3a^2} \sqrt{ax + b} + C.$$

$$22. \int \frac{x^2dx}{\sqrt{ax + b}} = \frac{2}{15a^3} (3a^2x^2 - 4abx + 8b^2)\sqrt{ax + b} + C.$$

$$23. \int \frac{dx}{x\sqrt{ax + b}} = \begin{cases} \frac{1}{\sqrt{b}} \ln \frac{\sqrt{ax + b} - \sqrt{b}}{\sqrt{ax + b} + \sqrt{b}} + C, & b > 0, \\ \frac{2}{\sqrt{-b}} \operatorname{arctg} \sqrt{\frac{ax + b}{-b}} + C, & b < 0. \end{cases}$$

$$24. \int \frac{dx}{x^2\sqrt{ax + b}} = -\frac{\sqrt{ax + b}}{bx} - \frac{a}{2b} \int \frac{dx}{x\sqrt{ax + b}}.$$

$$25. \int \frac{\sqrt{ax + b}}{x} dx = 2\sqrt{ax + b} + b \int \frac{dx}{x\sqrt{ax + b}}.$$

$$26. \int \frac{\sqrt{ax + b}}{x^2} dx = -\frac{\sqrt{ax + b}}{x} + \frac{a}{2} \int \frac{dx}{x\sqrt{ax + b}}.$$

В формулах 27—46 предполагаем, что $a > 0$, $b > 0$.

$$27. \int \frac{dx}{\sqrt{a^2x^2 \pm b^2}} = \frac{1}{a} \ln |ax + \sqrt{a^2x^2 \pm b^2}| + C.$$

$$28. \int \frac{dx}{\sqrt{b^2 - a^2x^2}} = \frac{1}{a} \arcsin \frac{ax}{b} + C.$$

$$29. \int \frac{xdx}{\sqrt{a^2x^2 \pm b^2}} = \frac{1}{a^2} \sqrt{a^2x^2 \pm b^2} + C.$$

$$30. \int \frac{xdx}{\sqrt{b^2 - a^2x^2}} = -\frac{1}{a^2} \sqrt{b^2 - a^2x^2} + C.$$

$$31. \int \frac{x^2dx}{\sqrt{a^2x^2 \pm b^2}} = \frac{1}{2a^3} (ax\sqrt{a^2x^2 \pm b^2} \mp b^2 \ln |ax + \sqrt{a^2x^2 \pm b^2}|) + C.$$

$$32. \int \frac{x^2dx}{\sqrt{b^2 - a^2x^2}} = \frac{1}{2a^3} \left(-ax\sqrt{b^2 - a^2x^2} + b^2 \arcsin \frac{ax}{b} \right) + C.$$

$$33. \int \sqrt{a^2x^2 \pm b^2} dx = \frac{1}{2} \left(x\sqrt{a^2x^2 \pm b^2} \pm \frac{b^2}{a} \ln |ax + \sqrt{a^2x^2 \pm b^2}| \right) + C.$$

$$34. \int \sqrt{b^2 - a^2x^2} dx = \frac{1}{2} \left(x\sqrt{b^2 - a^2x^2} + \frac{b^2}{a} \arcsin \frac{ax}{b} \right) + C.$$

35. $\int x\sqrt{a^2x^2 \pm b^2} dx = \frac{1}{3a^2} \sqrt{(a^2x^2 \pm b^2)^3} + C.$
36. $\int x\sqrt{b^2 - a^2x^2} dx = -\frac{1}{3a^2} \sqrt{(b^2 - a^2x^2)^3} + C.$
37. $\int x^2\sqrt{a^2x^2 \pm b^2} dx = \frac{1}{8a^3} (ax(2a^2x^2 \pm b^2)\sqrt{a^2x^2 \pm b^2} - b^4 \ln|ax + \sqrt{a^2x^2 \pm b^2}|) + C.$
38. $\int x^2\sqrt{b^2 - a^2x^2} dx = \frac{1}{8a^3} \left(ax(2a^2x^2 - b^2)\sqrt{b^2 - a^2x^2} + b^4 \arcsin \frac{ax}{b} \right) + C.$
39. $\int \frac{\sqrt{a^2x^2 + b^2}}{x} dx = \sqrt{a^2x^2 + b^2} + \frac{b}{2} \ln \frac{\sqrt{a^2x^2 + b^2} - b}{\sqrt{a^2x^2 + b^2} + b} + C.$
40. $\int \frac{\sqrt{b^2 - a^2x^2}}{x} dx = \sqrt{b^2 - a^2x^2} - b \ln \left| \frac{b - \sqrt{b^2 - a^2x^2}}{x} \right| + C.$
41. $\int \frac{\sqrt{a^2x^2 \pm b^2}}{x^2} dx = -\frac{\sqrt{a^2x^2 \pm b^2}}{x} + a \ln|ax + \sqrt{a^2x^2 \pm b^2}| + C.$
42. $\int \frac{\sqrt{b^2 - a^2x^2}}{x^2} dx = -\frac{\sqrt{b^2 - a^2x^2}}{x} - a \arcsin \frac{ax}{b} + C.$
43. $\int \frac{dx}{x\sqrt{a^2x^2 + b^2}} = \frac{1}{b} \ln \left| \frac{x}{b + \sqrt{a^2x^2 + b^2}} \right| + C.$
44. $\int \frac{dx}{x\sqrt{b^2 - a^2x^2}} = \frac{1}{b} \ln \left| \frac{x}{b + \sqrt{b^2 - a^2x^2}} \right| + C.$
45. $\int \frac{dx}{x^2\sqrt{a^2x^2 \pm b^2}} = \mp \frac{\sqrt{a^2x^2 \pm b^2}}{b^2x} + C.$
46. $\int \frac{dx}{x^2\sqrt{b^2 - a^2x^2}} = -\frac{\sqrt{b^2 - a^2x^2}}{b^2x} + C.$
47. $\int \sqrt{\frac{a+x}{b+x}} dx = \sqrt{(a+x)(b+x)} + (a-b) \ln(\sqrt{a+x} + \sqrt{b+x}) + C.$
48. $\int \sqrt{\frac{a-x}{b+x}} dx = \sqrt{(a-x)(b+x)} + (a+b) \arcsin \sqrt{\frac{b+x}{a+b}} + C.$
49. $\int \sqrt{\frac{a+x}{b-x}} dx = -\sqrt{(a+x)(b-x)} - (a+b) \arcsin \sqrt{\frac{b-x}{a+b}} + C.$
50. $\int \sqrt{\frac{1+x}{1-x}} dx = -\sqrt{1-x^2} + \arcsin x + C.$
51. $\int \frac{dx}{\sqrt{(x-a)(x-b)}} = \ln \left| \frac{\sqrt{x-a} + \sqrt{x-b}}{\sqrt{x-a} - \sqrt{x-b}} \right| + C.$
52. $\int \frac{dx}{\sqrt{(x-a)(b-x)}} = 2 \arcsin \sqrt{\frac{x-a}{b-a}} + C.$
53. $\int x^m(ax^n + b)^{\frac{r}{s}} dx$, где m, n, r, s — целые числа, $s > 1$. Применяется подстановка $ax^n + b = u^s$,

если $\frac{m+1}{m}$ — число целое, и подстановка $a + bx^{-n} = u^s$, если $\frac{m+1}{m} + \frac{r}{s}$ — число целое. В других случаях интеграл не выражается элементарной функцией (П. Л. Чебышев).

$$54. \int x^n e^x dx = e^x (x^n - nx^{n-1} + n(n-1)x^{n-2} - \dots + (-1)^n n!) + C, \quad n = 1, 2, 3, \dots$$

$$55. \int \ln^n x dx = x(\ln^n x - n \ln^{n-1} x + n(n-1) \ln^{n-2} x - \dots + (-1)^n n!) + C, \quad n = 1, 2, 3, \dots$$

$$56. \int x^n \ln x dx = \frac{x^{n+1}}{n+1} \ln x - \frac{x^{n+1}}{(n+1)^2} + C, \quad n \neq -1.$$

$$57. \int \frac{\ln^n x}{x} dx = \frac{\ln^{n+1} x}{n+1} + C, \quad n \neq -1.$$

$$58. \int \frac{dx}{x \ln x} = \ln |\ln x| + C.$$

$$59. \int \sin mx \cos nxdx = -\frac{\cos(m+n)x}{2(m+n)} - \frac{\cos(m-n)x}{2(m-n)} + C, \quad m \neq n.$$

$$60. \int \sin mx \sin nxdx = \frac{\sin(m-n)x}{2(m-n)} - \frac{\sin(m+n)x}{2(m+n)} + C, \quad m \neq n.$$

$$61. \int \cos mx \cos nxdx = \frac{\sin(m-n)x}{2(m-n)} + \frac{\sin(m+n)x}{2(m+n)} + C, \quad m \neq n.$$

$$62. \int \sin x \cos x dx = \frac{1}{2} \sin^2 x + C.$$

$$63. \int \sin^2 x dx = \frac{1}{2} x - \frac{1}{4} \sin 2x + C.$$

$$64. \int \cos^2 x dx = \frac{1}{2} x + \frac{1}{4} \sin 2x + C.$$

$$65. \int \frac{dx}{\cos x} = \ln \left| \operatorname{tg} \left(\frac{\pi}{4} + \frac{x}{2} \right) \right| + C.$$

$$66. \int \sin^m x \cos^n x dx = \begin{cases} -\frac{\sin^{m-1} x \cos^{n+1} x}{m+n} + \frac{m-1}{m+n} \int \sin^{m-2} x \cos^n x dx, & m \neq -n, \\ \frac{\sin^{m+1} x \cos^{n-1} x}{m+n} + \frac{n-1}{m+n} \int \sin^m x \cos^{n-2} x dx, & m \neq -n. \end{cases}$$

$$67. \int \frac{dx}{\sin^m x \cos^n x} = \begin{cases} -\frac{1}{m-1} \frac{1}{\sin^{m-1} x \cos^{n-1} x} + \frac{m+n-2}{m-1} \int \frac{dx}{\sin^{m-2} x \cos^n x}, & m \neq 1, \\ \frac{1}{n-1} \frac{1}{\sin^{m-1} x \cos^{n-1} x} + \frac{m+n-2}{n-1} \int \frac{dx}{\sin^m x \cos^{n-2} x}, & n \neq 1. \end{cases}$$

$$68. \int \operatorname{tg}^n x dx = \frac{\operatorname{tg}^{n-1} x}{n-1} - \int \operatorname{tg}^{n-2} x dx, \quad n \neq 1.$$

$$69. \int \operatorname{ctg}^n x dx = -\frac{\operatorname{ctg}^{n-1} x}{n-1} - \int \operatorname{ctg}^{n-2} x dx, \quad n \neq 1.$$

$$70. \int x^n \sin x dx = -x^n \cos x + n \int x^{n-1} \cos x dx.$$

$$71. \int x^n \cos x dx = x^n \sin x - n \int x^{n-1} \sin x dx.$$

$$72. \int \frac{dx}{a + b \cos x} = \begin{cases} \frac{2}{\sqrt{a^2 - b^2}} \operatorname{arctg} \left(\sqrt{\frac{a-b}{a+b}} \operatorname{tg} \frac{x}{2} \right) + C, & a^2 > b^2, \\ \frac{1}{\sqrt{b^2 - a^2}} \ln \left| \frac{\sqrt{b^2 - a^2} \operatorname{tg} \frac{x}{2} + a + b}{\sqrt{b^2 - a^2} \operatorname{tg} \frac{x}{2} - a - b} \right| + C, & a^2 < b^2. \end{cases}$$

$$73. \int \frac{dx}{a + b \sin x} = \begin{cases} \frac{2}{\sqrt{a^2 - b^2}} \operatorname{arctg} \frac{a \operatorname{tg} \frac{x}{2} + b}{\sqrt{a^2 - b^2}} + C, & a^2 > b^2, \\ \frac{1}{\sqrt{b^2 - a^2}} \ln \left| \frac{a \operatorname{tg} \frac{x}{2} + b - \sqrt{b^2 - a^2}}{a \operatorname{tg} \frac{x}{2} + b + \sqrt{b^2 - a^2}} \right| + C, & a^2 < b^2. \end{cases}$$

$$74. \int \frac{dx}{a \cos x + b \sin x} = \frac{1}{\sqrt{a^2 + b^2}} \ln \left| \frac{a \operatorname{tg} \frac{x}{2} - b + \sqrt{a^2 + b^2}}{a \operatorname{tg} \frac{x}{2} - b - \sqrt{a^2 + b^2}} \right| + C.$$

В частности,

$$75. \int \frac{dx}{\cos x + \sin x} = \frac{1}{\sqrt{2}} \ln \left| \operatorname{tg} \left(\frac{x}{2} + \frac{\pi}{8} \right) \right| + C.$$

$$76. \int \frac{dx}{a^2 \sin^2 x + b^2 \cos^2 x} = \frac{1}{ab} \operatorname{arctg} \left(\frac{a}{b} \operatorname{tg} x \right) + C.$$

$$77. \int e^{ax} \sin nx dx = \frac{e^{ax} (a \sin nx - n \cos nx)}{a^2 + n^2} + C.$$

$$78. \int e^{ax} \cos nx dx = \frac{e^{ax} (n \sin nx + a \cos nx)}{a^2 + n^2} + C.$$

$$79. \int e^{ax} \sin^n x dx = \frac{e^{ax} \sin^{n-1} x (a \sin x - n \cos x)}{a^2 + n^2} + \frac{n(n-1)}{a^2 + n^2} \int e^{ax} \sin^{n-2} x dx.$$

$$80. \int e^{ax} \cos^n x dx = \frac{e^{ax} \cos^{n-1} x (a \cos x + n \sin x)}{a^2 + n^2} + \frac{n(n-1)}{a^2 + n^2} \int e^{ax} \cos^{n-2} x dx.$$

$$81. \int \arcsin x dx = x \arcsin x + \sqrt{1-x^2} + C.$$

$$82. \int \arccos x dx = x \arccos x - \sqrt{1-x^2} + C.$$

$$83. \int x \arcsin x dx = \frac{1}{2} \left(x^2 - \frac{1}{2} \right) \arcsin x + \frac{x \sqrt{1-x^2}}{4} + C.$$

$$84. \int x \arccos x dx = \frac{1}{2} \left(x^2 - \frac{1}{2} \right) \arccos x - \frac{x \sqrt{1-x^2}}{4} + C.$$

$$85. \int x \operatorname{arctg} x dx = \frac{1}{2} (x^2 + 1) \operatorname{arctg} x - \frac{x}{2} + C.$$

$$86. \int x \operatorname{arcctg} x dx = \frac{1}{2} (x^2 + 1) \operatorname{arcctg} x + \frac{x}{2} + C.$$

Приложение 3

Таблица значений функции $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

x	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3652	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3,1	0033	0032	0031	0030	0029	0028	0027	0036	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0004	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

Таблица значений функции $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.0000	0.26	0.1026	0.52	0.1985	0.78	0.2823	1.04	0.3508	1.29	0.4015	1.54	0.4382	1.79	0.4633
0.01	0.0040	0.27	0.1064	0.53	0.2019	0.79	0.2852	1.05	0.3531	1.30	0.4032	1.55	0.4394	1.80	0.4641
0.02	0.0080	0.28	0.1103	0.54	0.2054	0.80	0.2881	1.06	0.3554	1.31	0.4049	1.56	0.4406	1.81	0.4649
0.03	0.0120	0.29	0.1141	0.55	0.2088	0.81	0.2910	1.07	0.3577	1.32	0.4066	1.57	0.4418	1.82	0.4656
0.04	0.0160	0.30	0.1179	0.56	0.2123	0.82	0.2939	1.08	0.3599	1.33	0.4082	1.58	0.4429	1.83	0.4664
0.05	0.0199	0.31	0.1217	0.57	0.2157	0.83	0.2967	1.09	0.3621	1.34	0.4099	1.59	0.4441	1.84	0.4671
0.06	0.0239	0.32	0.1255	0.58	0.2190	0.84	0.2995	1.10	0.3643	1.35	0.4115	1.60	0.4452	1.85	0.4678
0.07	0.0279	0.33	0.1293	0.59	0.2224	0.85	0.3023	1.11	0.3665	1.36	0.4131	1.61	0.4463	1.86	0.4686
0.08	0.0319	0.34	0.1331	0.60	0.2257	0.86	0.3051	1.12	0.3686	1.37	0.4147	1.62	0.4474	1.87	0.4693
0.09	0.0359	0.35	0.1368	0.61	0.2291	0.87	0.3078	1.13	0.3708	1.38	0.4162	1.63	0.4484	1.88	0.4699
0.10	0.0398	0.36	0.1406	0.62	0.2324	0.88	0.3106	1.14	0.3729	1.39	0.4177	1.64	0.4495	1.89	0.4706
0.11	0.0438	0.37	0.1443	0.63	0.2357	0.89	0.3133	1.15	0.3749	1.40	0.4192	1.65	0.4505	1.90	0.4713
0.12	0.0478	0.38	0.1480	0.64	0.2389	0.90	0.3159	1.16	0.3770	1.41	0.4207	1.66	0.4515	1.91	0.4719
0.13	0.0517	0.39	0.1517	0.65	0.2422	0.91	0.3186	1.17	0.3790	1.42	0.4222	1.67	0.4525	1.92	0.4726
0.14	0.0557	0.40	0.1554	0.66	0.2454	0.92	0.3212	1.18	0.3810	1.43	0.4236	1.68	0.4535	1.93	0.4732
0.15	0.0596	0.41	0.1591	0.67	0.2486	0.93	0.3238	1.19	0.3830	1.44	0.4251	1.69	0.4545	1.94	0.4738
0.16	0.0636	0.42	0.1628	0.68	0.2517	0.94	0.3264	1.20	0.3849	1.45	0.4265	1.70	0.4554	1.95	0.4744
0.17	0.0675	0.43	0.1664	0.69	0.2549	0.95	0.3289	1.21	0.3869	1.46	0.4279	1.71	0.4564	1.96	0.4750
0.18	0.0714	0.44	0.1700	0.70	0.2580	0.96	0.3315	1.22	0.3883	1.47	0.4292	1.72	0.4573	1.97	0.4756
0.19	0.0753	0.45	0.1736	0.71	0.2611	0.97	0.3340	1.23	0.3907	1.48	0.4306	1.73	0.4582	1.98	0.4761
0.20	0.0793	0.46	0.1772	0.72	0.2642	0.98	0.3365	1.24	0.3925	1.49	0.4319	1.74	0.4591	1.99	0.4767
0.21	0.0832	0.47	0.1808	0.73	0.2673	0.99	0.3389	1.25	0.3944	1.50	0.4332	1.75	0.4599	2.00	0.4772
0.22	0.0871	0.48	0.1844	0.74	0.2703	1.00	0.3413	1.26	0.3962	1.51	0.4345	1.76	0.4608	2.02	0.4783
0.23	0.0910	0.49	0.1879	0.75	0.2734	1.01	0.3438	1.27	0.3980	1.52	0.4357	1.77	0.4616	2.04	0.479

Приложение 5

Таблица значений $t_\gamma = t(\gamma, n)$

n	γ			n	γ		
	0,95	0,99	0,999		0,95	0,99	0,999
5	2,78	4,60	8,61	20	2,093	2,861	3,883
6	2,57	4,03	6,86	25	2,064	2,797	3,745
7	2,45	3,71	5,96	30	2,045	2,756	3,659
8	2,37	3,50	5,41	35	2,032	2,720	3,600
9	2,31	3,36	5,04	40	2,023	2,708	3,558
10	2,26	3,25	4,78	45	2,016	2,692	3,527
11	2,23	3,17	4,59	50	2,009	2,679	3,502
12	2,20	3,11	4,44	60	2,001	2,662	3,464
13	2,18	3,06	4,32	70	1,996	2,649	3,439
14	2,16	3,01	4,22	80	1,991	2,640	3,418
15	2,15	2,98	4,14	90	1,987	2,633	3,403
16	2,13	2,95	4,07	100	1,984	2,627	3,392
17	2,12	2,92	4,02	120	1,980	2,617	3,374
18	2,11	2,90	3,97	∞	1,960	2,576	3,291
19	2,10	2,88	3,92				

Приложение 6

Таблица значений $q = q(\gamma, n)$

$n \backslash \gamma$	0,95	0,99	0,999	$n \backslash \gamma$	0,95	0,99	0,999
5	1,37	2,67	5,64	20	0,37	0,58	0,88
6	1,09	2,01	3,88	25	0,32	0,49	0,73
7	0,92	1,62	2,98	30	0,28	0,43	0,63
8	0,80	1,38	2,42	35	0,26	0,38	0,56
9	0,71	1,20	2,06	40	0,24	0,35	0,50
10	0,65	1,08	1,80	45	0,22	0,32	0,46
11	0,59	0,98	1,60	50	0,21	0,30	0,43
12	0,55	0,90	1,45	60	0,188	0,269	0,38
13	0,52	0,83	1,33	70	0,174	0,245	0,34
14	0,48	0,78	1,23	80	0,161	0,226	0,31
15	0,46	0,73	1,15	90	0,151	0,211	0,29
16	0,44	0,70	1,07	100	0,143	0,198	0,27
17	0,42	0,66	1,01	150	0,115	0,160	0,221
18	0,40	0,63	0,96	200	0,099	0,136	0,185
19	0,39	0,60	0,92	250	0,089	0,120	0,162

Приложение 7

Таблица значений χ^2 в зависимости от p и k

$k \backslash p$	0,05	0,01	0,001
1	3,84	6,64	10,83
2	5,99	9,21	13,82
3	7,82	11,34	16,27
4	9,49	13,28	18,46
5	11,07	15,09	20,5
6	12,59	16,81	22,5
7	14,07	18,48	24,3
8	15,51	20,1	26,1
9	16,92	24,7	27,9
10	18,31	23,2	29,6
11	19,68	24,7	31,3
12	21,0	26,2	32,9
13	22,4	27,7	34,6
14	23,7	29,1	36,1
15	25,0	30,6	37,7
16	26,3	32,0	39,3
17	27,6	33,4	40,8
18	28,9	34,8	42,3
19	30,1	36,2	43,8
20	31,4	37,6	45,3
21	32,7	38,9	46,8
22	33,9	40,3	48,3
23	35,2	41,6	49,7
24	36,4	43,0	51,2
25	37,7	44,3	52,6
26	38,9	45,6	54,1
27	40,1	47,0	55,5
28	41,3	48,3	56,9
29	42,6	49,6	58,3
30	43,8	50,9	59,7

Приложение 8

Некоторые постоянные

Величина	x	$\lg x$	Величина	x	$\lg x$
π	3,14159	0,49715	$\frac{1}{e}$	0,36788	$\bar{1},56571$
2π	6,28318	0,79818	e^2	7,38906	0,86859
$\frac{\pi}{2}$	1,57080	0,19612	$\sqrt{e}$	1,64872	0,21715

Некоторые постоянные (продолжение)

Величина	x	$\lg x$	Величина	x	$\lg x$
$\frac{\pi}{4}$	0,78540	1,89509	$\sqrt[3]{e}$	1,39561	0,14476
$\frac{1}{\pi}$	0,31831	$\bar{1},50285$	$M = \lg e$	0,43429	$\bar{1},63778$
π^2	9,86960	0,99430	$\frac{1}{M} = \ln 10$	2,30258	0,36222
$\sqrt{\pi}$	1,77245	0,24857	1 радиан	$57^{\circ}17'45''$	
$\sqrt[3]{\pi}$	1,46459	0,16572	$\arcsin 1^{\circ}$	0,01745	$\bar{2},24188$
e	2,71828	0,43429	g	9,81	0,99167

Список литературы

- Александров П. С. Курс аналитической геометрии и линейной алгебры. — М., 1979.
- Будак Б. М., Фомин С. В. Кратные интегралы и ряды. — М., 1965.
- Владимиров В. С. Уравнения математической физики. — М., 1988.
- Гмурман В. Е. Теория вероятностей и математическая статистика. — М., 1997.
- Гнеденко Б. В. Курс теории вероятностей. — М., 1988.
- Ефимов Н. В. Краткий курс аналитической геометрии. — М., 1975.
- Ефимов Н. В. Квадратичные формы и матрицы. — М., 1963.
- Ильин В. А., Позняк Э. Г. Основы математического анализа. — М., 1971. — Ч. 1.
- Кудрявцев Л. Д. Математический анализ. — М., 1981. — Т. I, II.
- Курош А. Г. Курс высшей алгебры. — М., 1962.
- Никольский С. М. Курс математического анализа. — М., 1983. — Т. I.
- Понтрягин Л. С. Обыкновенные дифференциальные уравнения. — М., 1974.
- Привалов И. И. Введение в теорию функций комплексного переменного. — М., 1977.
- Пугачев В. С. Теория вероятностей и математическая статистика. — М., 1979.
- Смирнов В. И. Курс высшей математики. — М., 1974. — Т. I, II.
- Степанов В. В. Курс дифференциальных уравнений. — М., 1958.
- Тихонов А. П., Самарский А. А. Уравнения математической физики. — М., 1977.
- Феллер В. Введение в теорию вероятностей и ее приложения: Пер. с англ. — М., 1984. — Т. I.

Оглавление

Предисловие ко второму изданию	3
Раздел I. АНАЛИТИЧЕСКАЯ ГЕОМЕТРИЯ И ЛИНЕЙНАЯ АЛГЕБРА	4
Глава 1. СИСТЕМА КООРДИНАТ НА ПЛОСКОСТИ И ОСНОВНЫЕ ПОНЯТИЯ	4
§ 1.1. Декартова прямоугольная и полярная системы координат на плоскости	4
§ 1.2. Простейшие задачи на плоскости	7
§ 1.3. Геометрическое истолкование уравнения с двумя переменными	8
§ 1.4. Прямая линия	9
§ 1.5. Основные задачи на прямую	18
§ 1.6. Уравнение линии	19
Глава 2. ВЕКТОРНАЯ АЛГЕБРА	23
§ 2.1. Понятие вектора и линейные операции над векторами	23
§ 2.2. Нелинейные операции над векторами	37
Глава 3. ЛИНЕЙНАЯ АЛГЕБРА	44
§ 3.1. Матрицы и действия над ними	44
§ 3.2. Определители	48
§ 3.3. Выражение векторного и смешанного произведений векторов через координаты сомножителей	61
§ 3.4. Системы линейных уравнений	62
Глава 4. ПЛОСКОСТЬ И ПРЯМАЯ В ПРОСТРАНСТВЕ	75
§ 4.1. Плоскость	75
§ 4.2. Прямая в пространстве	80
§ 4.3. Основные задачи на плоскость и прямую в пространстве	85
Глава 5. КРИВЫЕ И ПОВЕРХНОСТИ ВТОРОГО ПОРЯДКА В КАНОНИЧЕСКОЙ ФОРМЕ	86
§ 5.1. Кривые второго порядка в канонической форме	86
§ 5.2. Изучение поверхностей второго порядка по их каноническим уравнениям	95
Глава 6. ОБЩЕЕ УРАВНЕНИЕ КРИВОЙ ВТОРОГО ПОРЯДКА	102
§ 6.1. Приведение матрицы квадратичной формы к диагональному виду	102
§ 6.2. Общее уравнение кривой второго порядка, его приведение к каноническому виду	105
§ 6.3. Инварианты кривых второго порядка	109
§ 6.4. Уравнение центра. Вырождение кривых второго порядка	111
Раздел II. МАТЕМАТИЧЕСКИЙ АНАЛИЗ	113
Глава 7. ВВЕДЕНИЕ В АНАЛИЗ	113
§ 7.1. Определение и способы задания функции	113
§ 7.2. Обзор элементарных функций и их графиков	118
§ 7.3. Предел функции	125
§ 7.4. Бесконечно малые и бесконечно большие величины	131
§ 7.5. Основные теоремы о пределах и их применение	134
§ 7.6. Непрерывность функции	142
§ 7.7. Комплексные числа	146
Глава 8. ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ	151
§ 8.1. Понятие производной и ее механический и геометрический смысл	151
§ 8.2. Правила дифференцирования функций и производные элементарных функций	155

§ 8.3.	Дифференциал функции	160
§ 8.4.	Производные и дифференциалы высших порядков	162
§ 8.5.	Параметрическое задание функции и ее дифференцирование	166
§ 8.6.	Свойства дифференцируемых функций	167
§ 8.7.	Возрастание и убывание функций. Максимум и минимум	173
§ 8.8.	Построение графиков функций	179
§ 8.9.	Формула Тейлора	181
Глава 9.	ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ	187
§ 9.1.	Первообразная функция и неопределенный интеграл	187
§ 9.2.	Основные методы интегрирования	190
§ 9.3.	Интегрирование дробно-рациональных функций	191
§ 9.4.	Интегрирование тригонометрических выражений	196
§ 9.5.	Интегрирование простейших иррациональностей	197
§ 9.6.	Понятие определенного интеграла	199
§ 9.7.	Основные свойства определенного интеграла	202
§ 9.8.	Виды несобственных интегралов, их сходимости	207
§ 9.9.	Геометрические приложения определенного интеграла	213
§ 9.10.	Физические приложения определенного интеграла	222
§ 9.11.	Вектор-функция скалярного аргумента	225
Глава 10.	РЯДЫ	236
§ 10.1.	Числовые ряды	236
§ 10.2.	Функциональные ряды	248
§ 10.3.	Степенные ряды в действительной области	250
§ 10.4.	Тригонометрические ряды	259
Глава 11.	ДИФФЕРЕНЦИАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ	266
§ 11.1.	Функции нескольких переменных. Предел и непрерывность функции	266
§ 11.2.	Частные производные. Полный дифференциал	272
§ 11.3.	Частные производные и дифференциалы высших порядков	280
§ 11.4.	Экстремум функций двух переменных	283
§ 11.5.	Метод наименьших квадратов	285
Глава 12.	ИНТЕГРАЛЬНОЕ ИСЧИСЛЕНИЕ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ	287
§ 12.1.	Двойные интегралы	287
§ 12.2.	Тройные интегралы	301
§ 12.3.	Криволинейные интегралы	306
§ 12.4.	Поверхностные интегралы	318
Глава 13.	ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ	327
§ 13.1.	Основные понятия о дифференциальных уравнениях	327
§ 13.2.	Дифференциальные уравнения первого порядка	327
§ 13.3.	Уравнения высших порядков	333
§ 13.4.	Линейные уравнения второго порядка	337
§ 13.5.	Системы линейных дифференциальных уравнений	346
Раздел III.	СПЕЦИАЛЬНЫЕ ГЛАВЫ	350
Глава 14.	ТЕОРИЯ ПОЛЯ	350
§ 14.1.	Скалярные поля	350
§ 14.2.	Векторные поля	354
§ 14.3.	Дифференциальные операции второго порядка и их приложения	371
Глава 15.	ЭЛЕМЕНТЫ ТЕОРИИ ФУНКЦИЙ КОМПЛЕКСНОЙ ПЕРЕМЕННОЙ	379
§ 15.1.	Функции комплексной переменной	379
§ 15.2.	Дифференцирование функций комплексной переменной	382
§ 15.3.	Интегралы по комплексному переменному	385
§ 15.4.	Равномерно сходящиеся ряды функций комплексной переменной	390
§ 15.5.	Элементарные функции комплексной переменной	395

§ 15.6.	Ряд Тейлора	400
§ 15.7.	Ряд Лорана	402
§ 15.8.	Изолированные особые точки аналитической функции	404
§ 15.9.	Вычеты	406
Глава 16.	ИНТЕГРАЛ ФУРЬЕ. ДЕЛЬТА-ФУНКЦИЯ	410
§ 16.1.	Интеграл Фурье	410
§ 16.2.	Дельта-функция	415
Глава 17.	ОСНОВНЫЕ УРАВНЕНИЯ И ОСНОВНЫЕ ЗАДАЧИ МАТЕМАТИЧЕСКОЙ ФИЗИКИ.	418
§ 17.1.	Основные понятия о дифференциальных уравнениях с частными производными второго порядка	418
§ 17.2.	Вывод уравнения колебаний струны	419
§ 17.3.	Вывод акустического уравнения	421
§ 17.4.	Вывод уравнения теплопроводности	423
§ 17.5.	Классификация задач математической физики	424
§ 17.6.	Задача Коши	426
§ 17.7.	Смешанная задача для одномерного однородного волнового уравнения и ее решение методом Фурье	429
§ 17.8.	Задача Дирихле для круга	435
Глава 18.	ТЕОРИЯ ВЕРОЯТНОСТЕЙ	438
§ 18.1.	Основные понятия. Определение вероятности	438
§ 18.2.	Свойства вероятности	443
§ 18.3.	Основные формулы комбинаторики	448
§ 18.4.	Дискретные случайные величины	449
§ 18.5.	Математическое ожидание дискретной случайной величины	450
§ 18.6.	Дисперсия дискретной случайной величины	453
§ 18.7.	Основные законы распределения дискретных случайных величин	455
§ 18.8.	Непрерывные случайные величины	461
§ 18.9.	Закон больших чисел	470
§ 18.10.	Использование теории вероятностей при обработке экспериментальных данных	473
§ 18.11.	Двумерные случайные величины	475
Глава 19.	ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ	476
§ 19.1.	Выборочный метод	476
§ 19.2.	Оценки параметров генеральной совокупности по ее выборке	480
§ 19.3.	Доверительные интервалы для параметров нормального распределения	491
§ 19.4.	Проверка статистических гипотез	497
§ 19.5.	Линейная корреляция	499
Глава 20.	ЭЛЕМЕНТЫ ВАРИАЦИОННОГО И ОПЕРАЦИОННОГО ИСЧИСЛЕНИЙ.	505
§ 20.1.	Элементы вариационного исчисления	505
§ 20.2.	Элементы операционного исчисления	511
Глава 21.	ЛИНЕЙНЫЕ ПРОСТРАНСТВА	522
§ 21.1.	Основные понятия	522
§ 21.2.	Евклидово пространство	526
§ 21.3.	Линейные операторы	529
Глава 22.	НЕКОТОРЫЕ ЧИСЛЕННЫЕ МЕТОДЫ	535
§ 22.1.	Численное интегрирование	535
§ 22.2.	Численное решение уравнений	538
ПРИЛОЖЕНИЯ		543
СПИСОК ЛИТЕРАТУРЫ		556

Учебное издание

Баврин Иван Иванович

КУРС ВЫСШЕЙ МАТЕМАТИКИ

*Учебник для студентов высших педагогических
учебных заведений*

Издание второе, переработанное и дополненное

Зав. редакцией С. В. Платонов

Зав. художественной редакцией И. А. Пшеничников

Художник обложки М. Б. Патрушева

Компьютерная верстка В. Г. Верхозин

Корректор Т. Я. Кокорева

Отпечатано с диапозитивов, изготовленных
ООО «Гуманитарный издательский центр ВЛАДОС».

Лицензия ИД № 03185 от 10.11.2000.

Санитарно-эпидемиологическое заключение

№ 77.99.02.953.Д.006153.08.03 от 18.08.2003.

Сдано в набор 17.08.03. Подписано в печать 18.12.03.

Формат 60×90/16. Печать офсетная. Бумага газетная. Усл. печ. л. 35,0.

Тираж 20 000 экз. (1-й завод 1—5 000 экз.). Заказ № .

Гуманитарный издательский центр ВЛАДОС.

119571, Москва, просп. Вернадского, 88,

Московский педагогический государственный университет.

Тел. 437-11-11, 437-25-52, 437-99-98; тел./факс 735-66-25.

E-mail: vlados@dol.ru

<http://www.vlados.ru>

ООО «Полиграфист».

160001, Россия, г. Вологда, ул. Челюскинцев, 3.