ВЫСШАЯ МАТЕМАТИКА

А.А.ГУСАК

2
том

ББК 22.1я73
Г96
УДК 51(075.8)

Рецензенты:

кафедра общей математики ЛГУ им. А. А. Жданова,
Л. Д. Кудрявцев,
доктор физико-математических наук

Гусак А. А.

Г96 Высшая математика: [Учеб. пособие для естеств. спец. ун-тов]. Т. 2. 2-е изд., перераб. и доп.— Мн.: Изд-во «Университетское», 1984.—383 с., черт.

В пер.: 80 к.

Второе издание пособия для студентов естественных специальностей университетов переработано и дополнено в соответствии с новой программой по высшей математике (1976).

Г $\frac{1702010000-056}{М317-84}$ 23—84 ББК 22.1я73

РАЗДЕЛ II

МАТЕМАТИЧЕСКИЙ АНАЛИЗ (продолжение)*)

Глава 18

ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

При исследовании различных вопросов естественных наук приходится рассматривать переменные величины, зависящие от многих других переменных величин. Для изучения такого рода зависимостей служит понятие функции нескольких переменных (см. § 10.4).

Глава посвящена дифференциальному исчислению функций нескольких переменных и его приложениям.

§ 18.1. Некоторые предварительные понятия

Введем определения понятий, которые понадобятся в дальнейшем.

Упорядоченную совокупность n действительных чисел $x_1, x_2, \ldots, x_n$ называют *точкой*, а сами числа — ее *координатами*. Запись $M(x_1, x_2, \ldots, x_n)$ означает, что точка M имеет координаты $x_1, x_2, \ldots, x_n$. Множество всех указанных точек называется *арифметическим n-мерным пространством* и обозначается символом A^n.

Арифметическое n-мерное пространство A^n называется *n-мерным евклидовым пространством*, если для любых двух точек $M'(x_1', x_2', \ldots, x_n')$, $M''(x_1'', x_2'', \ldots, x_n'')$, принадлежащих A^n, определено расстояние по формуле

$$\rho(M', M'') = \sqrt{(x_1'' - x_1')^2 + (x_2'' - x_2')^2 + \ldots + (x_n'' - x_n')^2}.$$

Евклидово n-мерное пространство обозначается E^n.

Приведем примеры множеств в n-мерном евклидовом пространстве E^n.

---

*) Начало см. в кн.: *Гусак А. А.* Высшая математика.— Минск, 1983, т. 1.—462 с.

1. Множество $\{M\}$ точек $M(x_1, x_2, \ldots, x_n)$, координаты которых удовлетворяют неравенству $\rho(M, M_0) \leqslant R$ или

$$(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2 + \ldots + (x_n - x_n^0)^2 \leqslant R^2,$$

называется *замкнутым n-мерным шаром* радиуса R с центром в точке $M_0(x_1^0, x_2^0, \ldots, x_n^0)$.

2. Если для координат всех точек множества $\{M\}$ выполняется строгое неравенство $\rho(M, M_0) < R$ или

$$(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2 + \ldots + (x_n - x_n^0) < R^2,$$

то $\{M\}$ называется *открытым n-мерным шаром.*

3. Множество $\{M\}$ точек $M(x_1, x_2, \ldots, x_n)$, для которых $\rho(M, M_0) = R$ или

$$(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2 + \ldots + (x_n - x_n^0)^2 = R^2,$$

называется *$(n-1)$-мерной сферой* радиуса R с центром в точке $M_0(x_1^0, x_2^0, \ldots, x_n^0)$.

4. Множество точек $M(x_1, x_2, \ldots, x_n)$, координаты которых заданы как непрерывные функции $x_i = x_i(t)$ $(i = 1, 2, \ldots, n)$, определенные на некотором отрезке $[a, b]$, называется *непрерывной кривой* в пространстве E^n. Аргумент t называется параметром кривой. Точка $A(x_1(a), x_2(a), \ldots, x_n(a))$ называется началом, точка $B(x_1(b), x_2(b), \ldots, x_n(b))$ — концом данной кривой.

Множество точек M n-мерного евклидова пространства E^n, для каждой из которых расстояние до фиксированной точки M_0 меньше $\varepsilon > 0$, называется *ε-окрестностью* точки M_0. Другими словами, ε-окрестностью точки M_0 называется n-мерный открытый шар радиуса ε с центром в точке M_0.

Окрестностью точки M_0 называется любой открытый n-мерный шар, содержащий эту точку.

Множество называется *ограниченным,* если все его точки принадлежат некоторому n-мерному шару. *Диаметром* ограниченного множества называется верхняя грань расстояний между его двумя любыми точками.

Пусть $\{M\}$ — некоторое множество точек n-мерного евклидова пространства E^n. Введем следующие определения.

Точка M называется *предельной точкой* множества $\{M\}$, если любая ее окрестность содержит по крайней мере одну точку множества $\{M\}$, отличную от M. Предельная точка множества может принадлежать или не принадлежать ему. Например, точки $x=1$, $x=2$ являются предельными для каждого из множеств $[1, 2]$, $]1, 2[$, но первому они принадлежат, а второму не принадлежат. Если существует окрестность точки M множества $\{M\}$, не содержащая никаких других точек этого множества кроме самой точки M, то эта точка называется *изолированной* точкой множества $\{M\}$. Множество, содержащее все свои предельные точки, называется *замкнутым*. Точка M множества $\{M\}$ называется *внутренней* точкой этого множества, если существует такая ее окрестность, все точки которой принадлежат множеству $\{M\}$. *Открытым множеством* называется множество, все точки которого — внутренние. Множество называется *связным*, если любые две его точки можно соединить в нем непрерывной кривой. *Областью* называется открытое связное множество. Точка M называется *граничной* точкой множества $\{M\}$, если любая ее окрестность содержит как точки этого множества, так и точки, не принадлежащие ему. Совокупность всех граничных точек множества $\{M\}$ называется его *границей*. Если к области присоединить ее границу, то полученное множество называется *замкнутой областью*. Замкнутая область, очевидно, является замкнутым множеством. Например, множество точек $M(x, y)$ плоскости Oxy, для которых $x^2+y^2\leqslant 1$, является замкнутой областью: к области, определяемой неравенством $x^2+y^2<1$, присоединены все ее граничные точки, т. е. точки окружности $x^2+y^2=1$.

Функция, определенная на некотором множестве n-мерного арифметического пространства A^n, называется функцией n переменных (см. § 10.4). В этом случае пишут

$$u=f(M), \quad \text{или} \quad u=f(x_1, x_2, \ldots, x_n), \tag{18.1}$$

где $M(x_1, x_2, \ldots, x_n)$ — точка указанного множества.

Замечание. Точку n-мерного пространства обозначают также буквой x, $x=(x_1, x_2, \ldots, x_n)$, тогда для функции n переменных $x_1, x_2, \ldots, x_n$ применяют запись $u=f(x)$.

§ 18.2. Предел и непрерывность функции нескольких переменных

Рассмотрим функцию $z=f(x, y)$, определенную в некоторой области. Фиксируем в этой области точки $M(x, y)$,

$M_1(x+\Delta x,\ y)$, $M_2(x,\ y+\Delta y)$, $M_3(x+\Delta x,\ y+\Delta y)$ и вычислим в них значения данной функции.

Полное приращение функции двух переменных

$$z=f(x,\ y) \tag{18.2}$$

в точке $M(x,\ y)$ определяется формулой

$$\Delta z=f(x+\Delta x,\ y+\Delta y)-f(x,\ y), \tag{18.3}$$

а ее *частные приращения* (по x и y соответственно) в той же точке — формулами

$$\Delta_x z=f(x+\Delta x,\ y)-f(x,\ y); \tag{18.4}$$

$$\Delta_y z=f(x,\ y+\Delta y)-f(x,\ y). \tag{18.5}$$

Аналогично определяются полное и частные приращения функции большего числа переменных.

Замечание 1. Частное приращение функции по одному из аргументов есть разность между двумя ее значениями, когда приращение получает только данный аргумент; полное приращение функции — разность между двумя значениями, когда приращения получают все ее аргументы.

Пусть дана функция $u=f(M)$, определенная на множестве $X\subset E^n$, и M_0 — предельная точка множества X (эта точка может как принадлежать множеству X, так и не принадлежать ему). Число a называется *пределом* функции $u=f(M)$ в точке M_0 (или при $M\to M_0$), если для любого числа $\varepsilon>0$ существует такое число $\delta>0$, что для каждой точки $M\in X$, удовлетворяющей условию

$$0<\rho(M,\ M_0)<\delta, \tag{18.6}$$

выполняется неравенство

$$|f(M)-a|<\varepsilon. \tag{18.7}$$

Обозначение предела функции $u=f(M)$ в точке M_0

$$\lim_{M\to M_0} f(M)=a. \tag{18.8}$$

Если число a является пределом функции двух переменных $u=f(x,\ y)$ в точке $M_0(x_0,\ y_0)$, то употребляется также запись

$$\lim_{\substack{x\to x_0\\ y\to y_0}} f(x,\ y)=a. \tag{18.9}$$

Функция $\alpha=\alpha(M)$ называется бесконечно малой в точке M_0, если

$$\lim_{M\to M_0} \alpha(M)=0. \tag{18.10}$$

Теорема 18.1. Для того чтобы функция $u=f(M)$ в точке M_0 имела предел, равный a, необходимо и достаточно, чтобы ее можно было представить в следующем виде:

$$f(M)=a+\alpha(M), \tag{18.11}$$

где $\alpha(M)$ — бесконечно малая функция в точке M_0.

Доказательство этой теоремы аналогично доказательству теоремы 10.4.

Рассуждая, как и в случае функции одной переменной, можно доказать также теорему о пределе суммы, произведения и частного функций нескольких переменных, аналогичную теореме 10.8.

Пример. Вычислить $\lim\limits_{\substack{x\to 1\\ y\to 2}} (3x^2+4y-1)$.

Принимая во внимание теорему о пределе суммы фуикций, получаем

$$\lim_{\substack{x\to 1\\ y\to 2}} (3x^2+4y-1)=\lim_{\substack{x\to 1\\ y\to 2}} 3x^2+\lim_{\substack{x\to 1\\ y\to 2}} 4y-1=$$
$$=\lim_{x\to 1} 3x^2+\lim_{y\to 2} 4y-1=3+8-1=10.$$

Рассмотрим функцию $u=f(M)$, областью определения которой является множество $X\subset E^n$, и точку $M_0\in X$. Функция $u=f(M)$ называется *непрерывной* в точке M_0, если для любого числа $\varepsilon>0$ существует такое число $\delta>0$, что для каждой точки $M\in X$, удовлетворяющей условию

$$\rho(M, M_0)<\delta, \tag{18.12}$$

выполняется неравенство

$$|f(M)-f(M_0)|<\varepsilon. \tag{18.13}$$

Замечание 2. Если M_0 — предельная точка множества X, то определение непрерывности функции в точке может быть сформулировано следующим образом.

Функция $u=f(M)$ называется непрерывной в точке M_0, если

$$\lim_{M\to M_0} f(M)=f(M_0). \tag{18.14}$$

Отметим, что функция, рассмотренная в предыдущем примере, непрерывна в точке $M_0(1, 2)$, так как

$$\lim_{\substack{x \to 1 \\ y \to 2}} f(x, y) = f(1, 2).$$

Необходимое и достаточное условие непрерывности функции $u = f(M)$ в точке M_0 выражается равенством

$$\lim_{\Delta\rho \to 0} \Delta u = 0, \text{ или } \lim_{\Delta\rho \to 0} [f(M) - f(M_0)] = 0, \qquad (18.15)$$

где $\Delta\rho = \rho(M, M_0)$, $\Delta u = f(M) - f(M_0)$.

Действительно, если выполняется равенство (18.14), то

$$\lim_{M \to M_0} [f(M) - f(M_0)] = 0, \text{ или } \lim_{\Delta\rho \to 0} \Delta u = 0.$$

Обратно, если выполнено равенство (18.15), то

$$\lim_{M \to M_0} f(M) = f(M_0).$$

Функция, непрерывная в каждой точке некоторого множества, называется непрерывной на этом множестве. Приведем без доказательства следующую теорему.

Теорема 18.2. Если функция $u = f(M)$ определена и непрерывна на ограниченном замкнутом множестве, то она ограничена и достигает своего наименьшего и наибольшего значения.

Точки области определения функции, в которых функция непрерывна, называются *точками непрерывности* этой функции. Те точки области определения функции, в которых она не является непрерывной, называются *точками разрыва* данной функции. Точками разрыва функции $u = f(M)$ называют также те точки, которые не принадлежат области ее определения, но являются предельными точками этой области.

§ 18.3. Частные производные функции нескольких переменных

Частной производной функции нескольких переменных по одной из них в фиксированной точке называется предел отношения соответствующего частного приращения этой функции к приращению данной переменной, когда последнее стремится к нулю.

Для функции $z = f(x, y)$ двух переменных частная производная по x в точке $M_0(x_0, y_0)$ определяется формулой

$$\left(\frac{\partial f}{\partial x}\right)_{M_0} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x,\ y_0) - f(x_0,\ y_0)}{\Delta x}. \qquad (18.16)$$

Заметим, что $\frac{\partial f}{\partial x}$ — единый символ для обозначения частной производной функции $f(x, y)$ по переменной x; ∂f и ∂x самостоятельного смысла не имеют. Запись $\left(\frac{\partial f}{\partial x}\right)_{M_0}$ означает, что частная производная вычислена в точке $M_0(x_0, y_0)$.

Частную производную функции $z=f(x, y)$ по переменной x в точке $M_0(x_0, y_0)$ обозначают и другими символами:

$$\left(\frac{\partial z}{\partial x}\right)_{M_0},\ z'_x(x_0,\ y_0),\ f'_x(x_0,\ y_0).$$

Частная производная функции $f(x, y)$ по переменной x выражает скорость изменения функции в данном направлении $(y=y_0)$, или скорость изменения функции $f(x, y_0)$ одной переменной x.

Частная производная функции $z=f(x, y)$ по переменной y в точке $M_0(x_0, y_0)$ определяется формулой

$$\left(\frac{\partial f}{\partial y}\right)_{M_0} = \lim_{\Delta y \to 0} \frac{f(x_0,\ y_0 + \Delta y) - f(x_0,\ y_0)}{\Delta y}. \qquad (18.17)$$

Она обозначается также и другими символами:

$$\left(\frac{\partial z}{\partial y}\right)_{M_0},\ z'_y(x_0,\ y_0),\ f'_y(x_0,\ y_0).$$

Частные производные функции $z=f(x, y)$ имеют следующую геометрическую интерпретацию:

$$f'_x(x_0,\ y_0) = \operatorname{tg}\alpha; \qquad (18.18)$$

$$f'_y(x_0,\ y_0) = \operatorname{tg}\beta, \qquad (18.19)$$

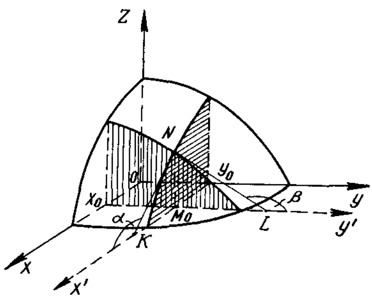

Рис. 18.1

где α — угол между осью Ox и касательной в точке $N(x_0, y_0, f(x_0, y_0))$ к линии пересечения поверхности $z=f(x, y)$ и плоскости $y = y_0$; β — угол между осью Oy и касательной в той же точке к линии пересечения данной поверхности с плоскостью $x=x_0$ (рис. 18.1).

Очевидно,

$$f'_x(x_0,\ y_0) = \left.\frac{df(x,\ y_0)}{dx}\right|_{x=x_0},$$

$$f'_y(x_0,\ y_0) = \left.\frac{df(x_0,\ y)}{dy}\right|_{y=y_0}, \qquad (18.20)$$

т. е. частная производная в данной точке равна производной функции одной переменной, вычисленной при соответствующем значении аргумента, поэтому при нахождении частных производных пользуются обычными правилами дифференцирования.

При переходе от точки $M_0(x_0;\ y_0)$ к точке $M(x,\ y)$ получим новые значения частных производных. Следовательно, частные производные функции $z=f(x,\ y)$ также являются некоторыми функциями двух переменных:

$$\frac{\partial z}{\partial x} = f'_x(x,\ y), \quad \frac{\partial z}{\partial y} = f'_y(x,\ y). \qquad (18.21)$$

Пример 1. Дана функция $z=x^2y^3$. Найти ее частные производные и их значения в точке $M_0(-2,\ 1)$.

Считая y постоянной и дифференцируя функцию по x, получаем

$$\frac{\partial z}{\partial x} = (x^2y^3)'_x = y^3(x^2)'_x = 2xy^3.$$

Считая x постоянной и дифференцируя функцию по y, находим

$$\frac{\partial z}{\partial y} = (x^2y^3)'_y = x^2(y^3)'_y = 3x^2y^2.$$

Вычисляем значения частных производных в точке M_0, т. е. при $x_0=-2$, $y_0=1$:

$$\left(\frac{\partial z}{\partial x}\right)_{M_0} = 2(-2)\,1^3 = -4, \quad \left(\frac{\partial z}{\partial y}\right)_{M_0} = 3(-2)^2 1^2 = 12.$$

В соответствии с определением для функции $u=f(x,\ y,\ z)$ и точки $M_0(x_0,\ y_0,\ z_0)$ получаем ее частные производные:

$$\left(\frac{\partial f}{\partial x}\right)_{M_0} = \lim_{\Delta x\to 0}\frac{f(x_0+\Delta x,\ y_0,\ z_0)-f(x_0,\ y_0,\ z_0)}{\Delta x}; \qquad (18.22)$$

$$\left(\frac{\partial f}{\partial y}\right)_{M_0} = \lim_{\Delta y\to 0}\frac{f(x_0,\ y_0+\Delta y,\ z_0)-f(x_0,\ y_0,\ z_0)}{\Delta y}; \qquad (18.23)$$

$$\left(\frac{\partial f}{\partial z}\right)_{M_0} = \lim_{\Delta z\to 0}\frac{f(x_0,\ y_0,\ z_0+\Delta z)-f(x_0,\ y_0,\ z_0)}{\Delta z}. \qquad (18.24)$$

Частные производные функции нескольких переменных называют также *частными производными первого порядка,* или *первыми частными производными.*

Частными производными второго порядка (или *вторыми частными производными*) данной функции называются соответствующие частные производные от ее первых частных производных.

Для функции $z=f(x, y)$ по определению имеем

$$\frac{\partial^2 z}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right)=[f'_x(x, y)]'_x=f''_{xx}(x, y);$$

$$\frac{\partial^2 z}{\partial y^2}=\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial y}\right)=[f'_y(x, y)]'_y=f''_{yy}(x, y);$$

$$\frac{\partial^2 z}{\partial x\partial y}=\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right)=[f'_x(x, y)]'_y=f''_{xy}(x, y);$$

$$\frac{\partial^2 z}{\partial y\partial x}=\frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right)=[f'_y(x, y)]'_x=f''_{yx}(x, y).$$

Пример 2. Дана функция $z=x^3+4x^2y-6xy^2+y^3$. Найти ее частные производные второго порядка.

Находим сначала первые частные производные:

$$\frac{\partial z}{\partial x}=3x^2+8xy-6y^2,\quad \frac{\partial z}{\partial y}=4x^2-12xy+3y^2.$$

Пользуясь определениями и правилами дифференцирования, получаем

$$\frac{\partial^2 z}{\partial x^2}=6x+8y,\quad \frac{\partial^2 z}{\partial x\partial y}=8x-12y;$$

$$\frac{\partial^2 z}{\partial y\partial x}=8x-12y,\quad \frac{\partial^2 z}{\partial y^2}=-12x+6y.$$

Вторые частные производные обозначаются также символами z''_{xx}, z''_{xy}, z''_{yx}, z''_{yy}. Производные z''_{xy}, z''_{yx} называются *смешанными частными производными.* В рассмотренном примере смешанные частные производные оказались равными, т. е. $z''_{xy}=z''_{yx}$. При выполнении некоторых условий последнее равенство всегда справедливо. Эти условия выражаются теоремой, приводимой здесь без доказательства.

Теорема 18.3. Если функция $z=f(x, y)$ и ее смешанные производные f''_{xy}, f''_{yx} определены в некоторой окрестности точки $M_0(x_0, y_0)$, причем производные непрерывны в этой точке, то

$$f''_{xy}(x_0, y_0)=f''_{yx}(x_0, y_0). \qquad (18.25)$$

Итак, если смешанные производные непрерывны, то они равны, т. е. результат дифференцирования не зависит от порядка дифференцирования.

Дифференцируя частные производные второго порядка по x и по y, получаем *частные производные третьего порядка,* или *третьи частные производные:*

$$\frac{\partial^3 z}{\partial x^3}, \quad \frac{\partial^3 z}{\partial x^2 \partial y}, \quad \frac{\partial^3 z}{\partial x \partial y \partial x}, \quad \frac{\partial^3 z}{\partial y \partial x^2}, \quad \frac{\partial^3 z}{\partial y \partial x \partial y}, \quad \frac{\partial^3 z}{\partial y^2 \partial x}, \quad \frac{\partial^3 z}{\partial y^3}.$$

Вообще, частная производная n-го порядка функции $z=f(x, y)$ есть первая частная производная от ее частной производной $(n-1)$-го порядка.

Аналогично определяются и вычисляются частные производные второго и высших порядков от функции трех и большего числа переменных.

Пример 3. Дана функция $u = xy \sin 2t + x^2z^5$. Найти $\frac{\partial^3 u}{\partial x \partial y \partial t}$, $\frac{\partial^4 u}{\partial x^2 \partial z^2}$, $\frac{\partial^5 u}{\partial z^5}$.

Дифференцируя по одной из переменных, считаем все другие постоянными:

$$\frac{\partial u}{\partial x} = y \sin 2t + 2xz^5, \quad \frac{\partial^2 u}{\partial x \partial y} = \sin 2t, \quad \frac{\partial^3 u}{\partial x \partial y \partial t} = 2 \cos 2t;$$

$$\frac{\partial^2 u}{\partial x^2} = 2z^5, \quad \frac{\partial^3 u}{\partial x^2 \partial z} = 10z^4, \quad \frac{\partial^4 u}{\partial x^2 \partial z^2} = 40z^3;$$

$$\frac{\partial u}{\partial z} = 5x^2z^4, \quad \frac{\partial^2 u}{\partial z^2} = 20x^2z^3, \quad \frac{\partial^3 u}{\partial z^3} = 60x^2z^2,$$

$$\frac{\partial^4 u}{\partial z^4} = 120x^2z, \quad \frac{\partial^5 u}{\partial z^5} = 120x^2.$$

§ 18.4. Полный дифференциал функции нескольких переменных

Рассмотрим функцию двух переменных $z=f(x, y)$ и ее полное приращение в точке $M_0(x_0, y_0)$:

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0). \qquad (18.26)$$

Если существуют числа P и Q такие, что полное приращение можно представить в виде

$$\Delta z = P\Delta x + Q\Delta y + \varepsilon \Delta \rho \quad (\Delta \rho \neq 0), \qquad (18.27)$$

где

$$\Delta\rho = \sqrt{\Delta x^2+\Delta y^2} \quad \text{и} \quad \varepsilon \to 0 \text{ при } \Delta\rho \to 0, \qquad (18.28)$$

то выражение $P\Delta x+Q\Delta y$ называется *полным дифференциалом* функции $z=f(x, y)$ в точке M_0. В этом случае полное приращение функции состоит из двух частей: первая часть $P\Delta x+Q\Delta y$ является линейной относительно Δx и Δy, вторая — бесконечно малой высшего порядка по сравнению с $\Delta\rho = \sqrt{\Delta x^2+\Delta y^2}$. Полный дифференциал функции $z=f(x, y)$ обозначается через dz, т. е.

$$dz=P\Delta x+Q\Delta y. \qquad (18.29)$$

Теорема 18.4. Если полное приращение функции $z=f(x, y)$ представимо формулой (18.27), то числа P и Q равны значениям ее соответствующих частных производных в точке M_0:

$$P = f'_x(x_0, y_0), \ Q = f'_y(x_0, y_0). \qquad (18.30)$$

Доказательство. Если $\Delta y=0$, формула (18.27) принимает вид $\Delta_x z=P\Delta x+\varepsilon|\Delta x|$, где $\varepsilon \to 0$ при $\Delta x \to 0$. Следовательно,

$$\frac{\Delta_x z}{\Delta x} = P \pm \varepsilon, \quad \lim_{\Delta x\to 0} \frac{\Delta_x z}{\Delta x} = P, \ f'_x(x_0, y_0) = P.$$

Если $\Delta x=0$, то $\Delta_y z=Q\Delta y+\varepsilon|\Delta y|$, где $\varepsilon \to 0$ при $\Delta y \to 0$, поэтому

$$\frac{\Delta_y z}{\Delta y} = Q \pm \varepsilon, \quad \lim_{\Delta y\to 0} \frac{\Delta_y z}{\Delta y} = Q, \ f'_y(x_0, y_0) = Q.$$

Доказанная теорема дает возможность записать формулу (18.29) по-другому:

$$dz = f'_x(x_0, y_0)\,\Delta x + f'_y(x_0, y_0)\,\Delta y. \qquad (18.31)$$

Теорема 18.5. Если первые частные производные функции $z=f(x, y)$ определены в некоторой окрестности точки $M_0(x_0, y_0)$ и непрерывны в самой точке M_0, то данная функция имеет полный дифференциал в этой точке.

Доказательство. Преобразуем формулу (18.26) прибавив с плюсом и минусом значение $f(x_0, y_0+\Delta y)$:

$$\begin{aligned}\Delta z = &[f(x_0+\Delta x, y_0+\Delta y)-f(x_0, y_0+\Delta y)]+\\ &+[f(x_0, y_0+\Delta y)-f(x_0, y_0)]. \end{aligned} \qquad (18.32)$$

В первой квадратной скобке записано приращение функции одной переменной x ($y_0+\Delta y$ фиксировано), во второй — приращение функции одной переменной y (x_0 фиксировано). К каждой из этих функций применим теорему Лагранжа о конечном приращении:

$$f(x_0+\Delta x,\ y_0+\Delta y)-f(x_0,\ y_0+\Delta y)=$$
$$=f'_x(\xi,\ y_0+\Delta y)\,\Delta x,\quad \xi\in[x_0,\ x_0+\Delta x];\qquad(18.33)$$
$$f(x_0,\ y_0+\Delta y)-f(x_0,\ y_0)=f'_y(x_0,\ \eta)\,\Delta y,$$
$$\eta\in[y_0,\ y_0+\Delta y],\qquad(18.34)$$

отметим, что $\xi\to x_0$, $\eta\to y_0$, когда $\Delta x\to 0$, $\Delta y\to 0$.

Так как по условию функции $f'_x(x,\ y)$, $f'_y(x,\ y)$ непрерывны в точке M_0, то (см. (18.11) и (18.14))

$$\left.\begin{aligned} f'_x(\xi,\ y_0+\Delta y)&=f'_x(x_0,\ y_0)+\alpha;\\ f'_y(x_0,\ \eta)&=f'_y(x_0,\ y_0)+\beta,\end{aligned}\right\}\qquad(18.35)$$

где $\alpha\to 0$, $\beta\to 0$ при $\Delta x\to 0$, $\Delta y\to 0$.

Формула (18.32) с учетом равенств (18.33)—(18.35) принимает вид

$$\Delta z=f'_x(x_0,\ y_0)\,\Delta x+f'_y(x_0,\ y_0)\,\Delta y+\alpha\Delta x+\beta\Delta y.\qquad(18.36)$$

Преобразуя сумму последних двух слагаемых, получаем

$$\alpha\Delta x+\beta\Delta y=\frac{\alpha\Delta x+\beta\Delta y}{\Delta\rho}\,\Delta\rho=\varepsilon\Delta\rho,\qquad(18.37)$$

где

$$\varepsilon=\frac{\alpha\Delta x+\beta\Delta y}{\Delta\rho},\quad \Delta\rho=\sqrt{\Delta x^2+\Delta y^2}.\qquad(18.38)$$

Поскольку $\left|\dfrac{\Delta x}{\Delta\rho}\right|\leqslant 1$, $\left|\dfrac{\Delta y}{\Delta\rho}\right|\leqslant 1$, то

$$|\varepsilon|=\left|\frac{\alpha\Delta x+\beta\Delta y}{\Delta\rho}\right|\leqslant|\alpha|\left|\frac{\Delta x}{\Delta\rho}\right|+|\beta|\left|\frac{\Delta y}{\Delta\rho}\right|\leqslant|\alpha|+|\beta|.$$

Так как $\alpha\to 0$, $\beta\to 0$ при $\Delta x\to 0$ и $\Delta y\to 0$, то и $\varepsilon\to 0$ при $\Delta x\to 0$, $\Delta y\to 0$, т. е. $\varepsilon\to 0$ при $\Delta\rho\to 0$.

Следовательно, формула (18.36) принимает вид

$$\Delta z=f'_x(x_0,\ y_0)\,\Delta x+f'_y(x_0,\ y_0)\,\Delta y+\varepsilon\Delta\rho,\qquad(18.39)$$

где $\varepsilon\to 0$ при $\Delta\rho\to 0$.

Сравнивая формулы (18.27) и (18.39), заключаем, что функция $z=f(x,\ y)$ имеет дифференциал в точке M_0, что и требовалось доказать.

Функция, имеющая полный дифференциал в данной точке, называется *дифференцируемой* в этой точке. Следовательно, если в данной точке и некоторой ее окрестности частные производные функции непрерывны, то функция дифференцируема в этой точке.

Замечание 1. Если функция дифференцируема в некоторой точке, то она и непрерывна в ней. Действительно, из формулы (18.39) следует $\lim\limits_{\Delta\rho\to 0}\Delta z=0$.

Замечание 2. Формулу (18.39) можно представить так:

$$\Delta z=dz+\varepsilon\Delta\rho,\ \varepsilon\to 0 \text{ при } \Delta\rho\to 0, \tag{18.40}$$

откуда

$$\Delta z\approx dz,\ f(x_0+\Delta x,\ y_0+\Delta y)-f(x_0,\ y_0)\approx f'_x(x_0,\ y_0)\Delta x+ \\ +f'_y(x_0,\ y_0)\Delta y \tag{18.41}$$

или

$$f(x_0+\Delta x,\ y_0+\Delta y)\approx f(x_0,\ y_0)+f'_x(x_0,\ y_0)\Delta x+f'_y(x_0,\ y_0)\Delta y.$$

Последние формулы применяются в приближенных вычислениях.

Перейдем от точки $M_0(x_0,\ y_0)$ к точке $M(x,\ y)$, в которой функция $z=f(x,\ y)$ определена и удовлетворяет в ней условиям теоремы 18.5; получим новое значение полного дифференциала:

$$dz=f'_x(x,\ y)\Delta x+f'_y(x,\ y)\Delta y. \tag{18.42}$$

Таким образом, полный дифференциал есть функция четырех переменных: $x,\ y,\ \Delta x,\ \Delta y$; при фиксированных значениях Δx и Δy полный дифференциал является функцией только x и y. Обозначим

$$dx=\Delta x,\ dy=\Delta y \tag{18.43}$$

и назовем эти величины дифференциалами независимых переменных. Формула (18.42) принимает вид

$$dz=f'_x(x,\ y)\,dx+f'_y(x,\ y)\,dy, \tag{18.44}$$

т. е. полный дифференциал функции равен сумме произведений первых частных производных на соответствующие дифференциалы аргументов. Каждое из произведений $f'_x(x,\ y)\,dx$, $f'_y(x,\ y)\,dy$ называют частным дифференциалом.

Формулу (18.44) можно записать и в другом виде:

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy, \; dz = z'_x dx + z'_y dy.$$

Аналогично определяется и выражается полный дифференциал функции трех и более переменных. Например, если для функции $u = f(x, y, z)$ и точки $M(x, y, z)$ существуют числа P, Q, R такие, что $\Delta u = P\Delta x + Q\Delta y + R\Delta z + \varepsilon\Delta\rho$ $(\Delta\rho \neq 0)$, где $\Delta\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$ и $\varepsilon \to 0$ при $\Delta\rho \to 0$, то полным дифференциалом этой функции в данной точке *называется сумма первых трех слагаемых*, т. е. $du = P\Delta x + Q\Delta y + R\Delta z$. Если первые частные производные функции непрерывны, то

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz,$$

где $dx = \Delta x$, $dy = \Delta y$, $dz = \Delta z$.

Если все первые частные производные функции n переменных $u = f(x_1, x_2, \ldots, x_n)$ непрерывны, то ее полный дифференциал выражается формулой

$$du = \frac{\partial u}{\partial x_1} dx_1 + \frac{\partial u}{\partial x_2} dx_2 + \ldots + \frac{\partial u}{\partial x_n} dx_n. \quad (18.45)$$

§ 18.5. Дифференциалы высших порядков

Полный дифференциал функции нескольких переменных называют также первым дифференциалом, или, короче, дифференциалом. Дифференциалы будем обозначать не только символом d, но и символом δ (например, δx, δy, δz).

Предположим, что функции $f'_x(x, y)$, $f'_y(x, y)$, входящие в формулу (18.44), дифференцируемы в точке $M(x, y)$. Следовательно, функции f''_{xx}, f''_{xy}, f''_{yx}, f''_{yy} непрерывны в этой точке и определены в некоторой ее окрестности (поэтому $f''_{xy} = f''_{yx}$ в силу теоремы 18.3). Будем считать dx и dy в формулах (18.44) постоянными множителями, тогда функция dz, определяемая указанной формулой, является функцией только двух переменных x и y. В силу предположения о функциях $f'_x(x, y)$ и $f'_y(x, y)$ функция dz дифференцируема в точке M. Дифференциал этой функции имеет вид

$$\delta(dz) = \delta(f'_x(x, y)\,dx + f'_y(x, y)\,dy).$$

Дифференциал $\delta(dz)$ от дифференциала dz в точке M, взятый при $\delta x=dx$, $\delta y=dy$, называют вторым дифференциалом функции $z=f(x, y)$ и обозначают d^2z:

$$d^2z=\delta(dz) \text{ при } \delta x=dx,\ \delta y=dy. \tag{18.46}$$

Аналогично определяется и обозначается третий дифференциал функции $z=f(x, y)$ в точке $M(x, y)$:

$$d^3z=\delta(d^2z) \text{ при } \delta x=dx,\ \delta y=dy. \tag{18.47}$$

Дифференциал $\delta(d^{n-1}z)$ от дифференциала $d^{n-1}z$, взятый при $\delta x=dx$, $\delta y=dy$, называется дифференциалом n-го порядка (или n-м дифференциалом) функции $z=f(x, y)$ в точке $M(x, y)$ и обозначается d^nz:

$$d^nz=\delta(d^{n-1}z) \text{ при } \delta x=dx,\ \delta y=dy. \tag{18.48}$$

Найдем выражение для дифференциала второго порядка функции $z=f(x, y)$ в точке $M(x, y)$:

$$\left.\begin{aligned} d^2z &= \delta(dz) = \delta(f'_x(x, y)\,dx + f'_y(x, y)\,dy) = \\ &= \delta(f'_x(x, y))\,dx + \delta(f'_y(x, y))\,dy = \\ &= (f''_{xx}(x, y)\,\delta x + f''_{xy}(x, y)\delta y)\,dx + (f''_{yx}(x, y)\,\delta x + \\ &+ f''_{yy}(x, y)\,\delta y)dy = f''_{xx}(x, y)\,\delta x dx + f''_{xy}(x, y)(dx\delta y + \\ &+ \delta x dy) + f''_{yy}(x, y)\,dy\delta y. \end{aligned}\right\} \tag{18.49}$$

Положив здесь $\delta x=dx$, $\delta y=dy$, получим

$$d^2z = f''_{xx}(x, y)\,dx^2 + 2f''_{xy}(x, y)\,dxdy + f''_{yy}(x, y)\,dy^2,$$

или

$$d^2z = \frac{\partial^2 z}{\partial x^2}dx^2 + 2\frac{\partial^2 z}{\partial x\partial y}dxdy + \frac{\partial^2 z}{\partial y^2}dy^2. \tag{18.50}$$

Замечание 1. При вычислении дифференциалов высших порядков обычно совмещаются оба шага: вычисление дифференциала от дифференциала и приравнивание дифференциалов аргументов ($\delta x = dx$, $\delta y = dy$). Например, если $z = f(x, y)$, $dz = z'_x dx + z'_y dy$, то

$$d^2z = d(dz) = d(z'_x dx + z'_y dy) = (z'_x dx + z'_y dy)'_x dx + (z'_x dx + $$
$$+ z'_y dy)'_y dy = z''_{xx}dx^2 + z''_{xy}dxdy + z''_{yx}dydx + z''_{yy}dy^2 = z''_{xx}dx^2 + $$
$$+ 2z''_{xy}dxdy + z''_{yy}dy^2 \ (z''_{xy} = z''_{yx}).$$

В случае непрерывности частных производных третьего порядка функции $z=f(x, y)$ аналогично получаем формулу для ее третьего дифференциала:

$$d^3z = \frac{\partial^3 z}{\partial x^3} dx^3 + 3\frac{\partial^3 z}{\partial x^2 \partial y} dx^2 dy + 3\frac{\partial^3 z}{\partial x \partial y^2} dx dy^2 + \frac{\partial^3 z}{\partial y^3} dy^3.$$

С помощью метода математической индукции можно доказать, что

$$d^n z = \sum_{k=0}^{n} C_n^k \frac{\partial^n z}{\partial x^{n-k} \partial y^k} dx^{n-k} dy^k \tag{18.51}$$

$$\left(C_n^k = \frac{n(n-1)\dots[n-(k-1)]}{1 \cdot 2 \dots k}\right).$$

Формулу (18.51) записывают и в следующем символическом виде:

$$d^n z = \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^{(n)} z.$$

Замечание 2. Определение дифференциалов высших порядков можно также ввести следующим образом.

Введем сначала определение билинейной формы. Функция $2n$ переменных $x_1, x_2, \dots, x_n$; $y_1, y_2, \dots, y_n$, или функция от упорядоченной пары точек n-мерного пространства $x=(x_1, x_2, \dots, x_n)$, $y=(y_1, y_2, \dots, y_n)$ вида

$$A(x, y) = A(x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n) = \sum_{i,k=1}^{n} a_{ik} x_i y_k,$$

где a_{ik} — заданные числа ($i, k=1, 2, \dots, n$), называется билинейной формой от x, y. Функция $A(x, x)$ называется квадратичной формой, соответствующей билинейной форме $A(x, y)$:

$$A(x, x) = A(x_1, x_2, \dots, x_n; x_1, x_2, \dots, x_n) = \sum_{i,k=1}^{n} a_{ik} x_i x_k$$

(см. гл. 8). В случае $a_{ik}=a_{ki}$ ($i, k=1, 2, \dots, n$) билинейная форма $A(x, y)$ и соответствующая ей квадратичная форма $A(x, x)$ называются симметричными.

Обратимся к формуле (18.49). Поскольку значения всех частных производных вычислены в заданной точке $M(x, y)$, то правая часть этой формулы — билинейная форма переменных dx, dy, δx, δy. В силу теоремы 18.3 ($z''_{xy} = z''_{yx}$) эта форма является симметричной. Положив в формуле (18.49) $\delta x=dx$, $\delta y=dy$, получим соответствующую данной билинейной форме квадратичную форму (18.50), которую назовем вторым дифференциалом функции $z=f(x, y)$ в точке $M(x, y)$.

Таким образом, введено следующее определение: вторым дифференциалом d^2z функции $z=f(x, y)$ в данной точке $M(x, y)$ называется квадратичная форма от дифференциалов dx и dy независимых переменных, соответствующая билинейной форме от первого дифференциала:

$$d^2z = \frac{\partial^2 z}{\partial x^2} dx^2 + 2\frac{\partial^2 z}{\partial x \partial y} dx dy + \frac{\partial^2 z}{\partial y^2} dy^2.$$

§ 18.6. Дифференцирование сложных функций

Рассмотрим функцию $z=f(x, y)$, определенную на открытом множестве G. Пусть функция

$$z=f(x, y) \tag{18.52}$$

дифференцируема в точке $(x, y)\in G$, а функции

$$x=x(t),\ y=y(t), \tag{18.53}$$

зависящие от скалярного параметра t, имеют производные по t. Предположим, что при каждом значении t из некоторого интервала соответствующая точка (x, y), где x и y вычислены по формулам (18.53), принадлежит области G. В этом случае имеет смысл сложная функция

$$z=f(x(t),\ y(t))=F(t). \tag{18.54}$$

Докажем, что производная сложной функции существует и вычисляется по формуле

$$F'(t)=f'_x(x(t),\ y(t))\,x'_t+f'_y(x(t),\ y(t))\,y'_t,$$

или, короче,

$$\frac{dz}{dt}=\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}. \tag{18.55}$$

Поскольку функция $z=f(x, y)$ дифференцируема в точке $(x, y)\in G$, то при любых достаточно малых приращениях Δx, Δy справедлива формула (18.39):

$$\Delta z=\frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y+\varepsilon\Delta\rho\quad(\Delta\rho\neq 0), \tag{18.56}$$

$$\varepsilon\to 0 \text{ при } \Delta\rho\to 0\quad(\Delta\rho=\sqrt{\Delta x^2+\Delta y^2}). \tag{18.57}$$

Значению t, которому по формулам (18.53) соответствуют координаты (x, y) данной точки, придадим приращение Δt. Оно вызовет приращения Δx, Δy функций (18.53). Подставляя эти приращения в формулу (18.56), находим приращение $\Delta z=\Delta F=F(t+\Delta t)-F(t)$ функции (18.54) в точке t. Так как функции (18.53) имеют производные и выполняется условие (18.57), т. е.

$$\lim_{\Delta t\to 0}\frac{\Delta x}{\Delta t}=\frac{dx}{dt},\quad \lim_{\Delta t\to 0}\frac{\Delta y}{\Delta t}=\frac{dy}{dt},$$

$$\lim_{\Delta t\to 0}\varepsilon\sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2+\left(\frac{\Delta y}{\Delta t}\right)^2}=0\cdot\sqrt{x_t'^2+y_t'^2}=0,$$

то, разделив это приращение на Δt и перейдя к пределу при $\Delta t \to 0$, получим формулу (18.55):

$$F'(t) = \lim_{\Delta t \to 0} \frac{\Delta F}{\Delta t} = \lim_{\Delta t \to 0} \left(\frac{\partial f}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y} \frac{\Delta y}{\Delta t} + \varepsilon \frac{\Delta \rho}{\Delta t} \right) =$$

$$= \lim_{\Delta t \to 0} \left(\frac{\partial f}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y} \frac{\Delta y}{\Delta t} + \varepsilon \sqrt{\left(\frac{\Delta x}{\Delta t} \right)^2 + \left(\frac{\Delta y}{\Delta t} \right)^2} \right) =$$

$$= \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}.$$

Замечание 1. Если аргументы x, y функции (18.52) являются дифференцируемыми функциями двух независимых переменных t и τ:

$$x = x(t, \tau), \quad y = y(t, \tau), \tag{18.58}$$

то для сложной функции

$$z = f(x(t, \tau), \ y(t, \tau)) = \varphi(t, \tau), \tag{18.59}$$

фиксируя сначала τ, а затем t, на основании (18.55) получаем

$$\frac{\partial z}{\partial t} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t}, \quad \frac{\partial z}{\partial \tau} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \tau} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \tau}. \tag{18.60}$$

Замечание 2. Аналогично можно доказать формулу для производной сложной функции трех и более аргументов. Если

$$u = f(x, y, z) \tag{18.61}$$

— дифференцируемая функция, а функции аргумента t

$$x = x(t), \quad y = y(t), \quad z = z(t) \tag{18.62}$$

имеют производные, то производная сложной функции

$$u = f(x(t), \ y(t), \ z(t)) = F(t) \tag{18.63}$$

существует и вычисляется по формуле

$$\frac{du}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt}. \tag{18.64}$$

Замечание 3. Если аргументы x, y, z функции (18.61) являются дифференцируемыми функциями двух независимых переменных t и τ:

$$x = x(t, \tau), \quad y = y(t, \tau), \quad z = z(t, \tau),$$

то для сложной функции $u = f(x(t, \tau), \ y(t, \tau), \ z(t, \tau)) = \varphi(t, \tau)$, фиксируя сначала τ, а затем t, на основании (18.64) получаем

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial t} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial t}; \\ \frac{\partial u}{\partial \tau} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial \tau} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \tau} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial \tau}. \end{aligned} \right\}$$

Замечание 4. Если аргументы x, y, z функции (18.61) являются дифференцируемыми функциями трех независимых переменных ξ, η, ζ:

$$x=x(\xi, \eta, \zeta),\ y=y(\xi, \eta, \zeta),\ z=z(\xi, \eta, \zeta),$$

то частные производные сложной функции от ξ, η, ζ $u=f(x(\xi, \eta, \zeta), y(\xi, \eta, \zeta), z(\xi, \eta, \zeta))$ вычисляются по формулам

$$\left.\begin{aligned}
\frac{\partial u}{\partial \xi} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \xi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \xi} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial \xi};\\
\frac{\partial u}{\partial \eta} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \eta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \eta} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial \eta};\\
\frac{\partial u}{\partial \xi} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \xi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \xi} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial \xi}.
\end{aligned}\right\}$$

§ 18.7. Дифференцирование неявных функций

Вопрос о существовании производной неявной функции решает

Теорема 18.6. Если $y=y(x)$ — непрерывная функция, заданная уравнением

$$F(x, y)=0, \tag{18.65}$$

где $F(x, y)$, $F'_x(x, y)$, $F'_y(x, y)$ — непрерывные функции в области, содержащей точку $M(x, y)$, в которой $F'_y(x, y)\neq 0$, то производная функции $y=y(x)$ в соответствующей точке существует и выражается формулой

$$\frac{dy}{dx} = -\frac{\partial F}{\partial x} : \frac{\partial F}{\partial y}. \tag{18.66}$$

Доказательство. $F(x, y)=0$ в точке $M(x, y)$. Аргументу x придадим приращение Δx, ему соответствует приращение Δy функции $y=y(x)$, поэтому $F(x+\Delta x, y+\Delta y)=0$. Приращение функции $F(x, y)$ также равно нулю, т. е. $F(x+\Delta x, y+\Delta y)-F(x, y)=0$.

Выражая полное приращение $\Delta F=F(x+\Delta x, y+\Delta y)-F(x, y)$ по формуле (18.36), получаем

$$\Delta F = \frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial y}\Delta y + \alpha\Delta x + \beta\Delta y,$$

где $\alpha\to 0$, $\beta\to 0$ при $\Delta x\to 0$ и $\Delta y\to 0$.

Последняя формула с учетом двух предыдущих принимает вид

$$\frac{\partial F}{\partial x}\Delta x + \frac{\partial F}{\partial y}\Delta y + \alpha\Delta x + \beta\Delta y = 0.$$

Разделив это равенство почленно на Δx и выразив $\frac{\Delta y}{\Delta x}$, найдем

$$\frac{\Delta y}{\Delta x} = -\left(\frac{\partial F}{\partial x} + \alpha\right) : \left(\frac{\partial F}{\partial y} + \beta\right).$$

Так как $\lim\limits_{\Delta x \to 0} \Delta y = 0$ и $\alpha \to 0$, $\beta \to 0$ при $\Delta x \to 0$, $\Delta y \to 0$, то, переходя к пределу при $\Delta x \to 0$, получаем

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -\frac{\partial F}{\partial x} : \frac{\partial F}{\partial y}, \quad y'_x = -\frac{\partial F}{\partial x} : \frac{\partial F}{\partial y}.$$

Замечание. Уравнение касательной к линии $F(x, y) = 0$ в точке $M_0(x_0, y_0)$ имеет вид

$$F'_x(x_0, y_0)(x - x_0) + F'_y(x_0, y_0)(y - y_0) = 0 \qquad (18.67)$$

и следует из формул (11.8) и (18.66).

Аналогично определяются неявные функции большего числа переменных и находятся их частные производные. Например, если $z = z(x, y)$ — функция, заданная уравнением

$$F(x, y, z) = 0, \qquad (18.68)$$

где $F(x, y, z)$, $F'_x(x, y, z)$, $F'_y(x, y, z)$, $F'_z(x, y, z)$ — непрерывные функции в области, содержащей точку $M(x, y, z)$, в которой $F'_z(x, y, z) \neq 0$, то

$$\frac{\partial z}{\partial x} = -\frac{\partial F}{\partial x} : \frac{\partial F}{\partial z}; \quad \frac{\partial z}{\partial y} = -\frac{\partial F}{\partial y} : \frac{\partial F}{\partial z}. \qquad (18.69)$$

§ 18.8. Касательная плоскость и нормаль к поверхности. Геометрический смысл полного дифференциала первого порядка

Если пространственная линия задана параметрическими уравнениями

$$x = x(t), \; y = y(t), \; z = z(t) \quad (\alpha < t < \beta), \qquad (18.70)$$

где $x(t)$, $y(t)$, $z(t)$ — дифференцируемые функции переменной t, то (см. § 13.6) уравнения касательной к ней в точке $M_0(x_0, y_0, z_0)$, $x_0 = x(t_0)$, $y_0 = y(t_0)$, $z_0 = z(t_0)$, имеют вид

$$\frac{x-x_0}{x'(t_0)}=\frac{y-y_0}{y'(t_0)}=\frac{z-z_0}{z'(t_0)} \quad (18.71)$$

(предполагается, что $x'(t_0)$, $y'(t_0)$, $z'(t_0)$ одновременно в нуль не обращаются). Направляющим вектором касательной к линии (18.70) в точке $M_0(x_0, y_0, z_0)$ является вектор

$$\tau=\{x'(t_0), y'(t_0), z'(t_0)\}. \quad (18.72)$$

Рассмотрим в пространстве поверхность, заданную уравнением

$$F(x, y, z)=0, \quad (18.73)$$

где $F(x, y, z)$ — дифференцируемая функция. Через точку $M_0(x_0, y_0, z_0)$ данной поверхности проведем линию, целиком расположенную на этой поверхности (рис. 18.2). Пусть указанная линия определяется уравнениями (18.70), тогда уравнения (18.71) будут уравнениями касательной к этой линии в точке M_0.

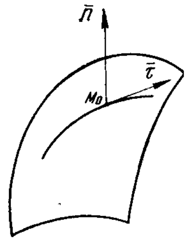

Рис. 18.2

Подставляя в уравнение (18.73) выражения (18.70), получаем соотношение $F[x(t), y(t), z(t)]=0$, являющееся тождеством относительно t для всех значений t, при которых данная точка $M(x, y, z)$ лежит на рассматриваемой линии. Продифференцируем это тождество, левая часть которого — сложная функция от t, по переменной t:

$$F'_x x'(t)+F'_y y'(t)+F'_z z'(t)=0.$$

При $t=t_0$ получаем

$$F'_x(x_0, y_0, z_0)\,x'(t_0)+F'_y(x_0, y_0, z_0)\,y'(t_0)+$$
$$+F'_z(x_0, y_0, z_0)\,z'(t_0)=0. \quad (18.74)$$

Введем в рассмотрение вектор

$$\mathbf{n}=\{F'_x(x_0, y_0, z_0),\ F'_y(x_0, y_0, z_0),\ F'_z(x_0, y_0, z_0)\}, \quad (18.75)$$

который полностью определяется поверхностью (18.73) и точкой M_0 и не зависит от выбора линии, проходящей через точку M_0.

Вектор $\mathbf{n}$, определяемый формулой (18.75), называется *вектором нормали* к поверхности (18.73) в точке $M_0(x_0, y_0, z_0)$.

Равенство (18.74), которое можно представить в виде

$$n\tau = 0, \tag{18.74a}$$

означает, что векторы (18.72) и (18.75) ортогональны.

Следовательно, какую бы линию ни провести на поверхности (18.73) через точку $M_0(x_0, y_0, z_0)$, вектор касательной τ к ней в этой точке будет перпендикулярен к вектору n. Другими словами, касательная к любой линии на поверхности (18.73), проходящей через точку M_0, лежит в плоскости, перпендикулярной к вектору n.

Касательной плоскостью к поверхности в данной ее точке называется плоскость, в которой лежит касательная прямая к любой линии на поверхности, проходящей через эту точку. Уравнение касательной плоскости к поверхности (18.73) в точке $M_0(x_0, y_0, z_0)$ имеет вид

$$F'_x(x_0, y_0, z_0)(x - x_0) + F'_y(x_0, y_0, z_0)(y - y_0) + \\ + F'_z(x_0, y_0, z_0)(z - z_0) = 0. \tag{18.76}$$

Уравнение (18.76) следует из (6.21).

Замечание. Эти рассуждения теряют смысл, если $F'_x = F'_y = F'_z = 0$. Точки поверхности (18.73), в которых все первые частные производные равны нулю, называются *особыми* (здесь они не рассматриваются).

Нормалью к поверхности в данной ее точке называется прямая, проходящая через эту точку перпендикулярно касательной плоскости к поверхности в этой точке. Направляющим вектором указанной прямой служит вектор нормали (см. (18.75)). Уравнения нормали к поверхности (18.73) в точке $M_0(x_0, y_0, z_0)$ имеют вид

$$\frac{x - x_0}{(F'_x)_0} = \frac{y - y_0}{(F'_y)_0} = \frac{z - z_0}{(F'_z)_0}, \tag{18.77}$$

где $(F'_x)_0 = F'_x(x_0, y_0, z_0)$, $(F'_y)_0 = F'_y(x_0, y_0, z_0)$, $(F'_z)_0 = F'_z(x_0, y_0, z_0)$. Эти уравнения получены из уравнений (6.41).

Пример. Написать уравнения нормали и касательной плоскости к поверхности $z = x^2 - y^2$ в точке $M_0(3, -2, 5)$.

Поскольку $F(x, y, z) = x^2 - y^2 - z$, $F'_x = 2x$, $F'_y = -2y$, $F'_z = -1$, $(F'_x)_0 = 2 \cdot 3 = 6$, $(F'_y)_0 = -2(-2) = 4$, $(F'_z)_0 = -1$, то на основании уравнений (18.76), (18.77) получаем

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

(уравнение касательной плоскости),

$$\frac{x-3}{6}=\frac{y+2}{4}=\frac{z-5}{-1}$$

(уравнения нормали).

Выясним геометрический смысл полного дифференциала первого порядка функции двух переменных. Если поверхность задана уравнением

$$z=f(x, y), \text{ или } z-f(x, y)=0,$$

то

$$\frac{\partial F}{\partial x}=-\frac{\partial f}{\partial x}, \quad \frac{\partial F}{\partial y}=-\frac{\partial f}{\partial y}, \quad \frac{\partial F}{\partial z}=1,$$

поэтому уравнение касательной плоскости к ней запишется так:

$$z-z_0=\frac{\partial f}{\partial x}(x-x_0)+\frac{\partial f}{\partial y}(y-y_0). \qquad (18.78)$$

Положив в этой формуле $x-x_0=\Delta x$, $y-y_0=\Delta y$, получим (см. (18.31)), что правая часть ее есть полный дифференциал функции $z=f(x, y)$:

$$\frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y=dz.$$

Следовательно,

$$z-z_0=dz, \qquad (18.79)$$

т. е. полный дифференциал функции двух переменных равен приращению аппликаты z касательной плоскости

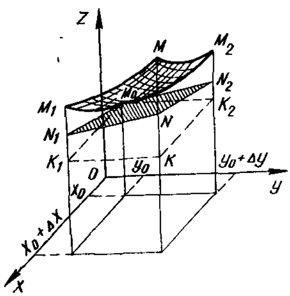

Рис. 18.3

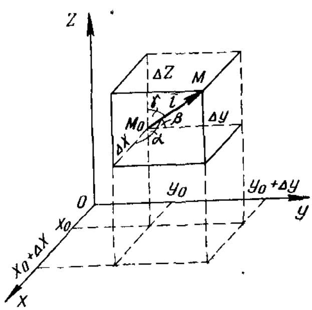

Рис. 18.4

в точке $M_0(x_0, y_0, z_0)$ к поверхности, являющейся графиком этой функции, когда аргументы x и y получают приращения Δx и Δy (рис. 18.3).

§ 18.9. Производная по направлению

Как уже отмечалось, частная производная выражает скорость изменения функции в некотором направлении. Поставим вопрос о скорости ее изменения в любом направлении.

Рассмотрим функцию $u = u(x, y, z)$, определенную и дифференцируемую в области S. Из точки $M_0(x_0, y_0, z_0)$ этой области отложим вектор $\overline{M_0M} = 1$, образующий с координатными осями углы α, β, γ соответственно (рис. 18.4). Предположим, что точка $M(x_0+\Delta x, y_0+\Delta y, z_0+\Delta z)$ также принадлежит области S.

Полное приращение функции $u=u(x, y, z)$ по аналогии с формулой (18.36) можно представить в виде

$$\Delta u = \left(\frac{\partial u}{\partial x}\right)_0 \Delta x + \left(\frac{\partial u}{\partial y}\right)_0 \Delta y + \left(\frac{\partial u}{\partial z}\right)_0 \Delta z + \\ + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y + \varepsilon_3 \Delta z, \qquad (18.80)$$

где $\varepsilon_i \to 0$ $(i=1, 2, 3)$ при $\Delta x \to 0$, $\Delta y \to 0$, $\Delta z \to 0$.

Разделим это равенство почленно на $\Delta l = \sqrt{\Delta x^2+\Delta y^2+\Delta z^2}$:

$$\frac{\Delta u}{\Delta l} = \left(\frac{\partial u}{\partial x}\right)_0 \frac{\Delta x}{\Delta l} + \left(\frac{\partial u}{\partial y}\right)_0 \frac{\Delta y}{\Delta l} + \left(\frac{\partial u}{\partial z}\right)_0 \frac{\Delta z}{\Delta l} + \\ + \varepsilon_1 \frac{\Delta x}{\Delta l} + \varepsilon_2 \frac{\Delta y}{\Delta l} + \varepsilon_3 \frac{\Delta z}{\Delta l}. \qquad (18.81)$$

Поскольку

$$\frac{\Delta x}{\Delta l} = \cos\alpha, \quad \frac{\Delta y}{\Delta l} = \cos\beta, \quad \frac{\Delta z}{\Delta l} = \cos\gamma, \quad \varepsilon_i \to 0 \text{ при } \Delta l \to 0,$$

то

$$\lim_{\Delta l \to 0} \frac{\Delta u}{\Delta l} = \left(\frac{\partial u}{\partial x}\right)_0 \cos\alpha + \left(\frac{\partial u}{\partial y}\right)_0 \cos\beta + \left(\frac{\partial u}{\partial z}\right)_0 \cos\gamma.$$

Предел отношения (18.81) при $\Delta l \to 0$ называется *производной функции $u=u(x, y, z)$ в точке $M_0(x_0, y_0, z_0)$ по направлению* вектора **1** и обозначается символом $\left(\frac{\partial u}{\partial l}\right)_0$. Следовательно,

$$\left(\frac{\partial u}{\partial l}\right)_0 = \left(\frac{\partial u}{\partial x}\right)_0 \cos\alpha + \left(\frac{\partial u}{\partial y}\right)_0 \cos\beta + \left(\frac{\partial u}{\partial z}\right)_0 \cos\gamma. \qquad (18.82)$$

Замечание 1. Производная функции $u=u(x, y, z)$ по направлению вектора **1** в области S выражается формулой

$$\frac{\partial u}{\partial l}=\frac{\partial u}{\partial x}\cos\alpha+\frac{\partial u}{\partial y}\cos\beta+\frac{\partial u}{\partial z}\cos\gamma, \qquad (18.83)$$

из которой видно, что $\partial u/\partial l$ не зависит от длины вектора, а только от направления.

Пример. Найти производную функции $u=x^2-2y^2+3z^2$ по направлению вектора $\mathbf{l}=\{1,\ 2,\ -2\}$ и ее значение в точке $M_0(9,\ 6,\ -1)$.

Так как $\frac{\partial u}{\partial x}=2x$, $\frac{\partial u}{\partial y}=-4y$, $\frac{\partial u}{\partial z}=6z$, $\cos\alpha=\frac{1}{3}$, $\cos\beta=$ $=\frac{2}{3}$, $\cos\gamma=-\frac{2}{3}$, то

$$\frac{\partial u}{\partial l}=2x\left(\frac{1}{3}\right)+(-4y)\frac{2}{3}+6z\left(-\frac{2}{3}\right)=\frac{2}{3}x-\frac{8}{3}y-4z,$$

$$\left(\frac{\partial u}{\partial l}\right)_0=\frac{2}{3}9-\frac{8}{3}6-4(-1)=-6.$$

Замечание 2. Частная производная функции нескольких переменных является частным случаем производной по направлению. Например, если $\mathbf{l}=\mathbf{i}=\{1,\ 0,\ 0\}$, т. е. $\alpha=0$, $\beta=\gamma=90°$, то

$$\frac{\partial u}{\partial l}=\frac{\partial u}{\partial x}\cos 0+\frac{\partial u}{\partial y}\cos\frac{\pi}{2}+\frac{\partial u}{\partial z}\cos\frac{\pi}{2}=\frac{\partial u}{\partial x}.$$

§ 18.10. Градиент скалярного поля

Если в каждой точке пространственной области определено значение некоторой величины, то говорят, что в данной области задано *поле* этой величины. В случае скалярной величины поле называют *скалярным,* в случае векторной — *векторным.* Примерами скалярных полей являются поле температур, поле плотностей, примерами векторных — поле скоростей, силовое поле. Поле называется *стационарным (установившимся),* если рассматриваемая величина не зависит от времени, и *нестационарным (неустановившимся)* — если она меняется с течением времени.

Если рассматриваемая величина задана в плоской области, то соответствующее поле называется *плоским.*

Плоское скалярное поле определяется функцией $u=u(x,\ y)$, или $u=u(M)$, где $M(x,\ y)$ — точка плоскости. Множество точек, в которых эта функция принимает одно и то же значение, называется *линией уровня.* Линии уровня определяются уравнением

$$u(x, y)=C \quad (C=\mathrm{const}), \tag{18.84}$$

где C принадлежит области значений функции. Скалярное поле в пространстве задается аналогично функцией $u=u(M)$, или функцией $u=u(x, y, z)$. Множество точек, в которых функция $u=u(x, y, z)$ принимает одно и то же значение, называется *поверхностью уровня*. Поверхности уровня определяются уравнением

$$u(x, y, z)=C \quad (C=\mathrm{const}). \tag{18.85}$$

Если в некоторой декартовой системе координат x, y, z функция u не зависит от одной из них, например z, то соответствующее скалярное поле называется *плоско-параллельным;* его можно рассматривать в плоскости Oxy. Для всех плоскостей, параллельных плоскости Oxy, это поле имеет одни и те же линии уровня (например, поле $u=x^2+y^2$ имеет линии уровня $x^2+y^2=C$).

В каждой точке пространственной области v, в которой задана дифференцируемая скалярная функция $u=u(x, y, z)$, определим вектор, координаты которого равны значениям частных производных этой функции в данной точке. Указанный вектор называют *градиентом* функции в этой точке и обозначают $\mathrm{grad}\, u$, т. е.

$$\mathrm{grad}\, u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}. \tag{18.86}$$

Следовательно, в области v определено векторное поле, а именно поле градиентов данной функции.

Связь между градиентом функции и производной по направлению устанавливает

Теорема 18.7. Производная функции $u=u(x, y, z)$ по направлению вектора $\mathbf{l}$ равна проекции градиента этой функции на вектор $\mathbf{l}$.

Доказательство. Пусть α, β, γ — углы, образованные вектором $\mathbf{l}$ с координатными осями, $\mathbf{l}_0$ — единичный вектор направления вектора $\mathbf{l}$, т. е.

$$\mathbf{l}_0=\mathbf{i}\cos\alpha+\mathbf{j}\cos\beta+\mathbf{k}\cos\gamma. \tag{18.87}$$

Скалярное произведение векторов (18.86) и (18.87) в координатах:

$$\mathrm{grad}\, u\cdot\mathbf{l}_0 = \frac{\partial u}{\partial x}\cos\alpha + \frac{\partial u}{\partial y}\cos\beta + \frac{\partial u}{\partial z}\cos\gamma. \tag{18.88}$$

Из равенств (18.83) и (18.88) следует

$$\frac{\partial u}{\partial l} = \operatorname{grad} u \cdot \mathbf{l}_0. \tag{18.89}$$

Обозначим через φ угол между векторами $\operatorname{grad} u$, $\mathbf{l}_0$. Поскольку

$$\operatorname{grad} u \cdot \mathbf{l}_0 = |\operatorname{grad} u|\,|\mathbf{l}_0| \cos\varphi = |\operatorname{grad} u| \cos\varphi = \text{пр}_{\mathbf{l}} \operatorname{grad} u,$$

то

$$\frac{\partial u}{\partial l} = |\operatorname{grad} u| \cos\varphi, \tag{18.90}$$

или $\dfrac{\partial u}{\partial l} = \text{пр}_{\mathbf{l}} \operatorname{grad} u$.

Следствие 1. Если $\operatorname{grad} u \neq 0$, то производная функции в точке по направлению вектора $\mathbf{l}$ имеет наибольшее значение, если направление l совпадает с направлением градиента данной функции. Это наибольшее значение равно модулю вектора $\operatorname{grad} u$.

Действительно, из равенства (18.90) видно, что наибольшее значение производной по направлению достигается при $\varphi=0$. В этом случае

$$\frac{\partial u}{\partial l} = |\operatorname{grad} u| \cos 0 = |\operatorname{grad} u| = \sqrt{u_x'^2 + u_y'^2 + u_z'^2}.$$

Другими словами, в направлении градиента скалярная функция изменяется быстрее, чем в других направлениях.

Если $\operatorname{grad} u = 0$, то $\dfrac{\partial u}{\partial l} = 0$ для любого вектора $\mathbf{l}$.

Следствие 2. Производная функции $u=u(x, y, z)$ по направлению вектора $\mathbf{l}$, касательного к поверхности уровня, определяемой уравнением $u(x, y, z)=C$, равна нулю.

В самом деле, уравнение поверхности уровня функции $u=u(x, y, z)$ можно записать так: $u(x, y, z)-C=0$, или $F(x, y, z)=0$, $F(x, y, z)=u(x, y, z)-C$. Вектор нормали $\mathbf{n}$ к этой поверхности совпадает с вектором $\operatorname{grad} u$: $n = (F_x', F_y', F_z') = (u_x', u_y', u_z') = \operatorname{grad} u$ (см. (18.75), (18.86)). В данном случае угол φ между вектором $\mathbf{l}$ и $\operatorname{grad} u$ равен $90°$, поэтому $\cos\varphi = 0$, $\dfrac{\partial u}{\partial l} = 0$.

§ 18.11. Формула Тейлора для функции нескольких переменных

Рассмотрим функцию $z=f(x, y)$, определенную в области S. Предположим, что в некоторой окрестности точки $M_0(a, b)$ этой области функция имеет непрерывные частные производные до порядка $n+1$ включительно $(n\geqslant 1)$.

Введем вспомогательную функцию $\varphi(t)=f(x, y)$, где $x=a+t\Delta x$, $y=b+t\Delta y$, $0\leqslant t\leqslant 1$. При $t=0$ получаем точку $M_0(a, b)$, при $t=1$ — точку $M(a+\Delta x, b+\Delta y)$; считаем, что точка M также принадлежит указанной окрестности точки M_0.

Как известно, формула Тейлора для функции $\varphi(t)$ имеет вид

$$\varphi(t)=\varphi(t_0)+\frac{\varphi'(t_0)}{1!}(t-t_0)+\frac{\varphi''(t_0)}{2!}(t-t_0)^2+\ldots$$

$$\ldots+\frac{\varphi^n(t_0)}{n!}(t-t_0)^n+\frac{\varphi^{n+1}(t_0+\theta t)}{(n+1)!}(t-t_0)^{n+1},$$

где $0<\theta<1$, в частности, при $t_0=0$

$$\varphi(t)=\varphi(0)+\frac{\varphi'(0)}{1!}t+\frac{\varphi''(0)}{2!}t^2+\ldots$$

$$\ldots+\frac{\varphi^n(0)}{n!}t^n+\frac{\varphi^{n+1}(\theta t)}{(n+1)!}t^{n+1}. \qquad (18.91)$$

Найдем производные функции $\varphi(t)=f(x, y)=f(a+t\Delta x, b+t\Delta y)$:

$$\varphi'(t)=f'_x x'_t+f'_y y'_t=f'_x\Delta x+f'_y\Delta y=df(x, y);$$

$$\varphi''(t)=f''_{xx}x'^2_t+f''_{xy}x'_t y'_t+f''_{yx}y'_t x'_t+f''_{yy}y'^2_t=$$

$$=f''_{xx}\Delta x^2+2f''_{xy}\Delta x\Delta y+f''_{xy}\Delta y^2=d^2f(x, y);$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\varphi^n(t)=d^nf(x, y);$$

$$\varphi^{n+1}(t)=d^{n+1}f(x, y).$$

Заменив в последнем равенстве t на θt и положив $t=0$ во всех предыдущих, найдем выражения для производных функции $\varphi(t)$, входящие в формулу (18.91):

$$\varphi'(0)=df(a, b);$$

$$\varphi''(0)=d^2f(a, b), \ldots, \varphi^n(0)=d^nf(a, b);$$

$$\varphi^{n+1}(\theta t)=d^{n+1}f(a+\theta t\Delta x, b+\theta t\Delta y).$$

Подставляя эти значения в равенство (18.91) и полагая $t=1$, получаем *формулу Тейлора* для функции двух переменных:

$$f(x, y) = f(a, b) + \sum_{k=1}^{n} \frac{d^k f(a, b)}{k!} + \\ + \frac{d^{n+1} f(a+\theta\Delta x, b+\theta\Delta y)}{(n+1)!} \qquad (18.92)$$

или

$$f(M) = f(M_0) + \sum_{k=1}^{n} \frac{d^k f(M_0)}{k!} + \frac{d^{n+1} f(M^*)}{(n+1)!}, \qquad (18.93)$$

где $M^*(a+\theta\Delta x, b+\theta\Delta y)$ — точка области S.

Формула Тейлора для функции большего числа переменных $u=f(M)$, $M(x_1, x_2, \ldots, x_n)$ аналогична формуле (18.93).

Замечание. При $n=1$ формула (18.92) принимает вид

$$f(x, y) = f(a, b) + [f'_x(a, b)\Delta x + f'_y(a, b)\Delta y] + \\ + \frac{1}{2}[f''_{xx}(\xi, \eta)\Delta x^2 + 2f''_{xy}(\xi, \eta)\Delta x\Delta y + f''_{yy}(\xi, \eta)\Delta y^2], \quad (18.92a)$$

где $\xi=a+\theta\Delta x$, $\eta=b+\theta\Delta y$, $0<\theta<1$.

§ 18.12. Экстремум функции нескольких переменных

Рассмотрим функцию $z=f(x, y)$, определенную в некоторой области.

Максимумом функции $z=f(x, y)$ называется такое ее значение $f(x_1, y_1)$, которое больше всех других значений, принимаемых в точках $M(x, y)$, достаточно близких к точке $M_1(x_1, y_1)$ и отличных от нее (рис. 18.5, *а*), т. е.

$$f(x_1, y_1) > f(x, y). \qquad (18.94)$$

Минимумом функции $z=f(x, y)$ называется такое ее значение $f(x_2, y_2)$, которое меньше всех других значений, принимаемых в точках $M(x, y)$, достаточно близких к точке $M_2(x_2, y_2)$ и отличных от нее (рис. 18.5, *б*), т. е.

$$f(x_2, y_2) < f(x, y). \qquad (18.95)$$

Максимум и минимум функции называют *экстремумом*. Точки, в которых достигается экстремум, называются *точками экстремума*. Аналогично определяется экстремум функции большего числа переменных.

Необходимое условие экстремума дифференцируемой функции нескольких переменных выражается следующей теоремой.

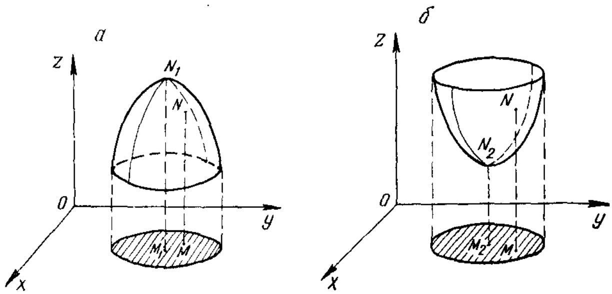

Рис. 18.5

Теорема 18.8. В точке экстремума дифференцируемой функции все ее первые частные производные равны нулю.

Доказательство проведем для случая функции двух переменных $z=f(x, y)$. Пусть $M_0(x_0, y_0)$ — точка экстремума данной функции. Предположим, что M_0 — точка максимума, тогда $f(x_0, y_0)>f(x, y)$ для всех точек $M(x, y)$, достаточно близких к точке $M_0(x_0, y_0)$, в частности,

$$f(x_0, y_0)>f(x_0+\Delta x, y_0).$$

Фиксируя $y=y_0$, получаем функцию $f(x, y_0)$ одной переменной x. Эта функция при $x=x_0$ имеет максимум (что видно из последнего неравенства), поэтому ее производная по x в точке x_0 равна нулю, т. е. $f'_x(x_0, y_0)=0$. Аналогично доказывается, что $f'_y(x_0, y_0)=0$. Итак, в точке максимума $M_0(x_0, y_0)$

$$f'_x(x_0, y_0)=0, \quad f'_y(x_0, y_0)=0. \tag{18.96}$$

Теорема 18.9. Пусть функция $z=f(x, y)$ имеет непрерывные первые и вторые частные производные в точке $M_0(a, b)$ и некоторой ее окрестности.

Если ее первые частные производные в точке M_0 равны нулю, а вторые принимают значения

$$f''_{xx}(a,\ b)=A,\ f''_{xy}(a,\ b)=B,\ f''_{yy}(a,\ b)=C, \tag{18.97}$$

то при

$$B^2-AC<0 \quad \text{и} \quad A>0 \tag{18.98}$$

точка M_0 является точкой минимума данной функции, а при

$$B^2-AC<0, \quad A<0 \tag{18.99}$$

— точкой максимума, при

$$B^2-AC>0 \tag{18.100}$$

в точке M_0 экстремума нет.

Доказательство. Воспользуемся формулой (18.92а). По условию $f'_x(a,\ b)=0,\ f'_y(a,\ b)=0$. В силу непрерывности частных производных второго порядка и равенств (18.97) получаем:

$$f''_{xx}(\xi,\ \eta)=f''_{xx}(a,\ b)+\alpha_1=A+\alpha_1,\ \alpha_1\to 0 \text{ при } \Delta\rho\to 0;$$

$$f''_{xy}(\xi,\ \eta)=f''_{xy}(a,\ b)+\alpha_2=B+\alpha_2,\ \alpha_2\to 0 \text{ при } \Delta\rho\to 0;$$

$$f''_{yy}(\xi,\ \eta)=f''_{yy}(a,\ b)+\alpha_3=C+\alpha_3,\ \alpha_3\to 0 \text{ при } \Delta\rho\to 0,$$

где $\Delta\rho=\sqrt{\Delta x^2+\Delta y^2}$.

Следовательно, формула (18.92а) принимает вид

$$f(x,\ y)-f(a,\ b)=\frac{1}{2}(A\Delta x^2+2B\Delta x\Delta y+C\Delta y^2)+$$
$$+\frac{1}{2}(\alpha_1\Delta x^2+2\alpha_2\Delta x\Delta y+\alpha_3\Delta y^2).$$

Преобразуя правую часть этой формулы, получаем при $\Delta y\neq 0$

$$f(x,\ y)-f(a,\ b)=\frac{\Delta y^2}{2}\left[A\left(\frac{\Delta x}{\Delta y}\right)^2+2B\frac{\Delta x}{\Delta y}+C\right]+$$
$$+\frac{\Delta\rho^2}{2}\left[\alpha_1\left(\frac{\Delta x}{\Delta\rho}\right)^2+2\alpha_2\frac{\Delta x}{\Delta\rho}\frac{\Delta y}{\Delta\rho}+\alpha_3\left(\frac{\Delta y}{\Delta\rho}\right)^2\right],$$

или

$$f(x,\ y)-f(a,\ b)=$$
$$=\frac{\Delta y^2}{2}\left[A\left(\frac{\Delta x}{\Delta y}\right)^2+2B\frac{\Delta x}{\Delta y}+C\right]+\alpha\Delta\rho^2, \tag{18.92б}$$

где

$$\alpha = \frac{1}{2}\left[\alpha_1\left(\frac{\Delta x}{\Delta\rho}\right)^2 + 2\alpha_2\left(\frac{\Delta x}{\Delta\rho}\right)\left(\frac{\Delta y}{\Delta\rho}\right) + \alpha_3\left(\frac{\Delta y}{\Delta\rho}\right)^2\right],$$

$$\alpha \to 0 \text{ при } \Delta\rho \to 0.$$

При достаточно малых $\Delta\rho$ знак правой части формулы (18.92б) определяется знаком выражения, стоящего в квадратной скобке, т. е. знаком квадратного трехчлена $At^2+2Bt+C$, где $t=\Delta x:\Delta y$. Известно, что при $A>0$ и $B^2-AC<0$ этот трехчлен будет положительным, при $A<0$ и $B^2-AC<0$ — отрицательным, при $B^2-AC>0$ он меняет знак.

Следовательно, если $A>0$ и $B^2-AC<0$, т. е. выполнено условие (18.98), то $f(x, y)-f(a, b)>0$ или $f(a, b)<$ $<f(x, y)$. Это означает, что $M_0(a, b)$ — точка минимума данной функции.

Если $A<0$, $B^2-AC<0$, т. е. выполнено условие (18.99), то $f(x, y)-f(a, b)<0$, или $f(x, y)<f(a, b)$, $M_0(a, b)$ — точка максимума.

Если $B^2-AC>0$, т. е. выполнено условие (18.100), то разность $f(x, y)-f(a, b)$ меняет знак, в точке $M_0(a, b)$ нет ни максимума, ни минимума. Теорема доказана.

Пример. Найти значение экстремума функции $f(x, y)=x^2+$ $+y^2-2x+4y+8$. Поскольку $f'_x=2x-2$, $f'_y=2y+4$, $f''_{xx}=2$, $f''_{xy}=0$, $f''_{yy}=2$, $f'_x=0$, $f'_y=0$ при $x=1$, $y=-2$, $B^2-AC=0-2\cdot2=-4<0$, $A=2$, то в точке $M_0(1, -2)$ функция имеет минимум, причем

$$\min f(x, y)=f(1, -2)=3.$$

Замечание 1. Точка, в которой $B^2-AC>0$, называется *точкой минимакса*. Например, точка $M_0(0, 0)$ является точкой минимакса функции $z=x^2-y^2$ (см. рис. 6.24).

Замечание 2. Случай $B^2-AC=0$ требует дополнительного исследования.

Замечание 3. Если дана функция большего числа переменных, то $f(M)-f(M_0)=\frac{1}{2}d^2f(M^*)$; M_0 — точка максимума, если $d^2f(M^*)<0$; M_0 — точка минимума, если $d^2f(M^*)>0$.

§ 18.13. Условный экстремум

Поставим задачу о нахождении экстремума функции

$$z=f(x, y) \qquad (18.101)$$

при условии, что переменные x и y связаны соотношением

$$\varphi(x, y)=0. \qquad (18.102)$$

[Т]акой экстремум называется условным в отличие от без[ус]ловного (см. § 18.12). Уравнение (18.102) называют [ур]*авнением связи.*

Пусть $y=y(x)$ — функция, определяемая уравнением [(1]8.102), тогда функция

$$z=f(x, y(x)) \tag{18.103}$$

[з]ависит от одной переменной x. Задача об условном [э]кстремуме функции двух переменных сводится к задаче [н]а отыскание экстремума функции одной переменной [(с]м. § 12.3, 12.4).

Разрешить уравнение (18.102) относительно y не всег[д]а представляется возможным, поэтому поставленную [з]адачу нужно уметь решать и другим способом.

Предположим, что $\varphi(x, y)$ — дифференцируемая функ[ц]ия, $\frac{\partial\varphi}{\partial y} \neq 0$, тогда, как известно (см. § 18.7),

$$y_x' = -\frac{\partial\varphi}{\partial x} : \frac{\partial\varphi}{\partial y}.$$

С другой стороны, дифференцируя функцию (18.103) [к]ак сложную, находим $z_x' = f_x' + f_y'y'$. Необходимое усло[ви]е экстремума этой функции выражается равенством [f_x'] $+ f_y'y' = 0$, откуда $y_x' = -\frac{\partial f}{\partial x} : \frac{\partial f}{\partial y}$.

Сравнивая два выражения для y_x', получаем

$$-\frac{\partial\varphi}{\partial x} : \frac{\partial\varphi}{\partial y} = -\frac{\partial f}{\partial x} : \frac{\partial f}{\partial y} \text{ или } \frac{\partial f}{\partial x} : \frac{\partial\varphi}{\partial x} = \frac{\partial f}{\partial y} : \frac{\partial\varphi}{\partial y}.$$

Обозначим последние равные отношения через $-\lambda$, [т]огда в точке условного экстремума $\frac{f_x'}{\varphi_x'} = \frac{f_y'}{\varphi_y'} = -\lambda$, [о]ткуда

$$f_x' + \lambda\varphi_x' = 0,\ f_y' + \lambda\varphi_y' = 0. \tag{18.104}$$

Введем вспомогательную функцию

$$F(x, y, \lambda) = f(x, y) + \lambda\varphi(x, y), \tag{18.105}$$

[г]де λ — параметр, называемый *множителем Лагранжа.*

Уравнения (18.104) можно записать так:

$$F_x'(x, y, \lambda) = 0,\ F_y'(x, y, \lambda) = 0. \tag{18.106}$$

Из уравнений (18.102) и (18.106) определяются координаты точки возможного условного экстремума и значение параметра λ, который играет лишь вспомогательную роль.

Следовательно, для нахождения точек возможного экстремума функции (18.101) с условием (18.102) составляют функцию (18.105), называемую *функцией Лагранжа*, находят ее частные производные по x, y, λ, приравнивают их к нулю и решают полученные уравнения относительно x, y, λ.

Заметим, что уравнения (18.102) и (18.106) выражают необходимый признак условного экстремума.

Аналогично определяется условный экстремум функции трех и более переменных при наличии уравнений связи (разумеется, их число меньше числа аргументов).

Если, например, требуется найти экстремум функции $u=f(x, y, z)$ при условиях

$$\varphi(x, y, z)=0, \ \psi(x, y, z)=0, \qquad (18.107)$$

то вводят функцию Лагранжа

$$F(x, y, z, \lambda, \mu)=f(x, y, z)+\lambda\varphi(x, y, z)+ \\ +\mu\psi(x, y, z) \qquad (18.108)$$

и к уравнениям (18.107) присоединяют еще три уравнения

$$F'_x(x, y, z, \lambda, \mu)=0, \ F'_y(x, y, z, \lambda, \mu)=0, \\ F'_z(x, y, z, \lambda, \mu)=0. \qquad (18.109)$$

Из полученной системы уравнений определяют значения x, y, z, λ, μ, где x, y, z — координаты точки возможного условного экстремума.

Уравнения (18.107) и (18.109) выражают необходимый признак условного экстремума функции $u=f(x, y, z)$.

§ 18.14. Семейства линий на плоскости. Огибающая однопараметрического семейства линий

Множество линий на плоскости, определяемое уравнением

$$F(x, y, C)=0, \qquad (18.110)$$

где C — параметр, называется *однопараметрическим* *семейством линий.* Примеры однопараметрических семейств линий: 1) $y=x^2+C$; 2) $y=(x-C)^2$; 3) $xy=C$ $(C\neq 0)$; 4) $(x-C)^2+(y-C)^2=C^2$, представляющие соответственно множества: 1) парабол с вершинами на оси Oy; 2) парабол с вершинами на оси Ox; 3) гипербол, для которых координатные оси являются асимптотами; 4) окружностей, касающихся координатных полуосей (одного знака). (Читателю предлагается построить несколько линий каждого семейства, фиксировав $C=1$, $C=2$, $C=-1$, $C=-2$, $C=-3$.)

Огибающей однопараметрического семейства линий на плоскости называется линия, которая в каждой своей точке касается некоторой линии этого семейства, причем в различных точках касается различных линий (рис. 18.6).

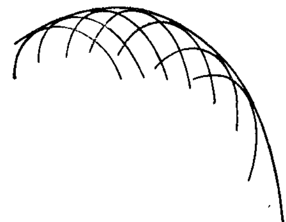

Рис. 18.6

Предположим, что семейство линий, определяемое уравнением (18.110), имеет огибающую, уравнение которой $y=y(x)$, где $y(x)$ — дифференцируемая функция. Пусть $M(x, y)$ — произвольная точка огибающей семейства (18.110), принадлежащая также некоторой линии данного семейства. Этой линии соответствует определенное значение параметра C, которое при фиксированных x и y определяется уравнением (18.110), т. е. $C=C(x, y)$, поэтому для всех точек огибающей выполняется равенство $F(x, y, C(x, y))=0$.

Если $y=y(x)$, последнее равенство обращается в тождество. Продифференцируем это тождество по x, предположив, что $C(x, y)$ — дифференцируемая функция своих аргументов, отличная от постоянной:

$$\frac{\partial F}{\partial x}+\frac{\partial F}{\partial y}y'+\frac{\partial F}{\partial C}\frac{\partial C}{\partial x}+\frac{\partial F}{\partial C}\frac{\partial C}{\partial y}y'=0$$

или

$$F'_x+F'_y y'+F'_c\left(\frac{\partial C}{\partial x}+\frac{\partial C}{\partial y}y'\right)=0. \qquad (18.111)$$

Из равенства (18.111) находим угловой коэффициент касательной к огибающей в точке $M(x, y)$.

Угловой коэффициент касательной к линии данного семейства в той же точке $M(x, y)$ найдем из уравнения

(18.110). Поскольку для линии семейства C является постоянной, то

$$F'_x + F'_y y' = 0. \qquad (18.112)$$

Так как угловые коэффициенты касательных к линии семейства и огибающей в одной и той же точке равны между собой (см. определение огибающей), то из уравнений (18.111) и (18.112) следует

$$F'_c\left(\frac{\partial C}{\partial x} + \frac{\partial C}{\partial y}\, y'\right) = 0.$$

Так как по предположению $C(x, y) \neq \text{const}$, то

$$\frac{\partial C}{\partial x} + \frac{\partial C}{\partial y}\, y' \neq 0.$$

Следовательно, для всех точек огибающей выполняется равенство $F'_c(x, y, C) = 0$.

Таким образом, для определения огибающей семейства (18.110) служат уравнения

$$F(x, y, C) = 0, \; F'_c(x, y, C) = 0. \qquad (18.113)$$

Замечание 1. *Особой точкой* линии $F(x, y) = 0$ называется такая ее точка, в которой обе частные производные первого порядка обращаются в нуль. Особая точка линии определяется системой уравнений

$$F(x, y) = 0, \; F'_x(x, y) = 0, \; F'_y(x, y) = 0.$$

Замечание 2. Если для семейства линий (18.110) некоторая функция $y = y(x)$ определяет множество особых точек, то координаты этих точек удовлетворяют уравнениям (18.113).

Следовательно, уравнения (18.113) определяют огибающую, либо множество особых точек, либо сочетание того и другого.

Множество всех точек, координаты которых удовлетворяют системе уравнений (18.113), называется *дискриминантной линией* семейства (18.110).

§ 18.15. Эмпирические формулы

При обработке результатов наблюдений часто встречаются со следующей задачей: получен ряд значений переменных x и y, однако характер функциональной зависимости между ними остается неизвестным. По полученным данным требуется найти аналитическое выражение зависимости между x и y. Формулы, полученные при решении подобного ряда задач, называются *эмпирическими*.

Эмпирическими формулами, составленными по результатам опытов и наблюдений, нередко приходится пользоваться в естествознании, в частности в химических, физических и других науках.

Задача о составлении эмпирической формулы заключается в следующем. Пусть результаты измерений имеют вид

x	x_1	x_2	x_3	...	x_k	x_{k+1}	...	x_n
y	y_1	y_2	y_3	...	y_k	y_{k+1}	...	y_n

и $y=\varphi(x, C_1^0, C_2^0, \ldots, C_m^0)$ — искомая эмпирическая формула, где $\varphi(x, C_1, C_2, \ldots, C_m)$ зависит от x и некоторых параметров $C_1, C_2, \ldots, C_m$. Разности $\varphi(x_k)-y_k=\varepsilon_k$ $(k=1, 2, \ldots, n)$, где y_k — числа из второй строки таблицы, $\varphi(x_k)$ — значения функции $\varphi(x, C_1, C_2, \ldots, C_m)$ при соответствующих значениях аргумента x_k — числах из первой строки таблицы, называются *уклонениями*, или *погрешностями*. Требуется так подобрать значения $C_1^0, C_2^0, \ldots, C_m^0$ параметров $C_1, C_2, \ldots, C_m$ функции $\varphi(x, C_1, C_2, \ldots, C_m)$, чтобы уклонения ε_k оказались наименьшими.

Наиболее распространенным критерием малости уклонений является критерий, лежащий в основе *метода наименьших квадратов:* параметры функции $\varphi(x, C_1, C_2, \ldots, C_m)$ необходимо выбрать так, чтобы оказалась минимальной сумма квадратов уклонений

$$\sum_{k=1}^{n} \varepsilon_k^2 = \varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_n^2. \qquad (18.114)$$

Рассмотрим вопрос об определении параметров эмпирической формулы по методу наименьших квадратов в случае линейной зависимости.

Обратимся к таблице. Будем рассматривать x_k, y_k как прямоугольные декартовы координаты точек на плоскости

$$M_1(x_1, y_1),\ M_2(x_2, y_2), \ldots, M_n(x_n, y_n).$$

Предположим, что эти точки лежат почти на некоторой прямой. Естественно предположить, что между x и y существует линейная зависимость, т. е.

$$y=ax+b, \tag{18.115}$$

где a и b — параметры, значения которых необходимо определить. Равенство (18.115) можно записать также в виде

$$ax+b-y=0. \tag{18.116}$$

Поскольку точки $M_k(x_k, y_k)$ только приблизительно лежат на прямой, определяемой уравнением (18.115), то и формула эта будет приближенной. Следовательно, подставляя в левую часть формулы (18.116) вместо x и y значения x_k, y_k ($k=1, 2, \ldots, n$), взятые из данной таблицы, получаем равенства

$$\left.\begin{aligned} ax_1+b-y_1&=\varepsilon_1;\\ ax_2+b-y_2&=\varepsilon_2;\\ &\ldots\ldots\\ ax_n+b-y_n&=\varepsilon_n,\end{aligned}\right\} \tag{18.117}$$

где $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ — уклонения, или погрешности.

Требуется подобрать коэффициенты a и b так, чтобы уклонения были по возможности малыми по абсолютной величине. Согласно методу наименьших квадратов подберем коэффициенты a и b так, чтобы сумма квадратов уклонений

$$u=\varepsilon_1^2+\varepsilon_2^2+\ldots+\varepsilon_n^2 \tag{18.118}$$

была наименьшей. Если эта сумма окажется достаточно малой, то и сами уклонения будут малыми по абсолютной величине.

Подставляя равенства (18.117) в формулу (18.118), получаем

$$u=(ax_1+b-y_1)^2+(ax_2+b-y_2)^2+\ldots$$
$$\ldots+(ax_n+b-y_n)^2. \tag{18.119}$$

Переменная величина u является функцией двух переменных a и b (a и b — неизвестные, подлежащие определению, x_k, y_k — числа, полученные в результате измерений). Подберем параметры a и b так, чтобы функция u принимала возможно меньшее значение. Для этого необходимо, чтобы $\frac{\partial u}{\partial a}=0$, $\frac{\partial u}{\partial b}=0$.

Находя частные производные функции u по a и b, приравнивая их нулю, получаем *нормальную систему* уравнений

$$\left.\begin{aligned} a\sum_{k=1}^{n} x_k^2 + b\sum_{k=1}^{n} x_k &= \sum_{k=1}^{n} x_k y_k; \\ a\sum_{k=1}^{n} x_k + bn &= \sum_{k=1}^{n} y_k, \end{aligned}\right\} \qquad (18.120)$$

из которой определяются параметры a и b эмпирической формулы (18.115).

Поставим задачу об определении параметров эмпирической формулы по методу наименьших квадратов в случае квадратичной зависимости. Обратимся снова к таблице. Пары значений x_k, y_k будем рассматривать как прямоугольные декартовы координаты точек на плоскости. Предположим, что точки с соответствующими координатами $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, ..., $M_n(x_n, y_n)$ почти лежат на некоторой параболе. В таком случае естественно предположить, что между x и y существует приближенная квадратичная зависимость, т. е.

$$y = ax^2 + bx + c, \qquad (18.121)$$

где a, b, c — параметры, значения которых подлежат определению.

Если в правую часть формулы (18.121) вместо x подставим значения x_k из таблицы, то получим числа $z_k = ax_k^2 + bx_k + c$.

Было бы идеальным подобрать значения параметров a, b, c так, чтобы при всех k оказалось $z_k = y_k$ ($k = 1, 2, \ldots, n$). Однако при $n > 3$ этого обычно сделать не удается, поскольку значения a, b, c, найденные из уравнений

$$y_1 = ax_1^2 + bx_1 + c, \quad y_2 = ax_2^2 + bx_2 + c, \quad y_3 = ax_3^2 + bx_3 + c,$$

как правило, не будут удовлетворять уравнениям

$$y_4 = ax_4^2 + bx_4 + c, \ \ldots, \ y_n = ax_n^2 + bx_n + c,$$

если $y(x)$ не являлась квадратным трехчленом. Другими словами, $z_k - y_k = \varepsilon_k$ ($k = 1, 2, \ldots, n$), где ε_k — уклонения, или погрешности.

Параметры a, b, c эмпирической формулы (18.121) выберем так, чтобы сумма квадратов уклонений

$$u = \varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_n^2 = (z_1 - y_1)^2 + (z_2 - y_2)^2 + \ldots$$
$$\ldots + (z_n - y_n)^2 = (ax_1^2 + bx_1 + c - y_1)^2 +$$
$$+ (ax_2^2 + bx_2 + c - y_2)^2 + \ldots + (ax_n^2 + bx_n + c - y_n)^2$$

была наименьшей. Для этого необходимо, чтобы $\frac{\partial u}{\partial a} = 0$, $\frac{\partial u}{\partial b} = 0$, $\frac{\partial u}{\partial c} = 0$.

Находя производные функции $u = u(a, b, c)$ по переменным a, b, c и приравнивая их нулю, получаем нормальную систему уравнений

$$\left.\begin{aligned}
&a\sum_{k=1}^{n} x_k^4 + b\sum_{k=1}^{n} x_k^3 + c\sum_{k=1}^{n} x_k^2 = \sum_{k=1}^{n} y_k x_k^2;\\
&a\sum_{k=1}^{n} x_k^3 + b\sum_{k=1}^{n} x_k^2 + c\sum_{k=1}^{n} x_k = \sum_{k=1}^{n} y_k x_k;\\
&a\sum_{k=1}^{n} x_k^2 + b\sum_{k=1}^{n} x_k + nc = \sum_{k=1}^{n} y_k,
\end{aligned}\right\} \quad (18.122)$$

из которой определяются значения параметров эмпирической формулы (18.121).

Глава 19

КРАТНЫЕ ИНТЕГРАЛЫ

Многие геометрические, физические и другие задачи приводят к понятию интеграла от функции двух, трех и большего числа переменных. В главе рассматриваются двойные и тройные интегралы, их приложения, вводится понятие о многократных интегралах.

§ 19.1. Задачи, приводящие к двойным интегралам

Задача об объеме цилиндроида. Рассмотрим тело с основанием S, лежащим в плоскости Oxy, ограниченное поверхностью $z = f(x, y)$ и цилиндрической поверхностью с образующей, параллельной оси Oz, и направляющей — линией γ, границей области S (рис. 19.1). Это тело называется *цилиндроидом* (*цилиндрическим брусом*, или *общим цилиндром*). Требуется вычислить объем цилиндроида.

Чтобы решить задачу, область S сетью дуг (рис. 19.2) разобьем на элементарные области $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$, площади которых также обозначим через ΔS_1, $\Delta S_2, \ldots, \Delta S_n$ соответственно. В каждой из элементарных областей ΔS_k $(k=1, 2, 3, \ldots, n)$ выберем произвольную

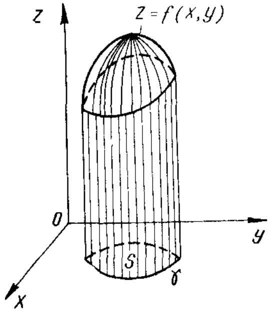

Рис. 19.1

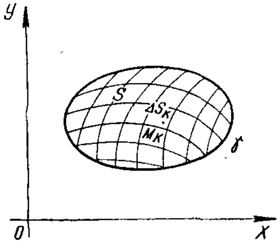

Рис. 19.2

точку $M_k(x_k, y_k)$ и значение функции в этой точке $f(x_k, y_k)$ умножим на площадь области ΔS_k. Произведение $f(x_k, y_k)\Delta S_k$ равно объему цилиндрического тела с площадью основания ΔS_k и высотой $h_k=f(x_k, y_k)$. Составим сумму всех таких произведений:

$$\sum_{k=1}^{n} f(x_k, y_k)\Delta S_k. \tag{19.1}$$

Эта сумма выражает объем V_n ступенчатого цилиндрического тела, приближенно заменяющего данный цилиндроид, т. е.

$$V_n = \sum_{k=1}^{n} f(x_k, y_k)\Delta S_k. \tag{19.2}$$

Обозначим диаметр элементарной области ΔS_k через d_k, а наибольший из этих диаметров — через λ_n. Очевидно, если $n \to \infty$, то $\lambda_n \to 0$.

Объемом общего цилиндра называется предел объема соответствующего ступенчатого тела при $\lambda_n \to 0$:

$$V = \lim_{n\to\infty} V_n = \lim_{\lambda_n\to 0} \sum_{k=1}^{n} f(x_k, y_k)\Delta S_k. \tag{19.3}$$

Задача о массе пластинки. Рассмотрим область S плоскости Oxy, ограниченную замкнутой линией γ (см. рис. 19.2), в которой распределено вещество с плотностью $p=f(x, y)$. Такую область называют *пластинкой*. Вычислим массу пластинки, предположив известной функцию $p=f(x, y)\geqslant 0$.

Область S сетью дуг разобьем на элементарные области $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$, площади которых обозначим теми же символами. Предположим, что в каждой элементарной области ΔS_k плотность постоянна и равна плотности в некоторой точке $M_k(x_k, y_k)$ этой области, т. е. $p_k=f(x_k, y_k)$. Тогда произведение $f(x_k, y_k)\Delta S_k$ выражает приближенную массу элементарной пластинки ΔS_k, а сумма всех таких произведений — приближенную массу m_n всей пластинки, т. е.

$$m_n = \sum_{k=1}^{n} f(x_k, y_k)\Delta S_k. \qquad (19.4)$$

Точное значение массы всей пластинки получим, перейдя к пределу при $\lambda_n \to 0$, где λ_n — наибольший из диаметров d_k области ΔS_k:

$$m = \lim_{n\to\infty} m_n = \lim_{\lambda_n\to 0} \sum_{k=1}^{n} f(x_k, y_k)\Delta S_k. \qquad (19.5)$$

Замечание. Обе задачи привели к необходимости рассмотрения двумерной интегральной суммы

$$\sum_{k=1}^{n} f(x_k, y_k)\Delta S_k$$

и ее предела при $\lambda_n \to 0$. Последний предел, если он существует и не зависит от способа разбиения этой области на элементарные и выбора точек M_k, называется *двойным интегралом* от функции $f(x, y)$ по области S и обозначается

$$\iint\limits_S f(x, y)\,dS = \lim_{\lambda_n\to 0} \sum_{k=1}^{n} f(x_k, y_k)\Delta S_k.$$

§ 19.2. Двойной интеграл

Рассмотрим функцию $z=f(x, y)$, определенную в области S, которая ограничена замкнутой линией γ (см. рис. 19.2). Область S сетью дуг разобьем на n элементарных областей $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$. Предполагается, что область S и элементарные области $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$ имеют площади, которые обозначим теми же символами. В каждой из элементарных областей ΔS_k $(k=1, 2, \ldots, n)$ произвольно выберем точку $M_k(x_k, y_k)$, значение функции $f(x_k, y_k)$ в этой точке умножим на площадь ΔS_k, составим сумму всех таких произведений:

$$I_n = \sum_{k=1}^{n} f(x_k, y_k)\Delta S_k. \tag{19.6}$$

Сумма (19.6) *называется интегральной суммой* для функции $f(x, y)$ по области S.

Обозначим через λ наибольший из диаметров элементарных областей ΔS_k $(k=1, 2, \ldots, n)$.

Число I называется *пределом интегральной суммы* I_n при $\lambda \to 0$, если для любого числа $\varepsilon > 0$ можно указать такое число $\delta > 0$, что при $\lambda < \delta$ выполняется неравенство

$$|I - I_n| < \varepsilon \tag{19.7}$$

независимо от выбора точек $M_k(x_k, y_k)$ в элементарных областях ΔS_k.

Двойным интегралом от функции $f(x, y)$ по области S называется предел ее интегральной суммы при $\lambda \to 0$:

$$\iint_S f(x, y)\, dS = \lim_{\lambda \to 0} \sum_{k=1}^{n} f(x_k, y_k)\, \Delta S_k, \tag{19.8}$$

функция $f(x, y)$ называется *подынтегральной* функцией, а область S — *областью интегрирования*. Двойной интеграл от функции $f(x, y)$ по области S обозначается также следующим образом: $\iint_S f(x, y)\, dxdy$, т. е.

$$\iint_S f(x, y)\, dxdy = \lim_{\lambda \to 0} \sum_{k=1}^{n} f(x_k, y_k)\, \Delta S_k. \tag{19.9}$$

Отметим без доказательства, что предел в правой части формулы (19.8) существует, если функция $z=f(x, y)$ непрерывна в замкнутой области, имеющей площадь. (Формулировка соответствующей теоремы и ее доказательство содержится, например, в [14].) Если существует предел (19.8), то функция $f(x, y)$ называется *интегрируемой* в области S. Следовательно, все непрерывные функции являются интегрируемыми, среди разрывных функций имеются интегрируемые и неинтегрируемые.

Из формул (19.3) и (19.8) следует, что

$$\iint\limits_S f(x, y)\, dS = V. \tag{19.10}$$

Формула (19.10) выражает *геометрический смысл двойного интеграла:* двойной интеграл от функции $f(x, y) \geqslant 0$ по области S равен объему цилиндроида с основанием S, который ограничен сверху поверхностью $z = f(x, y)$.

Из формул (19.5) и (19.8) следует, что

$$\iint\limits_S f(x, y)\, dS = m. \tag{19.11}$$

Формула (19.9) выражает *физический смысл двойного интеграла:* если неотрицательная функция $p = f(x, y)$ выражает поверхностную плотность пластинки S, то ее масса равна двойному интегралу от данной функции по данной области S.

§ 19.3. Свойства двойного интеграла

Двойной интеграл обладает свойствами, аналогичными свойствам определенного интеграла.

1. Если функции $f(x, y)$ и $\varphi(x, y)$ интегрируемы в области S, то интегрируемы в ней их сумма и разность, причем

$$\iint\limits_S [f(x, y) \pm \varphi(x, y)]\, dS = \iint\limits_S f(x, y)\, dS \pm \iint\limits_S \varphi(x, y)\, dS.$$

2. Постоянный множитель можно выносить за знак двойного интеграла:

$$\iint\limits_S cf(x, y)\, dS = c \iint\limits_S f(x, y)\, dS, \quad c = \text{const}.$$

3. Если $f(x, y)$ интегрируема в области S, а эта область разбита на две непересекающиеся области S_1 и S_2, то

$$\iint\limits_S f(x, y)\, dS = \iint\limits_{S_1} f(x, y)\, dS + \iint\limits_{S_2} f(x, y)\, dS.$$

4. Если $f(x, y)$ и $\varphi(x, y)$ интегрируемы в области S, в которой $f(x, y) \leqslant \varphi(x, y)$, то

$$\iint_S f_1(x, y)\, dS \leqslant \iint_S \varphi_1(x, y)\, dS.$$

5. Если $f(x, y)$ интегрируема в области S, то $|f(x, y)|$ также интегрируема в ней, причем

$$\left| \iint_S f(x, y)\, dS \right| \leqslant \iint_S |f(x, y)|\, dS.$$

6. Если в области S функция $f(x, y)$ удовлетворяет неравенствам $m \leqslant f(x, y) \leqslant M$, то

$$mS \leqslant \iint_S f(x, y)\, dS \leqslant MS,$$

где S — площадь области S.

Доказательства этих свойств приводятся, например, в [11, 14].

§ 19.4. Вычисление двойного интеграла в прямоугольных декартовых координатах

Пусть требуется вычислить двойной интеграл

$$\iint_R f(x, y)\, dxdy,$$

где область R — прямоугольник, определяемый неравенствами $a \leqslant x \leqslant b$, $c \leqslant y \leqslant d$ (рис. 19.3).

Предположим, что $f(x, y)$ непрерывна в этом прямоугольнике и принимает в нем неотрицательные значения, тогда данный двойной интеграл равен объему тела с основанием R, ограниченного сверху поверхностью $z = f(x, y)$, с боков — плоскостями $x = a$, $x = b$, $y = c$, $y = d$:

$$V = \iint_R f(x, y)\, dxdy.$$

С другой стороны, на основании формулы (17.16)

$$V = \int_a^b \sigma(x)\, dx,$$

где $\sigma(x)$ — площадь сечения данного тела плоскостью, проходящей через точку x и перпендикулярной к оси Ox. Так как рассматриваемое сечение является криволинейной трапецией, ограниченной сверху графиком функции

$z=f(x, y)$, где x фиксировано, $c \leqslant y \leqslant d$, то по формуле (16.8) будем иметь

$$\sigma(x)=\int_c^d f(x, y)\,dy.$$

Из этих трех равенств следует, что

$$\iint_R f(x, y)\,dxdy=\int_a^b\left[\int_c^d f(x, y)\,dy\right]dx. \qquad (19.12)$$

Итак, вычисление данного двойного интеграла свелось к вычислению двух определенных интегралов; при вычис-

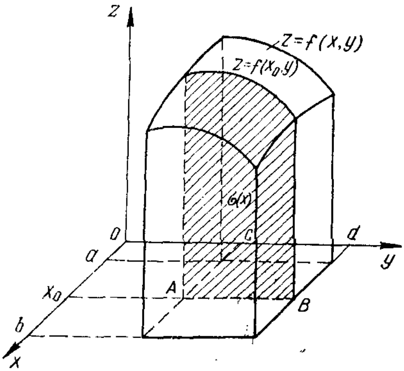

Рис. 19.3

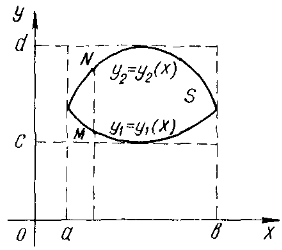

Рис. 19.4

лении «внутреннего интеграла» (записанного в квадратных скобках) x считается постоянным.

Замечание. Можно доказать, что формула (19.12) верна и при $f(x, y)<0$, а также в случае, когда функция $f(x, y)$ меняет знак в указанном прямоугольнике.

Правая часть формулы (19.12) называется *повторным* интегралом и обозначается так:

$$\int_a^b\left[\int_c^d f(x, y)\,dy\right]dx=\int_a^b dx\int_c^d f(x, y)\,dy. \qquad (19.13)$$

Аналогично можно показать, что

$$\iint_R f(x, y)\,dxdy=\int_c^d\left[\int_a^b f(x, y)\,dx\right]dy=$$

$$= \int_c^d dy \int_a^b f(x, y)\, dx. \qquad (19.14)$$

Из формул (19.12)—(19.14) получаем

$$\int_a^b dx \int_c^d f(x, y)\, dy = \int_c^d dy \int_a^b f(x, y)\, dx. \qquad (19.15)$$

Последнее равенство означает, что результат интегрирования не зависит от порядка интегрирования.

Чтобы рассмотреть более общий случай, введем понятие стандартной области. *Стандартной областью в направлении данной оси* называется такая область, для которой любая прямая, параллельная этой оси и имеющая с данной областью общие точки, пересекает границу области т о л ь к о в двух точках: M — «точке входа», N — «точке выхода» (рис. 19.4); другими словами, пересекает саму область и ее границу только по одному отрезку прямой.

Предположим, что ограниченная область S является стандартной в направлении оси Oy (рис. 19.4) и ограничена сверху графиком функции $y_2 = y_2(x)$, снизу — графиком функции $y_1 = y_1(x)$. Пусть $R\{a \leqslant x \leqslant b,\ c \leqslant y \leqslant d\}$ — минимальный прямоугольник, в котором заключена данная область S.

Пусть в области S определена и непрерывна функция $z = f(x, y)$. Введем новую функцию:

$$\varphi(x, y) = \begin{cases} f(x, y), & \text{если } (x, y) \in S, \\ 0, & \text{если } (x, y) \in R - S, \end{cases}$$

тогда в соответствии со свойством 3 двойного интеграла

$$\iint_S f(x, y)\, dxdy = \iint_R \varphi(x, y)\, dxdy. \qquad (19.16)$$

На основании формулы (19.12) получаем

$$\iint_R \varphi(x, y)\, dxdy = \int_a^b dx \int_c^d \varphi(x, y)\, dy.$$

Поскольку отрезок $[y_1(x),\ y_2(x)]$ целиком принадлежит области S и, значит, $\varphi(x, y) = f(x, y)$ при $y_1(x) \leqslant y \leqslant y_2(x)$, $\varphi(x, y) = 0$, если y лежит вне этого отрезка, то при фиксированном x

$$\int\limits_c^d \varphi(x, y)\,dy = \int\limits_c^{y_1(x)} \varphi(x, y)\,dy + \int\limits_{y_1(x)}^{y_2(x)} \varphi(x, y)\,dy +$$

$$+ \int\limits_{y_2(x)}^{d} \varphi(x, y)\,dy = \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy.$$

Следовательно,

$$\iint\limits_R \varphi(x, y)\,dxdy = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy. \tag{19.17}$$

Из формул (19.16) и (19.17) получаем

$$\iint\limits_S f(x, y)\,dxdy = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy. \tag{19.18}$$

Если область S является стандартной в направлении оси Ox и определяется неравенствами $c \leqslant y \leqslant d$, $x_1(y) \leqslant x \leqslant x_2(y)$, аналогично можно доказать, что

$$\iint\limits_S f(x, y)\,dxdy = \int\limits_c^d dy \int\limits_{x_1(y)}^{x_2(y)} f(x, y)\,dx. \tag{19.19}$$

Замечание. Для области S, стандартной в направлении осей Ox и Oy, будут выполнены равенства (19.18) и (19.19), поэтому

$$\int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} f(x, y)\,dy = \int\limits_c^d dy \int\limits_{x_1(y)}^{x_2(y)} f(x, y)\,dx. \tag{19.20}$$

По формуле (19.20) осуществляется изменение порядка интегрирования при вычислении соответствующего двойного интеграла.

§ 19.5. Замена переменных в двойном интеграле. Двойной интеграл в полярных координатах

В плоскости Oxy рассмотрим область S, ограниченную гладкой линией γ (рис. 19.5, *а*). Пусть

$$u=\varphi_1(x, y), \quad v=\psi_1(x, y) \tag{19.21}$$

— однозначные функции переменных x и y, непрерывные в области S и имеющие в ней непрерывные частные производные.

Предположим, что уравнения (19.21) однозначно разрешимы относительно x и y, т. е.

$$x = \varphi(u, v), \quad y = \psi(u, v), \tag{19.22}$$

где $\varphi(u, v)$, $\psi(u, v)$ — непрерывные функции в некоторой области S' плоскости Ouv и имеющие в ней непрерывные частные производные.

Каждой точке $M(x, y)$ области S формулы (19.21) ставят в соответствие единственную точку $M'(u, v)$ области S' (рис. 19.5, *б*). Обратно, каждой точке $M'(u, v) \in S'$

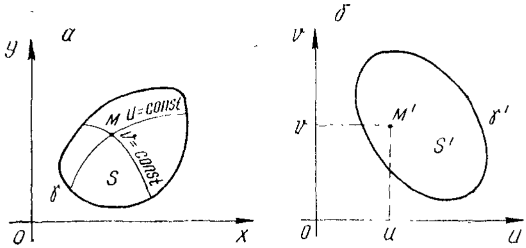

Рис. 19.5

формулы (19.22) ставят в соответствие единственную точку $M(x, y)$. Числа (u, v) называют *криволинейными координатами* точки M.

Следовательно, формулы (19.21) устанавливают взаимно-однозначное соответствие между точками областей S и S' или отображают область S на область S'. Область S' ограничена линией γ', в которую преобразуется линия γ.

Фиксируем значение $u = u_0$, тогда прямой $u = u_0$ в плоскости Ouv будет соответствовать в плоскости Oxy некоторая линия, параметрические уравнения которой $x = \varphi(u_0, v)$, $y = \psi(u_0, v)$ (получены из уравнений (19.22), роль параметра играет v).

Прямоугольнику $N_0N_1N_2N_3$ в плоскости Ouv, ограниченному прямыми $u = u_0$, $u = u_0 + \Delta u$, $v = v_0$, $v = v_0 + \Delta v$ (рис. 19.6, *а*), в плоскости Oxy соответствует криволинейный четырехугольник $M_0M_1M_2M_3$ (рис. 19.6, *б*), причем $M_0(x_0, y_0)$, $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$, где $x_0 = \varphi(u_0, v_0)$, $y_0 = \psi(u_0, v_0)$; $x_1 = \varphi(u_0 + \Delta u, v_0)$, $y_1 = \psi(u_0 + \Delta u, v_0)$; $x_2 = \varphi(u_0 + \Delta u, v_0 + \Delta v)$, $y_2 = \psi(u_0 + \Delta u, v_0 + \Delta v)$; $x_3 = \varphi(u_0, v_0 + \Delta v)$, $y_3 = \psi(u_0, v_0 + \Delta v)$.

С точностью до бесконечно малых высшего порядка площадь четырехугольника $M_0M_1M_2M_3$ равна площади параллелограмма, построенного на векторах $\overline{M_0M_1}$ и $\overline{M_0M_3}$. Площадь этого параллелограмма выражается формулой $\Delta S = |[\overline{M_0M_1}, \overline{M_0M_3}]|$.

Координаты векторов $\overline{M_0M_1}$, $\overline{M_0M_3}$ определяются по формулам (5.41)

$$\overline{M_0M_1} = \{\varphi(u_0+\Delta u, v_0) - \varphi(u_0, v_0);$$
$$\psi(u_0+\Delta u, v_0) - \psi(u_0, v_0)\},$$
$$\overline{M_0M_3} = \{\varphi(u_0, v_0+\Delta v) - \varphi(u_0, v_0);$$
$$\psi(u_0, v_0+\Delta v) - \psi(u_0, v_0)\}.$$

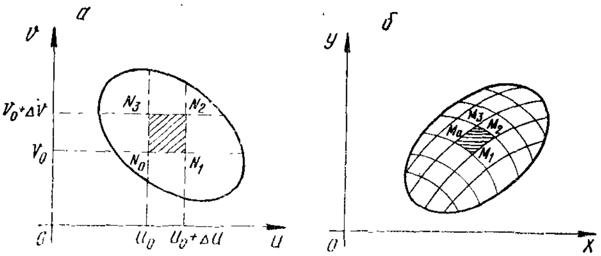

Рис. 19.6

Применяя теорему Лагранжа, получаем

$$\overline{M_0M_1} = \left\{\frac{\partial\varphi}{\partial u}\Delta u,\ \frac{\partial\psi}{\partial u}\Delta u\right\},\quad \overline{M_0M_3} = \left\{\frac{\partial\varphi}{\partial v}\Delta v,\ \frac{\partial\psi}{\partial v}\Delta v\right\}.$$

Следовательно,

$$\Delta S = |[\overline{M_0M_1}, \overline{M_0M_3}]| = \mathrm{mod}\begin{vmatrix} \dfrac{\partial\varphi}{\partial u} & \dfrac{\partial\psi}{\partial u} \\ \dfrac{\partial\varphi}{\partial v} & \dfrac{\partial\psi}{\partial v} \end{vmatrix}\Delta u\Delta v$$

или

$$\Delta S = |I|\Delta S',\quad \Delta S' = \Delta u\Delta v, \tag{19.23}$$

где

$$I = \begin{vmatrix} \dfrac{\partial\varphi}{\partial u} & \dfrac{\partial\psi}{\partial u} \\ \dfrac{\partial\varphi}{\partial v} & \dfrac{\partial\psi}{\partial v} \end{vmatrix} \quad \text{или} \quad I = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial x}{\partial v} & \dfrac{\partial y}{\partial v} \end{vmatrix}. \tag{19.24}$$

Определитель (19.24) называется *функциональным* *определителем* функций (19.22), или *якобианом преобразования* (19.22); предполагается, что он отличен от нуля.

Рассмотрим вопрос о замене переменных в двойном интеграле

$$\iint\limits_S f(x, y)\,dxdy.$$

Предположим, что функция $z=f(x, y)$ непрерывна в области S. Пусть в этой области заданы функции (19.22), удовлетворяющие указанным выше условиям, тогда

$$f(x, y)=f[\varphi(u, v), \psi(u, v)]=F(u, v).$$

Составим интегральную сумму для этой функции:

$$\sum_{k=1}^{n} f(x_k, y_k)\Delta S_k = \sum_{k=1}^{n} F(u_k, v_k)\Delta S_k.$$

Принимая во внимание равенство (19.23), получаем

$$\sum_{k=1}^{n} f(x_k, y_k)\Delta S_k \approx \sum_{k=1}^{n} F(u_k, v_k)\,|I|\,\Delta S_k'.$$

Переходя к пределу, находим искомую формулу

$$\iint\limits_S f(x, y)\,dxdy = \iint\limits_{S'} f[\varphi(u, v), \psi(u, v)]\,|I|\,dudv, \quad (19.25)$$

по которой осуществляется замена переменных в двойном интеграле.

В двойном интеграле перейдем к полярным координатам по формулам $x=\rho\cos\varphi$, $y=\rho\sin\varphi$.

Считая $u=\rho$, $v=\varphi$, находим якобиан данного преобразования:

$$I=\begin{vmatrix} \dfrac{\partial x}{\partial \rho} & \dfrac{\partial y}{\partial \rho} \\ \dfrac{\partial x}{\partial \varphi} & \dfrac{\partial y}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \cos\varphi & \sin\varphi \\ -\rho\sin\varphi & \rho\cos\varphi \end{vmatrix} = \rho.$$

В соответствии с формулой (19.25) имеем

$$\iint\limits_S f(x, y)\,dxdy = \iint\limits_{S'} f(\rho\cos\varphi, \rho\sin\varphi)\,\rho d\rho d\varphi. \quad (19.26)$$

§ 19.6. Приложения двойного интеграла

О некоторых приложениях двойного интеграла уже шла речь. Так, в § 19.2 приведены формулы (19.10) и (19.11), по которым соответственно вычисляются объем тела и масса пластинки:

$$V=\iint_S f(x,\ y)\,dxdy,\quad m=\iint_S p(x,\ y)\,dxdy.$$

С помощью двойного интеграла можно вычислять площадь плоской области. Действительно, положив в первой из этих формул $f(x,\ y)\equiv 1$, получим $V=1\cdot S=S$, т. е.

$$S=\iint_S dxdy,\quad S=\iint_S dS. \tag{19.27}$$

Центр тяжести. Пусть в точках $M_1(x_1, y_1), M_2(x_2, y_2), \ldots$ $\ldots, M_n(x_n,\ y_n)$ сосредоточены соответственно массы $m_1, m_2, \ldots, m_n$. Центр тяжести данной системы материальных точек, как известно (см. § 1.4), находится в точке, координаты которой выражаются формулами

$$x_c=\frac{\sum\limits_{k=1}^{n} x_k m_k}{\sum\limits_{k=1}^{n} m_k},\quad y_c=\frac{\sum\limits_{k=1}^{n} y_k m_k}{\sum\limits_{k=1}^{n} m_k}.$$

Найдем координаты центра тяжести пластинки, занимающей в плоскости Oxy некоторую область S. Пусть $p=p(x,\ y)$ — плотность этой пластинки в точке $M(x,\ y)$. Разобьем область на элементарные области ΔS_k $(k=1,$ $2, \ldots, n)$, площади которых обозначим теми же символами. Будем считать, что плотность в каждой области ΔS_k постоянна и равна $p_k=p(x_k,\ y_k)$ и вся ее масса $m_k=$ $=p_k\Delta S_k=p(x_k,\ y_k)\Delta S_k$ сосредоточена в точке $M_k(x_k,\ y_k)$. Координаты центра тяжести полученной системы из n материальных точек $M_k(x_k,\ y_k)$ определяются формулами

$$x=\frac{\sum\limits_{k=1}^{n} x_k p_k(x_k,\ y_k)\,\Delta S_k}{\sum\limits_{k=1}^{n} p(x_k,\ y_k)\,\Delta S_k},\quad y=\frac{\sum\limits_{k=1}^{n} y_k p_k(x_k,\ y_k)\,\Delta S_k}{\sum\limits_{k=1}^{n} p(x_k,\ y_k)\Delta S_k}.$$

Эти формулы приближенно выражают координаты центра тяжести пластинки.

Пусть λ — наибольший из диаметров областей ΔS_k ($k=1, 2, \ldots, n$). Перейдем к пределам этих сумм при $\lambda \to 0$. Эти пределы равны соответствующим двойным интегралам.

Следовательно, координаты центра тяжести рассматриваемой пластинки определяются формулами

$$x_c = \frac{\iint\limits_S xp(x, y)\,dxdy}{\iint\limits_S p(x, y)\,dxdy}, \quad y_c = \frac{\iint\limits_S yp(x, y)\,dxdy}{\iint\limits_S p(x, y)\,dxdy} \tag{19.28}$$

или

$$x_c = \frac{1}{m}\iint\limits_S xp(x, y)\,dxdy, \quad y_c = \frac{1}{m}\iint\limits_S yp(x, y)\,dxdy, \tag{19.29}$$

где m — масса пластинки.

Двойные интегралы, стоящие в правых частях последних формул, т. е.

$$M_y = \iint\limits_S xp(x, y)\,dxdy, \quad M_x = \iint\limits_S yp(x, y)\,dxdy, \tag{19.30}$$

называются *статическими моментами* пластинки S относительно осей Oy и Ox соответственно.

Если пластинка однородна, т. е. $p(x, y) = \text{const}$, то формулы (19.28) принимают вид

$$x_c = \frac{1}{S}\iint\limits_S xdxdy, \quad y_c = \frac{1}{S}\iint\limits_S ydxdy,$$

где S — площадь области S.

Момент инерции. *Моментом инерции* I_0 *материальной точки* M с массой m относительно точки O называется произведение массы на квадрат расстояния $r=\rho(O, M)$ между данными точками, т. е. $I_0 = mr^2$.

Моментом инерции системы материальных точек M_1, $M_2, \ldots, M_n$ с массами $m_1, m_2, \ldots, m_n$ относительно точки O называется сумма моментов инерции всех этих точек:

$$I_0 = \sum_{k=1}^{n} m_k r_k^2, \quad r_k = \rho(O, M_k).$$

Аналогично определяется момент инерции системы указанных материальных точек относительно оси Ou: $I_u = \sum_{k=1}^{n} m_k r_k^2$, где r_k — расстояние точки M_k до оси Ou.

Определим момент инерции пластинки S с плотностью $p=p(x, y)$ относительно точки O — начала прямоугольной декартовой системы координат.

Разбивая эту область на n элементарных областей ΔS_k $(k=1, 2, \ldots, n)$ и считая, что масса $m_k=p(x_k, y_k)\Delta S_k$ области ΔS_k сосредоточена в точке $M_k(x_k, y_k)\in\Delta S_k$, получаем систему n материальных точек. Поскольку

$$r_k = \rho(O, M_k) = \sqrt{x_k^2 + y_k^2},$$

то момент инерции этой системы относительно точки O выразится так:

$$I_{O_c} = \sum_{k=1}^{n} r_k^2 m_k = \sum_{k=1}^{n} (x_k^2 + y_k^2)\, p(x_k, y_k)\, \Delta S_k. \quad (19.31)$$

Моментом инерции пластинки относительно начала координат называется предел суммы (19.31) при $\lambda \to 0$ (λ — наибольший из диаметров областей ΔS_k, $k=1, 2, \ldots, n$):

$$I_0 = \lim_{\lambda\to 0} \sum_{k=1}^{n} (x_k^2 + y_k^2)\, p(x_k, y_k)\, \Delta S_k,$$

или

$$I_0 = \iint_S (x^2 + y^2)\, p(x, y)\, dxdy. \quad (19.32)$$

Моменты инерции этой пластинки относительно координатных осей Ox и Oy определяются формулами

$$I_x = \iint_S y^2 p(x, y)\, dxdy, \quad I_y = \iint_S x^2 p(x, y)\, dxdy. \quad (19.33)$$

Из формул (19.32) и (19.33) следует, что $I_0 = I_x + I_y$.

§ 19.7. Тройной интеграл

По аналогии с двойным интегралом вводится понятие тройного интеграла. Рассмотрим ограниченную замкнутую пространственную область V и определенную в ней непрерывную функцию $u=f(x, y, z)$. Область V разобьем на n элементарных пространственных областей ΔV_1, $\Delta V_2, \ldots, \Delta V_n$. Предполагается, что область V и элементарные области $\Delta V_1, \Delta V_2, \ldots, \Delta V_n$ имеют объемы, которые будем обозначать соответственно теми же символами. В каждой элементарной области ΔV_k $(k=1, 2, \ldots, n)$ выберем произвольную точку $M_k(x_k, y_k, z_k)$, значение функции в этой точке $f(x_k, y_k, z_k)$ умножим на объем элементарной области ΔV_k и составим сумму всех таких произведений:

$$\sigma_n = \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta V_k, \qquad (19.34)$$

которая называется *интегральной суммой* данной функции по данному объему.

Обозначим через d_k диаметр области ΔV_k. Пусть λ — наибольший из этих диаметров. В равенстве (19.34) перейдем к пределу при $\lambda \to 0$. Если предел интегральной суммы существует, то он и называется тройным интегралом от функции $f(x, y, z)$ по пространственной области V.

Итак, по определению

$$\iiint\limits_V f(x, y, z)\,dV = \lim_{\lambda\to 0} \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta V_k \qquad (19.35)$$

или

$$\iiint\limits_V f(x, y, z)\,dx dy dz = \lim_{\lambda\to 0} \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta V_k. \qquad (19.36)$$

Отметим без доказательства, что если функция $f(x, y, z)$ непрерывна в рассматриваемой замкнутой области V, то предел в правой части формулы (19.35) существует и не зависит от способа разбиения области V на элементарные и выбора точки M_k в элементарной области ΔV_k.

Предположим, что в области V распределено вещество, объемная плотность которого задана непрерывной функцией $f(x, y, z) \geqslant 0$, тогда произведение $f(x_k, y_k, z_k)\Delta V_k$ выражает приближенную массу элементарной области ΔV_k, интегральная сумма (19.34) — приближенную массу всей области V, а тройной интеграл (19.35) — точное значение этой массы, т. е.

$$m = \iiint\limits_V f(x, y, z)\, dV. \qquad (19.37)$$

Формулой (19.37) выражается *механический смысл тройного интеграла:* тройной интеграл представляет массу, заполняющую область интегрирования V.

Если в формуле (19.37) $f(x, y, z) \equiv 1$, то $m = V \cdot 1 = V$ и эта формула принимает вид

$$V = \iiint\limits_V dV, \quad V = \iiint\limits_V dxdydz. \qquad (19.38)$$

Тройной интеграл обладает свойствами, аналогичными свойствам двойного интеграла (см. § 19.3).

Переходим к вопросу о вычислении тройного интеграла в прямоугольных декартовых координатах. Предположим, что область V является стандартной в направлении оси Oz, т. е. удовлетворяющей следующим условиям:

1) всякая прямая, параллельная этой оси и имеющая с данной областью общие точки, пересекает границу области только в двух точках;

2) проекция S области V на плоскость Oxy представляет собой стандартную область в направлении оси Ox или оси Oy.

Пусть стандартная область V ограничена сверху поверхностью $z = z_2(x, y)$, снизу — поверхностью $z = z_1(x, y)$, тогда можно показать, что

$$\iiint\limits_V f(x, y, z) dxdydz = \iint\limits_S \left[\int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\, dz \right] dxdy. \qquad (19.39)$$

Если область S является стандартной в направлении оси Oy и определяется неравенствами $a \leqslant x \leqslant b$, $y_1(x) \leqslant y \leqslant y_2(x)$, то

$$\left.\begin{aligned} &\iint\limits_S \left[\int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z) dz \right] dxdy = \\ &= \int\limits_a^b \left\{ \int\limits_{y_1(x)}^{y_2(x)} \left[\int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z) dz \right] dy \right\} dx = \\ &= \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} dy \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z) dz. \end{aligned}\right\} \qquad (19.40)$$

Из формул (19.39) и (19.40) следует, что

$$\iiint\limits_V f(x, y, z)\,dxdydz = \int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} dy \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)dz. \quad (19.41)$$

Отметим частный случай этой формулы: если V — прямоугольный параллелепипед, определяемый неравенствами $a \leqslant x \leqslant A$, $b \leqslant y \leqslant B$, $c \leqslant z \leqslant C$, то

$$\iiint\limits_V f(x, y, z)dxdydz = \int\limits_a^A dx \int\limits_b^B dy \int\limits_c^C f(x, y, z)dz. \quad (19.42)$$

Замечание 1. Если область S является стандартной в направлении оси Ox и определяется неравенствами $x_1(y) \leqslant x \leqslant x_2(y)$, $c \leqslant y \leqslant d$, то

$$\iint\limits_S \left[\int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz\right] dxdy = \int\limits_c^d \left\{ \int\limits_{x_1(y)}^{x_2(y)} \left[\int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)\,dz\right] dx\right\} dy.$$

Следовательно,

$$\iiint\limits_V f(x, y, z)\,dxdydz = \int\limits_c^d dy \int\limits_{x_1(y)}^{x_2(y)} dx \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)dz. \quad (19.43)$$

Если область S является стандартной в направлении оси Ox и оси Oy, то из формул (19.41) и (19.43) получаем

$$\int\limits_a^b dx \int\limits_{y_1(x)}^{y_2(x)} dy \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)dz = \int\limits_c^d dy \int\limits_{x_1(y)}^{x_2(y)} dx \int\limits_{z_1(x, y)}^{z_2(x, y)} f(x, y, z)dz. \quad (19.44)$$

Замечание 2. Если область V является стандартной в направлении каждой координатной оси и ее проекции на координатные плоскости являются стандартными в направлении каждой соответствующей оси, то пределы интегрирования в трехкратном интеграле можно расставить шестью различными способами (два из них указаны формулой (19.44)).

§ 19.8. Замена переменных в тройном интеграле. Тройной интеграл в цилиндрических и сферических координатах

Пусть в области V заданы дифференцируемые функции

$$u = u(x, y, z), \; v = v(x, y, z), \; w = w(x, y, z), \quad (19.45)$$

где (x, y, z) — прямоугольные декартовы координаты точки $M(x, y, z) \in V$.

Предположим, что уравнения (19.45) однозначно разрешимы относительно x, y, z:

$$x=x(u, v, w),\ y=y(u, v, w),\ z=z(u, v, w), \tag{19.46}$$

причем $x(u, v, w)$, $y(u, v, w)$, $z(u, v, w)$ — дифференцируемые функции своих аргументов.

Каждой точке $M(x, y, z)$ области V ставится в соответствие точка $M'(u, v, w)$ области V'; (u, v, w) называются *криволинейными координатами* точки M.

Функции (19.45) осуществляют взаимно-однозначное отображение области V на область V'.

По аналогии с формулой (19.25) получаем

$$\iiint\limits_V f(x, y, z)\, dxdydz =$$

$$= \iiint\limits_{V'} f[x(u, v, w), y(u, v, w), z(u, v, w)]\,|I|\,dudvdw, \tag{19.47}$$

где

$$I = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{vmatrix} \tag{19.48}$$

Определитель (19.48) называется *функциональным определителем*, или *якобианом* функций (19.46). Предполагается, что в рассматриваемой области $I\neq 0$. По формуле (19.47) осуществляется замена переменных в тройном интеграле.

Выразим тройной интеграл в цилиндрических координатах (см. § 1.8). Принимая во внимание формулы $x=\rho\cos\varphi$, $y=\rho\sin\varphi$, $z=z$ и полагая $u=\rho$, $v=\varphi$, $w=z$, находим якобиан

$$I = \begin{vmatrix} \cos\varphi & -\rho\sin\varphi & 0 \\ \sin\varphi & \rho\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho.$$

Формула (19.47) запишется так:

$$\iiint\limits_V f(x, y, z)dxdydz = \iiint\limits_{V'} f(\rho\cos\varphi, \rho\sin\varphi, z)\,\rho d\rho d\varphi dz. \tag{19.49}$$

Перейдем к сферическим координатам (см. § 1.8). Поскольку $x=r\sin\theta\cos\varphi$, $y=r\sin\theta\sin\varphi$, $z=r\cos\theta$ $(r\geqslant 0,\ 0\leqslant\theta\leqslant\pi,\ 0\leqslant\varphi<2\pi)$, то, считая $u=r$, $v=\theta$, $w=\varphi$, получаем

$$I=\begin{vmatrix}\sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi\\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi\\ \cos\theta & -r\sin\theta & 0\end{vmatrix}=r^2\sin\theta.$$

Следовательно,

$$\iiint\limits_V f(x,\ y,\ z)\,dxdydz=$$

$$=\iiint\limits_{V'} f(r\sin\theta\cos\varphi,\ r\sin\theta\sin\varphi,\ r\cos\theta)\,r^2\sin\theta\,drd\theta d\varphi. \tag{19.50}$$

§ 19.9. Приложения тройного интеграла

В § 19.7 шла речь о некоторых приложениях тройного интеграла, в частности, были получены формулы (19.37), (19.38) для вычисления массы материальной области по заданной объемной плотности $p=p(x,\ y,\ z)$ и объема тела:

$$m=\iiint\limits_V p\,(x,\ y,\ z)\,dxdydz,\quad V=\iiint\limits_V dxdydz.$$

Тройной интеграл применяется также при вычислении *координат центра тяжести* $C(x_0,\ y_0,\ z_0)$ материальной области V, в которой распределено вещество с объемной плотностью $p=p(x,\ y,\ z)$. По аналогии с формулами (19.29) получаем

$$\left.\begin{aligned} x_0&=\frac{1}{m}\iiint\limits_V xp\,(x,\ y,\ z)\,dxdydz;\\ y_0&=\frac{1}{m}\iiint\limits_V yp\,(x,\ y,\ z)\,dxdydz;\\ z_0&=\frac{1}{m}\iiint\limits_V zp\,(x,\ y,\ z)\,dxdydz,\end{aligned}\right\} \tag{19.51}$$

где m — масса соответствующей области, определяемая первой из двух предыдущих формул.

Моменты инерции указанной материальной области относительно координатных осей Ox, Oy, Oz и начала координат определяются соответственно формулами

$$\left.\begin{aligned} I_x &= \iiint\limits_V (y^2+z^2)\,p(x,\ y,\ z)\,dxdydz; \\ I_y &= \iiint\limits_V (x^2+z^2)\,p(x,\ y,\ z)\,dxdydz; \\ I_z &= \iiint\limits_V (x^2+y^2)\,p(x,\ y,\ z)\,dxdydz; \\ I_0 &= \iiint\limits_V (x^2+y^2+z^2)\,p(x,\ y,\ z)dxdydz. \end{aligned}\right\} \quad (19.52)$$

§ 19.10. Понятие о многомерных интегралах

Понятие интеграла можно ввести и для функции n переменных, определенной в некоторой области G n-мерного пространства.

Одной из простейших областей этого пространства является n-мерный параллелепипед, т. е. множество точек $M(x_1, x_2, \ldots, x_n)$, для которых

$$a \leqslant x_1 \leqslant A,\ b \leqslant x_2 \leqslant B,\ \ldots,\ d \leqslant x_n \leqslant D,$$

где a, A, b, B, ..., d, D — постоянные.

По аналогии с формулой, выражающей объем параллелепипеда через координаты векторов, на которых он построен (см. § 5.10), объем n-мерного параллелепипеда также определяется как модуль определителя n-го порядка, составленного из координат соответствующих векторов.

Далее можно определить объем для многогранных n-мерных тел, а также для некоторого класса областей в n-мерном пространстве.

Интеграл от функции $u=f(x_1,\ x_2,\ \ldots,\ x_n)$ по некоторой n-мерной области G (n-мерный интеграл) вводится как предел соответствующей интегральной суммы и обозначается символом

$$\iint\limits_G \ldots \int f(x_1,\ x_2, \ldots,\ x_n)\,dx_1dx_2\ldots dx_n.$$

Если подынтегральная функция и область интегрирования удовлетворяют определенным условиям, то вычисление такого интеграла сводится к последовательному нахождению интегралов по формуле

$$\iint\limits_G \ldots \int f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n =$$

$$= \int\limits_a^A dx_1 \int\limits_{x_2'(x_1)}^{x_2^2(x_1)} dx_2 \ldots \int\limits_{x_n'(x_1, x_2, \ldots, x_{n-1})}^{x_n^2(x_1, x_2, \ldots, x_{n-1})} f(x_1, x_2, \ldots, x_n)\, dx_n.$$

Интеграл, стоящий в правой части этой формулы, называется n-кратным.

Формула замены переменных в n-мерном интеграле аналогична формулам (19.25) и (19.47).

Глава 20

КРИВОЛИНЕЙНЫЕ ИНТЕГРАЛЫ. ИНТЕГРАЛЫ ПО ПОВЕРХНОСТИ

При рассмотрении многих вопросов встречается необходимость введения понятия интеграла от функции, заданной на некоторой линии. Такие интегралы называются криволинейными, бывают двух типов и преобразуются один в другой.

Интегралы от функций, заданных на некоторой поверхности, называются *интегралами по поверхности* (или *поверхностными* интегралами). Они также бывают двух типов и связаны между собой соответствующей формулой.

§ 20.1. Задачи, приводящие к понятию криволинейного интеграла

Задача о вычислении массы дуги кривой. Дана дуга пространственной линии AB, по которой распределено вещество с плотностью $p=p(x, y, z)$. Требуется вычислить массу материальной дуги. Чтобы решить задачу, разобьем дугу на n элементарных дуг $A_{k-1}A_k$ ($k=1, 2, \ldots$ $\ldots, n$; $A_0=A$, $A_n=B$). Предположим, что на каждой элементарной дуге $A_{k-1}A_k$ средняя плотность вещества равна значению функции $p(x, y, z)$ в некоторой точке $M_k(x_k, y_k, z_k)$ этой дуги, т. е. $p_k=p(x_k, y_k, z_k)$. Умножив p_k на Δl_k — длину дуги $A_{k-1}A_k$, получим приближенное значение массы $m_k=p_k(x_k, y_k, z_k)\Delta l_k$ данной дуги. Сумма

$$\sum_{k=1}^{n} m_k = \sum_{k=1}^{n} p(x_k,\ y_k,\ z_k)\,\Delta l_k \qquad (20.1)$$

выражает приближенно массу всей дуги AB. Точное значение массы получим при переходе к пределу

$$m = \lim_{n\to\infty} \sum_{k=1}^{n} m_k = \lim_{\lambda\to 0} \sum_{k=1}^{n} p(x_k,\ y_k,\ z_k)\,\Delta l_k, \qquad (20.2)$$

где λ — наибольшее из чисел Δl_k.

Замечание. Сумма (20.1) называется *интегральной суммой* функции $p(x, y, z)$ по длине дуги. Предел интегральной суммы (20.1) называется *криволинейным интегралом первого рода* (от функции $p(x, y, z)$ по длине дуги).

Задача о вычислении работы переменной силы. Если сила $\mathbf{F}$ постоянна (по величине и по направлению), а путь $\overline{AB}=\mathbf{s}$ прямолинеен, то работа этой силы на заданном пути равна скалярному произведению $\mathbf{Fs}=|\mathbf{F}|\,|\mathbf{s}|\cos\varphi$, где φ — угол между векторами $\mathbf{F}$ и $\mathbf{s}$. Если векторы заданы своими координатами, скалярное произведение выражается формулой (5.59).

Пусть переменная сила

$$\mathbf{F}=P(x, y, z)\mathbf{i}+Q(x, y, z)\mathbf{j}+R(x, y, z)\mathbf{k} \qquad (20.3)$$

действует вдоль дуги AB пространственной линии. Требуется вычислить работу, производимую этой силой при перемещении точки приложения вдоль данной линии из A в B.

Дугу AB разобьем на n элементарных дуг $A_{k-1}A_k$ $(k=1, 2, \ldots, n;\ A_0=A,\ A_n=B)$. Будем считать, что на элементарной дуге $A_{k-1}A_k$ сила $\mathbf{F}$ постоянна и равна $\mathbf{F}_k=\mathbf{F}(M_k)$, где $M_k \in A_{k-1}A_k$, $M_k(\xi_k, \eta_k, \zeta_k)$, $A_k(x_k, y_k, z_k)$.

Путь $A_{k-1}A_k$ будем считать прямолинейным, тогда

$$\overline{A_{k-1}A_k}=\{\Delta x_k, \Delta y_k, \Delta z_k\},$$

где

$$\Delta x_k=x_k-x_{k-1},\ \Delta y_k=y_k-y_{k-1},\ \Delta z_k=z_k-z_{k-1}.$$

Следовательно, работа на этом элементарном пути выразится формулой

$$\mathbf{F}_k\overline{A_{k-1}A_k}=P(\xi_k, \eta_k, \zeta_k)\Delta x_k+Q(\xi_k, \eta_k, \zeta_k)\Delta y_k+ \\ +R(\xi_k, \eta_k, \zeta_k)\Delta z_k.$$

Вся работа силы (20.3) на криволинейном пути AB приближенно выражается формулой

$$W_n = \sum_{k=1}^{n} \mathbf{F}_k \overline{A_{k-1}A_k} = \sum_{k=1}^{n} (P_k \Delta x_k + Q_k \Delta y_k + R_k \Delta z_k), \qquad (20.4)$$

где

$$P_k = P(\xi_k, \eta_k, \zeta_k),\ Q_k = Q(\xi_k, \eta_k, \zeta_k),\ R_k = R(\xi_k, \eta_k, \zeta_k).$$

Переходя к пределу при $\lambda \to 0$, где λ — длина наибольшей из элементарных дуг $A_{k-1}A_k$, получаем точное значение работы

$$W = \lim_{\lambda \to 0} \sum_{k=1}^{n} [P(\xi_k, \eta_k, \zeta_k)\Delta x_k + Q(\xi_k, \eta_k, \zeta_k)\Delta y_k + R(\xi_k, \eta_k, \zeta_k)\Delta z_k]. \qquad (20.5)$$

Замечание. Сумма (20.4) называется интегральной суммой по координатам для вектора-функции (20.3), а ее предел при $\lambda \to 0$ — *криволинейным интегралом второго рода* (или криволинейным интегралом по координатам).

§ 20.2. Криволинейный интеграл первого рода

Рассмотрим функцию $u=f(x, y, z)$, определенную на дуге AB пространственной кусочно-гладкой линии (рис. 20.1). Разобьем эту дугу на n элементарных дуг $A_{k-1}A_k$ ($k=1, 2, \ldots, n$; $A_0=A$, $A_n=B$). Длину элементарной дуги $A_{k-1}A_k$ обозначим через Δl_k. На каждой элементарной дуге $A_{k-1}A_k$ выберем произвольно точку $M_k(x_k, y_k, z_k)$, значение функции в этой точке $f(x_k, y_k, z_k)$ умножим на длину дуги Δl_k и составим сумму всех таких произведений:

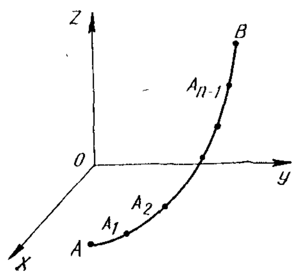

Рис. 20.1

$$T_n = \sum_{k=1}^{n} f(x_k, y_k, z_k)\Delta l_k. \qquad (20.6)$$

Сумма (20.6) называется *интегральной суммой* данной функции *по длине дуги*. Обозначим через λ длину наибольшей из элементарных дуг и перейдем к пределу при $\lambda \to 0$ (тогда $n \to \infty$). Если существует предел интегральной суммы (20.6), то он и называется криволинейным интегралом первого рода (или интегралом от функции по длине дуги). Следовательно, по определению

$$\int\limits_{AB} f(x,\ y,\ z)\,dl = \lim_{\lambda \to 0} \sum_{k=1}^{n} f(x_k,\ y_k,\ z_k)\,\Delta l_k. \qquad (20.7)$$

Отметим без доказательства, что предел интегральной суммы (20.6) существует, если функция $u=f(x,\ y,\ z)$ непрерывна.

Криволинейный интеграл первого рода обладает следующими свойствами:

1) не зависит от направления пути интегрирования:

$$\int\limits_{AB} f(x,\ y,\ z)\,dl = \int\limits_{BA} f(x,\ y,\ z)\,dl;$$

2) если функции $f(x,\ y,\ z)$ и $\varphi(x,\ y,\ z)$ интегрируемы на дуге AB, то на ней интегрируемы их сумма и разность, причем

$$\int\limits_{AB} [f(x,\ y,\ z) \pm \varphi(x,\ y,\ z)]dl = \int\limits_{AB} f(x,\ y,\ z)dl \pm \int\limits_{AB} \varphi(x,\ y,\ z)dl;$$

3) постоянный множитель можно выносить за знак криволинейного интеграла:

$$\int\limits_{AB} cf(x,\ y,\ z)\,dl = c\int\limits_{AB} f(x,\ y,\ z)\,dl,\ \ c = \text{const};$$

4) если дуга AB состоит из дуг AC и CB, а функция $f(x,\ y,\ z)$ интегрируема на AB, то она интегрируема на AC и CB, причем

$$\int\limits_{AB} f(x,\ y,\ z)dl = \int\limits_{AC} f(x,\ y,\ z)dl + \int\limits_{CB} f(x,\ y,\ z)\,dl.$$

Доказательства этих свойств приведены, например, в [11, 14].

Вычисление криволинейного интеграла первого рода сводится к вычислению определенного интеграла.

Если кривая задана параметрическими уравнениями

$$x=x(t),\ y=y(t),\ z=z(t),\ \alpha \leqslant t \leqslant \beta, \qquad (20.8)$$
$$A=(x,\ y,\ z)\,|_{t=\alpha},\ B=(x,\ y,\ z)\,|_{t=\beta}$$

и $x(t)$, $y(t)$, $z(t)$ — непрерывно дифференцируемые функции, то можно доказать, что

$$\int\limits_{\breve{AB}} f(x, y, z)\,dl = \int\limits_{\alpha}^{\beta} f[x(t), y(t), z(t)]\sqrt{x'^2+y'^2+z'^2}\,dt. \tag{20.9}$$

В частности, если дуга AB целиком лежит в плоскости Oxy $(z=0)$, последняя формула принимает вид

$$\int\limits_{\breve{AB}} f(x, y)\,dl = \int\limits_{\alpha}^{\beta} f[x(t), y(t)]\sqrt{x'^2+y'^2}\,dt. \tag{20.10}$$

Если дуга AB плоской линии задана уравнением $y=y(x)$, где $a \leqslant x \leqslant b$, $y(x)$ — непрерывно дифференцируемая функция, то

$$\int\limits_{\breve{AB}} f(x, y)\,dl = \int\limits_{a}^{b} f(x, y(x))\sqrt{1+y'^2}\,dx. \tag{20.11}$$

Если дуга AB плоской линии задана уравнением в полярных координатах $\rho=\rho(\varphi)$, где $\alpha \leqslant \varphi \leqslant \beta$, $\rho(\varphi)$ — непрерывно дифференцируемая функция, то

$$\int\limits_{\breve{AB}} f(x, y)\,dl = \int\limits_{\alpha}^{\beta} f(\rho\cos\varphi, \rho\sin\varphi)\sqrt{\rho^2+\rho'^2}\,d\varphi. \tag{20.12}$$

§ 20.3. Криволинейный интеграл второго рода

Рассмотрим дугу пространственной линии с началом в точке A и концом в точке B (см. рис. 20.1). Пусть на дуге AB определена и непрерывна вектор-функция

И

$$\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}. \tag{20.13}$$

Разобьем дугу AB на n элементарных дуг $A_{k-1}A_k$ $(k=1, 2, \ldots, n;\ A_0=A,\ A_n=B)$, на дуге $A_{k-1}A_k$ произвольно выберем точку $M_k(\xi_k, \eta_k, \zeta_k)$. Обозначим декартовы координаты точки A_k через x_k, y_k, z_k, т. е. $A_k(x_k, y_k, z_k)$, тогда проекции дуги $A_{k-1}A_k$ на координатные оси выразятся формулами

$$\Delta x_k = x_k - x_{k-1},\quad \Delta y_k = y_k - y_{k-1},\quad \Delta z_k = z_k - z_{k-1}.$$

Значения функций $P(x, y, z)$, $Q(x, y, z)$, $R(x, y, z)$ в точке $M_k(\xi_k, \eta_k, \zeta_k)$ умножим соответственно на Δx_k, Δy_k, Δz_k и сложим три произведения

$$\left.\begin{aligned} P(M_k)\,\Delta x_k + Q(M_k)\,\Delta y_k + R(M_k)\,\Delta z_k = \\ = P(\xi_k, \eta_k, \zeta_k)\,\Delta x_k + Q(\xi_k, \eta_k, \zeta_k)\,\Delta y_k + \\ + R(\xi_k, \eta_k, \zeta_k)\,\Delta z_k. \end{aligned}\right\}$$

Составим сумму всех таких выражений:

$$T_n = \sum_{k=1}^{n} [P(M_k)\,\Delta x_k + Q(M_k)\,\Delta y_k + R(M_k)\Delta z_k], \quad (20.14)$$

называемую *интегральной суммой по координатам* для вектор-функции (20.13).

Замечание. Отличие интегральной суммы (20.14) от интегральной суммы (20.6) состоит в том, что значение функции в фиксированной точке элементарной дуги умножается не на длину этой дуги, а на ее проекцию на соответствующую координатную ось.

Обозначим через λ длину наибольшей из дуг $A_{k-1}A_k$ и перейдем к пределу при $\lambda \to 0$. Если интегральная сумма (20.14) имеет предел при $\lambda \to 0$, то он и называется криволинейным интегралом второго рода от вектор-функции (20.13) и обозначается

$$\int\limits_{\breve{AB}} P(x, y, z)\,dx + Q(x, y, z)\,dy + R(x, y, z)\,dz$$

или

$$\int\limits_{\breve{AB}} P(M)dx + Q(M)dy + R(M)dz.$$

Если не указывать аргументов рассматриваемых функций, то по определению

$$\int\limits_{\breve{AB}} Pdx + Qdy + Rdz =$$

$$= \lim_{\lambda \to 0} \sum_{k=1}^{n} [P(M_k)\,\Delta x_k + Q(M_k)\,\Delta y_k + R(M_k)\,\Delta z_k]. \quad (20.15)$$

Очевидно,

$$\int\limits_{\breve{AB}} Pdx + Qdy + Rdz = -\int\limits_{\breve{BA}} Pdx + Qdy + Rdz, \quad (20.16)$$

т. е. при перемене направления интегрирования криволинейный интеграл второго рода меняет лишь знак, поскольку в этом случае меняются знаки проекций элементарной дуги на координатные оси.

Остальные свойства криволинейного интеграла второго рода аналогичны свойствам криволинейного интеграла первого рода.

Вычисление криволинейного интеграла второго рода также сводится к вычислению определенного интеграла. Если дуга AB задана параметрическими уравнениями (20.8), то

$$\left.\begin{aligned} &\int\limits_{AB} P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz = \\ &\qquad = \int\limits_{\alpha}^{\beta} \{P[x(t), y(t), z(t)]\,x'(t) + \\ &+Q[x(t), y(t), z(t)]y'(t) + R[x(t), y(t), z(t)]z'(t)\}\,dt. \end{aligned}\right\} \quad (20.17)$$

Если эта дуга лежит в плоскости Oxy $(z=0)$, последняя формула принимает вид

$$\int\limits_{AB} P(x, y)\,dx + Q(x, y)\,dy =$$

$$= \int\limits_{\alpha}^{\beta} \{P[x(t), y(t)]\,x'(t) + Q[x(t), y(t)]\,y'(t)\}\,dt. \quad (20.18)$$

Если дуга AB плоской кривой задана уравнением $y=y(x)$ $(a \leqslant x \leqslant b)$, где $y(x)$ — непрерывно дифференцируемая функция, то

$$\int\limits_{AB} P(x, y)\,dx + Q(x, y)\,dy =$$

$$= \int\limits_{a}^{b} [P(x, y(x)) + Q(x, y(x))\,y'(x)]\,dx. \quad (20.19)$$

§ 20.4. Связь между криволинейными интегралами первого и второго рода

Рассмотрим направленную дугу пространственной линии с началом в точке A и концом в точке B. Касательную в любой точке M дуги AB будем также считать направленной прямой (рис. 20.2). Углы, образуемые касательной с координатными осями Ox, Oy, Oz, обозначим соответственно через α, β, γ.

Вектор $\overline{dl} = \{dx, dy, dz\}$, где dl — дифференциал длины дуги, направлен по касательной, поэтому $dx = \cos\alpha\, dl$, $dy = \cos\beta\, dl$, $dz = \cos\gamma\, dl$.

Следовательно,

$$\int\limits_{\overset{\smile}{AB}} P dx + Q dy + R dz = \int\limits_{\overset{\smile}{AB}} (P\cos\alpha + Q\cos\beta + R\cos\gamma) dl. \quad (20.20)$$

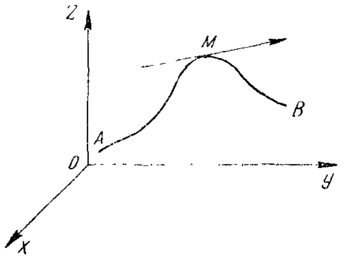

Рис. 20.2

Эта формула и выражает связь между криволинейными интегралами первого и второго рода.

Если дуга AB лежит в плоскости Oxy, то $z=0$ и формула (20.20) принимает вид

$$\int\limits_{\overset{\smile}{AB}} P dx + Q dy = \int\limits_{\overset{\smile}{AB}} (P\cos\alpha + Q\sin\alpha)\, dl,$$

где α — угол между касательной и осью Ox.

§ 20.5. Приложения криволинейных интегралов

Масса материальной дуги. Если $p = p(x, y, z)$ — плотность вещества, распределенного по дуге $\overset{\smile}{AB}$, то из равенств (20.2) и (20.7) следует, что масса m этой дуги выражается формулой

$$m = \int\limits_{\overset{\smile}{AB}} p(x, y, z)\, dl. \quad (20.21)$$

Длина дуги линии. Если $p(x, y, z) \equiv 1$, тогда масса дуги AB $m = 1 \cdot l$, $m = l$, поэтому из формулы (20.21) следует, что

$$l = \int\limits_{\overset{\smile}{AB}} dl. \quad (20.22)$$

Центр тяжести материальной дуги. Если $p = p(x, y, z)$ — плотность вещества, распределенного по дуге $\overset{\smile}{AB}$, то прямоугольные декартовы координаты ее центра тяжести $C_0(x_0, y_0, z_0)$ выражаются формулами

$$x_0 = \frac{1}{m}\int\limits_{\overset{\smile}{AB}} xp(x, y, z) dl, \quad y_0 = \frac{1}{m}\int\limits_{\overset{\smile}{AB}} yp(x, y, z)\, dl,$$

$$z_0 = \frac{1}{m}\int\limits_{\overset{\smile}{AB}} zp(x, y, z)\, dl, \quad (20.23)$$

где m — масса дуги. Формулы (20.23) получаются аналогично формулам (19.29).

Момент инерции материальной дуги. Моменты инерции материальной дуги AB, по которой распределено вещество с плотностью $p=p(x, y, z)$, относительно координатных осей и начала координат выражаются соответственно формулами

$$\left.\begin{aligned} I_x &= \int\limits_{AB} (y^2+z^2)\, p(x, y, z)\, dl; \\ I_y &= \int\limits_{AB} (x^2+z^2)\, p(x, y, z)\, dl; \\ I_z &= \int\limits_{AB} (x^2+y^2)\, p(x, y, z)\, dl; \\ I_0 &= \int\limits_{AB} (x^2+y^2+z^2)\, p(x, y, z)\, dl. \end{aligned}\right\} \qquad (20.24)$$

Работа переменной силы. Если

$$\mathbf{F}=P(x, y, z)\,\mathbf{i}+Q(x, y, z)\,\mathbf{j}+R(x, y, z)\,\mathbf{k}$$

— *переменная сила*, совершающая работу W вдоль криволинейного пути AB, и функции $P=P(x, y, z)$, $Q=Q(x, y, z)$, $R=R(x, y, z)$ непрерывны, то

$$W=\int\limits_{AB} P(x, y, z)dx+Q(x, y, z)dy+R(x, y, z)dz. \qquad (20.25)$$

Эта формула следует из формул (20.5) и (20.15).

З а м е ч а н и е. Если дуга AB лежит в плоскости Oxy, то формулы (20.23)—(20.25) упрощаются (полагаем в них $dz=0$, а z пропускаем).

§ 20.6. Формула Грина

Формула Грина[1] устанавливает связь между двойным интегралом по области S и криволинейным интегралом по контуру Γ, ограничивающему эту область. Будем считать, что область S является стандартной (см. § 19.4) в направлении каждой координатной оси и снизу ограни-

[1] *Джордж Грин* (*George Green*, 1793—1841) — английский математик.

чена графиком функции $y=y_1(x)$ (дугой ACB), сверху — графиком функции $y=y_2(x)$ (дугой ADB), которые вместе составляют замкнутый контур Г (рис. 20.3).

Пусть в области S и на ее границе Г заданы функции $P(x, y)$, $Q(x, y)$, непрерывные вместе со своими частными производными $P'_y(x, y)$, $Q'_x(x, y)$, тогда

$$\iint_S \frac{\partial P}{\partial y}\,dxdy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} \frac{\partial P}{\partial y}\,dy = \int_a^b P(x, y)\Big|_{y=y_1(x)}^{y=y_2(x)} dx =$$

$$= \int_a^b P(x, y_2(x))\,dx - \int_a^b P(x, y_1(x))\,dx =$$

$$= \int_{ADB} P(x, y)\,dx - \int_{ACB} P(x, y)\,dx =$$

$$= -\int_{BDA} P(x, y)\,dx - \int_{ACB} P(x, y)\,dx = -\oint_{Г} P(x, y)\,dx,$$

где обход контура Г совершается в положительном направлении, т. е. против часовой стрелки (область S остается слева). Следовательно,

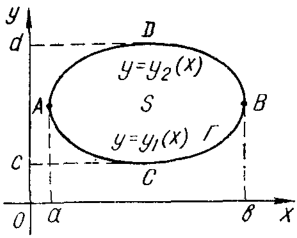

Рис. 20.3

$$\iint_S \frac{\partial P}{\partial y}\,dxdy = -\oint_{Г} P(x, y)\,dx. \qquad (20.26)$$

Аналогично получаем

$$\iint_S \frac{\partial Q}{\partial x}\,dxdy = \oint_{Г} Q(x, y)\,dy, \qquad (20.27)$$

где обход контура Г также совершается в положительном направлении.

Вычитая почленно (20.26) из (20.27), получаем формулу Грина

$$\iint_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy = \oint_{Г} Pdx + Qdy. \qquad (20.28)$$

Замечание 1. Если обход контура Г совершается в отрицательном направлении, т. е. по часовой стрелке (область S остается справа), то формула Грина принимает вид

$$\iint_S \left(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}\right) dxdy = \oint Pdx + Qdy.$$

Замечание 2. Формула Грина дает возможность вычислять площадь области с помощью криволинейного интеграла. Действительно, если $P(x, y) = -y$, $Q(x, y) = x$, то формула (20.28) перепишется так:

$$2\iint_S dS = \oint_\Gamma xdy - ydx,$$

откуда

$$S = \frac{1}{2}\oint_\Gamma xdy - ydx, \tag{20.29}$$

где обход контура Г совершается против часовой стрелки.

§ 20.7. Задачи, приводящие к понятиям интегралов по поверхности

Задача о массе материальной поверхности. Рассмотрим поверхность σ (рис. 20.4), на которой распределена масса с плотностью $p = p(x, y, z)$. Требуется вычислить массу этой материальной поверхности.

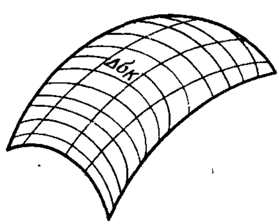

Рис. 20.4

Разобьем поверхность σ сетью дуг на n элементарных частей $\Delta\sigma_k$, площади которых также обозначим через $\Delta\sigma_k$ $(k = 1, 2, \ldots, n)$. Предположим, что в каждой элементарной части $\Delta\sigma_k$ плотность постоянна и равна $p_k = p(x_k, y_k, z_k)$, где $M_k(x_k, y_k, z_k) \in \Delta\sigma_k$. Произведение $p_k\Delta\sigma_k$ приближенно выражает массу элементарной части $\Delta\sigma_k$, а сумма всех таких произведений

$$T_n = \sum_{k=1}^{n} p(M_k)\,\Delta\sigma_k = \sum_{k=1}^{n} p(x_k, y_k, z_k)\Delta\sigma_k \tag{20.30}$$

приближенно выражает массу всей заданной материальной поверхности.

Обозначим через d_k диаметр области $\Delta\sigma_k$, а через λ_n — наибольший из диаметров d_k при заданном n.

Точное значение искомой массы выражается формулой

$$m = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} p(x_k, y_k, z_k)\Delta\sigma_k. \qquad (20.31)$$

Замечание 1. Предел вида (20.31) называют поверхностным интегралом первого рода и обозначают

$$\iint_{\sigma} p(x, y, z)\, d\sigma = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} p(x_k, y_k, z_k)\,\Delta\sigma_k. \qquad (20.32)$$

Задача о вычислении потока жидкости через поверхность. Пусть дана пространственная область, заполненная жидкостью, движущейся со скоростью $\mathbf{v} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$, где $\mathbf{v} = \mathbf{v}(x, y, z)$, $P = P(x, y, z)$, $Q = Q(x, y, z)$, $R = R(x, y, z)$.

Требуется вычислить количество жидкости П, протекающей в единицу времени через данную поверхность σ.

Разобьем эту поверхность на n элементарных частей $\Delta\sigma_k$ $(k = 1, 2, \ldots, n)$, площади которых также обозначим через $\Delta\sigma_k$. Будем считать, что для всех точек $\Delta\sigma_k$ скорость постоянна и равна скорости $\mathbf{v}_k$ в некоторой точке $M_k(x_k, y_k, z_k) \in \Delta\sigma_k$, т. е. $\mathbf{v}_k = \mathbf{v}(M_k) = \mathbf{v}(x_k, y_k, z_k)$.

Количество жидкости, протекающей через $\Delta\sigma_k$ за единицу времени, приближенно равно произведению $v_{n_k}\Delta\sigma_k$, где v_{n_k} — проекция скорости $\mathbf{v}_k$ на ось, определяемую единичным вектором нормали $\mathbf{n}_k$ к поверхности в точке M_k.

Так как

$$v_{n_k} = \mathbf{v}_k\mathbf{n}_k = [P(M_k)\cos\alpha_k + Q(M_k)\cos\beta_k + R(M_k)\cos\gamma_k],$$

где α_k, β_k, γ_k — углы, образованные нормалью $\mathbf{n}_k$ с координатными осями, то

$$v_{n_k}\Delta\sigma_k = [P(M_k)\cos\alpha_k + Q(M_k)\cos\beta_k + R(M_k)\cos\gamma_k]\,\Delta\sigma_k.$$

Количество жидкости, протекающей через всю поверхность за единицу времени, приближенно выразится формулой

$$\Pi_n = \sum_{k=1}^{n} [P(M_k)\cos\alpha_k + Q(M_k)\cos\beta_k + R(M_k)\cos\gamma_k]\,\Delta\sigma_k. \qquad (20.33)$$

Точное значение этого количества получим при переходе к пределу суммы (20.34), когда $\lambda_n \to 0$, где λ_n — наибольший из диаметров d_k области $\Delta\sigma_k$:

$$\Pi = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} [P(M_k) \cos \alpha_k + Q(M_k) \cos \beta_k + R(M_k) \cos \gamma_k] \Delta\sigma_k. \qquad (20.34)$$

Преобразуем сумму (20.33). Произведения $\cos \alpha_k \Delta\sigma_k$, $\cos \beta_k \Delta\sigma_k$, $\cos \gamma_k \Delta\sigma_k$ являются проекциями элементарной поверхности $\Delta\sigma_k$ соответственно на координатные плоскости Oyz, Oxz, Oxy. Обозначим эти проекции через $(\Delta S_{yz})_k$, $(\Delta S_{xz})_k$, $(\Delta S_{xy})_k$, т. е.

$$(\Delta S_{yz})_k = \Delta\sigma_k \cos \alpha_k, \quad (\Delta S_{xz})_k = \Delta\sigma_k \cos \beta_k, \quad (\Delta S_{xy})_k = \Delta\sigma_k \cos \gamma_k. \qquad (20.35)$$

Площади проекций берутся со знаком плюс или минус в зависимости от того, образует нормаль $\mathbf{n}_k$ с соответствующей осью острый или тупой угол.

С учетом (20.35) формулы (20.33) и (20.34) принимают вид

$$\Pi_n = \sum_{k=1}^{n} [P(M_k) (\Delta S_{zy})_k + Q(M_k) (\Delta S_{xz})_k + R(M_k) (\Delta S_{xy})_k], \qquad (20.36)$$

$$\Pi = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} [P(M_k)(\Delta S_{yz})_k + Q(M_k)(\Delta S_{xz})_k + R(M_k)(\Delta S_{xy})_k]. \qquad (20.37)$$

Замечание 2. Предел вида (20.37) называют поверхностным интегралом второго рода и обозначают

$$\iint_{\sigma} P dy dz + Q dx dz + R dx dy = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} [P(M_k)(\Delta S_{yz})_k + Q(M_k)(\Delta S_{xz})_k + R(M_k)(\Delta S_{xy})_k]. \qquad (20.38)$$

Замечание 3. Предел вида (20.34), как и предел вида (20.31), называют поверхностным интегралом первого рода и обозначают

$$\iint_{\sigma} (P \cos \alpha + Q \cos \beta + R \cos \gamma) \, d\sigma = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} [P(M_k) \cos \alpha_k + Q(M_k) \cos \beta_k + R(M_k) \cos \gamma_k] \Delta\sigma_k. \qquad (20.39)$$

Поскольку формулы (20.33) и (20.36) выражают одну и ту же сумму (но в разных формах), то их пределы (если они существуют) также равны, поэтому

$$\iint\limits_{\sigma} Pdydz + Qdxdz + Rdxdy = \iint\limits_{\sigma} (P\cos\alpha + Q\cos\beta + R\cos\gamma)d\sigma. \quad (20.40)$$

Формула (20.40) выражает связь между поверхностными интегралами первого и второго рода.

§ 20.8. Понятия интегралов по поверхности

Как и криволинейные интегралы, интегралы по поверхности бывают двух типов. Дадим определение каждого из них.

Поверхностный интеграл первого рода. Пусть на поверхности σ задана функция $u=f(x, y, z)$. Поверхность σ сетью линий разобьем на элементарные поверхности $\Delta\sigma_k$ (см. рис. 20.4). Площадь элементарной поверхности $\Delta\sigma_k$ обозначим теми же символами $\Delta\sigma_k$ $(k=1, 2, 3, \ldots, n)$ и наибольший из диаметров — через λ. На каждой элементарной поверхности $\Delta\sigma_k$ выберем произвольно точку $M_k(x_k, y_k, z_k)$, значение функции в этой точке $f(M_k) = f(x_k, y_k, z_k)$ умножим на площадь $\Delta\sigma_k$ и составим сумму всех таких произведений:

$$T_n = \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta\sigma_k. \quad (20.41)$$

Сумма (20.41) называется *интегральной суммой* функции $u=f(x, y, z)$ по поверхности σ.

Перейдем к пределу этой суммы при $\lambda \to 0$ (при этом $n \to \infty$). Если предел интегральной суммы существует, то он называется поверхностным интегралом первого рода.

Следовательно, по определению

$$\iint\limits_{\sigma} f(x, y, z)\,d\sigma = \lim_{\lambda\to 0} \sum_{k=1}^{n} f(x_k, y_k, z_k)\,\Delta\sigma_k. \quad (20.42)$$

Отметим без доказательства, что предел интегральной суммы (20.41) существует, если функция $f(x, y, z)$ непрерывна, а поверхность σ является гладкой[1] и ограниченной.

---

[1] *Гладкой* называется поверхность, имеющая в каждой своей точке касательную плоскость, положение которой меняется непрерывно при переходе от точки к точке.

Для определения поверхностного интеграла второго рода нам понадобятся некоторые вспомогательные понятия.

Двусторонняя поверхность. На поверхности σ фиксируем точку M_0 и одно из направлений нормали к ней в этой точке, указав единичный вектор $\mathbf{n}$, отложенный из M_0. Проведем через точку M_0 замкнутую линию Γ, целиком лежащую на поверхности σ и не имеющую общих точек с границей σ. Будем совершать обход линии Γ так, чтобы нормаль изменялась непрерывно, при этом вектор $\mathbf{n}$ в каждой точке M будет иметь вполне определенное направление (вообще говоря, отличное от направления в точке M_0). По возвращении в точку M_0 после совершения обхода может оказаться: 1) вектор $\mathbf{n}$ принял первоначальное направление; 2) вектор $\mathbf{n}$ изменил направление на противоположное.

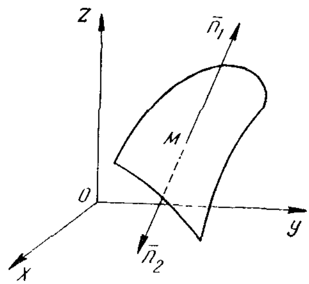

Рис. 20.5

Поверхность σ называется *двусторонней,* если обход по любой замкнутой линии, лежащей на этой поверхности и не имеющей общих точек с ее границей, не меняет направления нормали к поверхности.

Двусторонней поверхностью является всякая гладкая поверхность, определяемая уравнением $z=f(x, y)$. Действительно, выбрав направление нормального вектора $\mathbf{n}$ к ней так, чтобы он составил с осью Oz острый угол (рис. 20.5), получим одну сторону поверхности (верхнюю). Выбрав это направление так, чтобы вектор $\mathbf{n}$ составил с осью Oz тупой угол, получим другую сторону поверхности (нижнюю). В частности, плоскость и всякая ее часть (круг и т. п.) — двусторонняя поверхность. Любая замкнутая поверхность, не имеющая точек самопересечения (сфера, эллипсоид и др.), также является двусторонней. В самом деле, направив нормальный вектор внутрь объема, ограниченного этой поверхностью, получим одну сторону поверхности (внутреннюю), направив нормаль вне указанного объема,— другую сторону поверхности (внешнюю).

Двустороннюю поверхность называют также *ориенти-*

руемой, а выбор ее определенной стороны — *ориентацией поверхности.*

Если на поверхности существует замкнутая линия, обход по которой меняет направление нормали, то поверхность называется *односторонней.*

Простейшим примером односторонней поверхности является лист Мёбиуса[1] (рис. 20.6, *а*). Эту поверхность можно получить следующим образом: взяв полоску бу-

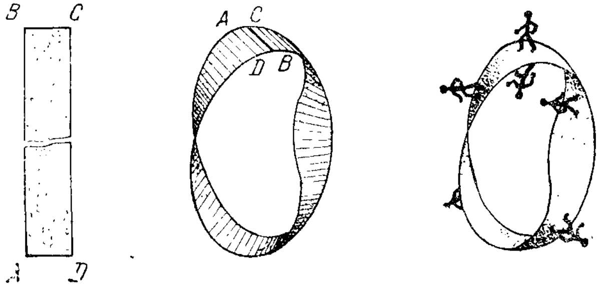

Рис. 20.6

маги $ABCD$ (рис. 20.6, *б*), склеим ее так, чтобы точка A совпала с точкой C, точка B — с точкой D, т. е. перед склеиванием повернуть на 180° (рис. 20.6, *в*). Читателю предлагается построить эту поверхность и убедиться в том, что при обходе листа Мёбиуса по его средней линии направление нормали к нему меняется на противоположное.

Поверхностный интеграл второго рода. Рассмотрим функцию $R=R(x, y, z)$, определенную и непрерывную на гладкой ориентируемой поверхности σ. Поверхность σ разобьем на n элементарных частей $\Delta\sigma_k$ $(k=1, 2, \ldots, n)$, выберем в каждой элементарной части точку $M_k(x_k, y_k, z_k)$ и составим интегральную сумму

$$T_n = \sum_{k=1}^{n} R(x_k, y_k, z_k)(\Delta S_{xy})_k, \qquad (20.43)$$

где $(\Delta S_{xy})_k$ — величина проекции $\Delta\sigma_k$ на плоскость Oxy; она равна площади области, в которую проектируется $\Delta\sigma_k$ на Oxy, взятой со знаком плюс, если нормаль к поверх-

[1] *Август Фердинанд Мёбиус (August Ferdinand Möbius,* 1790—1868) — немецкий геометр.

ности в точке M_k образует с осью Oz острый угол, и со знаком минус, если указанный угол — тупой.

Обозначим диаметр элементарной области $\Delta\sigma_k$ через d_k, наибольший из этих диаметров — через λ и перейдем к пределу интегральной суммы при $\lambda \to 0$ (при этом $n \to \infty$). Если предел интегральной суммы существует, то он называется *поверхностным интегралом второго рода*. Итак, по определению

$$\iint_\sigma R(x, y, z)dxdy = \lim_{\lambda \to 0} \sum_{k=1}^{n} R(x_k, y_k, z_k)(\Delta S_{xy})_k. \qquad (20.44)$$

Отметим без доказательства, что если ограниченная двусторонняя поверхность σ является гладкой и функция $R(x, y, z)$ непрерывна на ней, то предел суммы (20.43) существует и не зависит от способа разбиения поверхности σ на элементарные части $\Delta\sigma_k$ и выбора точки $M_k \in \Delta\sigma_k$.

Аналогичным образом определяются поверхностные интегралы второго рода

$$\iint_\sigma Q(x, y, z)dxdz = \lim_{\lambda \to 0} \sum_{k=1}^{n} Q(x_k, y_k, z_k)(\Delta S_{xz})_k,$$

где $(\Delta S_{xz})_k$ — величина проекции $\Delta\sigma_k$ на плоскость Oxz;

$$\iint_\sigma P(x, y, z)dydz = \lim_{\lambda \to 0} \sum_{k=1}^{n} P(x_k, y_k, z_k)(\Delta S_{yz})_k,$$

где $(\Delta S_{yz})_k$ — величина проекции $\Delta\sigma_k$ на плоскость Oyz, а также поверхностный интеграл общего вида

$$\iint_\sigma Pdydz + Qdxdz + Rdxdy =$$

$$= \iint_\sigma Pdydz + \iint_\sigma Qdxdz + \iint_\sigma Rdxdy. \qquad (20.45)$$

§ 20.9. Вычисление интегралов по поверхности

Пусть требуется вычислить поверхностный интеграл первого рода $\iint_\sigma \varphi(x, y, z)\, d\sigma$. Здесь σ — гладкая поверхность, заданная уравнением $z = f(x, y)$, где $f(x, y)$ — дифференцируемая функция.

Предположим, что поверхность σ однозначно проектируется на плоскость Oxy и S — ее проекция на эту плоскость. По определению

$$\iint_{\sigma} \varphi(x, y, z)\, d\sigma = \lim_{\lambda_n \to 0} \sum_{k=1}^{n} \varphi(x_k, y_k, z_k)\, \Delta\sigma_k.$$

Преобразуем эту интегральную сумму. Запишем уравнение поверхности в виде $F(x, y, z) = 0$, где $F(x, y, z) = z - f(x, y)$. Вектор нормали $\mathbf{n}$ к этой поверхности имеет координаты (см. § 18.8)

$$\frac{\partial F}{\partial x} = -\frac{\partial f}{\partial x}, \quad \frac{\partial F}{\partial y} = -\frac{\partial f}{\partial y}, \quad \frac{\partial F}{\partial z} = 1.$$

Введем обозначения:

$$\frac{\partial f}{\partial x} = p, \quad \frac{\partial f}{\partial y} = q, \quad \left(\frac{\partial f}{\partial x}\right)_{M_k} = p_k, \quad \left(\frac{\partial f}{\partial y}\right)_{M_k} = q_k. \qquad (20.46)$$

Направляющие косинусы нормали $\mathbf{n}$ к поверхности $z - f(x, y) = 0$ выразятся формулами

$$\cos\alpha = \frac{-p}{\sqrt{1 + p^2 + q^2}}, \quad \cos\beta = \frac{-q}{\sqrt{1 + p^2 + q^2}},$$

$$\cos\gamma = \frac{1}{\sqrt{1 + p^2 + q^2}}. \qquad (20.47)$$

Поскольку $\Delta\sigma_k \cos\gamma_k = (\Delta S_{xy})_k$ (см. формулы (20.35)), то

$$\Delta\sigma_k = \frac{(\Delta S_{xy})_k}{\cos\gamma_k} = \sqrt{1 + p_k^2 + q_k^2}\, (\Delta S_{xy})_k.$$

Следовательно, интегральная сумма принимает вид

$$\sum_{k=1}^{n} \varphi(x_k, y_k, z_k)\, \Delta\sigma_k =$$

$$= \sum_{k=1}^{n} \varphi[x_k, y_k, f(x_k, y_k)] \sqrt{1 + p_k^2 + q_k^2}\, (\Delta S_{xy})_k.$$

Переходя к пределу при $\lambda_n \to 0$, получаем

$$\iint_{\sigma} \varphi(x, y, z)\, d\sigma = \iint_{S} \varphi[x, y, f(x, y)] \sqrt{1 + p^2 + q^2}\, dx dy, \qquad (20.48)$$

где p и q определяются формулами (20.46).

Итак, вычисление поверхностного интеграла первого рода свелось к вычислению соответствующего двойного интеграла по области S — проекции поверхности σ на плоскость Oxy.

Замечание 1. Если гладкая поверхность σ задана уравнением $y=f(x, z)$ и S_1 — ее проекция на плоскость Oxz, то

$$\iint_\sigma \varphi(x, y, z)\, d\sigma =$$

$$= \iint_{S_1} \varphi[x, f(x, z), z] \sqrt{1+\left(\frac{\partial y}{\partial x}\right)^2+\left(\frac{\partial y}{\partial z}\right)^2}\, dxdz, \qquad (20.49)$$

Замечание 2. Если гладкая поверхность σ задана уравнением $x=f(y, z)$ и S_2 — ее проекция на плоскость Oyz, то

$$\iint_\sigma \varphi(x, y, z)\, d\sigma =$$

$$= \iint_{S_2} \varphi[f(y, z), y, z] \sqrt{1+\left(\frac{\partial x}{\partial y}\right)^2+\left(\frac{\partial x}{\partial z}\right)^2}\, dydz. \qquad (20.50)$$

Переходим к вопросу о вычислении поверхностного интеграла второго рода $\iint_\sigma F(x, y, z)\, dxdy$, где σ — гладкая поверхность, заданная уравнением $z=f(x, y)$.

Предположим, что σ взаимно-однозначно проектируется на область S плоскости Oxy.

По определению (см. формулу (20.44))

$$\iint_\sigma F(x, y, z)\, dxdy = \lim_{\lambda\to 0} \sum_{k=1}^{n} F(x_k, y_k, z_k)(\Delta S_{xy})_k,$$

где $(\Delta S_{xy})_k$ — величина проекции элементарной области $\Delta\sigma_k$ на плоскость Oxy, т. е. площадь области, в которую проектируется $\Delta\sigma_k$ на Oxy, взятая со знаком плюс, если нормаль к поверхности в точке $M_k(x_k, y_k, z_k)$ образует с осью Oz острый угол, и со знаком минус, если этот угол тупой.

Поскольку

$$\sum_{k=1}^{n} F(x_k, y_k, z_k)(\Delta S_{xy})_k = \sum_{k=1}^{n} F[x_k, y_k, f(x_k, y_k)]\,(\Delta S_{xy})_k,$$

то, переходя к пределу при $\lambda \to 0$, получаем

$$\iint_\sigma F(x,\ y,\ z)\,dxdy = \iint_S F[x,\ y,\ f(x,\ y)]\,dxdy \qquad (20.51)$$

для верхней стороны поверхности, т. е. в случае, когда $\cos\gamma > 0$, где γ — угол между нормалью к поверхности и осью Oz;

$$\iint_\sigma F(x,\ y,\ z)\,dxdy = -\iint_S F[x,\ y,\ f(x,\ y)]\,dxdy \qquad (20.52)$$

для нижней стороны поверхности, т. е. в случае, когда $\cos\gamma < 0$. Аналогично вычисляются поверхностные интегралы второго рода по координатам x, z и y, z.

§ 20.10. Приложения интегралов по поверхности

Масса материальной поверхности. Из формул (20.31) и (20.32) следует, что масса m поверхности, по которой распределено вещество с поверхностной плотностью $p = p(x,\ y,\ z)$, определяется формулой

$$m = \iint_\sigma p(x,\ y,\ z)\,d\sigma. \qquad (20.53)$$

Площадь поверхности. Если $p(x,\ y,\ z) \equiv 1$, то $m = 1\cdot\sigma = \sigma$, поэтому формула (20.53) принимает вид

$$\sigma = \iint_\sigma d\sigma, \qquad (20.54)$$

где σ — площадь поверхности σ.

Центр тяжести материальной поверхности. Координаты центра тяжести $M_0(x_0,\ y_0,\ z_0)$ материальной поверхности определяются формулами

$$\left.\begin{aligned} x_0 &= \frac{1}{m}\iint_\sigma xp(x,\ y,\ z)\,d\sigma;\\ y_0 &= \frac{1}{m}\iint_\sigma yp(x,\ y,\ z)\,d\sigma;\\ z_0 &= \frac{1}{m}\iint_\sigma zp(x,\ y,\ z)\,d\sigma, \end{aligned}\right\} \qquad (20.55)$$

где $p = p(x,\ y,\ z)$ — поверхностная плотность, а m — масса поверхности.

Вывод формул (20.55) аналогичен выводу формул

(19.29). Для однородной поверхности ($\rho(x, y, z) = \text{const}$) эти формулы упрощаются и принимают вид

$$x_0 = \frac{1}{\sigma}\iint\limits_{\sigma} x d\sigma, \quad y_0 = \frac{1}{\sigma}\iint\limits_{\sigma} y d\sigma, \quad z_0 = \frac{1}{\sigma}\iint\limits_{\sigma} z d\sigma, \tag{20.56}$$

где σ вычисляется по формуле (20.54).

Моменты инерции материальной поверхности. Моменты инерции материальной поверхности с плотностью $\rho = \rho(x, y, z)$ относительно координатных осей и начала прямоугольной декартовой системы координат определяются соответственно формулами

$$\left.\begin{aligned} I_x &= \iint\limits_{\sigma} (y^2 + z^2)\,\rho(x, y, z)\,d\sigma; \\ I_y &= \iint\limits_{\sigma} (x^2 + z^2)\,\rho(x, y, z)\,d\sigma; \\ I_z &= \iint\limits_{\sigma} (x^2 + y^2)\,\rho(x, y, z)\,d\sigma; \\ I_0 &= \iint\limits_{\sigma} (x^2 + y^2 + z^2)\,\rho(x, y, z)\,d\sigma. \end{aligned}\right\} \tag{20.57}$$

Вывод формул (20.57) аналогичен выводу формулы (19.32).

§ 20.11. Формула Стокса

Формула Стокса[1]) устанавливает связь между интегралом по поверхности и криволинейным интегралом по контуру, ограничивающему эту поверхность.

Предположим, что поверхность σ, ограниченная контуром L (рис. 20.7), задана уравнением $z = f(x, y)$ и взаимно-однозначно проектируется на область S плоскости Oxy, ограниченную линией Γ — проекцией линии L на ту же плоскость.

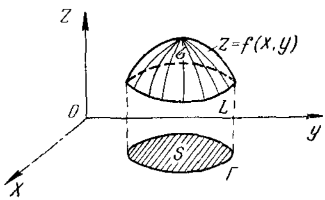

Рис. 20.7

Криволинейный интеграл по контуру L преобразуем сначала в криволинейный интег-

---

1) *Джордж Габриель Стокс* (*George Gabriel Stokes*, 1819—1903) — английский физик.

рал по контуру Г, полученный интеграл — в двойной интеграл по области S, наконец, последний — в интеграл по поверхности σ.

Так как контур L лежит на поверхности σ, заданной уравнением $z=f(x, y)$, то

$$\oint_L P(x, y, z)\,dx = \oint_\Gamma P[x, y, f(x, y)]\,dx. \qquad (20.58)$$

Воспользовавшись формулой (20.26), получим

$$\oint_\Gamma P[x, y, f(x, y)]\,dx = -\iint_S \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial y}\right)dxdy \qquad (20.59)$$

(при нахождении производной от P по y учтено то, что P — сложная функция x и y).

Обозначим через α, β, γ углы, образованные нормалью $\mathbf{n}$ к поверхности $z=f(x, y)$, или $z-f(x, y)=0$. Из формул (20.46) и (20.47) получаем

$$\frac{\partial f}{\partial y} = -\frac{\cos\beta}{\cos\gamma} \text{ или } \frac{\partial z}{\partial y} = -\frac{\cos\beta}{\cos\gamma}. \qquad (20.60)$$

Из формул (20.59) и (20.60) следует, что

$$\oint_\Gamma P[x, y, f(x, y)]\,dx = -\iint_S \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\frac{\cos\beta}{\cos\gamma}\right)dxdy.$$

Принимая во внимание формулы (20.40) и (20.51), находим

$$-\iint_S \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\frac{\cos\beta}{\cos\gamma}\right)dx\,dy =$$

$$= -\iint_\sigma \left(\frac{\partial P}{\partial y} - \frac{\partial P}{\partial z}\frac{\cos\beta}{\cos\gamma}\right)\cos\gamma d\sigma =$$

$$= \iint_\sigma \left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right)d\sigma.$$

Следовательно,

$$\oint_\Gamma P[x, y, f(x, y)]\,dx = \iint_\sigma \left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right)d\sigma.$$

Из последней формулы и формулы (20.58) имеем

$$\oint_L P(x, y, z)\,dx = \iint_\sigma \left(\frac{\partial P}{\partial z}\cos\beta - \frac{\partial P}{\partial y}\cos\gamma\right)d\sigma.$$

При соответствующих предположениях аналогично находим, что

$$\oint_L Q(x, y, z)\,dy = \iint_\sigma \left(\frac{\partial Q}{\partial x}\cos\gamma - \frac{\partial Q}{\partial z}\cos\alpha\right) d\sigma,$$

$$\oint_L R(x, y, z)\,dz = \iint_\sigma \left(\frac{\partial R}{\partial y}\cos\alpha - \frac{\partial R}{\partial x}\cos\beta\right) d\sigma.$$

Складывая почленно последние три равенства, получаем формулу Стокса

$$\oint_L Pdx+Qdy+Rdz = \iint_\sigma \left[\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\cos\gamma + \right.$$
$$\left. + \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\cos\alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\cos\beta\right] d\sigma. \quad (20.61)$$

Эту формулу можно переписать в виде

$$\oint_L Pdx + Qdy + Rdz = \iint_\sigma \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy +$$
$$+ \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) dydz + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) dzdx.$$

Замечание 1. Последнюю формулу легко запомнить, заметив, что первое слагаемое в правой части то же, что и под знаком двойного интеграла в формуле Грина, второе и третье получаются круговой перестановкой (см. § 5.8) функций P, Q, R и координат x, y, z.

Замечание 2. Если поверхность σ является плоской областью, лежащей в плоскости Oxy, то формула Стокса переходит в формулу Грина, поскольку интеграл по dz (в левой части) и интегралы по $dydz$, $dzdx$ (в правой части) обращаются в нуль.

§ 20.12. Формула Остроградского

Формула Остроградского[1] выражает связь между тройным интегралом по пространственной области и интегралом по поверхности, ограничивающей эту область.

Рассмотрим стандартную (см. § 19.7) пространственную область, ограниченную сверху поверхностью $z = z_2(x, y)$, снизу — поверхностью $z = z_1(x, y)$, а также боковой цилиндрической поверхностью с образующей, параллельной оси Oz. Эта цилиндрическая поверхность вырезает на плоскости Oxy область S — проекцию V на данную плоскость (рис. 20.8).

[1] *М. В. Остроградский* (1801—1861) — русский математик.

Пусть функции $R(x, y, z)$ и $R'_z(x, y, z)$ определены и непрерывны в области V и на ее границе.

Поскольку (см. формулу (19.39))

$$\iiint\limits_V \frac{\partial R}{\partial z}\,dx\,dy\,dz = \iint\limits_S \left[\int\limits_{z_1(x,\,y)}^{z_2(x,\,y)} \frac{\partial R}{\partial z}\,dz\right] dxdy$$

и

$$\int\limits_{z_1(x,\,y)}^{z_2(x,\,y)} \frac{\partial R}{\partial z}\,dz = R(x,\ y,\ z)\Big|_{z=z_1(x,\,y)}^{z=z_2(x,\,y)} =$$

$$= R[x,\ y,\ z_2(x,\ y)] - R[x,\ y,\ z_1(x,\ y)],$$

то

$$\iiint\limits_V \frac{\partial R}{\partial z}\,dxdydz = \iint\limits_S R[x,\ y,\ z_2(x,\ y)]\,dxdy -$$

$$- \iint\limits_S R[x,\ y,\ z_1(x,\ y)]\,dxdy.$$

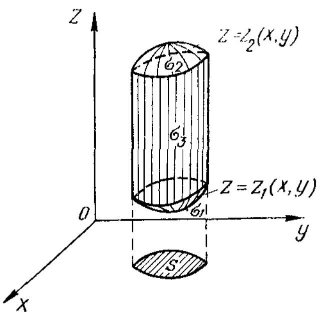

Рис. 20.8

Интегралы в правой части последнего равенства преобразуются в поверхностные: первый — по верхней стороне поверхности σ_2, заданной уравнением $z = z_2(x,\ y)$, второй — по верхней стороне поверхности σ_1, заданной уравнением $z = z_1(x,\ y)$, или со знаком минус — по нижней стороне поверхности σ_1. Следовательно,

$$\iiint\limits_V \frac{\partial R}{\partial z}\,dxdydz = \iint\limits_{\sigma_2} R(x,\ y,\ z)dxdy + \iint\limits_{\sigma_1} R(x,\ y,\ z)\,dxdy,$$

где первый интеграл берется по верхней стороне поверхности σ_2, второй — по нижней стороне поверхности σ_1.

Прибавив еще один интеграл по цилиндрической поверхности σ_3, ограничивающей область V:

$$\iint\limits_{\sigma_3} R(x,\ y,\ z)\,dxdy = \iint\limits_{\sigma_3} R(x,\ y,\ z)\cos\gamma d\sigma = 0$$

(на поверхности σ_3 вектор нормали $\mathbf{n}$ перпендикулярен к оси Oz, $\cos\gamma=0$), получим

$$\iiint\limits_V \frac{\partial R}{\partial z}\,dxdydz = \iint\limits_{\sigma_1} R(x,\ y,\ z)\,dxdy +$$

$$+\iint\limits_{\sigma_2} R(x,\ y,\ z)\,dxdy + \iint\limits_{\sigma_3} R(x,\ y,\ z)\,dxdy$$

или

$$\iiint\limits_V \frac{\partial R}{\partial z}\,dxdydz = \iint\limits_{\sigma} R(x,\ y,\ z)\,dxdy = \iint\limits_{\sigma} R(x,\ y,\ z)\cos\gamma d\sigma,$$

где $\sigma=\sigma_1+\sigma_2+\sigma_3$ — поверхность, ограничивающая объем V; интеграл берется по внешней стороне этой поверхности.

При соответствующих предположениях относительно области V и функций $Q(x,\ y,\ z), Q'_y(x,\ y,\ z)$, $P(x,\ y,\ z)$, $P'_x(x,\ y,\ z)$ аналогично находим

$$\iiint\limits_V \frac{\partial Q}{\partial y}\,dxdydz = \iint\limits_{\sigma} Q(x,\ y,\ z)dxdz = \iint\limits_{\sigma} Q(x,\ y,\ z)\cos\beta d\sigma,$$

$$\iiint\limits_V \frac{\partial P}{\partial x}\,dxdydz = \iint\limits_{\sigma} P(x,\ y,\ z)dydz = \iint\limits_{\sigma} P(x,\ y,\ z)\cos\alpha d\sigma.$$

Складывая почленно последние три равенства, получаем формулу Остроградского

$$\iiint\limits_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dxdydz =$$

$$= \iint\limits_{\sigma} Pdydz + Qdxdz + Rdxdy \qquad (20.62)$$

или

$$\iiint\limits_V \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dxdydz =$$

$$= \iint\limits_{\sigma} (P\cos\alpha + Q\cos\beta + R\cos\gamma)\,d\sigma. \qquad (20.63)$$

Замечание. С помощью формулы Остроградского можно найти выражение объема через интеграл по поверхности, ограничивающей этот объем. Действительно, подберем функции P, Q, R так, чтобы $\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = 1$, тогда

$$\iiint\limits_V \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)dxdydz=\iiint\limits_V dxdydz=V.$$

Взяв, например, $P=\frac{1}{3}x$, $Q=\frac{1}{3}y$, $R=\frac{1}{3}z$, получим

$$V=\frac{1}{3}\iint\limits_\sigma xdydz+ydxdz+zdxdy. \qquad (20.64)$$

§ 20.13. Поток, расходимость, циркуляция, вихрь. Векторная формулировка теорем Остроградского и Стокса

Поток векторного поля через поверхность. Обратимся снова к задаче о вычислении потока жидкости через данную поверхность (см. § 20.7).

Из формул (20.34) и (20.39) следует, что количество жидкости, протекающей в единицу времени через поверхность σ, выражается формулой

$$\Pi=\iint\limits_\sigma (P\cos\alpha+Q\cos\beta+R\cos\gamma)\,d\sigma. \qquad (20.65)$$

Величина П называется потоком жидкости через поверхность σ. Поскольку P, Q, R — координаты вектора скорости, $\mathbf{F}=\{P, Q, R\}$, $\cos\alpha$, $\cos\beta$, $\cos\gamma$ — координаты единичного вектора нормали $\mathbf{n}$ к поверхности и $P\cos\alpha+$ $+Q\cos\beta+R\cos\gamma=\mathbf{Fn}$, формулу (20.65) можно записать в виде

$$\Pi=\iint\limits_\sigma \mathbf{Fn}\,d\sigma \qquad (20.66)$$

или

$$\Pi=\iint\limits_\sigma F_n\,d\sigma, \qquad (20.67)$$

где F_n — нормальная составляющая вектора скорости, т. е. проекция вектора $\mathbf{F}$ на нормаль $\mathbf{n}$:

$$F_n=\mathbf{Fn}=|\mathbf{F}|\,|\mathbf{n}|\cos\varphi=|\mathbf{F}|\cos\varphi$$

(φ — угол между векторами $\mathbf{F}$ и $\mathbf{n}$; $|\mathbf{n}|=1$, так как $\mathbf{n}$ — единичный вектор).

Потоком векторного поля $\mathbf{F}$ через поверхность σ называется поверхностный интеграл $\iint\limits_{\sigma} F_n\,d\sigma$, где F_n — проекция вектора **F** на нормаль **n** к поверхности σ.

Если **F** — скорость движения жидкости, то поток векторного поля **F** через некоторую поверхность равен количеству жидкости, протекающей через эту поверхность в единицу времени.

Для векторного поля иной природы поток, разумеется, имеет другой физический смысл.

Расходимость векторного поля. *Расходимостью,* или *дивергенцией,* векторного поля $\mathbf{F}=\{P, Q, R\}$ называется скалярная функция, обозначаемая через $\operatorname{div}\mathbf{F}$ и определяемая формулой

$$\operatorname{div}\mathbf{F}=\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}. \qquad (20.68)$$

Пусть $\mathbf{F}=\{P, Q, R\}$ — вектор скорости жидкости, протекающей через область V. Интеграл в правой части формулы (20.63) выражает количество жидкости, вытекающей из области V через поверхность σ в единицу времени (или втекающей в область V, если этот интеграл отрицателен). Это количество жидкости выражается через тройной интеграл по области V от дивергенции вектора **F**.

Если $\operatorname{div}\mathbf{F}=0$, то и соответствующий интеграл по поверхности также равен нулю, т. е. количество вытекающей (или втекающей) жидкости через любую замкнутую поверхность σ равно нулю; это значит, что отсутствуют источники (или стоки).

Пользуясь понятиями потока и расходимости векторного поля, формулу Остроградского можно представить в виде

$$\iiint\limits_{V}\operatorname{div}\mathbf{F}\,dV=\iint\limits_{\sigma}\mathbf{F}\mathbf{n}\,d\sigma. \qquad (20.69)$$

Итак, тройной интеграл от дивергенции векторного поля по некоторому объему равен потоку этого векторного поля через поверхность, ограничивающую данный объем.

Циркуляция векторного поля. Пусть векторное поле задано вектор-функцией

$$\mathbf{F}=P(x, y, z)\mathbf{i}+Q(x, y, z)\mathbf{j}+R(x, y, z)\mathbf{k},$$

а Г — гладкая или кусочно-гладкая линия, на которой функции $P=P(x, y, z)$, $Q=Q(x, y, z)$, $R=R(x, y, z)$ непрерывны. Циркуляцией векторного поля $\mathbf{F}$ вдоль линии Г называется криволинейный интеграл

$$\int_{\Gamma} P dx + Q dy + R dz = \int_{\Gamma} F_{\tau}\, dl, \tag{20.70}$$

где F_{τ} — тангенциальная составляющая вектора $\mathbf{F}$ на контуре Г, т. е. $F_{\tau}=\mathbf{F}\boldsymbol{\tau}=P\cos\alpha+Q\cos\beta+R\cos\gamma$ ($\boldsymbol{\tau}$ — единичный вектор касательной к линии Г, образующий с осями координат углы α, β, γ).

Если поле $\mathbf{F}=\{P, Q, R\}$ является силовым, то его циркуляция вдоль линии Г представляет работу этого силового поля вдоль пути $\mathbf{F}$ (см. формулу (20.25)).

Вихрь векторного поля. Обратимся к формуле Стокса, т. е. к формуле (20.61). Правая часть этой формулы представляет собой поток вектора

$$\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}$$

через поверхность σ. Этот вектор называют *вихрем* (или *ротором*) векторного поля $\mathbf{F}=\{P, Q, R\}$ и обозначают символом $\operatorname{rot}\mathbf{F}$.

Итак, по определению

$$\operatorname{rot}\mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}. \tag{20.71}$$

Чтобы запомнить это выражение для вихря векторного поля $\mathbf{F}=\{P, Q, R\}$, представим его в виде символического определителя

$$\operatorname{rot}\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}, \tag{20.72}$$

где $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ — орты координатных осей, а под «умножением» символа второй строки на некоторую функцию третьей строки понимается выполнение соответствующей операции дифференцирования $\left(\text{например, } \frac{\partial}{\partial y}P \text{ означает } \frac{\partial P}{\partial y}.\right)$

Разложив этот определитель по элементам первой строки, получим правую часть равенства (20.71).

Пользуясь понятиями циркуляции и вихря, можно представить формулу Стокса в виде

$$\oint_{\Gamma} \mathbf{F}d\mathbf{l} = \iint_{\sigma} \mathbf{n}\,\mathrm{rot}\,\mathbf{F}\,d\sigma \tag{20.73}$$

или

$$\oint_{\Gamma} F_{\tau}dl = \iint_{\sigma} (\mathrm{rot}\,\mathbf{F})_n d\sigma. \tag{20.74}$$

Таким образом, циркуляция векторного поля **F** вдоль некоторого замкнутого контура Г равна потоку вихря этого векторного поля через поверхность, ограниченную данным контуром.

§ 20.14. Оператор «набла». Потенциальное и соленоидальное поле

В § 18.10 было определено понятие градиента скалярной функции. Переход от скалярного поля $u=u(x, y, z)$ к векторному полю $\mathrm{grad}\,u$ можно рассматривать как некоторую операцию, во многом аналогичную по своим свойствам операции дифференцирования с той, однако, разницей, что в данном случае скаляру ставится в соответствие вектор (а при дифференцировании скаляру ставится в соответствие скаляр).

Операция перехода от u к $\mathrm{grad}\,u$ обозначается символом ∇ (читается «набла») и называется *оператором «набла»* (∇-оператором), или оператором Гамильтона[1]. Итак, по определению

$$\nabla u = \mathrm{grad}\,u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}. \tag{20.75}$$

Во многих случаях оператор ∇ удобно рассматривать как символический вектор с координатами $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$:

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}, \tag{20.76}$$

[1] *Уильям Роуан Гамильтон* (*William Rowan Hamilton*, 1805—1865) — ирландский математик и механик.

а применение его к скалярной функции — как умножение скаляра на этот вектор [1]:

$$\nabla u = \left(\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}\right)u. \tag{20.77}$$

Найдем выражения для некоторых операций векторного анализа с помощью оператора «набла». Рассмотрим дифференцируемую вектор-функцию

$$\mathbf{F} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}. \tag{20.78}$$

Принимая во внимание определение дивергенции и сказанное выше, получаем

$$\operatorname{div}\mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} =$$

$$= \left(\frac{\partial}{\partial x}P + \frac{\partial}{\partial y}Q + \frac{\partial}{\partial z}R\right) = \nabla\mathbf{F},$$

$$\operatorname{div}\mathbf{F} = \nabla\mathbf{F}, \tag{20.79}$$

т. е. дивергенция вектора $\mathbf{F}$ равна скалярному произведению символического вектора ∇ на вектор $\mathbf{F}$.

Выражение для вихря вектора (20.78) можно переписать так:

$$\operatorname{rot}\mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} +$$

$$+ \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k} =$$

$$= \begin{vmatrix} \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ Q & R \end{vmatrix}\mathbf{i} + \begin{vmatrix} \frac{\partial}{\partial z} & \frac{\partial}{\partial x} \\ R & P \end{vmatrix}\mathbf{j} + \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ P & Q \end{vmatrix}\mathbf{k} = [\nabla, \mathbf{F}].$$

Следовательно,

$$\operatorname{rot}\mathbf{F} = [\nabla, \mathbf{F}], \tag{20.80}$$

т. е. вихрь вектора равен векторному произведению символического вектора ∇ на данный вектор.

Считая, что функции P, Q, R, входящие в формулу

---

[1] Как уже отмечалось, в связи с символической формулой (20.72) умножение символа вида $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$ на некоторую функцию означает дифференцирование этой функции по соответствующей переменной.

(20.78), имеют непрерывные частные производные второго порядка, находим $\operatorname{div}\operatorname{rot}\mathbf{F}$:

$$\operatorname{div}\operatorname{rot}\mathbf{F}=\frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)+$$
$$+\frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)=\frac{\partial^2 R}{\partial y\partial x}-\frac{\partial^2 Q}{\partial z\partial x}+\frac{\partial^2 P}{\partial z\partial y}-\frac{\partial^2 R}{\partial x\partial y}+$$
$$+\frac{\partial^2 Q}{\partial x\partial z}-\frac{\partial^2 P}{\partial y\partial z}=0.$$

Итак,

$$\operatorname{div}\operatorname{rot}\mathbf{F}=0. \tag{20.81}$$

Применяя формулу (20.80), находим

$$\operatorname{div}\operatorname{rot}\mathbf{F}=\nabla\operatorname{rot}\mathbf{F}=\nabla[\nabla,\ \mathbf{F}]=\nabla\nabla\mathbf{F}. \tag{20.82}$$

Из двух последних формул следует, что

$$\nabla[\nabla,\ \mathbf{F}]=\nabla\nabla\mathbf{F}=0. \tag{20.83}$$

Этот результат можно интерпретировать так: смешанное произведение трех векторов, среди которых два одинаковых, равно нулю, что согласуется с аналогичным равенством для обычных векторов.

Находим теперь выражение для $\operatorname{rot}\operatorname{grad}u$, считая, что функция $u=u(x,\ y,\ z)$ имеет непрерывные частные производные второго порядка. На основании формул (18.86) и (20.71) получаем (полагая в последней формуле $P=\frac{\partial u}{\partial x},\ Q=\frac{\partial u}{\partial y},\ R=\frac{\partial u}{\partial z}$)

$$\operatorname{rot}\operatorname{grad}u=\left[\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial z}\right)-\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial y}\right)\right]\mathbf{i}+$$
$$+\left[\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial x}\right)-\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial z}\right)\right]\mathbf{j}+$$
$$+\left[\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right)-\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right)\right]\mathbf{k}=\left(\frac{\partial^2 u}{\partial z\partial y}-\frac{\partial^2 u}{\partial y\partial z}\right)\mathbf{i}+$$
$$+\left(\frac{\partial^2 u}{\partial x\partial z}-\frac{\partial^2 u}{\partial z\partial x}\right)\mathbf{j}+\left(\frac{\partial^2 u}{\partial y\partial x}-\frac{\partial^2 u}{\partial x\partial y}\right)\mathbf{k}=\mathbf{0},$$
$$\operatorname{rot}\operatorname{grad}u=\mathbf{0}. \tag{20.84}$$

С другой стороны, посредством формул (20.75) и (20.80) находим выражение для $\operatorname{rot}\operatorname{grad}u$ через символический вектор ∇:

$$\operatorname{rot}\operatorname{grad} u=[\nabla,\ \operatorname{grad} u]=[\nabla,\ \nabla u]. \qquad (20.85)$$

Из двух последних формул следует, что

$$[\nabla,\ \nabla u]=\mathbf{0}. \qquad (20.86)$$

Равенство (20.86) означает, что векторное произведение двух символических векторов, отличающихся скалярным множителем, равно нулю. Этот факт также соответствует аналогичному факту для обычных векторов.

Пусть дана скалярная функция $u=u(x,\ y,\ z)$, имеющая непрерывные частные производные второго порядка, и векторное поле ее градиента:

$$\mathbf{F}=\operatorname{grad} u=\frac{\partial u}{\partial x}\mathbf{i}+\frac{\partial u}{\partial y}\mathbf{j}+\frac{\partial u}{\partial z}\mathbf{k}.$$

Найдем выражение для $\operatorname{div}(\operatorname{grad} u)$. В соответствии с формулой (20.68) получаем

$$\operatorname{div}(\operatorname{grad} u)=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right)+\frac{\partial}{\partial z}\left(\frac{\partial u}{\partial z}\right)=$$

$$=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2},$$

$$\operatorname{div}(\operatorname{grad} u)=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}. \qquad (20.87)$$

Правая часть этой формулы называется *оператором Лапласа* [1] от функции u и обозначается

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}. \qquad (20.88)$$

С учетом (20.75), (20.79) и (20.88) формулу (20.87) можно записать так:

$$\nabla\nabla u=\Delta u \quad (\Delta=\nabla^2). \qquad (20.89)$$

Замечание. Уравнение

$$\Delta u=0, \text{ или } \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=0, \qquad (20.90)$$

называется *уравнением Лапласа*. Функция $u=u(x,\ y,\ z)$, удовлетворяющая этому уравнению, называется *гармонической*.

Векторное поле $\mathbf{F}=P\mathbf{i}+Q\mathbf{j}+R\mathbf{k}$ называется *потенциальным*, если вектор $\mathbf{F}$ является градиентом некоторой скалярной функции, т. е.

---

[1] *Пьер Симон Лаплас* (*Pierre Simon de Laplace*, 1749—1827) — выдающийся французский математик, физик, астроном.

$$\mathbf{F} = \operatorname{grad} u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}. \tag{20.91}$$

Из двух приведенных выражений для $\mathbf{F}$ следует, что

$$P = \frac{\partial u}{\partial x}, \quad Q = \frac{\partial u}{\partial y}, \quad R = \frac{\partial u}{\partial z}.$$

В свою очередь, предполагая, что функция $u = u(x, y, z)$ имеет непрерывные частные производные второго порядка, из этих равенств получаем

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}, \quad \frac{\partial R}{\partial y} = \frac{\partial Q}{\partial z},$$

или

$$\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} = 0, \quad \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} = 0, \quad \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} = 0,$$

т. е.

$$\operatorname{rot} \mathbf{F} = 0. \tag{20.92}$$

Векторное поле $\mathbf{F}$, для которого выполнено равенство (20.92), называется *безвихревым.* Следовательно, всякое потенциальное поле будет безвихревым.

Справедливо и обратное заключение: всякое безвихревое поле является потенциальным.

Векторное поле $\mathbf{F} = \mathbf{F}(x, y, z)$ называется *соленоидальным,* или *трубчатым,* если

$$\operatorname{div} \mathbf{F} = 0. \tag{20.93}$$

Как уже отмечалось выше (см. § 20.13), векторное поле, для которого выполнено равенство (20.93), свободно от источников (или стоков).

Из равенства (20.81) следует, что поле вихрей данного векторного поля является соленоидальным; оно свободно от источников.

§ 20.15. Признак полного дифференциала

Рассмотрим криволинейный интеграл

$$I = \int_l P dx + Q dy + R dz. \tag{20.94}$$

Введем предварительно понятие односвязной пространственной области. Трехмерная область V называется *поверхностно односвязной,* если для любого замкнутого контура γ, принадлежащего этой области, найдется поверхность, целиком лежащая внутри V, имеющая данный контур своей границей. К таким областям относится шар; область, ограниченная эллипсоидом; область, заключенная между двумя концентрическими сферами, и т. п. Примером *поверхностно неодносвязной* области может служить шар, из которого вырезан цилиндр с осью, проходящей через центр шара (причем $r<R$, где r — радиус цилиндра, R — радиус шара).

Условия независимости криволинейного интеграла (20.94) от пути интегрирования и условия того, что выражение

$$Pdx+Qdy+Rdz \tag{20.95}$$

является полным дифференциалом некоторой функции $U(x, y, z)$, т. е.

$$dU=Pdx+Qdy+Rdz, \tag{20.96}$$

выражаются следующей теоремой.

Теорема 20.1. Если функции $P=P(x, y, z)$, $Q=Q(x, y, z)$, $R=R(x, y, z)$ и их частные производные первого и второго порядка непрерывны в некоторой замкнутой ограниченной поверхностно односвязной области V, то следующие четыре утверждения равносильны.

1. Криволинейный интеграл по любому замкнутому контуру, лежащему внутри V, равен нулю:

$$\oint Pdx+Qdy+Rdz=0.$$

2. Криволинейный интеграл не зависит от выбора пути, соединяющего точки A и B области V:

$$\int\limits_{ACB} Pdx+Qdy+Rdz=\int\limits_{ADB} Pdx+Qdy+Rdz.$$

3. Выражение $Pdx+Qdy+Rdz$ является полным дифференциалом, т. е.

$$Pdx+Qdy+Rdz=dU,$$

где $U=U(x, y, z)$ — некоторая функция, определенная в области V.

4. Выполняются равенства

$$\frac{\partial Q}{\partial x}=\frac{\partial P}{\partial y},\ \frac{\partial R}{\partial y}=\frac{\partial Q}{\partial z},\ \frac{\partial P}{\partial z}=\frac{\partial R}{\partial x}. \qquad (20.97)$$

Доказательство теоремы проведем по схеме $1\Rightarrow 2\Rightarrow 3\Rightarrow 4\Rightarrow 1$.

а) $1\Rightarrow 2$. В области V рассмотрим два произвольных пути, соединяющих точки A и B, например ACB и ADB. Они составляют в сумме замкнутый путь $ACBDA$. В соответствии с условием интеграл по любому замкнутому пути, лежащему внутри V, равен нулю, поэтому

$$\int\limits_{ACBDA} Pdx+Qdy+Rdz=0.$$

Поскольку

$$\int\limits_{ACBDA} Pdx+Qdy+Rdz=\int\limits_{ACB} Pdx+Qdy+Rdz+$$

$$+\int\limits_{BDA} Pdx+Qdy+Rdz=\int\limits_{ACB} Pdx+Qdy+Rdz-$$

$$-\int\limits_{ADB} Pdx+Qdy+Rdz,$$

то

$$\int\limits_{ACB} Pdx+Qdy+Rdz=\int\limits_{ADB} Pdx+Qdy+Rdz.$$

Последнее равенство означает, что криволинейный интеграл не зависит от пути интегрирования (зависит только от начальной и конечной точек его).

б) $2\Rightarrow 3$. Пусть интеграл (20.94) не зависит от пути интегрирования. Если фиксировать точку $A(x_0,\ y_0,\ z_0)$, то интеграл будет функцией переменных x, y, z — координат точки $B(x,\ y,\ z)$:

$$\int\limits_{AB} Pdx+Qdy+Rdz=U(x,\ y,\ z).$$

Покажем, что эта функция дифференцируема и выполняется равенство (20.96). Для этого достаточно доказать, что первые частные производные функции $U(x,\ y,\ z)$ существуют и соответственно равны

$$\frac{\partial U}{\partial x}=P(x,\ y,\ z),\quad \frac{\partial U}{\partial y}=Q(x,\ y,\ z),\quad \frac{\partial U}{\partial z}=R(x,\ y,\ z).$$

По определению

$$\frac{\partial U}{\partial x} = \lim_{\Delta x \to 0} \frac{U(x+\Delta x,\ y,\ z) - U(x,\ y,\ z)}{\Delta x}.$$

Величина $\Delta_x U = U(x+\Delta x,\ y,\ z) - U(x,\ y,\ z)$ представляет собой интеграл от выражения (20.95), взятого по пути, который соединяет точки $B(x,\ y,\ z)$ и $B_1(x+\Delta x,\ y,\ z)$. Поскольку по условию этот интеграл не зависит от пути интегрирования, то путь можно считать отрезком BB_1, параллельным оси Ox. Таким образом,

$$\frac{\Delta_x U}{\Delta x} = \frac{1}{\Delta x} \int\limits_{BB_1} P dx + Q dy + R dz =$$

$$= \frac{1}{\Delta x} \int\limits_{(x,\ y,\ z)}^{(x+\Delta x,\ y,\ z)} P(x,\ y,\ z) dx = P(x + \theta \Delta x,\ y,\ z).$$

(Последнее равенство получено на основании теоремы о среднем.)

Следовательно,

$$\frac{\partial U}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta_x U}{\Delta x} = \lim_{\Delta x \to 0} P(x + \theta \Delta x,\ y,\ z) = P(x,\ y,\ z),$$

поскольку функция $P(x,\ y,\ z)$ непрерывна.

Аналогично доказывается, что

$$\frac{\partial U}{\partial y} = Q(x,\ y,\ z), \quad \frac{\partial U}{\partial z} = R(x,\ y,\ z).$$

в) $3 \Rightarrow 4$. Если выражение (20.95) является полным дифференциалом некоторой функции, т. е. выполнено равенство (20.96), то

$$\frac{\partial U}{\partial x} = P\,(x,\ y,\ z), \quad \frac{\partial U}{\partial y} = Q\,(x,\ y,\ z), \quad \frac{\partial U}{\partial z} = R(x,\ y,\ z). \tag{20.98}$$

Последние равенства следуют из равенств (18.45) и (20.96).

По теореме о смешанных производных отсюда получаем

$$\frac{\partial Q}{\partial x} = \frac{\partial^2 U}{\partial y \partial x} = \frac{\partial^2 U}{\partial x \partial y} = \frac{\partial P}{\partial y}, \quad \frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}.$$

Аналогично доказывается, что

$$\frac{\partial R}{\partial y} = \frac{\partial Q}{\partial z}; \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}.$$

г) $4 \Rightarrow 1$. Пусть выполнены равенства (20.97), L — произвольный замкнутый контур, принадлежащий области V, σ — поверхность, целиком лежащая внутри V, для которой L является границей, тогда из формулы Стокса (см. (20.61)) следует, что

$$\int_L P dx + Q dy + R dz = 0.$$

Теорема доказана.

Замечание. Для криволинейного интеграла $\int_l P\,dx + Q\,dy$ условие (20.97) принимает вид

$$\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}. \qquad (20.99)$$

Глава 21

ЧИСЛОВЫЕ РЯДЫ

Рядом называется выражение вида

$$\sum_{k=1}^{\infty} u_k = u_1 + u_2 + u_3 + \ldots + u_k + \ldots,$$

где $\{u_k\}$ — последовательность чисел или функций; u_1, u_2, u_3, ..., u_k, ... называются *членами* ряда. Если все члены ряда — числа, то ряд называется *числовым*, если члены ряда — функции, то ряд *функциональный*.

§ 21.1. Сходимость и расходимость числовых рядов

Рассмотрим числовой ряд

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \ldots + a_k + \ldots, \qquad (21.1)$$

где $\{a_k\}$ — последовательность чисел. В дальнейшем слово «числовой» будем опускать и говорить просто ряд.

Ряд (21.1) задан, если известен его *общий член*

$$a_k = \varphi(k) \quad (k = 1, 2, 3, \ldots), \qquad (21.2)$$

т. е. задана функция натурального аргумента k.

Сумма n первых членов ряда называется его n-й *частичной суммой*. Если n-ю частичную сумму ряда (21.1) обозначить через S_n, то по определению $S_1=a_1$, $S_2=a_1+a_2$, $S_3=a_1+a_2+a_3$, ...,

$$S_n=a_1+a_2+a_3+\ldots+a_n. \tag{21.3}$$

Конечный или бесконечный предел последовательности частичных сумм ряда называется *суммой* данного ряда. Обозначим эту сумму через S, тогда по определению

$$S=\lim_{n\to\infty} S_n. \tag{21.4}$$

Ряд, имеющий *конечную* сумму, называется сходящимся. Для сходящегося ряда (21.1) с суммой S употребляется запись

$$S=\sum_{k=1}^{\infty} a_k. \tag{21.5}$$

Если предел последовательности частичных сумм бесконечен или не существует, то ряд называется *расходящимся*.

Необходимое и достаточное условие сходимости числового ряда выражается следующей теоремой, приводимой здесь без доказательства.

Теорема 21.1 (критерий Коши). Ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится тогда и только тогда, когда для любого $\varepsilon>0$ существует такой номер $N=N(\varepsilon)$, что при всех $n>N$ и любом целом $p\geqslant 0$ выполняется неравенство

$$|a_n+a_{n+1}+\ldots+a_{n+p}|<\varepsilon \quad \text{или} \quad \left|\sum_{k=n}^{n+p} a_k\right|<\varepsilon.$$

Пример. Исследовать, при каких значениях q сходится ряд $\sum\limits_{n=1}^{\infty} aq^{n-1}$.

Члены этого ряда образуют геометрическую прогрессию со знаменателем q. Будем кратко называть этот ряд *геометрической прогрессией*, или *геометрическим рядом*. Поскольку

$$S_n=a+aq+aq^2+\ldots+aq^{n-1}, \quad qS_n=aq+aq^2+aq^3+\ldots+aq^n,$$
$$S_n-qS_n=a-aq^n, \quad S_n(1-q)=a(1-q^n),$$

то при $q \neq 1$

$$S_n = \frac{a(1-q^n)}{1-q}.$$

Из последнего равенства видно, что наличие предела n-й частичной суммы ряда зависит от величины q, а именно:

1) при $|q| < 1$ $\lim\limits_{n\to\infty} q^n = 0$, $\lim\limits_{n\to\infty} S_n = \frac{a}{1-q}$, $S = \frac{a}{1-q}$;

2) при $|q| > 1$ $\lim\limits_{n\to\infty} q^n = \infty$, $\lim\limits_{n\to\infty} S_n = \infty$; ряд расходится;

3) при $q=1$ ряд приннмает вид $a+a+\ldots+a+\ldots$, для него $S_n = na$, $\lim\limits_{n\to\infty} S_n = \infty$; ряд расходнтся;

4) при $q=-1$ ряд принимает вид $a-a+a-\ldots+(-1)^{n-1}a+\ldots$, для него $S_{2m}=0$, $S_{2m+1}=a$, последовательность частичных сумм не имеет предела; ряд расходится. Следовательно, геометрическая прогрессия сходится тогда и только тогда, когда $|q|<1$; ее сумма выражается формулой $S=a:(1-q)$.

Остатком ряда после n-го члена (или n-м остатком) называется ряд, полученный из данного отбрасыванием его первых n членов.

Теорема 21.2. Ряд и любой его остаток одновременно сходятся или расходятся.

Доказательство. Рассмотрим ряд (21.1) и его остаток после n-го члена, т. е. ряд

$$a_{n+1}+a_{n+2}+a_{n+3}+\ldots. \tag{21.6}$$

Пусть

$$S_n = a_1+a_2+\ldots+a_n \quad (n=1, 2, \ldots)$$

есть частичные суммы ряда (21.1),

$$\sigma_m = a_{n+1}+a_{n+2}+\ldots+a_{n+m} \quad (m=1, 2, \ldots)$$

есть частичные суммы ряда (21.6). Очевидно, что

$$S_{n+m} = S_n + \sigma_m. \tag{21.7}$$

Из этого равенства следует, что при фиксированном n конечный предел $\lim\limits_{m\to\infty} S_{n+m}$ существует тогда и только тогда, когда существует конечный предел $\lim\limits_{m\to\infty} \sigma_m$. Это означает, что ряд (21.1) сходится тогда и только тогда, когда сходится ряд (21.6).

Следствие. При исследовании вопроса о сходимости ряда можно отбросить конечное число его первых членов.

Пусть ряд (21.1) сходится, тогда сходится и остаток данного ряда после n-го члена. Сумму остатка обозначим через r_n:

$$r_n = a_{n+1} + a_{n+2} + a_{n+3} + \ldots.$$

Если в равенстве (21.7) перейти к пределу при $m \to \infty$, то получим $S = S_n + r_n$. Отсюда следует, что

$$\lim_{n\to\infty} r_n = 0. \tag{21.8}$$

§ 21.2. Необходимый признак сходимости ряда

Необходимый признак сходимости ряда выражается следующей теоремой.

Теорема 21.3. Если ряд (21.1) сходится, то его n-й член стремится к нулю при $n \to \infty$, т. е.

$$\lim_{n\to\infty} a_n = 0. \tag{21.9}$$

Доказательство. Рассмотрим частичные суммы ряда (21.1):

$$S_{n-1} = a_1 + a_2 + a_3 + \ldots + a_{n-1},$$
$$S_n = a_1 + a_2 + a_3 + \ldots + a_{n-1} + a_n.$$

По условию ряд сходится. Обозначим его сумму через S, тогда $\lim\limits_{n\to\infty} S_{n-1} = S$, $\lim\limits_{n\to\infty} S_n = S$.

Так как $a_n = S_n - S_{n-1}$, то

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = S - S = 0,$$

что и требовалось доказать.

Следствие. Если n-й член ряда не стремится к нулю при $n \to \infty$, то ряд расходится.

Допустив противное (ряд сходится), получим по теореме, что $\lim\limits_{n\to\infty} a_n = 0$. Наличие противоречия доказывает следствие.

Пример. Выяснить, сходится ли ряд $\sum\limits_{n=1}^{\infty} \frac{n+3}{2n+1}$.

Поскольку

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n+3}{2n+1} = \frac{1}{2},$$

т. е. n-й член ряда не стремится к нулю при $n\to\infty$, то ряд расходится.

Замечание. Признак, выраженный условием (21.9), не является достаточным. Другими словами, может оказаться, что $\lim\limits_{n\to\infty} a_n=0$, а ряд расходится. Рассмотрим, например, так называемый гармонический ряд, т. е. ряд

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n} + \ldots. \qquad (21.10)$$

Для этого ряда $a_n = \frac{1}{n}$, $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} \frac{1}{n} = 0$, т. е. выполнено условие (21.9), но ряд расходится. Предположим противное, т. е. что ряд сходится. Обозначим его сумму через S, тогда

$$\lim_{n\to\infty} S_n = S, \ \lim_{n\to\infty} S_{2n} = S, \ \lim_{n\to\infty} (S_{2n} - S_n) = S - S = 0,$$

что противоречит неравенству

$$S_{2n} - S_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} > n\cdot\frac{1}{2n} = \frac{1}{2}.$$

Следовательно, гармонический ряд расходится.

§ 21.3. Признаки сходимости рядов с положительными членами

Будем рассматривать числовые ряды с положительными членами. Установим признаки, с помощью которых можно исследовать, сходится ли данный ряд на основании сравнения его с другим сходящимся или расходящимся рядом. Эти признаки выражаются следующими теоремами.

Теорема 21.4. Если каждый член ряда с положительными членами

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \ldots + a_n + \ldots \qquad (21.11)$$

не превосходит соответствующий член ряда

$$\sum_{n=1}^{\infty} b_n = b_1 + b_2 + \ldots + b_n + \ldots \qquad (21.12)$$

и ряд (21.12) сходится, то сходится и ряд (21.11); если ряд (21.11) расходится, то расходится и ряд (21.12).

Доказательство. По условию

$$0<a_n\leqslant b_n \quad (n=1, 2, 3, \ldots). \qquad (21.13)$$

Обозначим через S_n и σ_n n-е частичные суммы рядов (21.11) и (21.12) соответственно:

$$S_n=a_1+a_2+\ldots+a_n, \quad \sigma_n=b_1+b_2+\ldots+b_n. \qquad (21.14)$$

Так как ряд (21.12) сходится, то последовательность $\{\sigma_n\}$ его частичных сумм имеет предел и поэтому ограничена сверху: это означает, что существует такое число σ, что $\sigma_n\leqslant\sigma$ для всех n. Поскольку $S_n\leqslant\sigma_n$, то $S_n\leqslant\sigma$ для всех n, т. е. последовательность $\{S_n\}$ также ограничена сверху. Эта последовательность является возрастающей (члены ряда положительны, с увеличением числа слагаемых возрастает сумма), поэтому $\{S_n\}$ имеет предел (см. теорему 10.16), т. е. $\lim\limits_{n\to\infty} S_n=S$, что и требовалось доказать.

Если ряд (21.11) расходится, то ряд (21.12) также будет расходящимся. Допуская противное (ряд (21.12) сходится), получаем, что должен сходиться и ряд (21.11), а это противоречит условию.

Замечание. Теорема верна и в случае, если условие (21.13) выполняется начиная с некоторого номера n_0 (так как отбрасывание конечного числа членов ряда не влияет на его сходимость, см. § 21.1).

Теорема 21.5. Если существует предел

$$\lim_{n\to\infty}\frac{a_n}{b_n}=c \quad (0<c<+\infty), \qquad (21.15)$$

то ряды с положительными членами (21.11) и (21.12) одновременно сходятся или расходятся.

Доказательство. Равенство (21.15) означает, что для любого $\varepsilon>0$ существует такой номер N, что при всех $n>N$ выполняется неравенство

$$\left|\frac{a_n}{b_n}-c\right|<\varepsilon,$$

откуда

$$-\varepsilon<\frac{a_n}{b_n}-c<\varepsilon, \quad c-\varepsilon<\frac{a_n}{b_n}<c+\varepsilon \quad (n>N).$$

Взяв $\varepsilon<c$ и обозначив $c-\varepsilon=m$ $(m>0)$, $c+\varepsilon=M$ $(M>0)$, получим при $n>N$

$$m<\frac{a_n}{b_n}<M, \text{ или } mb_n<a_n<Mb_n. \qquad (21.16)$$

Пусть ряд (21.12) сходится, тогда, очевидно, сходится и ряд

$$\sum_{n=1}^{\infty} Mb_n=Mb_1+Mb_2+Mb_3+\ldots. \qquad (21.17)$$

Отсюда из неравенства $a_n<Mb_n$ (см. (21.16)) по теореме 21.4 получаем, что сходится и ряд (21.11). Если ряд (21.11) сходится, то из неравенства $mb_n<a_n$ и теоремы 21.4 следует, что сходится ряд

$$\sum_{n=1}^{\infty} mb_n=mb_1+mb_2+mb_3+\ldots, \qquad (21.18)$$

тогда, очевидно, сходится и ряд (21.12).

Таким образом, доказано, что из сходимости ряда (21.12) следует сходимость ряда (21.11) и обратно. Утверждение теоремы о расходимости рядов доказывается методом от противного.

§ 21.4. Признак Д'Аламбера. Признак Коши

В этом параграфе также рассматриваются ряды с положительными членами.

Признак Д'Аламбера[1] выражается следующей теоремой.

Теорема 21.6. Если существует предел

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n}=l, \qquad (21.19)$$

то ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится при $l<1$ и расходится при $l>1$.

Доказательство. Равенство (21.19) означает, что для любого $\varepsilon>0$ существует такой номер N, что при всех $n>N$ выполняется неравенство

$$\left|\frac{a_{n+1}}{a_n}-l\right|<\varepsilon$$

[1] *Жан Лерон Д'Аламбер* (*Jean le Rond d'Alembert*, 1717—1783) — французский математик и философ.

или

$$l-\varepsilon<\frac{a_{n+1}}{a_n}<l+\varepsilon. \qquad (21.20)$$

Пусть $l<1$. Выберем $\varepsilon>0$ так, чтобы $l+\varepsilon<1$, и обозначим $l+\varepsilon=q$, $q<1$. Из неравенств (21.20) следует, что $\frac{a_{n+1}}{a_n}<q$, или $a_{n+1}<a_n q$, при $n>N$. Полагая в последнем неравенстве $n=N+1$, $n=N+2$, $n=N+3$, ..., получаем

$$a_{N+2}<a_{N+1}q,\ a_{N+3}<a_{N+2}q<a_{N+1}q^2,$$
$$a_{N+4}<a_{N+3}q<a_{N+1}q^3,\ \ldots.$$

Следовательно, начиная с номера $N+2$ все члены данного ряда не превосходят соответствующих членов геометрической прогрессии. Поскольку эта прогрессия сходится ($q<1$), то сходится остаток после $(N+1)$-го члена; на основании теоремы 21.2 сходится и данный ряд.

Пусть $l>1$. Выберем $\varepsilon>0$ так, чтобы $l-\varepsilon>1$, и обозначим $l-\varepsilon=q$, $q>1$. Из неравенств (21.20) следует, что

$$q<\frac{a_{n+1}}{a_n},\ a_{n+1}>a_n q,\ \ n=N+1,\ N+2,\ \ldots.$$

Это означает, что начиная с номера $N+1$ члены ряда возрастают. В этом случае не выполняется необходимый признак сходимости и ряд расходится.

Признак Коши выражается следующей теоремой.

Теорема 21.7. Если существует предел

$$\lim_{n\to\infty}\sqrt[n]{a_n}=l, \qquad (21.21)$$

то ряд $\sum\limits_{n=1}^{\infty}a_n$ сходится при $l<1$ и расходится при $l>1$.

Доказательство. Условие (21.21) означает: для любого $\varepsilon>0$ существует такой номер N, что при всех $n\geqslant N$ выполняется неравенство $|\sqrt[n]{a_n}-l|<\varepsilon$ или

$$l-\varepsilon<\sqrt[n]{a_n}<l+\varepsilon. \qquad (21.22)$$

Пусть $l<1$. Выберем $\varepsilon>0$ так, чтобы $l+\varepsilon<1$. Обозначим $l+\varepsilon=q$, тогда $q<1$. Из неравенств (21.22) следует, что

$$\sqrt[n]{a_n}<q,\ a_n<q^n \quad (q<1,\ n\geqslant N).$$

Итак, каждый член ряда начиная с номера N меньше соответствующего члена сходящейся геометрической прогрессии ($q<1$), поэтому данный ряд сходится.

Пусть $l>1$. Выберем $\varepsilon>0$ так, чтобы $l-\varepsilon>1$, тогда $q=l-\varepsilon>1$. Из неравенств (21.22) следует, что

$$q<\sqrt[n]{a_n},\ q^n<a_n \quad (q>1,\ n\geqslant N).$$

В этом случае не выполняется необходимый признак сходимости и ряд расходится.

Пример 1. Исследовать, сходится ли ряд $\sum\limits_{n=1}^{\infty}\frac{n^n}{2^n n!}$.

Так как

$$a_n=\frac{n^n}{2^n n!},\ a_{n+1}=\frac{(n+1)^{n+1}}{2^{n+1}(n+1)!},$$

$$\frac{a_{n+1}}{a_n}=\frac{(n+1)^{n+1}}{2^{n+1}(n+1)!}:\frac{n^n}{2^n n!}=\frac{2^n(n+1)^{n+1}n!}{2^{n+1}n^n(n+1)!}=\frac{1}{2}\left(\frac{n+1}{n}\right)^n,$$

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\frac{1}{2}\left(\frac{n+1}{n}\right)^n=\frac{1}{2}\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n=\frac{e}{2}>1,$$

то данный ряд расходится.

Пример 2. Исследовать, сходится ли ряд $\sum\limits_{n=1}^{\infty}\left(\frac{n}{3n+2}\right)^n$.

Поскольку

$$\lim_{n\to\infty}\sqrt[n]{a_n}=\lim_{n\to\infty}\sqrt[n]{\left(\frac{n}{3n+2}\right)^n}=\lim_{n\to\infty}\frac{n}{3n+2}=\frac{1}{3}<1,$$

то ряд сходится.

§ 21.5. Интегральный признак сходимости

Интегральный признак сходимости выражается следующей теоремой.

Теорема 21.8. Дан ряд с положительными членами

$$\sum_{n=1}^{\infty}b_n \qquad (21.23)$$

и несобственный интеграл

$$\int_1^{\infty} f(x)\,dx. \qquad (21.24)$$

Если при $x \geqslant 1$ функция $f(x)$ непрерывна, положительна и не возрастает, а в точках $x=n$ принимает значения $f(n)=b_n$, то ряд (21.23) и несобственный интеграл (21.24) одновременно сходятся или расходятся.

Доказательство. Если $n<x<n+1$, то $f(n) \geqslant f(x) \geqslant f(n+1)$, откуда

$$\int_n^{n+1} f(n)\,dx \geqslant \int_n^{n+1} f(x)\,dx \geqslant \int_n^{n+1} f(n+1)\,dx;$$

$$f(n) \geqslant \int_n^{n+1} f(x)\,dx \geqslant f(n+1). \qquad (21.25)$$

Суммируя члены неравенств (21.25) от $n=1$ до $n=m$, получаем

$$\sum_{n=1}^{m} f(n) \geqslant \int_1^{m+1} f(x)\,dx \geqslant \sum_{n=1}^{m} f(n+1);$$

$$\sum_{n=1}^{m} f(n) \geqslant \int_1^{m+1} f(x)\,dx \geqslant \sum_{n=1}^{m+1} f(n) - f(1),$$

т. е.

$$\sum_{n=1}^{m} b_n \geqslant \int_1^{m+1} f(x)\,dx \geqslant \sum_{n=1}^{m+1} b_n - b_1$$

или

$$S_m \geqslant \int_1^{m+1} f(x)\,dx \geqslant S_{m+1} - b_1. \qquad (21.26)$$

Если интеграл (21.24) сходится и $\int_1^{\infty} f(x)\,dx = I$, то $\int_1^{m+1} f(x)\,dx \leqslant I$ при любом натуральном m. Следовательно,

$$S_{m+1} \leqslant b_1 + \int_1^{m+1} f(x)\,dx \leqslant b_1 + I = C$$

или

$$S_n \leqslant C \quad (n=1, 2, 3, \ldots, C>0).$$

Так как $\{S_n\}$ — монотонно возрастающая и ограниченная последовательность, то существует $\lim\limits_{n\to\infty} S_n=S$, т. е. ряд (21.23) также сходится. Если ряд (21.23) сходится и $\sum\limits_{n=1}^{\infty} b_n=B$, то $S_m \leqslant B$ при любом m. Из неравенств (21.26) следует, что $\int\limits_1^{m+1} f(x)\,dx \leqslant B$ при любом m. Несобственный интеграл также сходится.

С помощью интегрального признака можно доказать, что ряд

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = 1+\frac{1}{2^p}+\frac{1}{3^p}+\ldots+\frac{1}{n^p}+\ldots, \qquad (21.27)$$

где p — любое вещественное число, сходится при $p>1$ и расходится при $p \leqslant 1$.

Действительно, $\int\limits_1^{\infty} \frac{dx}{x^p}$ сходится при $p>1$ и расходится при $p \leqslant 1$ (см. § 16.10). Этот ряд называется *рядом Дирихле*.

Замечание. Теорема верна и в случае, если нижний предел в несобственном интеграле (21.24) равен a $(a>0)$ и равенство $f(n)=b_n$ выполняется для $n \geqslant m$ $(m \geqslant a)$.

§ 21.6. Знакочередующиеся ряды. Признак Лейбница

Знакочередующимся рядом называется ряд, у которого любых два члена с номерами n и $n+1$ $(n=1, 2, 3, \ldots)$ имеют противоположные знаки, т. е. ряд вида

$$\sum_{n=1}^{\infty} (-1)^{n+1} a_n = a_1-a_2+a_3-a_4+\ldots+(-1)^{n+1}a_n+\ldots, \qquad (21.28)$$

где $a_n>0$ $(n=1, 2, 3, \ldots)$.

Теорема 21.9 (признак Лейбница). Знакочередующийся ряд (21.28) сходится, если модули его членов убывают с возрастанием n и общий член стремится к нулю, т. е.

$$a_{n+1}<a_n \quad (n=1, 2, 3, \ldots) \tag{21.29}$$

и

$$\lim_{n\to\infty} a_n=0. \tag{21.30}$$

Доказательство. Рассмотрим частичные суммы ряда (21.28) с четными и нечетными номерами:

$$S_{2m}=a_1-a_2+a_3-\ldots+a_{2m-1}-a_{2m},$$
$$S_{2m+1}=a_1-a_2+a_3-\ldots+a_{2m-1}-a_{2m}+a_{2m+1}.$$

Преобразуем первую из этих сумм:

$$S_{2m}=(a_1-a_2)+(a_3-a_4)+\ldots+(a_{2m-1}-a_{2m}),$$
$$S_{2m}=a_1-(a_2-a_3)-(a_4-a_5)-\ldots-(a_{2m-2}-a_{2m+1})-a_{2m}.$$

В силу условия (21.29) разность в каждой скобке положительна, поэтому $S_{2m}>S_{2m-2}$ и $S_{2m}<a_1$ для всех m. Итак, последовательность четных частичных сумм $\{S_{2m}\}$ является монотонно возрастающей и ограниченной. Она имеет предел, который обозначим через S, т. е. $\lim\limits_{m\to\infty} S_{2m}=S$.

Поскольку $S_{2m+1}=S_{2m}+a_{2m+1}$, то, принимая во внимание предыдущее равенство и условие (21.30), получаем

$$\lim_{m\to\infty} S_{2m+1}=\lim_{m\to\infty}(S_{2m}+a_{2m+1})=$$
$$=\lim_{m\to\infty} S_{2m}+\lim_{m\to\infty} a_{2m+1}=S+0=S.$$

Итак, последовательности частичных сумм данного ряда соответственно с четными и нечетными номерами имеют один и тот же предел S. Отсюда следует, что последовательность всех частичных сумм ряда имеет предел S; $\lim\limits_{n\to\alpha} S_n=S$, т. е. ряд сходится.

Пример. Исследовать, сходится ли ряд

$$\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots+\frac{(-1)^{n+1}}{n}+\ldots. \tag{21.31}$$

Этот ряд является знакочередующимся. Он сходится, поскольку удовлетворяет условиям теоремы

$$\frac{1}{n+1}<\frac{1}{n} \quad (n=1, 2, \ldots), \quad \lim_{n\to\infty}\frac{1}{n}=0.$$

Оценка остатка знакочередующегося ряда определяется с помощью следующей теоремы.

Теорема 21.10. Сумма остатка знакочередующегося ряда, удовлетворяющего условиям теоремы Лейбница, имеет знак первого оставшегося члена и не превосходит его по модулю.

Доказательство. Рассмотрим остаток ряда (21.28) после $2m$ членов. Пусть r_{2m} — его сумма, σ_n — n-я частичная сумма, тогда

$$r_{2m}=a_{2m+1}-a_{2m+2}+a_{2m+3}-a_{2m+4}+a_{2m+5}-\ldots,$$
$$\sigma_{2n}=(a_{2m+1}-a_{2m+2})+(a_{2m+3}-a_{2m+4})+\ldots+$$
$$+(a_{2m+2n-1}-a_{2m+2n}),$$
$$\sigma_{2n}=a_{2m+1}-(a_{2m+2}-a_{2m+3})-\ldots-a_{2m+2n}.$$

Так как выполнены условия теоремы 21.9, то $\sigma_{2n}>0$ и $\sigma_{2n}<a_{2m+1}$ при всех n, т. е. $0<\sigma_{2n}<a_{2m+1}$ $(n=1, 2, 3, \ldots)$, откуда

$$\lim_{n\to\infty}\sigma_{2n}\leqslant a_{2m+1} \quad \text{или} \quad r_{2m}\leqslant a_{2m+1}.$$

Аналогично доказывается, что сумма r_{2m-1} остатка ряда после $2m-1$ членов $(r_{2m-1}=-a_{2m}+a_{2m+1}-a_{2m+2}+$ $+a_{2m+3}-\ldots)$ удовлетворяет условиям $0<-r_{2m-1}<a_{2m}$, т. е. $r_{2m-1}<0$ и $|r_{2m-1}|\leqslant a_{2m}$.

Следовательно, независимо от четности или нечетности n

$$|r_n|\leqslant a_{n+1}. \tag{21.32}$$

§ 21.7. Абсолютная сходимость рядов

В этом параграфе будем изучать ряды, члены которых являются действительными числами любого знака. Пусть дан такой ряд

$$\sum_{k=1}^{\infty}u_k=u_1+u_2+u_3+\ldots+u_k+\ldots. \tag{21.33}$$

Рассмотрим ряд, составленный из модулей членов данного ряда:

$$\sum_{k=1}^{\infty}|u_k|=|u_1|+|u_2|+|u_3|+\ldots+|u_k|+\ldots. \tag{21.34}$$

Теорема 21.11. Если ряд (21.34) сходится, то сходится и ряд (21.33).

Доказательство. Поскольку ряд (21.34) сходится, то в силу критерия Коши (теорема 21.1) для любого $\varepsilon>0$ существует такой номер $N=N(\varepsilon)$, что при всех $n>N$ и любом целом $p\geqslant 0$ выполняется неравенство

$$\left|\sum_{k=n}^{n+p}|u_k|\right|=\sum_{k=n}^{n+p}|u_k|<\varepsilon.$$

Так как

$$\left|\sum_{k=n}^{n+p}u_k\right|\leqslant\sum_{k=n}^{n+p}|u_k|,$$

то $\left|\sum\limits_{k=n}^{n+p}u_k\right|<\varepsilon$. Это означает, что ряд (21.33) также сходится.

Замечание. Из сходимости ряда (21.33) не следует сходимость ряда (21.34). Например, ряд $\sum\limits_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}$ сходится (см. § 21.6), а ряд из модулей его членов расходится (гармонический ряд, см. § 21.2).

Знакопеременный ряд называется *абсолютно сходящимся*, если сходится ряд из модулей его членов. Например, ряд

$$\sum_{k=0}^{\infty}\frac{(-1)^{k+1}}{2^k}=1-\frac{1}{2}+\frac{1}{2^2}-\frac{1}{2^3}+\ldots$$

является абсолютно сходящимся, поскольку сходится ряд из модулей его членов, т. е. ряд $\sum\limits_{k=0}^{\infty}\frac{1}{2^k}$ (геометрическая прогрессия со знаменателем $q=0,5$, $|q|<1$).

Знакопеременный ряд называется *неабсолютно сходящимся (условно сходящимся)*, если он сходится, а ряд из модулей его членов расходится. Например, ряд $\sum\limits_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}$ является неабсолютно сходящимся (см. замечание).

§ 21.8. Действия над рядами

Произведением ряда

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \ldots + a_n + \ldots \tag{21.35}$$

на число c называется ряд

$$\sum_{n=1}^{\infty} ca_n = ca_1 + ca_2 + ca_3 + \ldots + ca_n + \ldots . \tag{21.36}$$

Теорема 21.12. Если ряд (21.35) сходится, то ряд (21.36) также сходится, причем

$$\sum_{n=1}^{\infty} ca_n = c \sum_{n=1}^{\infty} a_n. \tag{21.37}$$

Доказательство. Обозначим через s_n и σ_n n-е частичные суммы рядов (21.35) и (21.36), т. е.

$$s_n = a_1 + a_2 + \ldots + a_n, \quad \sigma_n = ca_1 + ca_2 + \ldots + ca_n.$$

Очевидно, $\sigma_n = cs_n$. Если ряд (21.35) сходится и его сумма равна A, т. е. $\lim\limits_{n\to\infty} s_n = A$, $\sum\limits_{n=1}^{\infty} a_n = A$, то

$$\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} (cs_n) = c \lim_{n\to\infty} s_n = cA,$$

т. е.

$$\sum_{n=1}^{\infty} ca_n = cA = c \sum_{n=1}^{\infty} a_n.$$

Кроме ряда (21.35) рассмотрим ряд

$$\sum_{n=1}^{\infty} b_n = b_1 + b_2 + \ldots + b_n + \ldots . \tag{21.38}$$

Суммой двух рядов (21.35) и (21.38) называется ряд

$$\sum_{n=1}^{\infty} (a_n + b_n) = (a_1 + b_1) + (a_2 + b_2) + \ldots$$
$$\ldots + (a_n + b_n) + \ldots . \tag{21.39}$$

Аналогично определяется *разность* рядов (21.35) и (21.38):

$$\sum_{n=1}^{\infty} (a_n - b_n) = (a_1 - b_1) + (a_2 - b_2) + \ldots$$

$$\ldots + (a_n - b_n) + \ldots . \qquad (21.40)$$

Теорема 21.13. Если сходятся ряды (21.35) и (21.38), то сходятся их сумма и разность, причем

$$\left. \begin{aligned} \sum_{n=1}^{\infty} (a_n + b_n) &= \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n; \\ \sum_{n=1}^{\infty} (a_n - b_n) &= \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} b_n. \end{aligned} \right\} \qquad (21.41)$$

Доказательство. Обозначим соответственно через s_n, s_n', σ_n n-е частичные суммы рядов (21.35), (21.38), (21.39):

$$s_n = a_1 + a_2 + \ldots + a_n, \quad s_n' = b_1 + b_2 + \ldots + b_n,$$
$$\sigma_n = (a_1 + b_1) + (a_2 + b_2) + \ldots + (a_n + b_n).$$

Очевидно, $\sigma_n = s_n + s_n'$. Если ряды (21.35) и (21.38) сходятся, т. е.

$$\lim_{n\to\infty} s_n = A, \ A = \sum_{n=1}^{\infty} a_n, \ \lim_{n\to\infty} s_n' = B, \ B = \sum_{n=1}^{\infty} b_n,$$

то

$$\lim_{n\to\infty} \sigma_n = \lim_{n\to\infty} s_n + \lim_{n\to\infty} s_n' = A + B, \ A + B = \sum_{n=1}^{\infty} (a_n + b_n)$$

или

$$\sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n.$$

Аналогично доказывается второе из равенств (21.41).

Произведением двух рядов (21.35) и (21.38) называется ряд

$$\sum_{n=1}^{\infty} c_n = c_1 + c_2 + \ldots + c_n + \ldots , \qquad (21.42)$$

где

$$c_n = a_1 b_n + a_2 b_{n-1} + \ldots + a_n b_1 \ (n = 1, 2, 3, \ldots). \qquad (21.43)$$

Отметим без доказательства, что если ряды (21.35) и (21.38) сходятся абсолютно, то их произведение, т. е. ряд

$$\sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} (a_1 b_n + a_2 b_{n-1} + \ldots + a_n b_1), \qquad (21.44)$$

также сходится абсолютно и его сумма равна AB.

Доказательство этого утверждения приведено, например, в [14].

Замечание. Правила действий над рядами не всегда совпадают с правилами действий над конечными суммами. В частности, в конечных суммах можно произвольно менять порядок слагаемых, как угодно группировать члены, сумма от этого не изменится. Слагаемые конечной суммы можно складывать в обратном порядке, для ряда такой возможности нет, ибо у него не существует последнего члена.

В ряде не всегда можно группировать члены. Например, ряд

$$1-1+1-1+\ldots+(-1)^{k-1}+\ldots$$

является расходящимся, так как

$$s_{2m}=0 \quad (m=1,\,2,\,3,\,\ldots),\quad s_{2m+1}=1 \quad (m=0,\,1,\,2,\,\ldots)$$

и нет предела его частичных сумм. После группировки членов

$$(1-1)+(1-1)+\ldots+(1-1)+\ldots=0+0+\ldots+0+\ldots=0$$

получаем сходящийся ряд, его сумма равна нулю. При другой группировке членов

$$1-(1-1)-(1-1)-\ldots-(1-1)-\ldots=1-0-0-\ldots-0-\ldots=1$$

получаем сходящийся ряд, сумма которого равна единице.

Приведем без доказательства две теоремы (доказательство их можно найти, например, в книгах [11, 14]).

Теорема 21.14. Перестановка членов абсолютно сходящегося ряда не нарушает его сходимости, сумма ряда при этом остается прежней.

Теорема 21.15. Если ряд сходится неабсолютно, то путем надлежащей перестановки его членов всегда можно придать сумме ряда произвольное значение и даже сделать ряд расходящимся.

Глава 22

ФУНКЦИОНАЛЬНЫЕ РЯДЫ

В этой главе изучаются функциональные ряды, степенные ряды, их свойства и приложения.

§ 22.1. Сходимость функциональных последовательностей и рядов

Рассмотрим последовательность функций

$$f_1(x),\ f_2(x),\ \ldots,\ f_n(x),\ \ldots\,. \tag{22.1}$$

Зафиксировав $x=x_0$, получим числовую последовательность

$$f_1(x_0),\ f_2(x_0),\ \ldots,\ f_n(x_0),\ \ldots. \qquad (22.2)$$

Функциональная последовательность $\{f_n(x)\}$ называется *сходящейся в точке* x_0, если сходится соответствующая числовая последовательность $\{f_n(x_0)\}$.

Последовательность (22.1) называется *сходящейся к функции* $f(x)$ *на множестве* X, если для всех $x \in X$

$$\lim_{n\to\infty} f_n(x) = f(x). \qquad (22.3)$$

Последнее равенство означает следующее: для каждого $x \in X$ и любого числа $\varepsilon > 0$ можно указать такой номер N, что при всех $n > N$ выполняется неравенство

$$|f(x) - f_n(x)| < \varepsilon. \qquad (22.4)$$

Очевидно, номер N зависит от значения $x \in X$ и числа $\varepsilon > 0$, т. е. $N = N(\varepsilon, x)$.

Пусть дан функциональный ряд

$$\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots. \qquad (22.5)$$

Зафиксируем некоторое значение $x = x_0$, получим числовой ряд

$$\sum_{n=1}^{\infty} u_n(x_0) = u_1(x_0) + u_2(x_0) + \ldots + u_n(x_0) + \ldots. \qquad (22.6)$$

Если числовой ряд (22.6) сходится, то значение x_0 называется *точкой сходимости* функционального ряда (22.5). Множество всех точек сходимости функционального ряда называется *областью его сходимости.* Если числовой ряд (22.6) расходится, то x_0 называется *точкой расходимости* функционального ряда (22.1).

Функциональный ряд называется *сходящимся на некотором множестве,* если он сходится в любой точке этого множества.

Функциональный ряд (22.5) называется *абсолютно сходящимся* на множестве X, если на нем сходится ряд из модулей его членов:

$$\sum_{n=1}^{\infty} |u_n(x)| = |u_1(x)| + |u_2(x)| + \ldots + |u_n(x)| + \ldots. \qquad (22.7)$$

Поскольку каждой точке x_0 сходимости ряда (22.5) ставится в соответствие определенное значение суммы ряда (22.6), то сумма сходящегося на множестве X функционального ряда (22.5) является функцией переменной x. Обозначим эту функцию через $S(x)$, тогда

$$S(x)=\lim_{n\to\infty} S_n(x), \qquad (22.8)$$

где $S_n(x)$ — n-я частичная сумма ряда (22.5), т. е.

$$S_n(x)=u_1(x)+u_2(x)+\ldots+u_n(x), \qquad (22.9)$$

$$S(x)=\sum_{n=1}^{\infty} u_n(x).$$

Остатком функционального ряда (22.5) после n-го члена (или n-м остатком) называется ряд, полученный из данного отбрасыванием n его первых членов.

Отметим, что функциональный ряд (22.5) и любой его остаток на множестве X одновременно сходятся или расходятся.

Пусть функциональный ряд (22.5) сходится на множестве X, $S(x)$ — его сумма, S_n — n-я частичная сумма. Тогда на множестве X сходится и остаток данного ряда после n-го члена. Сумму остатка обозначим через $r_n(x)$:

$$r_n(x)=u_{n+1}(x)+u_{n+2}(x)+u_{n+3}(x)+\ldots. \qquad (22.10)$$

Очевидно, для всех $x\in X$

$$S(x)=S_n(x)+r_n(x). \qquad (22.11)$$

Отсюда следует, что для всех $x\in X$

$$\lim_{n\to\infty} r_n(x)=0. \qquad (22.12)$$

Замечание. Каждому функциональному ряду (22.5) можно поставить в соответствие функциональную последовательность, т. е. последовательность его частичных сумм (22.9):

$$S_n(x)=\sum_{k=1}^{n} u_k(x), \quad n=1,\ 2,\ 3,\ \ldots.$$

Обратно, каждой функциональной последовательности (22.1) соответствует функциональный ряд

$$f_1(x)+[f_2(x)-f_1(x)]+\ldots+[f_n(x)-f_{n-1}(x)]+\ldots, \qquad (22.13)$$

для которого частичные суммы будут равны соответствующим членам последовательности $\{f_n(x)\}$, т. е. $S_n(x)=f_n(x)$.

Это дает возможность каждую теорему, доказанную для функциональных рядов, перефразировать в соответствующую теорему для функциональных последовательностей, и наоборот.

§ 22.2. Равномерная сходимость функциональных последовательностей и рядов

Введем два эквивалентных определения равномерной сходимости последовательности функций $\{f_n(x)\}$ к функции $f(x)$ на множестве X. Обозначим через ρ_n верхнюю грань модуля разности $f(x)-f_n(x)$ на этом множестве, т. е.

$$\rho_n = \sup_{x\in X} |f(x)-f_n(x)|. \qquad (22.14)$$

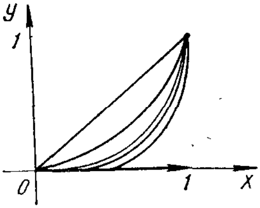

Рис. 22.1

Функциональная последовательность $\{f_n(x)\}$ называется *равномерно сходящейся* к $f(x)$ на множестве X, если для любого числа $\varepsilon>0$ можно указать такой номер $N=N(\varepsilon)$ (не зависящий от x, а только от ε), что при $n>N$ и всех $x\in X$ выполняется неравенство

$$|f(x)-f_n(x)|<\varepsilon. \qquad (22.15)$$

Функциональная последовательность $\{f_n(x)\}$ называется *равномерно сходящейся* к $f(x)$ на множестве X, если

$$\rho_n\to 0 \text{ при } n\to\infty, \qquad (22.16)$$

где ρ_n определяется формулой (22.14).

Легко видеть, что из первого определения следует второе и обратно.

Пример 1. Исследовать сходимость функциональиой последовательности $f_n(x)=x^n$, $n=1, 2, 3, \ldots$, на отрезке $[0, 1]$.

Если $0\leqslant x<1$, то $x^n\to 0$ при $n\to\infty$; если $x=1$, то $f_n(1)=1$. Последовательность $\{x^n\}$ сходится к функции

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

На отрезке $[0, 1]$ сходимость будет неравномерной, так как ρ_n не стремится к нулю при $n\to\infty$. Действительно,

$$\rho_n=\sup_{x\in[0,1]}|f(x)-f_n(x)|=1,$$

при $n\to\infty$ ρ_n не может стремиться к нулю.

На любом отрезке $[0, q]$, где $q<1$, последовательность $\{x^n\}$ равномерно стремится к функции $f(x)\equiv 0$. В самом деле,

$$\rho_n=\sup_{x\in[0,q]}|f(x)-f_n(x)|=\sup_{x\in[0,q]}|x^n|=q^n,\ \rho_n=q^n;$$

$q^n \to 0$ при $n \to \infty$; $\rho_n \to 0$ при $n \to \infty$. На рис. 22.1 изображены графики функций $f_n(x) = x^n$ ($n = 1, 2, 3, 4, 5$) и график функции $f(x) \equiv 0$.

Рассмотрим сходящийся на множестве X функциональный ряд

$$\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + \ldots + u_n(x) + \ldots, \tag{22.17}$$

для которого

$$\left.\begin{aligned} & s_n(x) = u_1(x) + u_2(x) + \ldots + u_n(x), \\ & s(x) = \lim_{n\to\infty} s_n(x), \\ & s(x) = s_n(x) + r_n(x), \\ & r_n(x) = s(x) - s_n(x); \end{aligned}\right\} \tag{22.18}$$

$$\rho_n = \sup_{x\in X} |r_n(x)| = \sup_{x\in X} |s(x) - s_n(x)|. \tag{22.19}$$

Функциональный ряд называется *равномерно сходящимся* на некотором множестве, если последовательность его частичных сумм сходится равномерно на этом множестве.

Принимая во внимание определения равномерной сходимости функциональной последовательности, можно дать два эквивалентных определения равномерной сходимости функционального ряда.

Функциональный ряд (22.17) называется равномерно сходящимся на множестве X, если для любого числа $\varepsilon > 0$ можно указать такой номер $N = N(\varepsilon)$, что при $n > N$ и всех $x \in X$ выполняется неравенство

$$|s(x) - s_n(x)| < \varepsilon \text{ или } |r_n(x)| < \varepsilon. \tag{22.20}$$

Функциональный ряд (22.17) называется равномерно сходящимся на множестве X, если

$$\rho_n \to 0 \text{ при } n \to \infty, \tag{22.21}$$

где ρ_n определяется формулой (22.19).

Пример 2. Исследовать, равномерно ли сходится ряд $\sum\limits_{n=1}^{\infty} (x^{n-1} - x^n)$ на отрезке $[0, 1]$.

Поскольку

$$s_n(x) = (1-x) + (x - x^2) + \ldots + (x^{n-1} - x^n), \quad s_n(x) = 1 - x^n,$$

то

$$\lim_{n\to\infty} s_n(x) = \begin{cases} 0 & (x=1), \\ 1 & (0 \leqslant x < 1), \end{cases} \qquad s(x) = \begin{cases} 0 & (x=1), \\ 1 & (0 \leqslant x < 1). \end{cases}$$

Ряд сходится на отрезке $[0, 1]$, но неравномерно. Действительно,

$$\rho_n = \sup_{x\in[0,\,1]} |s(x) - s_n(x)| = 1$$

и не может стремиться к нулю при $n \to \infty$.

Отметим, что ряд сходится равномерно на любом отрезке $[0, q]$, где $q < 1$. В самом деле,

$$\rho_n = \sup_{x\in[0,\,q]} |s(x) - s_n(x)| = \sup_{x\in[0,\,q]} |x^n| = q^n, \quad \rho_n = q^n;$$

$q^n \to 0$ при $n \to \infty$; $\rho_n \to 0$ при $n \to \infty$.

Критерий равномерной сходимости функционального ряда (критерий Коши) выражается теоремой, приводимой здесь без доказательства.

Теорема 22.1. Функциональный ряд $\sum\limits_{n=1}^{\infty} u_n(x)$ равномерно сходится на множестве X тогда и только тогда, когда для любого числа $\varepsilon > 0$ существует такой номер $N = N(\varepsilon)$, что при $n > N$, любом натуральном p и всех $x \in X$ выполняется неравенство

$$|u_n(x) + u_{n+1}(x) + \ldots + u_{n+p}(x)| < \varepsilon$$

или

$$\left| \sum_{k=n}^{n+p} u_k(x) \right| < \varepsilon. \tag{22.22}$$

Доказательство этой теоремы имеется, например, в [14].

Достаточный признак равномерной сходимости функционального ряда выражается следующей теоремой.

Теорема 22.2. Если члены функционального ряда $\sum\limits_{n=1}^{\infty} u_n(x)$ определены на множестве X и по модулю не превосходят соответствующих членов сходящегося числового ряда с положительными членами $\sum\limits_{n=1}^{\infty} a_n$ $(a_n > 0)$, т. е. для всех $x \in X$

$$|u_n(x)| \leqslant a_n \quad (n = 1, 2, 3, \ldots), \tag{22.23}$$

то этот функциональный ряд равномерно сходится на множестве X.

Доказательство. Если ряд $\sum\limits_{n=1}^{\infty} a_n$ $(a_n > 0)$ сходится, то в соответствии с критерием Коши для любого $\varepsilon > 0$ существует такой номер N, что $\sum\limits_{k=n}^{n+p} a_k < \varepsilon$ для всех $n > N$ и любых целых $p \geqslant 0$. Отсюда и из условия (22.23) следует, что для всех $n > N$ и $x \in X$

$$\left|\sum_{k=n}^{n+p} u_k(x)\right| \leqslant \sum_{k=n}^{n+p} |u_k(x)| \leqslant \sum_{k=n}^{n+p} a_k < \varepsilon.$$

Это означает, что ряд $\sum\limits_{n=1}^{\infty} u_n(x)$ сходится равномерно на множестве X.

Доказанная теорема называется теоремой Вейерштрасса[1].

Замечание. Функциональный ряд, удовлетворяющий условиям теоремы Вейерштрасса, называется *мажорируемым,* соответствующий числовой ряд — *мажорантным.*

Пример 3. Исследовать, равномерно ли сходится ряд $\sum\limits_{n=1}^{\infty} \dfrac{\sin nx}{n^3}$.

Так как $\left|\dfrac{\sin nx}{n^3}\right| \leqslant \dfrac{1}{n^3}$ $(n = 1, 2, 3, \ldots)$ при любом x и сходится числовой ряд $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^3}$ (ряд Дирихле, см. (21.27)), то данный функциональный ряд сходится равномерно в интервале $]-\infty, +\infty[$.

§ 22.3. Свойства равномерно сходящихся рядов

Равномерно сходящиеся ряды обладают важными свойствами, которые выражаются рассматриваемыми в этом параграфе теоремами.

Теорема 22.3. Если ряд $\sum\limits_{n=1}^{\infty} u_n(x)$ сходится равномерно в промежутке X, на котором его члены $u_n(x)$ непрерывны, то и сумма ряда $s(x)$ непрерывна в этом промежутке.

[1] *Карл Вейерштрасс* (*Karl Weierstraß,* 1815—1897) — немецкий математик.

Доказательство. Зафиксируем произвольное значение $x_0 \in X$. Если $s_n(x)$ — n-я частичная сумма данного ряда, $r_n(x)$ — остаток ряда после n-го члена, то

$$s(x) = s_n(x) + r_n(x), \quad s(x_0) = s_n(x_0) + r_n(x_0).$$

Вычитая почленно второе равенство из первого, получаем

$$s(x) - s(x_0) = s_n(x) - s_n(x_0) + r_n(x) - r_n(x_0),$$

откуда

$$|s(x) - s(x_0)| \leqslant |s_n(x) - s_n(x_0)| + \\ + |r_n(x)| + |r_n(x_0)|. \tag{22.24}$$

Поскольку $u_n(x)$ — функции, непрерывные в промежутке X, то и любая их конечная сумма $s_n(x)$ также непрерывна в этом промежутке; задав любое число $\varepsilon > 0$, можно указать такое $\delta > 0$, что при $|x - x_0| < \delta$ будет выполняться неравенство $|s_n(x) - s_n(x_0)| < \varepsilon/3$. Так как ряд сходится равномерно, то для любого $\varepsilon > 0$ можно указать такой номер $N = N(\varepsilon)$, что при всех $n > N$ и всех $x \in X$ выполняется неравенство $|r_n(x)| < \varepsilon/3$, в частности, $|r_n(x_0)| < \varepsilon/3$.

Итак, для любого $\varepsilon > 0$ можно указать такое $\delta > 0$, что при всех x, удовлетворяющих условию $|x - x_0| < \delta$, выполняется неравенство

$$|s(x) - s(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \quad |s(x) - s(x_0)| < \varepsilon.$$

Это означает, что функция $s(x)$ непрерывна в точке x_0.

Замечание. Доказанное утверждение можно выразить формулой

$$\lim_{x \to x_0} \sum_{n=1}^{\infty} u_n(x) = \lim_{x \to x_0} s(x) = s(x_0) = \sum_{n=1}^{\infty} u_n(x_0) = \sum_{n=1}^{\infty} \lim_{x \to x_0} u_n(x),$$

$$\lim_{x \to x_0} \sum_{n=1}^{\infty} u_n(x) = \sum_{n=1}^{\infty} \lim_{x \to x_0} u_n(x). \tag{22.25}$$

Эта формула означает, что в условиях теоремы возможен почленный переход к пределу.

Теорема 22.4. Если функции $u_n(x)$ $(n = 1, 2, 3, \ldots)$ непрерывны на отрезке $[a, b]$ и ряд $\sum\limits_{n=1}^{\infty} u_n(x) = s(x)$ сходится равномерно на $[a, b]$, то ряд, полученный интегрированием членов данного ряда, также сходится равномерно на $[a, b]$, причем

$$\int_{x_0}^{x} s(t)\,dt = \sum_{n=1}^{\infty} \int_{x_0}^{x} u_n(t)\,dt$$

или

$$\int_{x_0}^{x} \sum_{n=1}^{\infty} u_n(t)\,dt = \sum_{n=1}^{\infty} \int_{x_0}^{x} u_n(t)\,dt, \tag{22.26}$$

где $a \leqslant x_0 \leqslant x \leqslant b$.

Доказательство. Пусть

$$s_n(x) = \sum_{k=1}^{n} u_k(x), \quad r_n(x) = s(x) - s_n(x), \quad \rho_n = \sup_{x \in [a,b]} |r_n(x)|,$$

тогда в соответствии с условием $\rho_n \to 0$ при $n \to \infty$.

Так как

$$\left| \int_{x_0}^{x} s(t)\,dt - \sum_{k=1}^{n} \int_{x_0}^{x} u_k(t)\,dt \right| = \left| \int_{x_0}^{x} s(t)\,dt - \int_{x_0}^{x} \sum_{k=1}^{n} u_k(t)\,dt \right| =$$

$$= \left| \int_{x_0}^{x} s(t)\,dt - \int_{x_0}^{x} s_n(t)\,dt \right| = \left| \int_{x_0}^{x} [s(t) - s_n(t)]\,dt \right| \leqslant$$

$$\leqslant \int_{a}^{b} |s(t) - s_n(t)|\,dt \leqslant (b-a)\,\rho_n \to 0 \text{ при } n \to \infty,$$

то выполняется (22.26).

Теорема 22.5. Пусть функции $u_n(x)$ $(n=1, 2, 3, \ldots)$ определены на отрезке $[a, b]$ и имеют на нем непрерывные производные $u_n'(x)$. Если на этом отрезке сходится ряд

$$\sum_{n=1}^{\infty} u_n(x) = u_1(x) + u_2(x) + u_3(x) + \ldots \tag{22.27}$$

и равномерно сходится ряд, составленный из производных:

$$\sum_{n=1}^{\infty} u_n'(x) = u_1'(x) + u_2'(x) + u_3'(x) + \ldots, \tag{22.28}$$

то сумма $S(x)$ ряда (22.27) имеет производную, равную сумме ряда (22.28), т. е.

$$S'(x) = \sum_{n=1}^{\infty} u_n'(x) \text{ или } \left[\sum_{n=1}^{\infty} u_n(x)\right]' = \sum_{n=1}^{\infty} u_n'(x). \quad (22.29)$$

Доказательство. Сумму ряда (22.28) обозначим через $f(x)$:

$$f(x) = \sum_{n=1}^{\infty} u_n'(x). \quad (22.30)$$

Так как ряд (22.28) сходится равномерно на отрезке $[a, b]$, то $f(x)$ — непрерывная функция. В соответствии с теоремой 22.4 ряд (22.28) можно интегрировать почленно:

$$\int_a^x f(t)\,dt = \sum_{n=1}^{\infty} \int_a^x u_n'(t)\,dt \quad (a \leqslant x \leqslant b).$$

Применяя теорему Ньютона — Лейбница, получаем

$$\int_a^x f(t)\,dt = \sum_{n=1}^{\infty} [u_n(x) - u_n(a)] = S(x) - S(a),$$

откуда

$$S(x) = \int_a^x f(t)\,dt + S(a).$$

На основании теоремы 16.3 (см. § 16.6) находим $S'(x) = f(x)$. Отсюда и из равенства (22.30) следует равенство (22.29).

Замечание. Равенство (22.29) означает, что ряд можно почленно дифференцировать.

§ 22.4. Степенные ряды. Теорема Абеля. Интервал и радиус сходимости степенного ряда

Степенным рядом называется функциональный ряд вида

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \ldots, \quad (22.31)$$

где a_n — вещественные числа, называемые *коэффициентами ряда.*

Степенным рядом называют также ряд

$$\sum_{n=0}^{\infty} a_n (x-a)^n = a_0 + a_1(x-a) + a_2(x-a)^2 + \ldots \quad (22.32)$$

Ряд (22.32) сводится к ряду (22.31) заменой переменной по формуле $x-a=X$.

Теорема 22.6. Если степенной ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится при некотором значении $x_0 \neq 0$, то он абсолютно сходится при любом x, для которого $|x| < |x_0|$.

Доказательство. По условию $\sum_{n=0}^{\infty} a_n x_0^n$ сходится, поэтому $\lim\limits_{n\to\infty} a_n x_0^n = 0$. Следовательно, существует такое число $C > 0$, что для всех n выполняется неравенство

$$|a_n x_0^n| < C. \quad (22.33)$$

Так как

$$|a_n x^n| = \left| a_n x_0^n \left(\frac{x}{x_0}\right)^n \right| = |a_n x_0^n| \left|\frac{x}{x_0}\right|^n < Cq^n,$$

$$q = \left|\frac{x}{x_0}\right|$$

и ряд $\sum_{n=0}^{\infty} Cq^n$ сходится при $|q| < 1$, то сходится абсолютно и данный ряд при $\left|\frac{x}{x_0}\right| < 1$ или при $|x| < |x_0|$. Доказанная теорема называется теоремой Абеля[1].

Следствие. Если степенной ряд расходится при некотором значении x_1, то он расходится и при любом x, для которого $|x| > |x_1|$. Действительно, допустив противное (ряд сходится при значении x таком, что $|x| > |x_1|$), по теореме Абеля получим, что ряд сходится и при значении x_1, что противоречит условию.

Из теоремы Абеля следует, что если степенной ряд сходится при $x_0 \neq 0$, то он сходится при любом x из интервала $]-|x_0|, |x_0|[$; если расходится при $x = x_1$, то расходится вне интервала $]-|x_1|, |x_1|[$, т. е. при $x < -|x_1|$ и $x > |x_1|$.

Радиусом сходимости степенного ряда (22.31) назы-

[1] *Нильс Хенрик Абель* (*Niels Henrik Abel*, 1802—1829) — норвежский математик.

вается число R такое, что при $|x|<R$ ряд сходится, а при $|x|>R$ расходится. *Интервалом сходимости* ряда (22.31) называется интервал $]-R,\ R[$, где R — радиус сходимости. Существование радиуса сходимости можно доказать с помощью теоремы Абеля.

Замечание. Если ряд (22.31) сходится в единственной точке, то считают $R=0$; если ряд сходится при любом x, то полагают $R=\infty$. Например, ряд $\sum\limits_{n=1}^{\infty} n!x^n$ сходится только при $x=0$, а ряд $\sum\limits_{n=1}^{\infty} x^n/n!$ — при всех x.

Найдем выражение радиуса сходимости степенного ряда (22.31) через его коэффициенты. Применим признак Д'Аламбера к исследованию сходимости ряда $\sum\limits_{n=0}^{\infty} |a_n x^n|$, составленного из модулей членов ряда (22.31). Предположим, что $a_n \neq 0$ $(n=0,\ 1,\ 2,\ \ldots)$ и существует предел

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{R}, \tag{22.34}$$

тогда

$$\lim_{n\to\infty}\frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} = \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| |x| = \frac{|x|}{R}.$$

Следовательно, ряд сходится при $|x|<R$ и расходится при $|x|>R$, т. е. R — радиус сходимости данного ряда. Из соотношения (22.34) следует, что радиус сходимости степенного ряда (22.31) определяется формулой

$$R = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|, \tag{22.35}$$

если этот предел существует.

Формулой (22.35) выражается и радиус сходимости ряда (22.32), интервалом сходимости этого ряда является интервал $]a-R,\ a+R[$.

§ 22.5. Непрерывность суммы степенного ряда. Интегрирование и дифференцирование степенных рядов

Теорема 22.7. Степенной ряд $\sum\limits_{n=0}^{\infty} a_n x^n$ сходится равномерно на любом отрезке $[\alpha,\ \beta]$, целиком принадлежащем его интервалу сходимости.

Доказательство. Пусть $[-\rho, \rho]$ — отрезок, содержащий отрезок $[\alpha, \beta]$ и принадлежащий интервалу $]-R, R[$, тогда абсолютно сходится числовой ряд

$$\sum_{n=0}^{\infty} a_n \rho^n = a_0 + a_1\rho + a_2\rho^2 + \ldots . \tag{22.36}$$

Поскольку для всех $x \in [\alpha, \beta]$ $|a_n x^n| \leqslant |a_n| \rho^n$, то по теореме Вейерштрасса на отрезке $[\alpha, \beta]$ сходится равномерно ряд

$$\sum_{n=0}^{\infty} a_n x^n. \tag{22.37}$$

Следствие 1. Сумма степенного ряда является непрерывной функцией на любом отрезке, целиком принадлежащем его интервалу сходимости.

Следствие 2. Степенной ряд можно почленно интегрировать по любому отрезку $[\alpha, \beta]$, принадлежащему его интервалу сходимости $]-R, R[$.

Теорема 22.8. Если степенной ряд

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \ldots \tag{22.38}$$

имеет интервал сходимости $]-R, R[$ и $S(x)$ — его сумма, то ряд

$$\sum_{n=0}^{\infty} n a_n x^{n-1} = a_1 + 2a_2 x + 3a_3 x^2 + \ldots , \tag{22.39}$$

полученный почленным дифференцированием ряда (22.38), имеет тот же интервал сходимости, причем для всех $x \in]-R, R[$

$$S'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}$$

или

$$\left(\sum_{n=0}^{\infty} a_n x^n\right)' = \sum_{n=0}^{\infty} n a x^{n-1}. \tag{22.40}$$

Доказательство. Докажем сначала, что ряд (22.39) сходится на любом отрезке $[-\rho, \rho]$, целиком принадлежащем интервалу $]-R, R[$. Зафиксируем точку

$x_0 \in]-R, R[$ такую, что $\rho < x_0 < R$. В точке x_0 ряд (22.38) сходится, т. е. сходится числовой ряд $\sum_{n=0}^{\infty} a_n x_0^n$, поэтому можно указать такое число $C > 0$, что для всех n $|a_n x_0^n| < C$. Если $|x| \leqslant \rho$, то

$$|na_n x^{n-1}| \leqslant |na_n \rho^{n-1}| \leqslant n |a_n x_0^{n-1}| \left|\frac{\rho}{x_0}\right|^{n-1} <$$

$$< n \frac{C}{x_0} q^{n-1} \left(q = \frac{\rho}{x_0} < 1\right). \tag{22.41}$$

Следовательно, члены ряда (22.39) при $|x| \leqslant \rho$ по модулю не превосходят соответствующих членов числового положительного ряда

$$\sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} \frac{C}{x_0} nq^{n-1} =$$

$$= \frac{C}{x_0}(1 + 2q + 3q^2 + \ldots + nq^{n-1} + \ldots). \tag{22.42}$$

Так как

$$\lim_{n\to\infty} \frac{b_{n+1}}{b_n} = \lim_{n\to\infty} \frac{(n+1)q^n}{nq^{n-1}} = q < 1,$$

то ряд 22.42 сходится. На основании теоремы Вейерштрасса заключаем, что ряд (22.39) сходится равномерно, поэтому ряд (22.38) можно дифференцировать почленно (в силу теоремы 22.5), т. е. равенство (22.40) выполняется для всех $x \in [-\rho, \rho]$. Поскольку любую точку интервала $[-R, R[$ можно заключить в некоторый отрезок $[-\rho, \rho]$, то отсюда следует, что ряд (22.39) сходится в каждой точке интервала $]-R, R[$ и выполняется равенство (22.40).

Докажем, что вне интервала $]-R, R[$ ряд (22.39) расходится. Допустим противное, при $x_2 > R$ этот ряд сходится. Проинтегрировав его по отрезку $[0, x_1]$, где $R < x_1 < x_2$, получим ряд (22.38); он должен сходиться в точке $x_1 > R$, что противоречит условию теоремы. Итак, ряд (22.39) сходится при $|x| < R$ и расходится при $|x| > R$, т. е. имеет тот же интервал сходимости $[-R, R[$, что и ряд (22.38).

С л е д с т в и е. Степенной ряд можно почленно дифференцировать любое число раз в интервале его сходимости.

§ 22.6. Разложение в степенные ряды некоторых функций

Прежде всего отметим, что

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \ldots + x^k + \ldots . \quad (22.43)$$

Ряд $\sum_{k=0}^{\infty} x^k$ сходится при $|x| < 1$ и представляет собой геометрическую прогрессию со знаменателем $q = x$. Сумма его вычисляется по формуле

$$S = \frac{a}{1-q}, \quad S = \frac{1}{1-x} \quad (a = 1, \ q = x).$$

Следовательно, формула (22.43) представляет собой разложение функции $f(x) = 1 : (1 - x)$ в степенной ряд; радиус сходимости $R = 1$.

Записав в равенстве (22.43) $(-t)$ вместо x, получим разложение в степенной ряд функции $f(t) = 1 : (1 + t)$:

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \ldots + (-1)^k t^k + \ldots . \quad (22.44)$$

Разумеется, равенство (22.44) можно получить непосредственно, поскольку ряд $\sum_{k=0}^{\infty} (-1)^k t^k$ сходится при $|t| < 1$ и его сумма равна $1 : (1 + t)$.

Интегрируя ряд (22.44) по отрезку $[0, x]$, где $|x| < 1$, получаем

$$\int_0^x \frac{dt}{1+t} = \int_0^x [1 - t + t^2 - t^3 + t^4 - \ldots + (-1)^k t^k + \ldots] dt,$$

или

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} + \ldots +$$
$$+ (-1)^k \frac{x^{k+1}}{k+1} + \ldots \quad (|x| < 1). \quad (22.45)$$

В соответствии с теоремой 22.7 ряд (22.45) сходится при $|x| < 1$. При $x = 1$ ряд также сходится (см. пример § 21.6), при $x = -1$ ряд расходится (получен из гармонического ряда умножением на -1). Следовательно, ряд (22.45) сходится при $-1 < x \leqslant 1$.

Заменив в равенстве (22.45) x на $(-x)$, получим ряд

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5} - \ldots$$
$$(|x|<1).$$

Вычитая почленно последний ряд из ряда (22.45), находим

$$\ln(1+x) - \ln(1-x) =$$
$$= 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \ldots + \frac{x^{2m-1}}{2m-1} + \ldots\right)$$

или

$$\ln\frac{1+x}{1-x} = 2\left(x + \frac{x^3}{3} + \ldots + \frac{x^{2m-1}}{2m-1} + \ldots\right). \quad (22.46)$$

Это разложение имеет место при $|x|<1$.

§ 22.7. Ряд Тейлора

Предположим, что в интервале $]a-R, a+R[$ функция $f(x)$ разлагается в ряд по степеням $(x-a)$, т. е.

$$f(x) = A_0 + A_1(x-a) + A_2(x-a)^2 + \ldots + A_k(x-a)^k + \ldots$$
$$(|x-a|<R). \quad (22.47)$$

Покажем, как выражаются коэффициенты этого ряда через значения функции и ее производных.

Дифференцируя этот ряд в интервале его сходимости, получаем

$$f'(x) = A_1 + 2A_2(x-a) + 3A_3(x-a)^2 +$$
$$+ 4A_4(x-a)^3 + \ldots + kA_k(x-a)^{k-1} + \ldots ;$$
$$f''(x) = 2A_2 + 2\cdot 3A_3(x-a) + 3\cdot 4A_4(x-a)^2 + \ldots +$$
$$+ k(k-1)A_k(x-a)^{k-2} + \ldots ;$$
$$f'''(x) = 2\cdot 3A_3 + 2\cdot 3\cdot 4A_4(x-a) + \ldots +$$
$$+ k(k-1)(k-2)A_k(x-a)^{k-3} + \ldots ;$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f^{(k)}(x) = k(k-1)(k-2)\ldots 2A_k +$$
$$+ (k+1)k\ldots 2A_{k+1}(x-a) + \ldots .$$

Положив во всех этих равенствах $x=a$, найдем

$$f(a) = A_0,\ f'(a) = A,\ f''(a) = 2A_2,$$
$$f'''(a) = 2\cdot 3A_3,\ \ldots,\ f^{(k)}(a) = 2\cdot 3\ldots(k-1)kA_k,$$

откуда

$$A_0=f(a),\ A_1=\frac{f'(a)}{1},\ A_2=\frac{f''(a)}{1\cdot 2}=\frac{f''(a)}{2!},$$

$$A_3=\frac{f'''(a)}{1\cdot 2\cdot 3}=\frac{f'''(a)}{3!}$$

или

$$A_0=f(a),\ A_k=\frac{f^{(k)}(a)}{k!}\quad (k=1,\ 2,\ 3,\ \ldots). \qquad (22.48)$$

Подставив эти выражения для коэффициентов в формулу (22.47), получим

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\ldots+$$
$$+\frac{f^{(k)}(a)}{k!}(x-a)^k+\ldots. \qquad (22.49)$$

Ряд в правой части равенства (22.49) называется *рядом Тейлора*. Отметим его частный случай, когда $a=0$:

$$f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+$$
$$+\frac{f^{(k)}(0)}{k!}x^k+\ldots. \qquad (22.50)$$

Последний ряд иногда называют *рядом Маклорена*.

Чтобы установить необходимое и достаточное условие представления функции степенным рядом, обратимся к формуле Тейлора (см. § 11.13). Формулу (11.82) можно переписать в виде $f(x)=s_n(x)+r_n(x)$, где $s_n(x)$ — n-я частичная сумма ряда Тейлора; $r_n(x)$ — остаточный член формулы Тейлора, причем

$$r_n(x)=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}\quad (\xi\in]a-R,\ a+R[). \qquad (22.51)$$

Из равенства $f(x)=s_n(x)+r_n(x)$ следует, что ряд Тейлора сходится к $f(x)$ тогда и только тогда, когда

$$\lim_{n\to\infty} r_n(x)=0. \qquad (22.52)$$

Практически важное достаточное условие представления функции ее рядом Тейлора выражается следующей теоремой.

Теорема 22.9. Если при любых x, удовлетворяющих неравенству $|x-a|<R$, производные функции $f(x)$ всех порядков ограничены одним и тем же числом $C>0$, т. е.

$$|f^{(n)}(x)|<C \quad (n=1, 2, 3, \ldots), \tag{22.53}$$

то ряд Тейлора для этой функции сходится в интервале $]a-R, a+R[$ и его сумма равна $f(x)$.

Доказательство. Принимая во внимание формулу (22.51), из условия теоремы получаем

$$|r_n(x)|=|f^{(n+1)}(\xi)|\frac{|x-a|^{n+1}}{(n+1)!}<C\frac{R^{n+1}}{(n+1)!} \tag{22.54}$$
$$(|x-a|<R).$$

Так как ряд $\sum\limits_{n=0}^{\infty}\frac{CR^{n+1}}{(n+1)!}$ сходится (в чем можно убедиться с помощью признака Д'Аламбера), то его общий член стремится к нулю: задав любое $\varepsilon>0$, можно указать такое $N=N(\varepsilon)$, что для всех $n>N$ выполняется неравенство $\frac{R^{n+1}}{(n+1)!}<\frac{\varepsilon}{C}$.

Следовательно, $|r_n(x)|<\varepsilon$ для $n>N$ и всех $x\in]a-R, a+R[$. Это означает, что выполняется равенство (22.52), которое выражает необходимое и достаточное условие сходимости к функции $f(x)$ ее ряда Тейлора.

Отметим без доказательства, что функция $f(x)$ разлагается в ряд Тейлора единственным образом, т. е. если для одной и той же функции $f(x)=\sum\limits_{n=0}^{\infty}a_n(x-a)^n$, $f(x)=\sum\limits_{n=0}^{\infty}b_n(x-a)^n$ $(|x-a|<R,\ R>0)$, то $a_n=b_n$ $(n=1, 2, 3, \ldots)$.

В § 22.6 уже были приведены примеры разложения функций в ряд Маклорена без использования формул (22.48). Приведем пример разложения функции с помощью указанных формул (при $a=0$). Рассмотрим так называемый биномиальный ряд, т. е. разложение в степенной ряд функции $f(x)=(1+x)^{\alpha}$, где α — действительное число.

Поскольку

$$f'(x)=\alpha(1+x)^{\alpha-1};$$
$$f''(x)=\alpha(\alpha-1)(1+x)^{\alpha-2};$$

$$f'''(x)=\alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3};$$
$$f^{IV}(x)=\alpha(\alpha-1)(\alpha-2)(\alpha-3)(1+x)^{\alpha-4}, \dots ;$$
$$f^{(n)}(x)=\alpha(\alpha-1)\dots[\alpha-(n-1)](1+x)^{\alpha-n}$$

и

$$f(0)=1, f'(0)=\alpha, f''(0)=\alpha(\alpha-1);$$
$$f'''(0)=\alpha(\alpha-1)(\alpha-2);$$
$$f^{IV}(0)=\alpha(\alpha-1)(\alpha-2)(\alpha-3), \dots ;$$
$$f^{(n)}(0)=\alpha(\alpha-1)\dots[\alpha-(n-1)],$$

то функция $f(x)=(1+x)^\alpha$ имеет следующий ряд Тейлора:

$$1+\sum_{n=1}^{\infty}\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n. \tag{22.55}$$

Можно доказать, что этот ряд сходится в интервале $]-1, 1[$ и его сумма равна $(1+x)^\alpha$, т. е. при $|x|<1$

$$(1+x)^\alpha=1+\frac{\alpha}{1}x+\frac{\alpha(\alpha-1)}{2!}x^2+$$
$$+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\dots+$$
$$+\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}x^n+\dots . \tag{22.56}$$

Ряд (22.56) сходится при $|x|<1$, его радиус сходимости $R=1$. Действительно, по формуле (22.35) находим

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|=\lim_{n\to\infty}\left|\frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}:\right.$$
$$\left.:\frac{\alpha(\alpha-1)\dots(\alpha-n+1)(\alpha-n)}{(n+1)!}\right|=$$
$$=\lim_{n\to\infty}\left|\frac{\alpha(\alpha-1)\dots(\alpha-n+1)(n+1)!}{\alpha(\alpha-1)\dots(\alpha-n+1)(\alpha-n)n!}\right|=$$
$$=\lim_{n\to\infty}\left|\frac{n+1}{\alpha-n}\right|=1,\ R=1.$$

Замечание. Формула (22.56) содержит разложения в ряды многих функций. Например, при $\alpha_1=1/2$, $\alpha_2=-1/2$, $\alpha_3=1/3$ из нее получаем разложения в степенные ряды соответственно для функций

$$f_1(x)=(1+x)^{1/2}=\sqrt{1+x},\quad f_2(x)=(1+x)^{-1/2}=\frac{1}{\sqrt{1+x}},$$

$$f_3(x) = (1+x)^{1/3} = \sqrt[3]{1+x}.$$

Кроме того, записав в этой формуле x^2 вместо x, получим разложения в ряды многих других функций, например,

$$\varphi_1(x) = \sqrt{1+x^2},\quad \varphi_2(x) = 1 : \sqrt{1+x^2},\quad \varphi_3(x) = \sqrt[3]{1+x^2}.$$

Приведем разложения в ряды следующих функций:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + \ldots; \tag{22.57}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \ldots; \tag{22.58}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots + (-1)^n \frac{x^{2n}}{2n!} + \ldots; \tag{22.59}$$

$$\operatorname{sh} x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + \frac{x^{2n+1}}{(2n+1)!} + \ldots; \tag{22.60}$$

$$\operatorname{ch} x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + \frac{x^{2n}}{2n!} + \ldots. \tag{22.61}$$

Ряды (22.57)—(22.61) сходятся при всех x, т. е. на бесконечном интервале $]-\infty, +\infty[$.

Докажем, например, первую из приведенных формул. Так как для функции $f(x) = e^x$ в любом интервале $]-R, R[$ выполнено условие

$$|f^{(n)}(x)| = e^x < e^R \quad (n = 1, 2, 3, \ldots),$$

то по теореме 22.9 эта функция разлагается в ряд Маклорена на интервале $]-\infty, +\infty[$. Поскольку

$$f(x) = e^x,\ f'(x) = e^x, \ldots, f^{(n)}(x) = e^x, \ldots, f(0) = 1,$$
$$f'(0) = 1, \ldots, f^{(n)}(0) = 1, \ldots,$$

то ряд Маклорена для функции $f(x) = e^x$ имеет вид (22.57). Это разложение данной функции в степенной ряд является единственным.

§ 22.8. Приложения рядов

Ряды имеют самое широкое применение. В частности, они используются в приближенных вычислениях. С помощью рядов вычисляют приближенные значения функций. При вычислении значений функций e^x, $\sin x$, $\cos x$, $\operatorname{sh} x$, $\operatorname{ch} x$ применяются соответственно формулы (22.57) — (22.61). Значения корней из чисел находятся посредством биномиального ряда (22.56), для чего предварительно преобразуются корни.

Пример. Вычислить $\sqrt[3]{124}$ с точностью до 0,0001.

Преобразуя корень, получаем

$$\sqrt[3]{124}=\frac{5}{\sqrt[3]{\frac{125}{124}}}=5\left(1+\frac{1}{124}\right)^{-1/3}.$$

Применяя формулу (22.56), находим ряд

$$\sqrt[3]{124}=5\left(1+\frac{1}{124}\right)^{-1/3}=5\left[1+\left(-\frac{1}{3}\right)\frac{1}{124}+\right.$$

$$\left.+\frac{\left(-\frac{1}{3}\right)\left(-\frac{1}{3}-1\right)}{1\cdot 2}\left(\frac{1}{124}\right)^2+\ldots\right]=$$

$$=5\left(1-\frac{1}{3\cdot 124}+\frac{2}{3^2\cdot 124^2}-\ldots\right)=5-\frac{5}{372}+\frac{5}{9\cdot 62\cdot 124}-\ldots.$$

Поскольку полученный ряд является рядом лейбницевского типа и

$$\frac{5}{9\cdot 62\cdot 124}=\frac{5}{69192}<\frac{1}{10000}=0{,}0001,$$

то $\sqrt[3]{124}\approx 5-\frac{5}{372}\approx 5-0{,}0134=4{,}9866.$

Разложение (22.46) может быть использовано для вычисления натуральных логарифмов чисел. Положим в этом разложении $x=1:(2n+1)$, где n — натуральное число, тогда $0<x<1$ при любом n.

Следовательно,

$$\ln\frac{1+x}{1-x}=\ln\frac{n+1}{n}=$$

$$=2\left[\frac{1}{2n+1}+\frac{1}{3(2n+1)^3}+\frac{1}{5(2n+1)^5}+\ldots\right]$$

или

$$\ln(n+1)-\ln n=$$

$$=2\left[\frac{1}{2n+1}+\frac{1}{3(2n+1)^3}+\frac{1}{5(2n+1)^5}+\ldots\right].$$

Пользуясь этой формулой, вычисляют $\ln 2$ (при $n=1$), $\ln 3$ (при $n=2$) и т. д. Оценка погрешности приближенных вычислений по этой формуле и примеры приведены в [7, ч. 2].

С помощью рядов вычисляют многие определенные интегралы, для чего предварительно подынтегральную функцию разлагают в степенной ряд и затем пользуются почленным интегрированием ряда. Соответствующие примеры также имеются в [7, ч. 2].

В заключение отметим, что ряды широко используются при интегрировании дифференциальных уравнений (см. § 24.6).

§ 22.9. Числовые ряды с комплексными членами

Рассмотрим числовой ряд с комплексными членами, т. е. ряд

$$\sum_{k=1}^{\infty} w_k = w_1 + w_2 + w_3 + \ldots + w_k + \ldots, \tag{22.62}$$

$$w_k = u_k + iv_k \quad (k=1, 2, 3, \ldots), \tag{22.63}$$

где u_k, v_k — действительные числа, $i=\sqrt{-1}$.

Для этого ряда, как и для ряда с действительными членами, вводятся понятия n-й частичной суммы

$$s_n = w_1 + w_2 + \ldots + w_n \tag{22.64}$$

и суммы ряда

$$s = \lim_{n\to\infty} s_n. \tag{22.65}$$

Замечание. Здесь имеется в виду предел последовательности комплексных чисел. Комплексное число c называется пределом последовательности $\{c_n\}$ комплексных чисел, если при любом $\varepsilon>0$ можно указать такой номер N, что для всех $n \geqslant N$ выполняется неравенство $|c-c_n|<\varepsilon$. Модуль комплексного числа $c=a+ib$, как известно, является действительным числом $|c|=\sqrt{a^2+b^2}$.

Если предел (22.65) является конечным, то ряд (22.62) называется *сходящимся*. В этом случае пишут

$$s = \sum_{k=1}^{\infty} w_k, \text{ или } s = w_1 + w_2 + \ldots + w_k + \ldots. \tag{22.66}$$

Если предел частичных сумм бесконечен или не существует, то ряд называется *расходящимся*.

О сходимости ряда (22.62) можно судить по сходимости двух рядов с действительными членами

$$\sum_{k=1}^{\infty} u_k = u_1 + u_2 + \ldots + u_k + \ldots ; \tag{22.67}$$

$$\sum_{k=1}^{\infty} v_k = v_1 + v_2 + \ldots + v_k + \ldots \tag{22.68}$$

на основании следующей теоремы, приводимой без доказательства.

Теорема 22.10. Ряд с комплексными членами сходится тогда и только тогда, когда сходятся ряды из действительных и мнимых частей его членов. Если A и B — соответственно суммы двух последних рядов, то сумма исходного ряда $S = A + iB$.

Рассмотрим ряд, составленный из модулей членов ряда (22.62), т. е. ряд с действительными (положительными) членами

$$\sum_{k=1}^{\infty} |w_k| = |w_1| + |w_2| + |w_3| + \ldots . \tag{22.69}$$

Теорема 22.11. Если сходится ряд из модулей членов данного ряда, то сходится и сам ряд с комплексными членами, т. е. из сходимости ряда (22.69) следует сходимость ряда (22.62).

Доказательство. Поскольку

$$|u_k| \leqslant \sqrt{u_k^2 + v_k^2} = |w_k|, \quad |u_k| \leqslant |w_k|,$$

$$|v_k| \leqslant \sqrt{u_k^2 + v_k^2} = |w_k|, \quad |v_k| \leqslant |w_k|$$

и ряд (22.69) сходится, то по признаку сходимости рядов с положительными членами сходятся ряды

$$\sum_{k=1}^{\infty} |u_k| = |u_1| + |u_2| + \ldots , \quad \sum_{k=1}^{\infty} |v_k| = |v_1| + |v_2| + \ldots ,$$

а следовательно, и ряды (22.67), (22.68). На основании предыдущей теоремы отсюда следует, что ряд (22.62) также сходится.

Как и в случае рядов с действительными членами, ряд (22.62) называется абсолютно сходящимся, если сходится ряд из модулей его членов.

§ 22.10. Степенные ряды в комплексной области. Формулы Эйлера

Степенным рядом с комплексными членами называется ряд вида

$$\sum_{k=0}^{\infty} c_k (z-c)^k = c_0 + c_1(z-c) + c_2(z-c)^2 + \ldots + c_k(z-c)^k + \ldots, \qquad (22.70)$$

где $c=a+ib$, $c_k=a_k+ib_k$ $(k=0, 1, 2, \ldots)$, $z=x+iy$, a, b, a_k, b_k $(k=0, 1, 2, \ldots)$ — действительные числа, $i=\sqrt{-1}$; x, y — действительные переменные.

Если $c=0$, ряд имеет вид

$$\sum_{k=0}^{\infty} c_k z^k = c_0 + c_1 z + c_2 z^2 + \ldots + c_k z^k + \ldots . \qquad (22.71)$$

Отметим, что ряд (22.70) сводится к ряду (22.71) с помощью введения новой переменной Z по формуле $Z=z-c$.

Теорема Абеля. Если степенной ряд (22.71) сходится при некотором значении $z_0 \neq 0$, то он сходится при любом z, для которого $|z| < |z_0|$.

Доказательство этой теоремы аналогично доказательству соответствующей теоремы для степенных рядов с действительными членами.

Утверждение теоремы геометрически означает следующее: если ряд (22.71) сходится в некоторой точке z_0 плоскости комплексной переменной, то он сходится в любой точке, лежащей в круге радиуса $|z_0|$ с центром в начале координат.

Из теоремы Абеля следует, что область сходимости ряда (22.71) представляет собой круг радиуса R с центром в начале координат: внутри этого круга ряд сходится абсолютно, на границе круга, т. е. в точках соответствующей окружности, ряд может сходиться или расходиться. Указанный круг называется *кругом сходимости* ряда (22.71), а его радиус R — радиусом сходимости. Если ряд сходится в единственной точке, то считают $R=0$; если ряд сходится при любом z, то полагают $R=\infty$.

Радиус сходимости R степенного ряда (22.71) можно найти, например, по формуле

$$R = \lim_{n\to\infty}\left|\frac{c_n}{c_{n+1}}\right|. \qquad (22.72)$$

Для комплексной переменной z определим функцию e^z посредством формулы

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots + \frac{z^n}{n!} + \ldots . \qquad (22.73)$$

Можно показать (применяя, например, признак Д'Аламбера), что последний ряд сходится при любом z.

Очевидно, при $z=x$ эта формула обращается в формулу (22.57).

Полагая в (22.73) $z=ix$, получаем

$$e^{ix} = 1 + \frac{ix}{1!} + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \frac{(ix)^5}{5!} + \ldots =$$

$$= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots\right).$$

Следовательно, с учетом (22.58) и (22.59)

$$e^{ix} = \cos x + i\sin x. \qquad (22.74)$$

Аналогично находим, что

$$e^{-ix} = \cos x - i\sin x. \qquad (22.75)$$

Формулы (22.74) и (22.75) называются *формулами Эйлера.* Из этих формул находим выражения для $\cos x$ и $\sin x$:

$$\cos x = \frac{e^{ix} + e^{-ix}}{2}, \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}. \qquad (22.76)$$

Замечание. Если $z=x+iy$, где $i=\sqrt{-1}$, x, y — действительные переменные, то можно показать, что $e^z=e^{x+iy}=e^xe^{iy}$, поэтому

$$e^z = e^x(\cos y + i\sin y). \qquad (22.77)$$

Глава 23

РЯДЫ ФУРЬЕ

В классе функциональных рядов особое значение имеют ряды Фурье, рассмотрению которых посвящена эта глава.

§ 23.1. Тригонометрический ряд Фурье

Тригонометрическим рядом называется функциональный ряд вида

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{23.1}$$

где a_0, a_n, b_n $(n=1, 2, 3, \ldots)$ — действительные числа, называемые коэффициентами ряда.

Тригонометрической системой функций называется множество функций

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos nx, \sin nx, \ldots . \tag{23.2}$$

Лемма 23.1. Основная тригонометрическая система функций (23.2) обладает следующими свойствами:

1) интеграл по отрезку $[-\pi, \pi]$ от произведения любых двух различных функций равен нулю, т. е.

$$\left.\begin{aligned}
&\int_{-\pi}^{\pi} 1 \cdot \sin mx dx = \int_{-\pi}^{\pi} \sin mx dx = 0;\\
&\int_{-\pi}^{\pi} 1 \cdot \cos nx dx = \int_{-\pi}^{\pi} \cos nx dx = 0;\\
&\int_{-\pi}^{\pi} \cos mx \cos nx dx = 0 \quad (m \neq n);\\
&\int_{-\pi}^{\pi} \sin mx \sin nx dx = 0 \quad (m \neq n);\\
&\int_{-\pi}^{\pi} \sin mx \cos nx dx = 0 \quad (m,\ n=1,\ 2,\ 3,\ \ldots);
\end{aligned}\right\} \tag{23.3}$$

2) интеграл по отрезку $[-\pi, \pi]$ от квадрата любой функции отличен от нуля, причем

$$\int_{-\pi}^{\pi} \cos^2 mx dx = \int_{-\pi}^{\pi} \sin^2 mx dx = \pi, \quad \int_{-\pi}^{\pi} 1^2 dx = 2\pi. \tag{23.4}$$

Доказательство. Докажем первые три равенства (23.3):

$$\int_{-\pi}^{\pi} \cos nx dx = \frac{1}{n} \int_{-\pi}^{\pi} \cos nx d(nx) = \frac{1}{n} \sin nx \Big|_{-\pi}^{\pi} = 0,$$

$$\int_{-\pi}^{\pi} \sin mx dx = -\frac{\cos mx}{m} \Big|_{-\pi}^{\pi} = 0,$$

$$\int_{-\pi}^{\pi} \cos mx \cos nx dx =$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} [\cos (m-n)x + \cos (m+n) x]\, dx = 0 \quad (m \neq n).$$

Остальные два равенства (23.3) доказываются аналогично. Далее,

$$\int_{-\pi}^{\pi} \sin^2 mx dx = \int_{-\pi}^{\pi} \frac{1-\cos 2mx}{2}\, dx =$$

$$= \frac{1}{2} \int_{-\pi}^{\pi} dx - \frac{1}{2} \int_{-\pi}^{\pi} \cos 2mx dx = \pi.$$

Аналогично доказываются два других равенства.

Теорема 23.1. Если

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{23.5}$$

и ряд (23.5) сходится равномерно на отрезке $[-\pi, \pi]$, то его коэффициенты определяются формулами

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx,$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \quad (n = 1,\ 2,\ 3,\ \ldots). \tag{23.6}$$

Доказательство. Интегрируя почленно ряд (23.5) и принимая во внимание формулы (23.3), находим

$$\int_{-\pi}^{\pi} f(x)\, dx = \int_{-\pi}^{\pi} \frac{a_0}{2} dx + \int_{-\pi}^{\pi} \left[\sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx\right] dx =$$

$$= \frac{a_0}{2} \int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty} a_n \int_{-\pi}^{\pi} \cos nx dx + \sum_{n=1}^{\infty} b_n \int_{-\pi}^{\pi} \sin nx dx =$$

$$= \frac{a_0}{2} \int_{-\pi}^{\pi} dx = a_0 \pi,$$

$$\int_{-\pi}^{\pi} f(x)\, dx = a_0 \pi, \quad a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx.$$

Отметим, что при почленном умножении ряда (23.5) на $\sin kx$ или $\cos kx$ получаем ряд, равномерно сходящийся на отрезке $[-\pi, \pi]$, так как члены этого ряда по модулю не превосходят членов данного ряда ($|\sin kx| \leqslant 1$, $|\cos kx| \leqslant 1$ при любом k).

Умножая поочередно ряд (23.5) на $\cos kx$, $\sin kx$, интегрируя почленно полученные ряды и принимая во внимание формулы (23.3) и (23.4), соответственно находим

$$\int_{-\pi}^{\pi} f(x) \cos kx dx = \int_{-\pi}^{\pi} \frac{a_0}{2} \cos kx dx +$$

$$+ \int_{-\pi}^{\pi} \left[\sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \cos kx\right] dx =$$

$$= \frac{a_0}{2} \int_{-\pi}^{\pi} \cos kx dx + \sum_{n=1}^{\infty} \int_{-\pi}^{\pi} a_n \cos nx \cos kx dx +$$

$$+ \sum_{n=1}^{\infty} \int_{-\pi}^{\pi} b_n \sin nx \cos kx dx = \int_{-\pi}^{\pi} a_k \cos^2 kx dx = a_k \pi,$$

$$\int_{-\pi}^{\pi} f(x) \cos kx dx = a_k \pi, \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx,$$

$$\int_{-\pi}^{\pi} f(x) \sin kx dx = b_k \pi, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx.$$

Итак,

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx, \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx,$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx \quad (k = 1,\ 2,\ 3,\ \ldots),$$

что и требовалось доказать.

Тригонометрический ряд (23.5), коэффициенты которого определяются формулами (23.6), называется *рядом Фурье* [1], а числа a_n, b_n — коэффициентами Фурье функ-

---

[1] *Жозеф Фурье* (*Joseph Fourier*, 1768—1830) — французский ученый.

ции $f(x)$. Ряд Фурье, построенный для функции $f(x)$, обозначают так:

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx).$$

Чтобы сформулировать теорему о сходимости ряда Фурье к функции, для которой он составлен, введем понятие кусочно-дифференцируемой функции.

§ 23.2. Сходимость ряда Фурье для кусочно-дифференцируемой функции

Функция $f(x)$ называется *кусочно-дифференцируемой* на отрезке $[a, b]$, если этот отрезок можно разбить точками $a=x_0<x_1<x_2<\ldots<x_N=b$ так, что $f(x)$ и $f'(x)$ непрерывны в каждом интервале $]x_k, x_{k+1}[$ $(k=0, 1, \ldots, N-1)$ и, кроме того, существуют конечные односторонние пределы $f(x)$ и $f'(x)$ в концевых точках этих интервалов. Функция, кусочно-дифференцируемая на отрезке $[a, b]$, может быть непрерывной на нем или иметь конечное число точек разрыва первого рода. Приведем примеры кусочно-дифференцируемых функций.

1. Функция $f(x)=|x|$ — кусочно-дифференцируема на отрезке $[-a, a]$ $(a>0)$: она непрерывна на этом отрезке, производная функции непрерывна в интервалах $]-a, 0[$, $]0, a[$; $f(x)$ и $f'(x)$ имеют соответствующие односторонние пределы на концах указанных интервалов (отметим, что функция $f(x)=|x|$ не является дифференцируемой на отрезке $[-a, a]$, так как в точке $x=0$ не имеет производной).

2. Функция $f(x)=E(x)$ ($E(x)$ — наибольшее целое число, не превышающее x) является кусочно-дифференцируемой на любом отрезке $[a, b]$; эта функция разрывна, точки $x_n=n$ $(n=0, \pm1, \pm2, \ldots)$ являются точками разрыва первого рода (см. [7, ч. 1], § 8.2); в этих точках существуют конечные односторонние пределы функции и ее производной.

Теорема 23.2. Если функция $f(x)$ кусочно-дифференцируема на отрезке $[-\pi, \pi]$, то ее ряд Фурье сходится в каждой точке $x_0 \in]-\pi, \pi[$ и имеет сумму

$$S(x_0) = \frac{f(x_0+0) + f(x_0-0)}{2}, \qquad (23.7)$$

в частности, если x_0 — точка непрерывности $f(x)$, то

$$f(x_0) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx_0 + b_n \sin nx_0). \qquad (23.8)$$

На концах отрезка сумма ряда Фурье определяется формулой

$$S(\pm\pi) = \frac{f(-\pi+0)+f(\pi-0)}{2}. \qquad (23.9)$$

Доказательство этой теоремы приведено, например, в [11, 14].

§ 23.3. Ряды Фурье для четных и нечетных функций

Прежде всего отметим, что для четных и нечетных интегрируемых функций выполняются соответственно равенства

$$\int_{-a}^{a} f(x)\,dx = 2\int_{0}^{a} f(x)\,dx, \quad \int_{-a}^{a} f(x)\,dx = 0 \quad (a>0), \qquad (23.10)$$

которые следуют из равенств

$$\int_{-a}^{a} f(x)\,dx = \int_{-a}^{0} f(x)\,dx + \int_{0}^{a} f(x)\,dx =$$

$$= \int_{a}^{0} f(-u)\,d(-u) + \int_{0}^{a} f(x)\,dx =$$

$$= -\int_{a}^{0} f(-u)\,du + \int_{0}^{a} f(x)\,dx;$$

$$\int_{a}^{0} f(-u)\,du = \int_{a}^{0} f(u)\,du = -\int_{0}^{a} f(u)\,du = -\int_{0}^{a} f(x)\,dx$$

(для четной функции);

$$\int_{a}^{0} f(-u)\,du = -\int_{a}^{0} f(u)\,du = \int_{0}^{a} f(u)\,du = \int_{0}^{a} f(x)\,dx$$

(для нечетной функции).

Очевидно, произведение двух четных и двух нечетных функций есть функция четная, произведение четной и нечетной функций является нечетной функцией; как известно, $\sin nx$ — нечетная, $\cos nx$ — четная функция. Принимая это во внимание, на основании формул (23.10) заключаем, что если $f(x)$ — четная функция, то

$$\left.\begin{aligned} &a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx = \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos nx dx \\ &(n=0,\ 1,\ 2,\ \ldots); \\ &b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx dx = 0 \quad (n = 1,\ 2,\ 3,\ \ldots), \end{aligned}\right\} \qquad (23.11)$$

т. е. ряд Фурье для четной функции содержит только косинусы. Если $f(x)$ — функция, кусочно-дифференцируемая на отрезке $[-\pi, \pi]$, то для ее точек непрерывности $x \in]-\pi, \pi[$

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx. \qquad (23.12)$$

В случае нечетной функции $f(x)$ получаем

$$\left.\begin{aligned} &a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx dx = 0 \quad (n = 0,\ 1,\ 2,\ \ldots); \\ &b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx dx = \frac{2}{\pi}\int_{0}^{\pi} f(x)\sin nx dx, \end{aligned}\right\} \qquad (23.13)$$

т. е. ряд Фурье для нечетной функции $f(x)$ содержит только синусы. Если функция $f(x)$ кусочно-дифференцируема на отрезке $[-\pi, \pi]$, то для ее точек непрерывности $x \in]-\pi, \pi[$

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx, \qquad (23.14)$$

где b_n определяются второй из формул (23.13).

Замечание. Если функция $f(x)$ определена на отрезке $[0, \pi]$, то, продолжив ее на промежуток $[-\pi, 0[$ четным или нечетным образом (т. е. положив $f(-x) = f(x)$ или $f(-x) = -f(x)$), можно разложить функцию в ряд Фурье только по косинусам (в первом случае) и только по синусам (во втором случае).

§ 23.4. Ряд Фурье для функции, заданной на отрезке $[-l, l]$

Рассмотрим функцию $f(x)$, определенную и кусочно-дифференцируемую на отрезке $[-l, l]$. Введем новую переменную t по формуле

$$t = \frac{\pi x}{l}, \tag{23.15}$$

тогда

$$x = \frac{lt}{\pi}, \quad dx = \frac{l}{\pi} dt, \quad dt = \frac{\pi}{l} dx.$$

Очевидно, если $x \in [-l, l]$, то $t \in [-\pi, \pi]$. Получаем функцию

$$f(x) = f\left(\frac{lt}{\pi}\right) \equiv \varphi(t), \tag{23.16}$$

для которой в точках непрерывности

$$\varphi(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt + b_n \sin nt; \tag{23.17}$$

$$\left.\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \varphi(t) \cos nt\,dt, \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} \varphi(t) \sin nt\,dt \quad (n = 0,\ 1,\ 2,\ \ldots), \end{aligned}\right. \tag{23.18}$$

а в точках разрыва сумма ряда Фурье вычисляется по формуле (23.7).

Возвращаясь к переменной x, получаем

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi x}{l} + b_n \sin \frac{n\pi x}{l}, \tag{23.19}$$

где

$$\left.\begin{aligned} a_n &= \frac{1}{l} \int_{-l}^{l} f(x) \cos \frac{n\pi x}{l}\,dx, \\ b_n &= \frac{1}{l} \int_{-l}^{l} f(x) \sin \frac{n\pi x}{l}\,dx \quad (n = 0,\ 1,\ 2,\ \ldots). \end{aligned}\right. \tag{23.20}$$

Ряд (23.19), коэффициенты a_n, b_n которого вычисляются по формулам (23.20), называется *рядом Фурье* функции $f(x)$, заданной на отрезке $[-l, l]$. Условия сходимости ряда Фурье к функции $f(x)$ выражаются теоремой, аналогичной теореме 23.2. Равенство (23.19) выполняется в точках непрерывности функции $f(x)$. В точках разрыва сумма ряда Фурье определяется формулой (23.7).

Замечание. Формулы (23.20) можно получить тем же путем, что и формулы (23.6). Для этого нужно рассмотреть общую тригонометрическую систему функций:

$$1, \cos\frac{\pi x}{l}, \sin\frac{\pi x}{l}, \cos\frac{2\pi x}{l}, \sin\frac{2\pi x}{l}, \ldots,$$
$$\cos\frac{n\pi x}{l}, \sin\frac{n\pi x}{l}, \ldots, \tag{23.21}$$

Система функций (23.21) на отрезке $[-l, l]$ обладает теми же свойствами, что и основная тригонометрическая система функций (23.2) на отрезке $[-\pi, \pi]$:

$$\left.\begin{aligned}
&\int_{-l}^{l}\cos\frac{m\pi x}{l}\cos\frac{n\pi x}{l}\,dx=0,\\
&\int_{-l}^{l}\sin\frac{m\pi x}{l}\sin\frac{n\pi x}{l}\,dx=0 \quad (m\neq n),\\
&\int_{-l}^{l}\sin\frac{m\pi x}{l}\cos\frac{n\pi x}{l}\,dx=0 \ (m, n=0, 1, 2, \ldots);
\end{aligned}\right\} \tag{23.22}$$

$$\int_{-l}^{l}1^2dx=2l, \quad \int_{-l}^{l}\cos^2\frac{m\pi x}{l}\,dx=\int_{-l}^{l}\sin^2\frac{m\pi x}{l}\,dx=l. \tag{23.23}$$

§ 23.5. Ряд Фурье по произвольной ортогональной системе функций

Две функции $f(x)$ и $\varphi(x)$, определенные и интегрируемые на отрезке $[a, b]$, называются *ортогональными* на этом отрезке, если

$$\int_a^b f(x)\,\varphi(x)\,dx=0. \tag{23.24}$$

Система функций

$$\varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x),\ \ldots, \qquad (23.25)$$

определенных и интегрируемых на отрезке $[a, b]$, называется *ортогональной* на этом отрезке, если выполняется условие

$$\int_a^b \varphi_k(x)\,\varphi_m(x)\,dx = 0 \quad (m \neq k). \qquad (23.26)$$

В дальнейшем будем считать, что функции $\varphi_k(x)$ $(k=1, 2, \ldots, n, \ldots)$ непрерывны на отрезке $[a, b]$ и $\varphi_k(x) \not\equiv 0$, т. е.

$$\int_a^b \varphi_k^2(x)\,dx \neq 0 \quad (k = 1,\ 2,\ 3,\ \ldots). \qquad (23.27)$$

Примерами ортогональных систем функций являются основная тригонометрическая система (23.2) и общая тригонометрическая система (23.21); условия (23.26) и (23.27) для них выполнены (см. равенства (23.3) и (23.4), (23.22) и (23.23)).

Пусть функция $f(x)$ разлагается в ряд по ортогональной системе функций (23.25):

$$f(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x). \qquad (23.28)$$

Предположим, что на отрезке $[a, b]$ сходится равномерно ряд (23.28). Принимая во внимание условия (23.26) и (23.27), получаем

$$\int_a^b f(x)\,\varphi_k(x)\,dx = \int_a^b \sum_{n=1}^{\infty} c_n \varphi_n(x)\,\varphi_k(x)\,dx =$$

$$= \sum_{n=1}^{\infty} c_n \int_a^b \varphi_n(x)\,\varphi_k(x)\,dx = c_k \int_a^b \varphi_k^2(x)\,dx,$$

откуда

$$c_k = \left(\int_a^b f(x)\,\varphi_k(x)\,dx\right) : \left(\int_a^b \varphi_k^2(x)\,dx\right) \qquad (23.29)$$

$$(k=1,\ 2,\ 3,\ \ldots).$$

Ряд (23.28), для которого c_k определяются формулами (23.29), называется *обобщенным рядом Фурье* функции $f(x)$; числа c_k называются *коэффициентами Фурье*.

§ 23.6. Комплексная форма ряда Фурье

Пусть

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{23.30}$$

где

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \quad (n = 0,\ 1,\ 2,\ \ldots),$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \quad (n = 1,\ 2,\ \ldots).$$

Принимая во внимание формулы Эйлера

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2},$$

$$\sin nx = \frac{e^{inx} - e^{-inx}}{2i} = i\,\frac{e^{-inx} - e^{inx}}{2},$$

преобразуем равенство (23.30):

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \frac{e^{inx} + e^{-inx}}{2} + ib_n \frac{e^{-inx} - e^{inx}}{2}\right] =$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[\left(\frac{a_n - ib_n}{2}\right) e^{inx} + \left(\frac{a_n + ib_n}{2}\right) e^{-inx}\right].$$

Введем обозначения:

$$\frac{a_0}{2} = c_0, \quad \frac{a_n - ib_n}{2} = c_n, \quad \frac{a_n + ib_n}{2} = c_{-n},$$

тогда

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}, \tag{23.31}$$

где

$$c_n = \frac{a_n - ib_n}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}[f(x)\cos nx - if(x)\sin nx]\,dx =$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,e^{-inx}dx,$$

$$c_0 = \frac{a_0}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx,$$

$$c_{-n} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,e^{inx}dx \quad (n = 1,\ 2,\ 3,\ \ldots).$$

Эти формулы можно объединить в одну:

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,e^{-inx}dx \quad (n = 0,\ \pm 1,\ \pm 2,\ \ldots). \tag{23.32}$$

Формула (23.31) выражает ряд Фурье (23.30) в комплексной форме, а формула (23.32) — коэффициенты этого ряда.

Комплексная форма ряда Фурье (23.19) находится аналогично:

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega x + b_n \sin n\omega x \equiv$$

$$\equiv \sum_{n=-\infty}^{\infty} c_n e^{in\omega x} \qquad \left(\omega = \frac{\pi}{l}\right), \tag{23.33}$$

где

$$a_n = \frac{1}{l}\int_{-l}^{l} f(t)\cos n\omega t\,dt,$$

$$b_n = \frac{1}{l}\int_{-l}^{l} f(t)\sin n\omega t\,dt \quad (n = 0,\ 1,\ 2,\ \ldots),$$

$$c_n = \overline{c_{-n}} = \frac{1}{2}(a_n - ib_n) = \frac{1}{2l}\int_{-l}^{l} f(t)\,e^{-in\omega t}dt \tag{23.34}$$

$$(n = 0,\ \pm 1,\ \pm 2,\ \ldots).$$

РАЗДЕЛ III

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Дифференциальные уравнения — большая и важная область современной математики, интенсивно развиваемая в трудах советских и зарубежных ученых.

Дифференциальные уравнения широко используются при решении разнообразных задач науки и техники. Следует отметнть, что во многих случаях различные явления описываются одними и теми же дифференциальными уравнениями.

Глава 24

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ПЕРВОГО ПОРЯДКА

В этой главе сообщаются основные сведения о дифференциальных уравнениях, рассматриваются методы интегрирования дифференциальных уравнений первого порядка.

§ 24.1. Основные сведения о дифференциальных уравнениях

Дифференциальным уравнением называется уравнение относительно неизвестной функции и ее производных различных порядков. Порядком дифференциального уравнения называется порядок старшей производной, входящей в это уравнение.

Если искомая функция зависит от одной переменной, то соответствующее дифференциальное уравнение называется *обыкновенным*. Если искомая функция зависит от нескольких переменных, то соответствующее дифференциальное уравнение называется *уравнением с частными производными*. В главах 24 и 25 рассматриваются обыкновенные дифференциальные уравнения.

Обыкновенное дифференциальное уравнение n-го порядка в общем виде можно записать так:

$$F(x, y, y', y'', \ldots, y^{(n)})=0, \qquad (24.1)$$

где x — независимая переменная; $y=y(x)$ — искомая функция переменной x; $y', y'', \ldots, y^{(n)}$ — ее производные; $F(x, y, y', y'', \ldots, y^{(n)})$ — заданная функция своих аргументов. Отметим, что функция F может не содержать некоторых своих аргументов, но непременно должна зависеть от $y^{(n)}$ (когда речь идет об уравнении n-го порядка).

Если уравнение (24.1) разрешимо относительно производной n-го порядка, его можно представить в виде

$$y^{(n)}=f(x, y, y', y'', \ldots, y^{(n-1)}). \qquad (24.2)$$

Функция $y=\varphi(x)$, определенная и непрерывно дифференцируемая n раз в интервале $]a, b[$, называется *решением* дифференциального уравнения (24.1) в этом интервале, если она обращает указанное уравнение в тождество, т. е.

$$F(x, \varphi(x), \varphi'(x), \varphi''(x), \ldots, \varphi^{(n)}(x))\equiv 0$$

для всех $x\in]a, b[$.

График решения дифференциального уравнения n-го порядка называется *интегральной линией* (или *интегральной кривой*).

Задача Коши для дифференциального уравнения n-го порядка состоит в следующем: найти решение $y=y(x)$ уравнения (24.1), удовлетворяющее условиям

$$y=y_0,\ y'=y_0',\ y''=y_0'',\ \ldots,\ y^{(n-1)}=y_0^{(n-1)} \qquad (24.3)$$

$$\text{при}\ x=x_0,$$

где $x_0, y_0, y_0', \ldots, y_0^{(n-1)}$ — заданные числа, называемые *начальными данными* решения. Равенства (24.3), которые называют *начальными условиями*, можно записать и в таком виде:

$$y(x_0)=y_0,\ y'(x_0)=y_0',\ y''(x_0)=y_0'',\ \ldots,\ y^{(n-1)}(x_0)=y_0^{(n-1)}.$$

Условия существования и единственности решения задачи Коши для уравнения (24.2) определяются следующей теоремой, приводимой здесь без доказательства.

Теорема 24.1. Если в уравнении $y^{(n)}=f(x, \ldots, y^{(n-1)})$

функция $f(x, y, y', \ldots, y^{(n-1)})$ и ее частные производные по $y, y', \ldots, y^{(n-1)}$ непрерывны в некоторой замкнутой области G, определяемой неравенствами

$$|x-x_0| \leqslant a, \ |y-y_0| \leqslant b, \ |y'-y_0'| \leqslant b, \ \ldots,$$
$$|y^{(n-1)}-y_0^{(n-1)}| \leqslant b \quad (a>0, \ b>0),$$

и, следовательно, ограничены в ней, т. е.

$$|f(x, y, y', \ldots, y^{(n-1)})| \leqslant C, \ \left|\frac{\partial f(x, y, y', \ldots, y^{(n-1)})}{\partial y^{(k)}}\right| \leqslant C_1$$
$$(k=0, 1, 2, \ldots, n-1; \ y^0 \equiv y),$$

где

$$C>0, \ C_1>0, \ M(x, y, y', \ldots, y^{(n-1)}) \in G,$$
$$M_0(x_0, y_0, y_0', \ldots, y_0^{(n-1)}) \in G,$$

то существует единственное решение $y=y(x)$ данного уравнения, удовлетворяющее условиям $y=y_0$, $y'=y_0'$, ..., $y^{(n-1)}=y_0^{(n-1)}$ при $x=x_0$. Это решение определено и непрерывно вместе с производными до порядка n включительно в промежутке $|x-x_0| \leqslant h$, где

$$h=\min\left(a, \frac{b}{\max\limits_G(C, |y'|, \ldots, |y^{(n-1)}|)}\right).$$

Общим решением дифференциального уравнения n-го порядка (24.1) называется функция

$$y=\varphi(x, C_1, C_2, \ldots, C_n), \tag{24.4}$$

обладающая следующими свойствами: 1) при любых значениях произвольных постоянных $C_1, C_2, \ldots, C_n$ она обращает уравнение (24.1) в тождество; 2) значения постоянных $C_1, C_2, \ldots, C_n$ можно подобрать так, чтобы она удовлетворяла условиям (24.3).

Частным решением дифференциального уравнения n-го порядка называется решение, получающееся из общего решения (24.4) при фиксированных значениях произвольных постоянных, т. е. функция

$$y=\varphi(x, C_1^0, C_2^0, \ldots, C_n^0),$$

где $C_1^0, C_2^0, \ldots, C_n^0$ — некоторые числа.

Решение дифференциального уравнения n-го порядка, в каждой точке которого нарушается единственность решения задачи Коши, называется *особым*.

Общим интегралом дифференциального уравнения n-го порядка называется соотношение вида

$$\Phi(x, y, C_1, C_2, \ldots, C_n)=0, \qquad (24.5)$$

неявно определяющее общее решение $y=\varphi(x, C_1, \ldots, C_n)$ этого уравнения. *Частным интегралом* дифференциального уравнения n-го порядка называется соотношение $\Phi(x, y, C_1^0, C_2^0, \ldots, C_n^0)=0$, полученное из общего интеграла путем фиксирования значений $C_1^0, C_2^0, \ldots, C_n^0$ произвольных постоянных.

§ 24.2. Дифференциальные уравнения первого порядка. Уравнения с разделяющимися переменными

Дифференциальное уравнение первого порядка в общем виде записывается так:

$$F(x, y, y')=0. \qquad (24.6)$$

Если это уравнение разрешимо относительно y', то

$$y'=f(x, y) \qquad (24.7)$$

или

$$dy-f(x, y)\,dx=0. \qquad (24.8)$$

Последнее уравнение является частным случаем уравнения

$$P(x, y)\,dx+Q(x, y)\,dy=0. \qquad (24.9)$$

Задача Коши для дифференциального уравнения первого порядка: найти решение $y=y(x)$ уравнения (24.6), удовлетворяющее условию

$$y=y_0 \text{ при } x=x_0, \text{ или } y(x_0)=y_0, \qquad (24.10)$$

где x_0, y_0 — заданные числа. Геометрически задача Коши означает следующее: найти интегральную линию, проходящую через точку $M_0(x_0, y_0)$. В соответствии с теоремой 24.1, если в уравнении $y'=f(x, y)$ функция $f(x, y)$ и ее частная производная по y, т. е. $f'_y(x, y)$, непрерывны в некоторой области G плоскости Oxy, содержащей точку

$M_0(x_0, y_0)$, то решение задачи Коши существует и является единственным. В этом случае через точку $M_0(x_0, y_0)$ проходит единственная интегральная линия. Если известно общее решение $y=\varphi(x, C)$ уравнения (24.7) или его общий интеграл $\Phi(x, y, C)=0$, то нахождение решения задачи Коши сводится к вычислению значения произвольной постоянной C из уравнения $y=\varphi(x_0, C)$ или уравнения $\Phi(x_0, y_0, C)=0$.

Будем считать, что дифференциальное уравнение первого порядка задано в виде (24.9). Рассмотрим частный случай, а именно, когда функции $P(x, y)$ и $Q(x, y)$ представляют собой произведения функции только от x на функцию только от y, т. е. $P(x, y)=f(x)\varphi(y)$, $Q(x, y)=f_1(x)\varphi_1(y)$. В этом случае уравнение принимает вид

$$f(x)\varphi(y)\,dx+f_1(x)\varphi_1(y)\,dy=0. \tag{24.11}$$

Дифференциальное уравнение первого порядка называется *уравнением с разделяющимися переменными*, если его можно привести к виду (24.11), где $f(x)$, $f_1(x)$ — функции только от x; $\varphi(y)$, $\varphi_1(y)$ — функции только от y.

Разделив почленно это уравнение на $f_1(x)\varphi(y)$ в предположении, что

$$f_1(x)\varphi(y)\neq 0, \tag{24.12}$$

получим уравнение

$$\frac{f(x)}{f_1(x)}\,dx+\frac{\varphi_1(y)}{\varphi(y)}\,dy=0. \tag{24.13}$$

Уравнение (24.13) называется уравнением с разделенными переменными: при dx находится функция только от x, при dy — только от y.

Взяв неопределенные интегралы от обеих частей уравнения, получим

$$\int\frac{f(x)}{f_1(x)}\,dx+\int\frac{\varphi_1(y)}{\varphi(y)}\,dy=C. \tag{24.14}$$

Равенством (24.14) выражается общий интеграл уравнения (24.11).

Разумеется, интегралы (или один из них) могут оказаться и «неберущимися», но уравнение (24.11) считается решенным; говорят, что решение найдено «в квадратурах».

Пример. Проинтегрировать уравнение $y'=3\sqrt[3]{y^2}$. Найти решение, удовлетворяющее условию $y=1$ при $x=1$.

Полагая $y \neq 0$, разделяя переменные и интегрируя, получаем

$$dy = 3y^{2/3}dx, \quad \frac{dy}{y^{2/3}} = 3dx, \quad y^{-2/3}dy = 3dx, \quad y^{1/3} = x + C,$$

$$y = (x + C)^3.$$

Подставляя начальные данные $x_0=1$, $y_0=1$ в формулу для общего решения $y=(x+C)^3$, находим значение C: $1=(1+C)^3$, $C=0$; $y=x^3$ — искомое частное решение. Очевидно, $y=0$ — также решение уравнения, это решение является особым: в каждой точке оси Ox нарушаются условия теоремы 24.1 (производная функции $f(x, y)=3y^{2/3}$ по y обращается в бесконечность). Через каждую точку $M_0(x_0, 0)$ оси Ox проходят два решения: $y=(x-x_0)^3$ и $y=0$; последнее решение нельзя получить из общего решения $y=(x+C)^3$ ни при каком численном значении C, включая $C=\pm\infty$ (а только при $C=C(x)=-x$).

§ 24.3. Однородные дифференциальные уравнения первого порядка

Функция $F(x, y)$ называется однородной измерения n, если при любом t выполняется тождество

$$F(tx, ty) = t^n F(x, y). \tag{24.15}$$

Например, функции

$$F_1(x, y) = x + 2y, \quad F_2(x, y) = x^2 \sin\frac{x}{y}, \quad F_3(x, y) = \frac{x+y}{x}$$

являются однородными функциями соответственно первого, второго, нулевого измерений. Действительно,

$$F_1(tx, ty) = tx + 2ty = t(x+2y) = tF_1(x, y);$$

$$F_2(tx, ty) = (tx)^2 \sin\frac{tx}{ty} = t^2x^2 \sin\frac{x}{y} = t^2F_2(x, y);$$

$$F_3(tx, ty) = \frac{tx+ty}{tx} = \frac{t(x+y)}{tx} = \frac{x+y}{x} = t^0F_3(x, y).$$

Дифференциальное уравнение первого порядка

$$P(x, y)\,dx + Q(x, y)\,dy = 0 \tag{24.16}$$

называется *однородным*, если $P(x, y)$ и $Q(x, y)$ — однородные функции одного и того же измерения n. В этом случае соотношение (24.15) для данных функций принимает вид

$$P(tx, ty) = t^nP(x, y), \quad Q(tx, ty) = t^nQ(x, y).$$

Полагая в последних равенствах $t=1:x$, $x\neq 0$, находим

$$P\left(1, \frac{y}{x}\right)=\frac{1}{x^n}P(x, y), \quad Q\left(1, \frac{y}{x}\right)=\frac{1}{x^n}Q(x, y),$$

откуда

$$P(x, y)=x^nP\left(1, \frac{y}{x}\right), \quad Q(x, y)=x^nQ\left(1, \frac{y}{x}\right).$$

Подставив эти выражения в уравнение (24.16), получим

$$x^nP\left(1, \frac{y}{x}\right)dx+x^nQ\left(1, \frac{y}{x}\right)dy=0$$

или

$$P\left(1, \frac{y}{x}\right)dx+Q\left(1, \frac{y}{x}\right)dy=0. \qquad (24.17)$$

Введем новую переменную u по формуле

$$u=\frac{y}{x}, \text{ или } y=ux. \qquad (24.18)$$

Поскольку в этом случае $dy=udx+xdu$, то уравнение (24.17) принимает вид $P(1, u)dx+Q(1, u)(udx+xdu)=0$, или

$$[P(1, u)+uQ(1, u)]dx+xQ(1, u)du=0. \qquad (24.19)$$

Последнее уравнение является уравнением с разделяющимися переменными x, u; из него определяется u, а из формулы (24.18) — искомая функция y.

Если

$$\Phi(x, u, c)=0 \qquad (24.20)$$

— общий интеграл уравнения (24.19), то

$$\Phi\left(x, \frac{y}{x}, c\right)=0 \qquad (24.21)$$

— общий интеграл уравнения (24.16).

§ 24.4. Линейные дифференциальные уравнения первого порядка

Линейным дифференциальным уравнением первого порядка называется уравнение вида

$$a(x)y'+b(x)y=c(x), \qquad (24.22)$$

где $y=y(x)$ — искомая функция; $a(x)$, $b(x)$, $c(x)$ — заданные функции. Будем считать, что они непрерывны на отрезке $[\alpha, \beta]$, причем $a(x)\neq 0$. Поскольку $a(x)\neq 0$ при любом $x\in[\alpha, \beta]$, то данное уравнение можно переписать так:

$$y'+p(x)y=f(x). \qquad (24.23)$$

Решение уравнения будем искать в виде произведения двух функций $v=v(x)$, $u=u(x)$

$$y=uv. \qquad (24.24)$$

Так как $y'=u'v+uv'$, то подстановка выражений для y и y' в уравнение (24.23) приводит его к виду

$$u'v+uv'+p(x)uv=f(x),$$

или

$$u'v+u[v'+p(x)v]=f(x). \qquad (24.25)$$

В качестве v выберем одну из функций, обращающих в нуль сумму в квадратных скобках, т. е. функцию, удовлетворяющую уравнению

$$v'+p(x)v=0. \qquad (24.26)$$

С учетом (24.26) уравнение (24.25) принимает вид

$$u'v=f(x). \qquad (24.27)$$

Уравнение (24.26) является уравнением с разделяющимися переменными x и v, из него определяется функция $v=v(x)$. Функцию $u=u(x)$ находят из уравнения (24.27), которое при $v=v(x)$ также является уравнением с разделяющимися переменными. Определив $u=u(x)$ и $v=v(x)$, по формуле (24.24) найдем y.

Действительно, из уравнения (24.26) получаем

$$v(x)=C_1e^{-\int p(x)\,dx}, \qquad (24.28)$$

а из уравнения (24.27) —

$$u(x) = \frac{1}{C_1} \int f(x)\, e^{\int p(x)dx} dx + C_2. \tag{24.29}$$

По формуле (24.24) находим общее решение линейного уравнения (24.23):

$$y = e^{-\int p(x)\,dx} \left[\int f(x)\, e^{\int p(x)dx} dx + C \right] \quad (C = C_1 C_2). \tag{24.30}$$

Пример. Проинтегрировать дифференциальное уравнение

$$y' \cos x + y \sin x = 1.$$

Это уравнение является линейным (содержит только первые степени y и y', не содержит их произведения). Положим $y = uv$, тогда $y' = u'v + uv'$, уравнение примет вид

$$(u'v + uv')\cos x + uv \sin x = 1, \text{ или } u'v \cos x + u(v' \cos x + v \sin x) = 1. \quad (1)$$

В качестве v выберем одну из функций, удовлетворяющих уравнению

$$v' \cos x + v \sin x = 0, \quad \frac{dv}{dx} \cos x + v \sin x = 0,$$

$$\frac{dv}{v} = -\frac{\sin x dx}{\cos x}, \quad \frac{dv}{v} = \frac{d(\cos x)}{\cos x}.$$

Очевидно, в качестве такой функции можно взять $v = \cos x$. Подставив эту функцию в уравнение (1), или $u'v \cos x = 1$, получим

$$u' \cos^2 x = 1, \quad du = \frac{dx}{\cos^2 x}, \quad u = \operatorname{tg} x + C.$$

Следовательно, общим решением данного дифференциального уравнения является функция $y = uv = (\operatorname{tg} x + C)\cos x$, или $y = C \cos x + \sin x$.

Замечание. *К линейному уравнению приводится уравнение Бернулли*

$$y' + p(x)\, y = f(x)\, y^{\alpha},$$

где α — действительное число. При $\alpha = 0$ и $\alpha = 1$ получаем линейное уравнение.

Пусть $\alpha \neq 0$, $\alpha \neq 1$; предположив, что $y \neq 0$, уравнение перепишем так: $y^{-\alpha} y' + p(x)\, y^{-\alpha+1} = f(x)$.

Полагая $y^{-\alpha+1} = z$, находим $z' = (-\alpha + 1)\, y^{-\alpha} y'$, откуда $y^{-\alpha} y' = z' : (-\alpha + 1)$; уравнение принимает вид

$$z' + (-\alpha + 1)\, p(x)\, z = (-\alpha + 1)\, f(x).$$

Последнее уравнение является линейным дифференциальным уравнением (относительно искомой функции $z = y^{1-\alpha}$).

§ 24.5. Уравнения в полных дифференциалах

Уравнением в полных дифференциалах называется уравнение

$$P(x,\ y)\,dx+Q(x,\ y)\,dy=0, \tag{24.31}$$

левая часть которого есть полный дифференциал некоторой функции $U=U(x,\ y)$, т. е.

$$P(x,\ y)\,dx+Q(x,\ y)\,dy=dU. \tag{24.32}$$

Переписав исходное уравнение в виде $dU=0$, заключим, что общий интеграл уравнения (24.31) определяется формулой

$$U(x,\ y)=C. \tag{24.33}$$

Как известно, полный дифференциал функции $U=U(x,\ y)$ выражается формулой

$$dU=\frac{\partial U}{\partial x}\,dx+\frac{\partial U}{\partial y}\,dy. \tag{24.34}$$

Из равенств (24.32) и (24.34) следует, что

$$P(x,\ y)=\frac{\partial U}{\partial x},\quad Q(x,\ y)=\frac{\partial U}{\partial y}. \tag{24.35}$$

Как известно (см. замечание к § 20.15), необходимое и достаточное условие того, что левая часть уравнения (24.31) является полным дифференциалом некоторой функции, выражается равенством

$$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}. \tag{24.36}$$

Функция $U=U(x,\ y)$, входящая в формулу (24.33), находится из уравнений (24.35).

Пример. Проинтегрировать дифференциальное уравнение $(2x-3y)\,dx+(2y-3x)\,dy=0$.

Для данного уравнения $P(x,\ y)=2x-3y$, $Q(x,\ y)=2y-3x$, $\frac{\partial P}{\partial y}=-3$, $\frac{\partial Q}{\partial x}=-3$.

Так как выполнено условие (24.36), то данное уравнение является уравнением в полных дифференциалах; следовательно,

$$\frac{\partial U}{\partial x}=2x-3y,\quad \frac{\partial U}{\partial y}=2y-3x. \tag{I}$$

Интегрируя первое из этих уравнений (y при этом считается постоянным), находим

$$U(x, y) = x^2 - 3xy + \varphi(y), \qquad \text{(II)}$$

где $\varphi(y)$ — функция, подлежащая определению.

Дифференцируя по y функцию $U = U(x, y)$ и принимая во внимание второе из равенств (I), получаем $-3x + \varphi'(y) = 2y - 3x$, откуда

$$\varphi'(y) = 2y, \quad \frac{d\varphi}{dy} = 2y, \quad d\varphi = 2y\,dy, \quad \varphi(y) = y^2 + C_1.$$

Подставив выражение для $\varphi(y)$ в равенство (II), найдем

$$U(x, y) = x^2 - 3xy + y^2 + C_1.$$

В соответствии с формулой (24.33) получаем

$$x^2 - 3xy + y^2 + C_1 = C_2, \text{ или } x^2 - 3xy + y^2 = C,$$

где $C = C_2 - C_1$.

Итак, $x^2 - 3xy + y^2 = C$ — общий интеграл данного уравнения.

Замечание. Это уравнение является также однородным, его можно проинтегрировать с помощью формулы (24.18).

§ 24.6. Приближенные методы интегрирования дифференциальных уравнений

Прежде всего нужно отметить, что не существует общих методов интегрирования дифференциальных уравнений. Дифференциальные уравнения, интегрируемые с помощью известных методов, в приложениях встречаются сравнительно редко. В связи с этим особое значение приобретают *приближенные методы* решения дифференциальных уравнений.

Среди приближенных методов различают *аналитические* и *численные*. Аналитические методы дают возможность найти решение дифференциального уравнения в виде некоторого аналитического выражения. При использовании численных методов решение дифференциального уравнения получают в виде таблицы численных значений искомой функции при заданных значениях ее аргумента.

К аналитическим методам относятся, например, метод интегрирования дифференциальных уравнений с помощью рядов, метод последовательных приближений и др.

Сущность метода интегрирования дифференциальных уравнений с помощью рядов поясним на примере применения его к линейному уравнению

$$p(x)y' + q(x)y = r(x). \qquad (24.37)$$

Если функции $p(x)$, $q(x)$, $r(x)$ разлагаются в ряды по степеням $(x-a)$, сходящиеся в некотором интервале $(a-R,\ a+R)$, то искомая функция $y=y(x)$ также разлагается в ряд

$$y(x)=c_0+c_1(x-a)+c_2(x-a)^2+\ldots+ \\ +c_n(x-a)^n+\ldots, \qquad (24.38)$$

сходящийся в том же интервале (доказательство см., например, в книге [13]).

В этом интервале будет сходиться и ряд для производной

$$y'(x)=c_1+2c_2(x-a)+\ldots+nc_n(x-a)^{n-1}+\ldots.$$

Подставляя ряды для $y(x)$, $y'(x)$, $p(x)$, $q(x)$, $r(x)$ в уравнение (24.37), сравнивая коэффициенты при одинаковых степенях $(x-a)$, получаем уравнения, из которых определяются коэффициенты c_n $(n=0,\ 1,\ 2,\ \ldots)$ ряда (24.38). Этот способ называется *способом неопределенных коэффициентов*.

Употребляют и другой способ, основанный на применении ряда Тейлора и последовательном дифференцировании данного дифференциального уравнения.

Если требуется найти решение дифференциального уравнения

$$F(x,\ y,\ y')=0, \qquad (24.39)$$

удовлетворяющее условию $y(a)=b$, его ищут в виде

$$y(x)=y(a)+\frac{y'(a)}{1!}(x-a)+\frac{y''(a)}{2!}(x-a)^2+\ldots+ \\ +\frac{y^{(n)}(a)}{n!}(x-a)^n+\ldots. \qquad (24.40)$$

Из начального условия определяется $y(a)$, из уравнения (24.39) находится $y'(a)$. Из уравнения, полученного дифференцированием уравнения (24.39), находят $y''(a)$. Аналогично вычисляют $y'''(a)$, $y^{\mathrm{IV}}(a)$ и т. д.

Замечание. Указанные способы применимы и к дифференциальным уравнениям высших порядков.

К простейшим численным методам относится *метод Эйлера*, основанный на использовании приближенной формулы $y'\approx\dfrac{\Delta y}{\Delta x}$.

Пусть требуется найти численные значения решения дифференциального уравнения

$$y'=f(x, y), \tag{24.41}$$

удовлетворяющего условию $y=y_0$ при $x=x_0$ или $y(x_0)=y_0$, в заданных точках отрезка $[x_0, b]$.

Отрезок $[x_0, b]$ разобьем на n равных частей точками $x_1, x_2, \ldots, x_n=b$, длину каждого элементарного отрезка $[x_k, x_{k+1}]$ обозначим через h, тогда

$$x_k=x_0+kh \quad (k=0, 1, 2, \ldots, n). \tag{24.42}$$

Этим значениям x_k будут соответствовать некоторые значения искомой функции $y=y(x)$. Найдем формулу для определения приближенных значений $y_k=y(x_k)$. Равенство (24.41) заменим приближенным равенством

$$\frac{\Delta y}{\Delta x}=f(x, y) \text{ или } \Delta y=f(x, y)\,\Delta x.$$

Для указанных точек последнее равенство запишется так:

$$\Delta y_k=f(x_k, y_k)\,\Delta x_k,$$

где

$$\Delta x_k=x_{k+1}-x_k=h, \quad \Delta y_k=y_{k+1}-y_k \quad (k=0, 1, \ldots, n-1).$$

Следовательно,

$$y_{k+1}-y_k=f(x_k, y_k)\,h, \text{ или } y_{k+1}=y_k+hf(x_k, y_k). \tag{24.43}$$

По этой формуле определяются приближенные значения искомой функции в заданных точках.

Краткие теоретические сведения о других численных методах интегрирования дифференциальных уравнений и соответствующие примеры имеются, в частности, в [8].

Глава 25

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ ВЫСШИХ ПОРЯДКОВ. СИСТЕМЫ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

В этой главе рассматриваются простейшие интегрируемые дифференциальные уравнения n-го порядка ($n>1$), вводятся основные понятия, связанные с системами дифференциальных уравнений.

§ 25.1. Некоторые интегрируемые типы дифференциальных уравнений n-го порядка. Уравнения, допускающие понижение порядка

Рассмотрим уравнение (24.2), правая часть которого является непрерывной функцией только одной переменной x, т. е. уравнение

$$y^{(n)}=f(x). \tag{25.1}$$

Общее решение уравнения (25.1) находится с помощью n-кратного интегрирования по следующей схеме:

$$y^{(n)}=\frac{dy^{(n-1)}}{dx},\quad \frac{dy^{(n-1)}}{dx}=f(x),$$

$$dy^{(n-1)}=f(x)\,dx,\quad y^{(n-1)}=\int f(x)\,dx+c_1,$$

$$y^{(n-2)}=\int\left[\int f(x)\,dx+c_1\right]dx+c_2=$$

$$=\int dx\int f(x)\,dx+c_1x+c_2,\ \ldots,$$

$$y=\int dx\int dx\ldots\int f(x)\,dx+C_1x^{n-1}+\ldots+C_{n-1}x+C_n.$$

Пример 1. Проинтегрировать уравнение $y'''=\sin x-\cos x$.

Так как

$$y'''=\frac{dy''}{dx},\quad \frac{dy''}{dx}=\sin x-\cos x,\quad dy''=(\sin x-\cos x)\,dx,$$

то $y''=\int(\sin x-\cos x)\,dx$, $y''=-\cos x-\sin x+c_1$.

Аналогично получаем

$$y'=-\sin x+\cos x+c_1x+C_2,\quad y=\cos x+\sin x+C_1x^2+C_2x+C_3$$

$$\left(C_1=\frac{c_1}{2}\right).$$

К числу уравнений, допускающих понижение порядка, относятся следующие:

$$F(y^{(n-1)},\ y^{(n)})=0; \tag{25.2}$$

$$F(x,\ y^{(k)},\ y^{(k+1)},\ \ldots,\ y^{(n)})=0; \tag{25.3}$$

$$F(y,\ y',\ y'',\ \ldots,\ y^{(n)})=0. \tag{25.4}$$

Уравнение (25.2) с помощью подстановки

$$y^{(n-1)}=z, \tag{25.5}$$

где z — новая неизвестная функция, приводится к уравнению первого порядка $F(z, z')=0$. Если $z=\varphi(x, C_1)$ — общее решение последнего уравнения, то с учетом (25.5) получаем $y^{(n-1)}=\varphi(x, C_1)$. Это уравнение вида (25.1), его общее решение может быть найдено с помощью $n-1$ интегрирований (при этом вводятся еще $n-1$ произвольных постоянных $C_2, C_3, \ldots, C_n$).

В случае уравнения (25.3) полагаем

$$y^{(k)}=z, \tag{25.6}$$

тогда $y^{(k+1)}=z'$, $y^{(k+2)}=z''$, $\ldots$, $y^{(n)}=z^{(n-k)}$. Уравнение (25.3) принимает вид $F(x, z, z', \ldots, z^{(n-k)})=0$. Если $z=\varphi(x, C_1, C_2, \ldots, C_{n-k})$ — общее решение этого уравнения (порядок которого равен $n-k$), то уравнение (25.6) принимает вид $y^{(k)}=\varphi(x, C_1, C_2, \ldots, C_{n-k})$. Из последнего уравнения искомая функция $y(x)$ находится путем k-кратного интегрирования. Отметим частный случай уравнения (25.3): при $k=1$ получаем

$$F(x, y', y'', \ldots, y^{(n)})=0. \tag{25.7}$$

Уравнение (25.7) подстановкой $y'=z$ приводится к уравнению $(n-1)$-го порядка $F(x, z, z', \ldots, z^{(n-1)})=0$.

С помощью той же подстановки $y'=z$ порядок уравнения (25.4) также понижается на единицу. Действительно,

$$y''=(y')'=z'=z'_y y'=z'_y z,$$

$$y'''=(y'')'=(z'_y z)'_y y'=z''\cdot z^2+z\cdot z'^2,$$

производная k-го порядка от y по x выражается через производные $(k-1)$-го порядка от z по y. Уравнение (25.4) принимает вид $F_1(y, z, z', \ldots, z^{(n-1)})=0$.

В заключение отметим, что уравнение

$$F(y^{(n-2)}, y^{(n)})=0 \tag{25.8}$$

приводится к уравнению второго порядка $F(z, z'')=0$, где $z=y^{(n-2)}$.

Пример 2. Проинтегрировать уравнение $xy^{\mathrm{IV}}-y'''=0$.

Это уравнение вида (25.3). Положим $y'''=z$, тогда $y^{\mathrm{IV}}=z'$ и уравнение примет вид $xz'-z=0$. Интегрируя полученное уравнение первого порядка, находим

$$x\frac{dz}{dx}-z=0,\quad \frac{dz}{z}=\frac{dx}{x},\quad \ln|z|=\ln|x|+\ln|C|,\quad z=Cx.$$

Подставляя эту функцию в уравнение $y'''=z$, получаем $y'''=Cx$, откуда $y=C_1x^4+C_2x^2+C_3x+C_4$.

§ 25.2. Линейные однородные уравнения n-го порядка

Свойства решений линейного однородного уравнения n-го порядка. *Линейным* дифференциальным уравнением n-го порядка называется уравнение вида

$$q_0(x)y^{(n)}+q_1(x)y^{(n-1)}+\ldots$$
$$\ldots+q_{n-1}(x)y'+q_n(x)y=f_1(x), \tag{25.9}$$

где $y=y(x)$ — искомая функция; $f_1(x)$, $q_k(x)$ $(k=0, 1, 2, \ldots, n)$ — заданные функции. Будем считать, что они определены и непрерывны на некотором отрезке $[a, b]$.

Если $f_1(x)\not\equiv 0$, то данное уравнение называется *линейным неоднородным*, если $f_1(x)\equiv 0$ — *линейным однородным*.

При $q_0(x)\neq 0$ линейное неоднородное уравнение после деления на $q_0(x)$ можно привести к виду

$$y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_{n-1}(x)y'+p_n(x)y=f(x). \tag{25.10}$$

Линейное однородное уравнение n-го порядка в этом случае $(q_0(x)\neq 0)$ приводится к виду

$$y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_{n-1}(x)y'+p_n(x)y=0. \tag{25.11}$$

Последнее уравнение кратко запишем так:

$$L[y]=0, \tag{25.12}$$

где

$$L[y]\equiv y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_{n-1}(x)y'+p_n(x)y. \tag{25.13}$$

Будем называть L *линейным дифференциальным оператором.*

Линейный дифференциальный оператор обладает следующими свойствами:

$$L[Cy]\equiv CL[y] \quad (C=\text{const}); \tag{25.14}$$

$$L[y_1+y_2]\equiv L[y_1]+L[y_2]. \tag{25.15}$$

Действительно,

$$L[Cy]\equiv (Cy)^{(n)}+p_1(x)(Cy)^{(n-1)}+\ldots+p_n(x)(Cy)\equiv$$
$$\equiv C[y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_n(x)y]\equiv CL[y],$$
$$L[y_1+y_2]\equiv (y_1+y_2)^{(n)}+p_1(x)(y_1+y_2)^{(n-1)}+\ldots$$
$$\ldots+p_n(x)(y_1+y_2)\equiv [y_1^{(n)}+p_1(x)y_1^{(n-1)}+\ldots+p_n(x)y_1]+$$
$$+[y_2^{(n)}+p_1(x)y_2^{(n-1)}+\ldots+p_n(x)y_2]\equiv L[y_1]+L[y_2].$$

С помощью формул (25.14), (25.15) докажем некоторые теоремы о свойствах решений линейного однородного уравнения n-го порядка.

Теорема 25.1. Если y_1 — решение линейного однородного уравнения $L[y]=0$, то Cy_1, где C — постоянная, также является решением этого уравнения.

Доказательство. По условию $L[y_1]\equiv 0$. В соответствии с формулой (25.14) получаем

$$L[Cy_1]\equiv CL[y_1]\equiv 0, \quad L[Cy_1]\equiv 0.$$

Это и означает, что Cy_1 — решение уравнения $L[y]=0$.

Теорема 25.2. Если y_1 и y_2 — два решения линейного однородного уравнения $L[y]=0$, то их сумма y_1+y_2 также является решением этого уравнения.

Доказательство. По условию $L[y_1]\equiv 0$, $L[y_2]\equiv 0$. Применяя формулу (25.15), получаем

$$L[y_1+y_2]\equiv L[y_1]+L[y_2]\equiv 0, \quad L[y_1+y_2]\equiv 0.$$

Итак, сумма y_1+y_2 — решение уравнения $L[y]=0$.

Следствие. Если $y_1, y_2, \ldots, y_n$ — решения уравнения $L[y]=0$, то функция

$$y=C_1y_1+C_2y_2+\ldots+C_ny_n \tag{25.16}$$

также является решением этого уравнения.

Это утверждение следует из теорем 25.1, 25.2.

Теорема 25.3. Если уравнение $L[y]=0$ с действительными коэффициентами $p_k(x)$ $(k=1, 2, \ldots, n)$ имеет комплексное решение $y(x)=u(x)+iv(x)$, то действительная $u(x)$ и мнимая $v(x)$ части также являются решениями этого уравнения.

Доказательство. По условию $L[u(x)+iv(x)]\equiv 0$. В соответствии с формулами (25.14) и (25.15) получаем

$$L[u(x)+iv(x)]=L[u(x)]+iL[v(x)]\equiv 0,$$

откуда $L[u(x)]\equiv 0$, $L[v(x)]\equiv 0$, так как комплексная функция действительной переменной тождественно равна нулю тогда и только тогда, когда тождественно равны нулю действительная и мнимая части.

Замечание. Свойства оператора L, выражаемые формулами (25.14) и (25.15), здесь применены к комплексной функции действительной переменной. Это допустимо, поскольку при выводе этих формул использовались свойства производных $(Cy)'=Cy'$, $(y_1+y_2)'=$

$=y_1'+y_2'$, остающиеся справедливыми и для комплексных функций действительной переменной. Производная комплексной функции действительной переменной определяется формулой

$$[u(x)+iv(x)]'=u'(x)+iv'(x). \tag{25.17}$$

Линейная зависимость и линейная независимость функций. Определитель Вронского[1]. Понятие линейной независимости функций вводится аналогично понятию линейной независимости векторов (см. § 5.11). Функции

$$y_1=y_1(x),\ y_2=y_2(x),\ \ldots,\ y_n=y_n(x) \tag{25.18}$$

называются *линейно-зависимыми* на отрезке $[a, b]$, если существуют действительные числа $\alpha_1, \alpha_2, \ldots, \alpha_n$, не все равные нулю ($\alpha_1^2+\alpha_2^2+\ldots+\alpha_n^2\neq 0$), такие, что для любых $x\in[a, b]$ выполняется равенство

$$\alpha_1 y_1+\alpha_2 y_2+\ldots+\alpha_n y_n=0. \tag{25.19}$$

Если равенство (25.19) выполняется лишь при

$$\alpha_1=\alpha_2=\ldots=\alpha_n=0, \tag{25.20}$$

то функции $y_1, y_2, \ldots, y_n$ называются *линейно-независимыми*. Отметим, что функции

$$y_1=1,\ y_2=x,\ y_3=x^2,\ \ldots,\ y_n=x^{n-1} \tag{25.21}$$

линейно-независимы на любом отрезке $[a, b]$. Действительно, равенство $\alpha_1+\alpha_2x+\alpha_3x^2+\ldots+\alpha_nx^{n-1}=0$ для всех $x\in[a, b]$ выполняется лишь в случае, когда $\alpha_1=\alpha_2=\ldots$ $\ldots=\alpha_n=0$ (так как многочлен степени $n-1$ имеет не более $n-1$ корней). Можно доказать, что если $k_i\neq k_j$ при $i\neq j$, то функции

$$y_1=e^{k_1x},\ y_2=e^{k_2x},\ \ldots,\ y_n=e^{k_nx}, \tag{25.22}$$

а также функции

$$e^{k_1x},\ xe^{k_1x},\ \ldots,\ x^{n_1}e^{k_1x},\ e^{k_2x},\ xe^{k_2x},\ \ldots,\ x^{n_2}e^{k_2x},$$
$$e^{k_px},\ xe^{k_px},\ \ldots,\ x^{n_p}e^{k_px} \tag{25.23}$$

линейно-независимы на любом отрезке $[a, b]$.

Очевидно, если одна из функций $y_1(x), y_2(x), \ldots$ $\ldots, y_n(x)$ тождественно равна нулю, то функции линейно-зависимы. Действительно, пусть, например, $y_1(x)\equiv 0$,

[1] *Гене-Вронский* (*Jósef Hoene Wroński*, 1778—1853) — польский математик и философ, работал в Париже.

тогда $1\cdot y_1+0\cdot y_2+\ldots+0\cdot y_n=0$, $\alpha_1=1\neq0$. Если среди n функций k $(k<n)$ линейно-зависимы, то и все функции линейно-зависимы. В самом деле, если $\alpha_1y_1+\alpha_2y_2+\ldots$ $\ldots+\alpha_ky_k=0$, где не все α_i $(i=1, 2, \ldots, k)$ равны нулю, например, $\alpha_1\neq0$, то

$$\alpha_1y_1+\alpha_2y_2+\ldots+\alpha_ky_k+0\cdot y_{k+1}+\ldots+0\cdot y_n=0,\ \alpha_1\neq0.$$

В случае двух функций $y_1=y_1(x)$, $y_2=y_2(x)$ $(y_1\not\equiv0,$ $y_2\not\equiv0)$ необходимым и достаточным условием линейной зависимости является их пропорциональность. Действительно, если

$$y_2=ky_1,\ y_2(x)=ky_1(x),\ k=\mathrm{const}, \tag{25.24}$$

то $ky_1-y_2=0$ или $\alpha_1y_1+\alpha_2y_2=0$, где $\alpha_2=-1\neq0$. Обратно, если $\alpha_1y_1+\alpha_2y_2=0$ и $\alpha_1^2+\alpha_2^2\neq0$, пусть $\alpha_2\neq0$, тогда $y_2=-(\alpha_1:\alpha_2)y_1$ или $y_2=ky_1$, где $k=-\alpha_1:\alpha_2$. Например, функции $y_1=x$, $y_2=2x$ линейно-зависимы на любом отрезке $[a, b]$; функции $y_1=e^{k_1x}$, $y_2=e^{k_2x}$ $(k_1\neq k_2)$ линейно-независимы. Линейно-независимыми на любом отрезке являются также функции

$$y_1=e^{\alpha x}\sin\beta x,\quad y_2=e^{\alpha x}\cos\beta x. \tag{25.25}$$

Теорема 25.4. Если функции $y_1=y_1(x)$, $y_2=y_2(x)$, $\ldots$ $\ldots$, $y_n=y_n(x)$ линейно-зависимы на отрезке $[a, b]$, то определитель

$$w(x)=w[y_1,\ y_2,\ \ldots,\ y_n]=$$

$$=\begin{vmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ y_1'' & y_2'' & \ldots & y_n'' \\ \cdot & \cdot & \cdot & \cdot \\ y_1^{(n-1)} & y_2^{(n-1)} & \ldots & y_n^{(n-1)} \end{vmatrix} \tag{25.26}$$

тождественно равен нулю на этом отрезке.

Доказательство. Так как функции $y_1(x)$, $y_2(x)$, $\ldots$, $y_n(x)$ линейно-зависимы на отрезке $[a, b]$, то по определению справедливо тождество

$$\alpha_1y_1+\alpha_2y_2+\ldots+\alpha_ny_n\equiv0,\quad x\in[a, b],$$

причем не все α_i равны нулю.

Продифференцируем это тождество $n-1$ раз:

$$\left.\begin{aligned} \alpha_1 y_1 + \alpha_2 y_2 + \ldots + \alpha_n y_n &\equiv 0; \\ \alpha_1 y_1' + \alpha_2 y_2' + \ldots + \alpha_n y_n' &\equiv 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ \alpha_1 y_1^{(n-1)} + \alpha_2 y_2^{(n-1)} + \ldots + \alpha_n y_n^{(n-1)} &\equiv 0. \end{aligned}\right\} \qquad (25.27)$$

При любом $x \in [a, b]$ получаем линейную однородную систему алгебраических уравнений относительно n неизвестных $\alpha_1, \alpha_2, \ldots, \alpha_n$. Поскольку эта система имеет ненулевое решение (не все α_i равны нулю), то, как известно, определитель ее равен нулю (см. § 4.2, следствие из теоремы Крамера).

Следовательно, определитель системы (25.27), являющийся определителем (25.26), равен нулю в каждой точке $x \in [a, b]$.

Определитель (25.26) называется определителем Вронского.

Проиллюстрируем теорему 25.4 на примере функций $y_1(x) = e^x$, $y_2(x) = \operatorname{sh} x$, $y_3(x) = \operatorname{ch} x$. Эти функции линейно-зависимы (см. формулы (10.42) и (10.43)), так как для любых x $\operatorname{sh} x + \operatorname{ch} x = e^x$, $\operatorname{sh} x + \operatorname{ch} x - e^x = 0$.

Определитель Вронского для этих функций содержит две одинаковые строки (первую и третью), поэтому равен нулю при любом x.

Теорема 25.5. Если $y_1 = y_1(x)$, $y_2 = y_2(x)$, $\ldots$ $\ldots$, $y_n = y_n(x)$ — линейно-независимые решения уравнения

$$y^{(n)} + p_1(x) y^{(n-1)} + \ldots + p_n(x) y = 0 \qquad (25.28)$$

с коэффициентами $p_k(x)$ $(k = 1, 2, \ldots, n)$, непрерывными на отрезке $[a, b]$, то определитель Вронского $w(x) = w[y_1, y_2, \ldots, y_n]$ не обращается в нуль ни в одной точке отрезка $[a, b]$.

Доказательство. Предположим противное, т. е. в некоторой точке $x_0 \in [a, b]$ определитель Вронского равен нулю: $w(x_0) = 0$.

Выберем числа $\alpha_1, \alpha_2, \ldots, \alpha_n$, среди которых есть отличные от нуля, так, чтобы они удовлетворяли системе алгебраических уравнений

$$\left.\begin{aligned} \alpha_1 y_1(x_0) + \alpha_2 y_2(x_0) + \ldots + \alpha_n y_n(x_0) &= 0; \\ \alpha_1 y_1'(x_0) + \alpha_2 y_2'(x_0) + \ldots + \alpha_n y_n'(x_0) &= 0; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots& \\ \alpha_1 y_1^{(n-1)}(x_0) + \alpha_2 y_2^{(n-1)}(x_0) + \ldots + \alpha_n y_n^{(n-1)}(x_0) &= 0. \end{aligned}\right\} \qquad (25.29)$$

Это можно сделать, поскольку определитель системы (25.29) является определителем Вронского и $w(x_0)=0$. Система в этом случае имеет ненулевое решение.

На основании следствия из теорем 25.1 и 25.2 функция

$$y=\alpha_1 y_1(x)+\alpha_2 y_2(x)+\ldots+\alpha_n y_n(x) \qquad (25.30)$$

есть решение уравнения (25.28). Это решение удовлетворяет начальным условиям

$$y(x_0)=0,\ y'(x_0)=0,\ \ldots,\ y^{(n-1)}(x_0)=0,$$

что следует из равенств (25.29). Но этим начальным условиям удовлетворяет и тривиальное решение $y\equiv 0$ уравнения (25.28).

В соответствии с теоремой 24.1 заключаем, что решение (25.30) тождественно равно нулю, т. е.

$$y(x)\equiv 0 \text{ или } \alpha_1 y_1(x)+\alpha_2 y_2(x)+\ldots+\alpha_n y_n(x)\equiv 0.$$

Последнее означает, что функции $y_1, y_2, \ldots, y_n$ линейно-зависимы на отрезке $[a, b]$, что противоречит условию.

Следовательно, определитель Вронского для линейно-независимых решений уравнения (26.28) не равен нулю ни в одной точке отрезка $[a, b]$.

Общее решение линейного однородного уравнения n-го порядка. Структуру общего решения линейного однородного уравнения n-го порядка выясняет следующая теорема.

Теорема 25.6. Если $y_1=y_1(x)$, $y_2=y_2(x)$, ..., $y_n=y_n(x)$ — линейно-независимые на отрезке $[a, b]$ решения однородного уравнения (25.28), коэффициенты которого непрерывны на этом отрезке, то его общее решение определяется формулой

$$y=C_1 y_1+C_2 y_2+\ldots+C_n y_n, \qquad (25.31)$$

где $C_1, C_2, \ldots, C_n$ — произвольные постоянные.

Доказательство. На основании следствия из теоремы (25.2) функция (25.31) обращает в тождество уравнение (25.28).

Далее, постоянные $C_1, C_2, \ldots, C_n$ можно подобрать так, чтобы получить решение, удовлетворяющее произвольным заданным условиям:

$$y(x_0)=y_0,\ y'(x_0)=y_0',\ \ldots,\ y^{(n-1)}(x_0)=y_0^{(n-1)},$$

т. е.

$$\left.\begin{aligned}&C_1y_1(x_0)+C_2y_2(x_0)+\ldots+C_ny_n(x_0)=y_0;\\&C_1y_1'(x_0)+C_2y_2'(x_0)+\ldots+C_ny_n'(x_0)=y_0';\\&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\&C_1y_1^{(n-1)}(x_0)+C_2y_2^{(n-1)}(x_0)+\ldots\\&\ldots+C_ny_n^{(n-1)}(x_0)=y_0^{(n-1)}.\end{aligned}\right\}\qquad(25.32)$$

Определитель системы (25.32) является определителем Вронского $w(x_0)$. Поскольку решения $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ линейно-независимы на отрезке $[a, b]$, то для любого $x_0 \in [a, b]$ $w(x_0) \neq 0$. Следовательно, система (25.32) имеет единственное решение C_1^0, C_2^0, ..., C_n^0.

Это означает, что формула (25.31) определяет общее решение уравнения (25.28).

Следствие. Максимальное число линейно-независимых решений линейного однородного уравнения равно порядку этого уравнения.

Определение. Любые n линейно-независимых решений линейного однородного уравнения n-го порядка называются его *фундаментальной системой решений.*

§ 25.3. Линейные однородные уравнения n-го порядка с постоянными коэффициентами

Линейным однородным уравнением n-го порядка с постоянными коэффициентами называется уравнение вида

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n-1}y'+a_ny=0,\qquad(25.33)$$

где a_1, a_2, ..., a_n — действительные числа.

Уравнение (25.33) является частным случаем уравнения (25.28), поэтому все сказанное в § 25.2 применимо к уравнению (25.33).

Решение уравнения (25.33) будем искать в виде

$$y=e^{kx}\quad(k=\text{const}).\qquad(25.34)$$

Подставляя эту функцию и ее производные

$$y'=ke^{kx},\quad y''=k^2e^{kx},\ \ldots,\ y^{(n)}=k^ne^{kx}$$

в уравнение (25.33), получаем

$$(k^ne^{kx}+a_1k^{n-1}e^{kx}+\ldots+a_ne^{kx})=0,$$
$$e^{kx}(k^n+a_1k^{n-1}+\ldots+a_{n-1}k+a_n)=0.$$

Функция (25.34) будет решением уравнения тогда и только тогда, когда

$$k^n+a_1k^{n-1}+\ldots+a_{n-1}k+a_n=0, \tag{25.35}$$

т. е. когда k — корень алгебраического уравнения (25.35).

Уравнение (25.35) называется *характеристическим уравнением*. Это алгебраическое уравнение n-й степени, оно имеет n корней (считая и равные корни), среди которых могут быть и комплексные.

Рассмотрим основные возможные случаи.

1. Характеристическое уравнение имеет различные действительные корни.

Обозначим эти корни через $k_1, k_2, \ldots, k_n$ ($k_i \neq k_j$ при $i \neq j$). Им соответствуют решения уравнения (25.33):

$$y_1=e^{k_1x},\ y_2=e^{k_2x},\ \ldots,\ y_n=e^{k_nx}. \tag{25.36}$$

Функции (25.36) в этом случае ($k_i \neq k_j$ при $i \neq j$) линейно-независимы на любом отрезке $[a, b]$ (см. (25.22)). В соответствии с формулой (25.31) получаем общее решение уравнения (25.33):

$$y=C_1e^{k_1x}+C_2e^{k_2x}+\ldots+C_ne^{k_nx}. \tag{25.37}$$

Пример 1. Проинтегрировать уравнение $y'''-2y''-y'+2y=0$. Это линейное однородное дифференциальное уравнение третьего порядка с постоянными коэффициентами. Ему соответствует характеристическое уравнение $k^3-2k^2-k+2=0$, имеющее корни $k_1=-1$, $k_2=1$, $k_3=2$.

Следовательно, общее решение данного дифференциального уравнения определяется формулой $y=C_1e^{-x}+C_2e^x+C_3e^{2x}$.

2. Характеристическое уравнение имеет действительные корни, среди которых m равны между собой.

Будем говорить в этом случае, что среди корней характеристического уравнения имеется один m-кратный корень. Обозначим его через k_1. Предположим, что все остальные действительные корни различны. Корням характеристического уравнения

$$k_1=k_2=\ldots=k_m,\ k_{m+1},\ k_{m+2},\ \ldots,\ k_n$$

соответствуют решения дифференциального уравнения

$$y_1=e^{k_1x},\ y_2=e^{k_1x},\ \ldots,\ y_m=e^{k_1x},$$
$$y_{m+1}=e^{k_{m+1}x},\ \ldots,\ y_n=e^{k_nx}.$$

Эти решения линейно-зависимы (так как линейно-зависимы первые m функций; они равны между собой), поэтому посредством их получить общее решение дифференциального уравнения не представляется возможным.

Можно показать, что корню k_1 кратности m будут соответствовать m линейно-независимых решений $y_1=e^{k_1x}$, $y_2=xe^{k_1x}, \ldots, y_m=x^{m-1}e^{k_1x}$. Рассмотрим n решений

$$\begin{gathered} y_1=e^{k_1x},\ y_2=xe^{k_1x},\ \ldots,\ y_m=x^{m-1}e^{k_1x}, \\ y_{m+1}=e^{k_{m+1}x},\ \ldots,\ y_n=e^{k_nx}. \end{gathered} \tag{25.38}$$

Решения (25.38) линейно-независимы на любом отрезке $[a, b]$ (см. (25.23)), поэтому общее решение уравнения (25.33) определяется формулой

$$y=C_1e^{k_1x}+C_2xe^{k_1x}+\ldots+C_mx^{m-1}e^{k_1x}+ \\ +C_{m+1}e^{k_{m+1}x}+\ldots+C_ne^{k_nx},$$

или

$$y=(C_1+C_2x+C_3x^2+\ldots+C_mx^{m-1})e^{k_1x}+ \\ +C_{m+1}e^{k_{m+1}x}+\ldots+C_ne^{k_nx}. \tag{25.39}$$

Пример 2. Проинтегрировать уравнение $y'''+3y''+3y'+y=0$.

Характеристическое уравнение $k^3+3k^2+3k+1=0$, или $(k+1)^3=0$, имеет трехкратный корень $k_1=k_2=k_3=-1$, поэтому общее решение данного дифференциального уравнения определяется формулой $y=(C_1+C_2x+C_3x^2)e^{-x}$.

3. Характеристическое уравнение имеет простые комплексно-сопряженные корни.

Обозначим эти корни $k_1=\alpha-i\beta$, $k_2=\alpha+i\beta$, где α, β — действительные числа, $i=\sqrt{-1}$. Им соответствуют комплексные решения дифференциального уравнения

$$y_1=e^{k_1x}=e^{(\alpha-i\beta)x}=e^{\alpha x}(\cos\beta x-i\sin\beta x), \\ y_2=e^{k_2x}=e^{(\alpha+i\beta)x}=e^{\alpha x}(\cos\beta x+i\sin\beta x).$$

На основании теоремы 25.3 уравнение будет также иметь и действительные решения $y_1=e^{\alpha x}\cos\beta x$, $y_2=e^{\alpha x}\sin\beta x$.

Если все остальные корни $k_3, k_4, \ldots, k_n$ являются действительными и различными, то общее решение уравнения (25.33) определяется формулой

$$y=(C_1\cos\beta x+C_2\sin\beta x)e^{\alpha x}+C_3e^{k_3x}+\ldots+C_ne^{k_nx}. \tag{25.40}$$

В частном случае, когда характеристическое уравнение имеет чисто мнимые корни $k_1=-i\beta$, $k_2=i\beta$ $(\alpha=0)$, формула (25.40) принимает вид

$$y=C_1\cos\beta x+C_2\sin\beta x+C_3e^{k_3x}+\ldots+C_ne^{k_nx}. \tag{25.41}$$

Если среди корней $k_3, k_4, \ldots, k_n$ имеются кратные, то общее решение записывается с учетом формулы (25.39).

Пример 3. Проинтегрировать уравнение $y'''-2y''+y'-2y=0$. Характеристическое уравнение

$$k^3-2k^2+k-2=0,\ k^2(k-2)+(k-2)=0,\ (k^2+1)(k-2)=0$$

имеет корни $k_1=-i$, $k_2=i$, $k_3=2$, поэтому в соответствии с формулой (25.41) получаем общее решение $y=C_1\cos x+C_2\sin x+C_3e^{2x}$ данного дифференциального уравнения.

4. Характеристическое уравнение имеет кратные комплексные корни.

Пусть комплексно-сопряженные корни $k_1=\alpha-i\beta$, $k_2=\alpha+i\beta$ являются m-кратными. Если остальные $n-2m$ корней $k_{2m+1}, \ldots, k_n$ — действительны и различны, можно показать, что общее решение дифференциального уравнения имеет вид

$$y=(C_1+C_2x+C_3x^2+\ldots+C_mx^{m-1})e^{\alpha x}\cos\beta x+ \\ +(C_{m+1}+C_{m+2}x+\ldots+C_{2m}x^{m-1})e^{\alpha x}\sin\beta x+ \\ +C_{2m+1}e^{k_{2m+1}x}+\ldots+C_ne^{k_nx}. \qquad (25.42)$$

В частном случае, когда характеристическое уравнение имеет кратные корни $k_1=-i\beta$, $k_2=i\beta$ $(\alpha=0)$, формула (25.42) принимает вид

$$y=(C_1+C_2x+C_3x^2+\ldots+C_mx^{m-1})\cos\beta x+ \\ +(C_{m+1}+C_{m+2}x+\ldots+C_{2m}x^{m-1})\sin\beta x+ \\ +C_{2m+1}e^{k_{2m+1}x}+\ldots+C_ne^{k_nx}. \qquad (25.43)$$

Если среди действительных корней имеются кратные, то общее решение записывается с учетом формулы (25.39).

Пример 4. Проинтегрировать уравнение $y^{\mathrm{IV}}+2y''+y=0$. Характеристическое уравнение $k^4+2k^2+1=0$, $(k^2+1)^2=0$ имеет двукратные корни $k_1=-i$, $k_2=i$, поэтому

$$y=(C_1+C_2x)\cos x+(C_3+C_4x)\sin x.$$

§ 25.4. Линейные неоднородные уравнения n-го порядка

Рассмотрим линейное неоднородное уравнение n-го порядка

$$y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_{n-1}(x)y'+p_n(x)y=f(x). \qquad (25.44)$$

Будем считать, что функции $p_k(x)$ $(k=1, 2, \ldots, n)$ и $f(x)$ непрерывны на отрезке $[a, b]$. Уравнение (25.44) кратко запишем так:

$$L[y]=f(x),$$
$$L[y]\equiv y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_{n-1}(x)y+p_n(x)y. \qquad (25.45)$$

Теорема 25.7. Если y_0 — решение однородного уравнения $L[y]=0$, y_1 — решение соответствующего неоднородного уравнения $L[y]=f(x)$, то сумма y_0+y_1 является решением этого неоднородного уравнения.

Доказательство. По условию $L[y_0]\equiv 0$, $L[y_1]\equiv f(x)$, поэтому

$$L[y_0+y_1]\equiv L[y_0]+L[y_1]\equiv 0+f(x),$$
$$L[y_0+y_1]\equiv f(x).$$

Это означает, что y_0+y_1 — решение неоднородного уравнения $L[y]=f(x)$.

Структура общего решения неоднородного уравнения (25.44) определяется следующей теоремой.

Теорема 25.8. Если Y — частное решение уравнения $L[y]=f(x)$ с непрерывными коэффициентами, $y_0=\sum\limits_{k=1}^{n} C_k y_k$ — общее решение соответствующего однородного уравнения $L[y]=0$, то общее решение данного неоднородного уравнения определяется формулой

$$y=y_0+Y, \text{ или } y=\sum_{k=1}^{n} C_k y_k+Y. \qquad (25.46)$$

Доказательство этой теоремы предоставляется читателю, оно аналогично доказательству теоремы 25.6.

Замечание. Чтобы записать общее решение линейного неоднородного уравнения, необходимо найти какое-нибудь частное решение этого уравнения и общее решение соответствующего однородного уравнения (для чего нужно знать n линейно-независимых его решений).

§ 25.5. Линейные неоднородные уравнения n-го порядка с постоянными коэффициентами

Рассмотрим линейное неоднородное уравнение n-го порядка с постоянными коэффициентами

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n-1}y'+a_ny=f(x), \qquad (25.47)$$

где $a_1, a_2, \ldots, a_n$ — действительные числа. Запишем соответствующее однородное уравнение, т. е. уравнение

$$y^{(n)}+a_1y^{(n-1)}+\ldots+a_{n-1}y'+a_ny=0, \qquad (25.48)$$

в котором постоянные $a_1, a_2, \ldots, a_n$ те же, что и в уравнении (25.47). Поскольку уравнение (25.47) есть частный случай уравнения (25.44), общее решение его определяется формулой (25.46).

Общее решение y_0 однородного уравнения (25.48) находится согласно § 25.3, частное решение Y неоднородного уравнения может быть найдено *методом неопределенных коэффициентов* в следующих простейших случаях.

1. $f(x)=e^{ax}P_n(x)$, где $P_n(x)$ — многочлен степени n.

Если a не является корнем соответствующего характеристического уравнения, то полагают $Y=e^{ax}Q_n(x)$, где $Q_n(x)$ — многочлен степени n с неопределенными коэффициентами. Если a — корень характеристического уравнения, то $Y=x^re^{ax}Q_n(x)$, где r — кратность корня a.

2. $f(x)=e^{ax}[P_n(x)\cos bx+R_m(x)\sin bx]$.

Если $a\pm ib$ не являются корнями характеристического уравнения, то полагают

$$Y=e^{ax}[Q_\nu(x)\cos bx+S_\nu(x)\sin bx],$$

где $Q_\nu(x)$, $S_\nu(x)$ — многочлены степени $\nu=\max\{n, m\}$ с неопределенными коэффициентами. Если $(a\pm ib)$ — корни характеристического уравнения, то

$$Y=x^re^{ax}[Q_\nu(x)\cos bx+S_\nu(x)\sin bx],$$

где r — кратность корней $a\pm ib$.

В общем случае при решении неоднородного уравнения (25.47) применяется метод вариации произвольных постоянных (см. § 25.6).

Пример. Проинтегрировать уравнение $y'''-5y''+8y'-6y=12e^{2x}$.

Найдем сначала общее решение соответствующего однородного уравнения $y'''-5y''+8y'-6y=0$. Так как характеристическое уравнение $k^3-5k^2+8k-6=0$, $k^3-3k^2-2k^2+6k+2k-6=0$, $k^2(k-3)-2k(k-3)+2(k-3)=0$, $(k-3)(k^2-2k+2)=0$ имеет корни $k_1=1-i$, $k_2=1+i$, $k_3=3$, то общее решение однородного уравнения определяется формулой

$$y_0=(C_1\cos x+C_2\sin x)e^x+C_3e^{3x}.$$

Частное решение неоднородного уравнения будем искать в виде

$Y=Ae^{2x}$. Подставляя в исходное уравнение выражения для функции Y и ее производных

$$Y'=2Ae^{2x},\quad Y''=4Ae^{2x},\quad Y'''=8Ae^{2x},$$

получаем

$$8Ae^{2x}-20Ae^{2x}+16Ae^{2x}-6Ae^{2x}\equiv 12e^{2x},\qquad -2Ae^{2x}\equiv 12e^{2x},$$

откуда $-2A=12$, $A=-6$. Итак, $Y=-6e^{2x}$.

В соответствии с формулой (25.46) находим общее решение данного неоднородного уравнения: $y=(C_1\cos x+C_2\sin x)e^x+C_3e^{3x}-6e^{2x}$.

§ 25.6. Метод вариации произвольных постоянных

Обратимся снова к линейному неоднородному уравнению n-го порядка с переменными коэффициентами

$$y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_n(x)y=f(x). \tag{25.49}$$

Будем считать, что функции $p_1(x), \ldots, p_n(x)$ непрерывны на некотором отрезке $[a, b]$.

Если нахождение частного решения этого уравнения оказывается затруднительным, но известно общее решение соответствующего однородного уравнения, то общее решение уравнения (25.49) можно найти *методом вариации произвольных постоянных.*

Пусть соответствующее однородное уравнение

$$y^{(n)}+p_1(x)y^{(n-1)}+\ldots+p_n(x)y=0 \tag{25.50}$$

имеет общее решение

$$y=C_1y_1+C_2y_2+\ldots+C_ny_n. \tag{25.51}$$

Общее решение уравнения (25.49) будем искать в виде

$$y=C_1(x)y_1(x)+C_2(x)y_2(x)+\ldots+C_n(x)y_n(x), \tag{25.52}$$

где $y_1=y_1(x)$, $y_2=y_2(x)$, $\ldots$, $y_n=y_n(x)$ — линейно-независимые решения уравнения (25.50), входящие в формулу (25.51), а $C_1(x)$, $C_2(x)$, $\ldots$, $C_n(x)$ — неизвестные функции.

Чтобы найти эти функции, подчиним их некоторым условиям.

Найдем производную функции (25.52):

$$y'=[C_1(x)y_1'+C_2(x)y_2'+\ldots+C_n(x)y_n']+$$
$$+[C_1'(x)y_1+C_2'(x)y_2+\ldots+C_n'(x)y_n].$$

Потребуем, чтобы сумма во второй квадратной скобке равнялась нулю, т. е.

$$C_1'y_1 + C_2'y_2 + \ldots + C_n'y_n = 0,$$

тогда $y' = C_1y_1' + C_2y_2' + \ldots + C_ny_n'$.

Найдем вторую производную

$$y'' = (C_1y_1'' + C_2y_2'' + \ldots + C_ny_n'') + \\ + (C_1'y_1' + C_2'y_2' + \ldots + C_n'y_n')$$

и потребуем, чтобы

$$C_1'y_1' + C_2'y_2' + \ldots + C_n'y_n' = 0,$$

тогда $y'' = C_1y_1'' + C_2y_2'' + \ldots + C_ny_n''$.

Продолжая аналогичный процесс, получаем

$$y^{(n)} = (C_1y_1^{(n)} + C_2y_2^{(n)} + \ldots + C_ny_n^{(n)}) + \\ + (C_1'y_1^{(n-1)} + C_2'y_2^{(n-1)} + \ldots + C_n'y_n^{(n-1)}).$$

В этом случае нельзя потребовать, чтобы сумма во второй скобке обратилась в нуль, так как функции $C_1(x)$, $C_2(x), \ldots, C_n(x)$ уже подчинены $n-1$ условиям, а нужно еще удовлетворить уравнению (25.49).

Подставляя в уравнение (25.49) выражения для функции (25.52) и ее производных

$$y' = \sum_{k=1}^{n} C_ky_k', \quad y'' = \sum_{k=1}^{n} C_ky_k'', \quad \ldots, \quad y^{(n-1)} = \sum_{k=1}^{n} C_ky_k^{(n-1)},$$

$$y^{(n)} = \sum_{k=1}^{n} C_ky_k^{(n)} + \sum_{k=1}^{n} C_k'y_k^{(n-1)},$$

получаем

$$\sum_{k=1}^{n} C_k'y_k^{(n-1)} + \sum_{k=1}^{n} C_ky_k^{(n)} + p_1(x)\sum_{k=1}^{n} C_ky_k^{(n-1)} + \\ + p_2(x)\sum_{k=1}^{n} C_ky_k^{(n-2)} + \ldots + p_{n-1}(x)\sum_{k=1}^{n} C_ky_k' + \\ + p_n(x)\sum_{k=1}^{n} C_ky_k = f(x),$$

или

$$\sum_{k=1}^{n} C_k' y_k^{(n-1)} + \sum_{k=1}^{n} C_k [y_k^{(n)} + p_1(x) y_k^{(n-1)} + \ldots$$
$$\ldots + p_{n-1}(x) y_k' + p_n(x) y_k] = f(x).$$

Поскольку $y_k = y_k(x)$ $(k=1, 2, \ldots, n)$ — решения уравнения (25.50), то

$$y_k^{(n)} + p_1(x) y_k^{(n-1)} + \ldots + p_{n-1} y_k' + p_n(x) y_k \equiv 0$$

и последнее уравнение принимает вид

$$\sum_{k=1}^{n} C_k' y_k^{(n-1)} = f(x).$$

Следовательно, для определения функций $C_k'(x)$ $(k=1, 2, \ldots, n)$ имеем систему линейных уравнений

$$\left.\begin{aligned} C_1' y_1 + C_2' y_2 + \ldots + C_n' y_n &= 0; \\ C_1' y_1' + C_2' y_2' + \ldots + C_n' y_n' &= 0; \\ \ldots \ldots \ldots \ldots \ldots & \\ C_1' y_1^{(n-2)} + C_2' y_2^{(n-2)} + \ldots + C_n' y_n^{(n-2)} &= 0; \\ C_1' y_1^{(n-1)} + C_2' y_2^{(n-1)} + \ldots + C_n' y_n^{(n-1)} &= f(x). \end{aligned}\right\} \qquad (25.53)$$

Определитель этой системы отличен от нуля, как определитель Вронского для линейно-независимых функций, поэтому система имеет единственное решение.

Из системы (25.53) находим $C_k' = \varphi_k(x)$ $(k = 1, 2, \ldots, n)$, а потом и сами функции

$$C_k(x) = \int \varphi_k(x) dx + \tilde{C}_k \quad (k=1, 2, \ldots, n), \qquad (25.54)$$

где $\tilde{C}_k$ — постоянные.

Подставляя выражения (25.54) в формулу (25.52), получаем искомое общее решение уравнения (25.49).

§ 25.7. Линейные дифференциальные уравнения второго порядка с постоянными коэффициентами

Рассмотрим линейное неоднородное дифференциальное уравнение второго порядка с постоянными коэффициентами

$$y'' + py' + qy = f(x) \qquad (25.55)$$

и соответствующее ему однородное уравнение

$$y''+py'+qy=0. \tag{25.56}$$

Уравнения (25.56), (25.55) являются частными случаями уравнений (25.33) и (25.47), исследование которых дано в § 25.3 и 25.5. Если характеристическое уравнение

$$k^2+pk+q=0 \tag{25.57}$$

имеет различные действительные корни k_1 и k_2 ($k_1 \neq k_2$), то общее решение уравнения (25.56) определяется формулой

$$y=C_1e^{k_1x}+C_2e^{k_2x}. \tag{25.58}$$

В случае кратного корня характеристического уравнения $k_1=k_2=-p/2$ общее решение уравнения (25.56) выражается формулой

$$y=(C_1+C_2x)\,e^{-px/2}. \tag{25.59}$$

Если характеристическое уравнение имеет комплексно-сопряженные корни $k_1=\alpha-i\beta$, $k_2=\alpha+i\beta$, то

$$y=e^{\alpha x}(C_1\cos\beta x+C_2\sin\beta x). \tag{25.60}$$

В частности, если уравнение (25.57) имеет чисто мнимые корни $k_1=-i\beta$, $k_2=i\beta$ $(\alpha=0)$, то

$$y=C_1\cos\beta x+C_2\sin\beta x. \tag{25.61}$$

Общее решение неоднородного уравнения (25.55) определяется формулой

$$y=y_0+Y, \tag{25.62}$$

где y_0 — общее решение уравнения (25.56); Y — частное решение уравнения (25.55). В простейших случаях частное решение можно найти методом неопределенных коэффициентов. Например, если $f(x)=ax^2+bx+c$, то частное решение ищут в виде функции

$$Y=Ax^2+Bx+C. \tag{25.63}$$

Подставляя функцию (25.63) и ее производные $Y'=2Ax+B$, $Y''=2A$ в уравнение (25.55), получаем

$$2A+p(2Ax+B)+q(Ax^2+Bx+C)\equiv ax^2+bx+c,$$

откуда

$$qA=a,\ 2Ap+Bq=b,\ 2A+Bp+qC=c. \tag{25.64}$$

Если $q \neq 0$, то из системы (25.64) определяются значения коэффициентов A, B, C. Если $q=0$, то система (25.64) противоречива; в этом случае нет частного решения вида (25.63). Частное решение следует искать в виде функции $Y=x(Ax^2+Bx+C)$, или $Y=Ax^3+Bx^2+Cx$. Примеры интегрирования дифференциальных уравнений второго порядка с постоянными коэффициентами приведены в [7, ч. 2; 13].

§ 25.8. Уравнение колебаний

Рассмотрим задачу о механических колебаниях. Пусть груз массы m покоится на упругой рессоре, закрепленной в некоторой точке A (рис. 25.1). Отклонение груза от положения равновесия обозначим буквой y. Будем считать положительным отклонение вниз и отрицательным — отклонение вверх. В положении равновесия сила веса уравновешивается упругостью рессоры. Предположим, что сила F_1, стремящаяся вернуть груз в положение равновесия (так называемая *восстанавливающая* сила), пропорциональна отклонению, т. е.

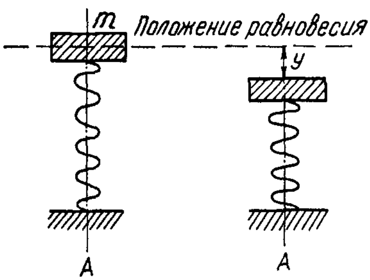

Рис. 25.1

$$F_1=-ky \quad (k=\mathrm{const},\ k>0) \qquad (25.65)$$

(k — положительная постоянная для данной рессоры, так называемая жесткость рессоры). Предположим еще, что движению груза препятствует сила сопротивления F_2, направленная в сторону, противоположную направлению движения и пропорциональная скорости v движения груза, т. е.

$$F_2=-\lambda v=-\lambda\frac{dy}{dt} \qquad (25.66)$$

$$(\lambda=\mathrm{const},\ \lambda>0).$$

Составим дифференциальное уравнение движения груза на рессоре. В соответствии со вторым законом Ньютона получаем $mw = F_1 + F_2$, где $w = d^2y/dt^2$, или

$$m\frac{d^2y}{dt^2} = -ky - \lambda\frac{dy}{dt} \quad (k > 0,\ \lambda > 0). \qquad (25.67)$$

Уравнение (25.67) является линейным дифференциальным уравнением второго порядка с постоянными коэффициентами, его можно записать в виде

$$\frac{d^2y}{dt^2} + p\frac{dy}{dt} + qy = 0, \qquad (25.68)$$

где

$$p = \frac{\lambda}{m}, \quad q = \frac{k}{m}. \qquad (25.69)$$

Уравнение (25.68) называется уравнением *свободных колебаний.*

Рассмотрим случай, когда точка A прикрепления рессоры совершает вертикальное движение по закону $z = \varphi(t)$. Это имеет место тогда, когда нижний конец рессоры прикреплен к катку, движущемуся вместе с рессорой и грузом по неровности (рис. 25.2).

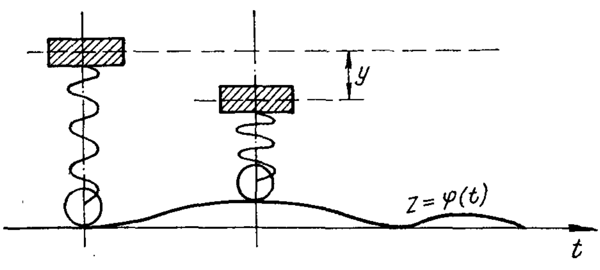

Рис. 25.2

При этом восстанавливающая сила F_1 будет выражаться формулой $F_1 = -k[y + \varphi(t)]$, а сила сопротивления F_2 — формулой $F_2 = -\lambda[y' + \varphi'(t)]$. Дифференциальное уравнение движения груза на рессоре принимает вид

$$m\frac{d^2y}{dt^2} = -k[y + \varphi(t)] - \lambda[y' + \varphi'(t)] \qquad (25.70)$$

или

$$\frac{d^2y}{dt^2} + p\frac{dy}{dt} + qy = f(t), \qquad (25.71)$$

где

$$p=\frac{\lambda}{m},\ q=\frac{k}{m},\ f(t)=-\frac{k\varphi(t)+\lambda\varphi'(t)}{m}. \qquad (25.72)$$

Уравнение (25.71) является линейным неоднородным дифференциальным уравнением второго порядка с постоянными коэффициентами. Это уравнение называется уравнением *вынужденных колебаний.*

Свободные колебания. Рассмотрим уравнение свободных колебаний, т. е. уравнение $\frac{d^2y}{dt^2}+p\frac{dy}{dt}+qy=0$. Соответствующее характеристическое уравнение $k^2+pk+q=0$ имеет корни

$$k_1=-\frac{p}{2}+\sqrt{\frac{p^2}{4}-q},\quad k_2=-\frac{p}{2}-\sqrt{\frac{p^2}{4}-q},$$

в зависимости от которых получаем общее решение уравнения свободных колебаний.

Рассмотрим сначала случай, когда отсутствует сила сопротивления, т. е. $F_2=0$, тогда $\lambda=0$ и $p=0$ (см. (25.66) и (25.69)).

Характеристическое уравнение в этом случае принимает вид $k^2+q=0$ и имеет комплексные корни $k_1=-i\beta$, $k_2=i\beta$, где $\beta=\sqrt{q}$, поэтому общее решение уравнения свободных колебаний $y''+qy=0$ определяется формулой

$$y=C_1\cos\beta t+C_2\sin\beta t. \qquad (25.73)$$

Преобразуем правую часть этой формулы, введя новые постоянные A и φ_0:

$$C_1=A\sin\varphi_0,\quad C_2=A\cos\varphi_0,$$

откуда

$$A=\sqrt{C_1^2+C_2^2},\quad \varphi_0=\operatorname{arctg}\frac{C_1}{C_2}.$$

Подставляя выражения для C_1 и C_2 в формулу (25.73), получаем общее решение

$$y=A\sin(\beta t+\varphi_0), \qquad (25.74)$$

где A и φ_0 — произвольные постоянные.

Колебания в данном случае называются *гармоническими.* Интегральными линиями являются синусоиды. *Периодом* колебаний называется промежуток времени T, за который аргумент синуса изменится на 2π. В рассматриваемом случае $T=2\pi/\beta$. *Частотой* колебаний называется число колебаний за время 2π (в данном случае частота равна β). *Амплитуда* колебаний есть величина наибольшего отклонения от положения равновесия. Из формулы (25.74) видно, что это наибольшее отклонение равно A. *Начальной фазой* называется величина φ_0.

Исследуем случай, когда $F_2\neq 0$, т. е. $p\neq 0$. Рассмотрим три возможности: $\frac{p^2}{4}<q$, $\frac{p^2}{4}>q$, $\frac{p^2}{4}=q$.

Если $\frac{p^2}{4}<q$, то характеристическое уравнение $k^2+pk+q=0$ имеет комплексные корни $k_1=\alpha+i\beta$, $k_2=\alpha-i\beta$, где $\alpha=-\frac{p}{2}<0$, $\beta=\sqrt{q-\frac{p^2}{4}}$.

Общее решение уравнения (25.68) определяется формулой (см. § 25.7)

$$y=e^{\alpha t}(C_1\cos\beta t+C_2\sin\beta t) \tag{25.75}$$

или

$$y=Ae^{\alpha t}\sin(\beta t+\varphi_0). \tag{25.76}$$

В качестве амплитуды здесь рассматривают величину $Ae^{\alpha t}$, зависящую от времени. Поскольку $\alpha<0$, то эта величина стремится к нулю при $t\to\infty$ и колебания называются затухающими.

Если $\frac{p^2}{4}>q$, то характеристическое уравнение имеет действительные различные отрицательные корни k_1 и k_2. Общее решение уравнения (25.50) определяется формулой

$$y=C_1e^{k_1t}+C_2e^{k_2t}\quad (k_1<0,\ k_2<0), \tag{25.77}$$

из которой видно, что $y\to 0$ при $t\to\infty$. Колебаний в данном случае не будет (сила сопротивления велика по сравнению с жесткостью рессоры).

Если $\frac{p^2}{4}=q$, то корни характеристического уравнения равны между собой ($k_1=k_2=-p/2<0$), поэтому общее решение уравнения (25.68) определяется формулой

$$y=(C_1+C_2t)\,e^{-pt/2}. \tag{25.78}$$

Здесь также $y\to 0$ при $t\to\infty$, однако y стремится медленнее, чем в предыдущем случае.

Вынужденные колебания. Обратимся к уравнению вынужденных колебаний, т. е. к уравнению (25.71). Рассмотрим практически важный случай, когда возмущающая внешняя сила является периодической и изменяется по закону $f(t)=a\sin\omega t$. В этом случае уравнение (25.71) запишется так:

$$\frac{d^2y}{dt^2}+p\frac{dy}{dt}+qy=a\sin\omega t. \qquad (25.79)$$

Исследуем уравнение (25.79) при $p=0$ (сила сопротивления отсутствует) и $p\neq 0$, $p^2/4<q$.

Если сила сопротивления отсутствует, т. е. $p=0$, то уравнение вынужденных колебаний принимает вид

$$\frac{d^2y}{dt^2}+qy=a\sin\omega t. \qquad (25.80)$$

Соответствующее однородное уравнение имеет общее решение

$$y_0=C_1\cos\beta t+C_2\sin\beta t \quad (\beta^2=q) \qquad (25.81)$$

или

$$y_0=A\sin(\beta t+\varphi_0). \qquad (25.82)$$

Предположим, что частота внешней силы не равна частоте собственных колебаний, т. е. $\beta\neq\omega$, тогда частное решение неоднородного уравнения (25.80) имеет вид

$$Y=P\cos\omega t+Q\sin\omega t. \qquad (25.83)$$

Подставляя эту функцию и ее вторую производную в уравнение (25.80), находим коэффициенты P и Q: $P=0$, $Q=\frac{a}{q-\omega^2}$.

Следовательно,

$$Y=\frac{a}{q-\omega^2}\sin\omega t. \qquad (25.84)$$

Общее решение уравнения (25.80) определяется формулой $y=y_0+Y$, т. е.

$$y=A\sin(\beta t+\varphi_0)+\frac{a}{q-\omega^2}\sin\omega t. \qquad (25.85)$$

Если частота собственных колебаний совпадает с частотой внешней силы, т. е. $\beta=\omega$, то частное решение уравнения (25.80) следует искать (см. § 25.5) в следующем виде:

$$Y=t(P\cos\omega t+Q\sin\omega t). \qquad (25.86)$$

Подставляя функцию (25.86) и ее вторую производную в уравнение (25.80), находим $P=-\frac{a}{2\omega}$, $Q=0$, т. е. $Y=-\frac{a}{2\omega}t\cos\omega t$.

Итак, общее решение уравнения (25.80) в этом случае имеет вид

$$y=A\sin(\beta t+\varphi_0)-\frac{a}{2\omega}t\cos\omega t,$$

или

$$y=A\sin(\beta t+\varphi_0)-\frac{a}{2\beta}t\cos\beta t, \tag{25.87}$$

поскольку $\omega=\beta$.

Из последней формулы видно (см. второе слагаемое алгебраической суммы), что амплитуда колебания неограниченно возрастает при неограниченном возрастании времени t. Рассматриваемое явление (частота собственных колебаний совпадает с частотой внешней силы) называется *резонансом*.

При $p\neq0$ и $\frac{p^2}{4}<q$ характеристическое уравнение $k^2+pk+q=0$ имеет комплексно-сопряженные корни $k_1=\alpha+i\beta$, $k_2=\alpha-i\beta$. Общее решение однородного уравнения $y''+py'+qy=0$, где $y=y(t)$, определяется формулой (см. § 25.7)

$$y_0=e^{\alpha t}(C_1\cos\beta t+C_2\sin\beta t)$$

или

$$y_0=Ae^{\alpha t}\sin(\beta t+\varphi_0). \tag{25.88}$$

Частное решение неоднородного уравнения (25.79) ищем в виде функции

$$Y=P\cos\omega t+Q\sin\omega t. \tag{25.89}$$

Подставляя функцию (25.89) и ее производные Y', Y'' в уравнение (25.79), находим коэффициенты P и Q:

$$P=\frac{-p\omega a}{(q-\omega^2)^2+p^2\omega^2},\quad Q=\frac{(q-\omega^2)a}{(q-\omega^2)^2+p^2\omega^2}. \tag{25.90}$$

Вводим новые постоянные A_1 и φ_1, полагая

$$P=A_1\sin\varphi_1,\quad Q=A_1\cos\varphi_1, \tag{25.91}$$

т. е.

$$A_1 = \sqrt{P^2 + Q^2} = \frac{a}{\sqrt{(q-\omega^2)^2 + p^2\omega^2}}, \quad \operatorname{tg}\varphi_1 = \frac{P}{Q}.$$

Формула (25.89) с учетом равенств (25.90) и (25.91) принимает вид

$$Y = \frac{a}{\sqrt{(q-\omega^2)^2 + p^2\omega^2}} \sin(\omega t + \varphi_1). \tag{25.92}$$

Общее решение уравнения (25.79) определяется формулой $y = y_0 + Y$ или

$$y = Ae^{\alpha t} \sin(\beta t + \varphi_0) + \frac{a}{\sqrt{(q-\omega^2)^2 + p^2\omega^2}} \sin(\omega t + \varphi_1). \tag{25.93}$$

Таким образом, отклонение y состоит из суммы двух слагаемых. Первое из них (решение соответствующего однородного уравнения) определяет затухающие колебания ($\alpha = -p/2 < 0$, ибо $p > 0$), это слагаемое убывает при возрастании t. Следовательно, через некоторый промежуток времени главное значение y будет определяться вторым слагаемым, представляющим вынужденные колебания. Частота ω этих колебаний равна частоте внешней силы: $f(t) = a \sin \omega t$. Амплитуда вынужденных колебаний тем больше, чем меньше p и разность $q - \omega^2$.

§ 25.9. Системы дифференциальных уравнений

Введем основные понятия, связанные с системами обыкновенных дифференциальных уравнений.

Нормальной системой дифференциальных уравнений называется совокупность уравнений вида

$$\left.\begin{array}{l} \dfrac{dy_1}{dx} = F_1(x, y_1, y_2, \ldots, y_n); \\ \dfrac{dy_2}{dx} = F_2(x, y_1, y_2, \ldots, y_n); \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \dfrac{dy_n}{dx} = F_n(x, y_1, y_2, \ldots, y_n), \end{array}\right\} \tag{25.94}$$

где $F_k(x, y_1, y_2, \ldots, y_n)$ $(k = 1, 2, \ldots, n)$ — заданные функции своих аргументов; $y_k = y_k(x)$ — искомые функции независимой переменной x.

Порядком нормальной системы называется число входящих в нее уравнений. Система (25.94) по определению является системой n-го порядка.

Всякая совокупность n функций

$$y_1=\varphi_1(x),\ y_2=\varphi_2(x),\ \ldots,\ y_n=\varphi_n(x), \tag{25.95}$$

определенных и непрерывно дифференцируемых в интервале $]a, b[$, называется *решением системы* (25.94) в этом интервале, если она обращает в тождество каждое из уравнений данной системы, т. е.

$$\varphi_k'(x) \equiv F_k(x,\ \varphi_1(x),\ \varphi_2(x),\ \ldots,\ \varphi_n(x))\ (k=1,\ 2,\ \ldots,\ n)$$

для всех $x \in]a, b[$.

Задача Коши для системы (25.94) состоит в следующем: среди всех решений системы найти такое решение

$$y_1=y_1(x),\ y_2=y_2(x),\ \ldots,\ y_n=y_n(x), \tag{25.96}$$

которое удовлетворяет условиям

$$y_1=y_1^0,\ y_2=y_2^0,\ \ldots,\ y_n=y_n^0 \text{ при } x=x_0, \tag{25.97}$$

где y_1^0, y_2^0, ..., y_n^0 — заданные числа, называемые *начальными значениями* искомых функций, или начальными значениями решения (25.96). Число x_0 называется *начальным значением* независимой переменной x; числа x_0, y_1^0, y_2^0, ..., y_n^0, вместе взятые, — *начальными данными* решения (25.96), а условия (25.97) — *начальными условиями* этого решения. Начальные условия можно записать так:

$$y_k(x_0)=y_k^0\ (k=1,\ 2,\ \ldots,\ n). \tag{25.98}$$

Совокупность n функций

$$\left.\begin{aligned} y_1&=f_1(x,\ C_1,\ C_2,\ \ldots,\ C_n);\\ y_2&=f_2(x,\ C_1,\ C_2,\ \ldots,\ C_n);\\ &\cdots\cdots\cdots\cdots\\ y_n&=f_n(x,\ C_1,\ C_2,\ \ldots,\ C_n) \end{aligned}\right\} \tag{25.99}$$

называется *общим решением* системы (25.94), если:

1) система (25.99) разрешима относительно произвольных постоянных

$$\left.\begin{aligned} \varphi_1(x,\ y_1,\ y_2,\ \ldots,\ y_n)&=C_1;\\ \varphi_2(x,\ y_1,\ y_2,\ \ldots,\ y_n)&=C_2;\\ &\cdots\cdots\cdots\cdots\\ \varphi_n(x,\ y_1,\ y_2,\ \ldots,\ y_n)&=C_n; \end{aligned}\right\} \tag{25.100}$$

2) совокупность функций (25.99) является решением

системы (25.94) при всех значениях произвольных постоянных, определяемых формулами (25.100).

Решение, получающееся из общего при фиксированных значениях произвольных постоянных, называется *частным*.

Рассмотрим любое из равенств (25.100)

$$\varphi_k(x, y_1, y_2, \ldots, y_n) = C_k. \tag{25.101}$$

Функция $\varphi_k(x, y_1, y_2, \ldots, y_n)$, входящая в равенство (25.101), обладает следующим свойством: она обращается в постоянную при замене $y_1, y_2, \ldots, y_n$ любым частным решением системы (25.94), т. е.

$$\varphi_k(x, f_1(x, C_1, C_2, \ldots, C_n), f_2(x, C_1, C_2, \ldots, C_n), \ldots, f_n(x, C_1, C_2, \ldots, C_n)) \equiv C_k.$$

Любая функция $\psi(x, y_1, y_2, \ldots, y_n)$, обладающая таким свойством, называется *интегралом* системы (25.94).

Равенство

$$\varphi(x, y_1, y_2, \ldots, y_n) = C, \tag{25.102}$$

где $\varphi(x, y_1, y_2, \ldots, y_n)$ — интеграл системы (25.94), а C — произвольная постоянная, называется *первым интегралом* этой системы.

Каждое из равенств (25.100) является первым интегралом системы (25.94).

Совокупность n первых интегралов системы (25.100) обладает тем свойством, что она разрешима относительно искомых функций $y_1, y_2, \ldots, y_n$, причем в результате получается общее решение (25.99).

Совокупность n первых интегралов, обладающую таким свойством, называют *общим интегралом* системы (25.98).

Первые интегралы (25.100), образующие общий интеграл системы (25.94), обладают тем свойством, что интегралы $\varphi_1, \varphi_2, \ldots, \varphi_n$ независимы, т. е. между функциями $\varphi_1, \varphi_2, \ldots, \varphi_n$ не существует соотношения вида $\Phi(\varphi_1, \varphi_2, \ldots, \varphi_n) = 0$ ни при каком выборе функции Φ.

Пример. Найти общий интеграл системы $\frac{dy}{dx} = \frac{y}{x}, \ \frac{dz}{dx} = \frac{z}{x}$.

Это нормальная система дифференциальных уравнений. Интегрируем ее:

$$\frac{dy}{y} = \frac{dx}{x}, \quad \frac{dz}{z} = \frac{dx}{x}, \quad \ln|y| = \ln|x| + \ln|C_1|, \quad \ln|z| = \ln|x| + \ln|C_2|,$$

$$y=C_1x, \quad z=C_2x. \tag{1}$$

Разрешив эти уравнения относительно произвольных постоянных, получим

$$\frac{y}{x}=C_1, \quad \frac{z}{x}=C_2. \tag{2}$$

Формулы (1) определяют общее решение системы, формулы (2) — ее первые интегралы. Левые части уравнений (2) являются интегралами системы.

Поскольку интегралы $\varphi_1=y:x$, $\varphi_2=z:x$ независимы, то общий интеграл системы определяется равенствами (2).

Замечание. Систему можно записать и в симметрической форме:

$$\frac{dx}{1}=\frac{dy}{\frac{y}{x}}=\frac{dz}{\frac{z}{x}}, \quad \frac{dx}{x}=\frac{dy}{y}=\frac{dz}{z}.$$

Из последней системы следует, что имеется еще один первый интеграл, $z:y=C_3$. Соответствующий интеграл $\varphi_3=z:y$ выражается через независимые интегралы $\varphi_1=y:x$, $\varphi_2=z:x$, а именно: $\varphi_3=\varphi_2:\varphi_1$.

В некоторых случаях систему дифференциальных уравнений (25.94) можно привести к одному дифференциальному уравнению n-го порядка. Соответствующие примеры решения систем дифференциальных уравнений приведены, в частности, в [7, ч. 2].

Глава 26

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ ПЕРВОГО И ВТОРОГО ПОРЯДКА

В этой главе сообщаются основные сведения о дифференциальных уравнениях с частными производными, изучаются линейные однородные дифференциальные уравнения с частными производными первого порядка, приводится классификация дифференциальных уравнений второго порядка, линейных относительно старших производных; рассматриваются вопросы приведения таких уравнений к каноническому виду.

§ 26.1. Основные сведения о дифференциальных уравнениях с частными производными

Дифференциальным уравнением с частными производными называется уравнение относительно неизвестной функции нескольких переменных, ее аргументов и ее частных производных различных порядков. Если искомая функция u зависит от n переменных $x_1, x_2, \ldots, x_n$, т. е. $u=u(x_1, x_2, \ldots, x_n)$, то дифференциальное уравнение с частными производными имеет вид

$$F\left(x_1, x_2, \ldots, x_n, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_n}, \ldots \right.$$
$$\left. \ldots, \frac{\partial^k u}{\partial x_1^{k_1}, \ldots, \partial x_n^{k_n}}\right) = 0, \qquad (26.1)$$

где F — заданная функция своих аргументов.

Порядком дифференциального уравнения с частными производными называется порядок старшей производной, входящей в это уравнение.

Решением дифференциального уравнения с частными производными (26.1) называется функция

$$u=u(x_1, x_2, \ldots, x_n),$$

имеющая соответствующие частные производные и обращающая это уравнение в тождество (при подстановке в уравнение вместо неизвестной функции и ее частных производных).

Рассмотрим дифференциальное уравнение второго порядка $\frac{\partial^2 z}{\partial y \partial x} = 0$ относительно неизвестной функции $z=z(x, y)$. Обозначим частную производную функции z по y через v, т. е. $\frac{\partial z}{\partial y} = v$, тогда

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial v}{\partial x}, \quad \frac{\partial v}{\partial x} = 0, \quad v = f(y),$$

где $f(y)$ — произвольная функция переменной y. Поскольку

$$\frac{\partial z}{\partial y} = v = f(y), \quad \frac{\partial z}{\partial y} = f(y),$$

то $z = \int f(y)dy + \psi(x)$, где $\psi(x)$ — произвольная функция аргумента x. Первое слагаемое последней формулы представляет собой произвольную функцию от y. Обозначим ее через $\varphi(y)$, тогда $z=\psi(x)+\varphi(y)$. Получено решение данного уравнения, это решение содержит две произвольные функции.

§ 26.2. Линейные однородные дифференциальные уравнения с частными производными первого порядка

Дифференциальное уравнение с частными производными первого порядка от функции $u=u(x_1, x_2, \ldots, x_n)$ n переменных $x_1, x_2, \ldots, x_n$ в общем виде можно записать так:

$$F\left(x_1, x_2, \ldots, x_n, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_n}\right)=0, \quad (26.2)$$

где F — заданная функция своих аргументов.

Линейным дифференциальным уравнением с частными производными первого порядка называется уравнение вида

$$X_1(x_1, x_2, \ldots, x_n)\frac{\partial u}{\partial x_1}+X_2(x_1, x_2, \ldots, x_n)\frac{\partial u}{\partial x_2}+\ldots$$
$$\ldots+X_n(x_1, x_2, \ldots, x_n)\frac{\partial u}{\partial x_n}=0, \quad (26.3)$$

где $X_1(x_1, x_2, \ldots, x_n)$, $X_2(x_1, x_2, \ldots, x_n)$, $\ldots$, $X_n(x_1, x_2, \ldots$ $\ldots, x_n)$ — заданные функции n аргументов $x_1, x_2, \ldots, x_n$; $u=u(x_1, x_2, \ldots, x_n)$ — неизвестная функция.

Очевидно, одним из решений этого уравнения является постоянная $u=c$, $c=\text{const}$.

Наряду с уравнением (26.3) будем рассматривать соответствующую ему систему дифференциальных уравнений, т. е. систему

$$\frac{dx_1}{X_1}=\frac{dx_2}{X_2}=\ldots=\frac{dx_n}{X_n}, \quad (26.4)$$

в которой функции $X_1=X_1(x_1, x_2, \ldots, x_n)$, $\ldots$, $X_n=X_n(x_1, x_2, \ldots, x_n)$ те же, что и в уравнении (26.3).

Система (26.4) обыкновенных дифференциальных уравнений называется системой в *симметрической форме.* Предположим, что в точке $M_0(x_1^0, x_2^0, \ldots, x_n^0)$ и некоторой ее окрестности функции $X_1, X_2, \ldots, X_n$ одновременно в нуль не обращаются. Пусть для определенности $X_n(x_1^0, x_2^0, \ldots, x_n^0)\neq 0$.

Систему (26.4) запишем так:

$$\frac{dx_1}{dx_n}=\frac{X_1}{X_n}, \quad \frac{dx_2}{dx_n}=\frac{X_2}{X_n}, \quad \ldots, \quad \frac{dx_{n-1}}{dx_n}=\frac{X_{n-1}}{X_n}; \quad (26.5)$$

система (26.5) называется *нормальной*. Рассмотрим общее решение этой системы

$$\left.\begin{aligned} x_1 &= f_1(x_n, C_1, C_2, \ldots, C_{n-1}); \\ x_2 &= f_2(x_n, C_1, C_2, \ldots, C_{n-1}); \\ &\cdots\cdots\cdots\cdots \\ x_{n-1} &= f_{n-1}(x_n, C_1, C_2, \ldots, C_{n-1}). \end{aligned}\right\} \tag{26.6}$$

Разрешим уравнения (26.6) относительно $C_1, C_2, \ldots, C_{n-1}$:

$$\left.\begin{aligned} \varphi_1(x_1, x_2, \ldots, x_n) &= C_1; \\ \varphi_2(x_1, x_2, \ldots, x_n) &= C_2; \\ &\cdots\cdots\cdots \\ \varphi_{n-1}(x_1, x_2, \ldots, x_n) &= C_{n-1}. \end{aligned}\right\} \tag{26.7}$$

Каждое из уравнений (26.7) называется *первым интегралом* системы (26.5), а каждая из функций $\varphi_k(x_1, x_2, \ldots, x_n)$ $(k=1, 2, \ldots, n-1)$ — *интегралом* той же системы.

Связь между решениями дифференциального уравнения с частными производными (26.3) и интегралами соответствующей системы обыкновенных дифференциальных уравнений (26.4) или (26.5) выражается следующими теоремами.

Теорема 26.1. Если $\varphi(x_1, x_2, \ldots, x_n)$ — интеграл системы (26.4) или (26.5), то функция $u=\varphi(x_1, x_2, \ldots, x_n)$ — решение уравнения (26.3).

Доказательство. Пусть $\varphi(x_1, x_2, \ldots, x_n)$ — интеграл системы (26.5), тогда $\varphi(x_1, x_2, \ldots, x_n)\equiv C$, где $x_1, x_2, \ldots, x_{n-1}$ — решения этой системы, определяемые равенствами (26.6), и полный дифференциал данной функции равен нулю, $d\varphi\equiv 0$. Подставляя в формулу (см. (18.45))

$$d\varphi = \frac{\partial\varphi}{\partial x_1}dx_1 + \frac{\partial\varphi}{\partial x_2}dx_2 + \ldots + \frac{\partial\varphi}{\partial x_n}dx_n$$

выражения для dx_k $(k=1, 2, \ldots, n-1)$, полученные из системы (26.5), находим, что

$$d\varphi = \frac{\partial\varphi}{\partial x_1}\frac{X_1}{X_n}dx_n + \frac{\partial\varphi}{\partial x_2}\frac{X_2}{X_n}dx_n + \ldots$$

$$\ldots + \frac{\partial\varphi}{\partial x_{n-1}}\frac{X_{n-1}}{X_n}dx_n + \frac{\partial\varphi}{\partial x_n}dx_n =$$

$$= \left(X_1\frac{\partial\varphi}{\partial x_1} + X_2\frac{\partial\varphi}{\partial x_2} + \ldots + X_{n-1}\frac{\partial\varphi}{\partial x_{n-1}} + \right.$$

$$+X_n\frac{\partial\varphi}{\partial x_n}\Big)\frac{1}{X_n}\,dx_n\equiv 0,$$

откуда

$$X_1\frac{\partial\varphi}{\partial x_1}+X_2\frac{\partial\varphi}{\partial x_2}+\ldots+X_n\frac{\partial\varphi}{\partial x_n}\equiv 0.$$

Последнее означает, что функция $u=\varphi(x_1, \ldots, x_n)$ — решение уравнения (26.3).

Теорема 26.2. Если $u=\varphi(x_1, x_2, \ldots, x_n)$ — решение уравнения (26.3), то $\varphi(x_1, x_2, \ldots, x_n)$ — интеграл системы (26.5).

Доказательство. По условию

$$X_1\frac{\partial\varphi}{\partial x_1}+X_2\frac{\partial\varphi}{\partial x_2}+\ldots+X_n\frac{\partial\varphi}{\partial x_n}\equiv 0.$$

Так как

$$d\varphi=\frac{\partial\varphi}{\partial x_1}dx_1+\frac{\partial\varphi}{\partial x_2}dx_2+\ldots+\frac{\partial\varphi}{\partial x_{n-1}}dx_{n-1}+\frac{\partial\varphi}{\partial x_n}dx_n=$$
$$=\frac{\partial\varphi}{\partial x_1}\,\frac{X_1}{X_n}dx_n+\frac{\partial\varphi}{\partial x_2}\,\frac{X_2}{X_n}dx_n+\ldots+\frac{\partial\varphi}{\partial x_{n-1}}\,\frac{X_{n-1}}{X_n}dx_n+$$
$$+\frac{\partial\varphi}{\partial x_n}\,\frac{X_n}{X_n}dx_n=\Big(X_1\frac{\partial\varphi}{\partial x_1}+X_2\frac{\partial\varphi}{\partial x_2}+\ldots$$
$$\ldots+X_{n-1}\frac{\partial\varphi}{\partial x_{n-1}}+X_n\frac{\partial\varphi}{\partial x_n}\Big)\frac{1}{X_n}dx_n,$$

то $d\varphi\equiv 0$, $\varphi(x_1, x_2, \ldots, x_n)\equiv C$, т. е. $\varphi(x_1, x_2, \ldots, x_n)$ — интеграл системы (26.5).

Пример. Найти решения уравнения

$$x\frac{\partial u}{\partial x}-y\frac{\partial u}{\partial y}-z\frac{\partial u}{\partial z}=0.$$

Система (26.4), соответствующая данному уравнению, имеет вид

$$\frac{dx}{x}=\frac{dy}{-y}=\frac{dz}{-z}.$$

Интегрируя уравнения этой системы

$$\frac{dx}{x}=\frac{dy}{-y},\quad \frac{dx}{x}=\frac{dz}{-z},$$

соответственно получаем

$$\ln|x|=-\ln|y|+\ln|C_1|,\quad \ln|x|=-\ln|z|+\ln|C_2|$$

или $xy=C_1$, $xz=C_2$.

Получены первые интегралы системы; интегралами системы являются функции $\varphi_1(x, y)=xy$, $\varphi_2(x, z)=xz$; $u_1=xy$, $u_2=xz$ — решения данного уравнения с частными производными.

Замечание. Пусть $\varphi_1(x_1, x_2, \ldots, x_n)$, $\varphi_2(x_1, x_2, \ldots, x_n)$, ..., $\varphi_{n-1}(x_1, x_2, \ldots, x_n)$ — независимые интегралы системы (26.4). Можно показать, что произвольная функция

$$u=F(\varphi_1, \varphi_2, \ldots, \varphi_{n-1}), \tag{26.8}$$

имеющая непрерывные производные по $\varphi_1, \varphi_2, \ldots, \varphi_{n-1}$, является решением линейного однородного уравнения с частными производными первого порядка (26.3).

§ 26.3. Классификация линейных дифференциальных уравнений с частными производными второго порядка

Будем рассматривать дифференциальные уравнения с частными производными второго порядка относительно функции $u=u(x, y)$ двух переменных.

Уравнение с частными производными второго порядка называется *линейным относительно старших производных*, если оно содержит эти производные лишь в первой степени, а коэффициенты при указанных производных являются функциями только переменных x и y, т. е. уравнение

$$A(x, y)\frac{\partial^2 u}{\partial x^2}+2B(x, y)\frac{\partial^2 u}{\partial x\partial y}+C(x, y)\frac{\partial^2 u}{\partial y^2}+F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right)=0. \tag{26.9}$$

Уравнение (26.9) называется *уравнением гиперболического типа* в данной области, если в этой области

$$B^2-AC>0; \tag{26.10}$$

уравнением параболического типа, если

$$B^2-AC=0; \tag{26.11}$$

уравнением эллиптического типа, если

$$B^2-AC<0. \tag{26.12}$$

Если выражение B^2-AC в данной области меняет знак, то уравнение (26.9) называется *уравнением смешанного типа.*

Например, уравнение

$$(2+3\sin^2 x)\frac{\partial^2 u}{\partial x^2}-6\cos x\frac{\partial^2 u}{\partial x\partial y}-4\frac{\partial^2 u}{\partial y^2}+5x\frac{\partial u}{\partial x}-7y\frac{\partial u}{\partial y}=0$$

является уравнением гиперболического типа на всей плоскости Oxy, так как для него

$$B^2-AC=(-3\cos x)^2-(2+3\sin^2 x)(-4)= \\ =17+3\sin^2 x>0.$$

Уравнение

$$x^2\frac{\partial^2 u}{\partial x^2}+y^2\frac{\partial^2 u}{\partial y^2}-2y\frac{\partial u}{\partial x}+3x\frac{\partial u}{\partial y}=0$$

является уравнением эллиптического типа в любой области, не содержащей осей координат ($x=0$, $y=0$), поскольку для него $B^2-AC=0-x^2y^2=-x^2y^2<0$, если $x\neq 0$, $y\neq 0$.

Уравнение

$$x^2\frac{\partial^2 u}{\partial x^2}+2xy\frac{\partial^2 u}{\partial x\partial y}+y^2\frac{\partial^2 u}{\partial y^2}+\frac{\partial u}{\partial x}-\frac{\partial u}{\partial y}=0$$

является уравнением параболического типа на всей плоскости Oxy, потому что $B^2-AC=(xy)^2-x^2y^2\equiv 0$.

Уравнение

$$(1-x^2)\frac{\partial^2 u}{\partial x^2}-2xy\frac{\partial^2 u}{\partial x\partial y}-(1+y^2)\frac{\partial^2 u}{\partial y^2}+3\frac{\partial u}{\partial x}-5\frac{\partial u}{\partial y}=0$$

на всей плоскости Oxy является уравнением смешанного типа, поскольку $B^2-AC=(-xy)^2+(1-x^2)(1+y^2)=$ $=1+y^2-x^2$.

Последнее выражение может быть положительным, отрицательным или нулем. В частности, $B^2-AC=0$ для всех точек гиперболы $x^2-y^2=1$, которая называется *линией параболичности* данного уравнения.

§ 26.4. Преобразование линейного уравнения с частными производными при переходе к новым переменным

Рассмотрим уравнение с частными производными второго порядка, линейное относительно старших производных, т. е. уравнение (26.9)

$$A(x,y)\frac{\partial^2 u}{\partial x^2}+2B(x,y)\frac{\partial^2 u}{\partial x\partial y}+C(x,y)\frac{\partial^2 u}{\partial y^2}+ \\ +F\left(x,y,u,\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}\right)=0.$$

Введем новые переменные ξ и η по формулам

$$\xi=\xi(x, y), \quad \eta=\eta(x, y), \tag{26.13}$$

где $\xi(x, y)$, $\eta(x, y)$ — дважды непрерывно дифференцируемые функции своих аргументов. Будем считать, что якобиан этих функций отличен от нуля, т. е.

$$I=\begin{vmatrix} \dfrac{\partial\xi}{\partial x} & \dfrac{\partial\xi}{\partial y} \\ \dfrac{\partial\eta}{\partial x} & \dfrac{\partial\eta}{\partial y} \end{vmatrix} \neq 0, \tag{26.14}$$

и уравнения (26.13) разрешимы относительно x и y:

$$x=x(\xi, \eta), \quad y=y(\xi, \eta).$$

Найдем выражения для частных производных функции u по старым переменным x, y через частные производные по новым переменным ξ и η:

$$\left.\begin{aligned} &u_x=u_\xi\xi_x+u_\eta\eta_x, \quad u_y=u_\xi\xi_y+u_\eta\eta_y; \\ &u_{xy}=u_{\xi\xi}\xi_x\xi_y+u_{\xi\eta}(\xi_x\eta_y+\xi_y\eta_x)+ \\ &\qquad +u_{\eta\eta}\eta_x\eta_y+u_\xi\xi_{xy}+u_\eta\eta_{xy}; \\ &u_{xx}=u_{\xi\xi}\xi_x^2+2u_{\xi\eta}\xi_x\eta_x+u_{\eta\eta}\eta_x^2+u_\xi\xi_{xx}+u_\eta\eta_{xx}; \\ &u_{yy}=u_{\xi\xi}\xi_y^2+2u_{\xi\eta}\xi_y\eta_y+u_{\eta\eta}\eta_y^2+u_\xi\xi_{yy}+u_\eta\eta_{yy}. \end{aligned}\right\} \tag{26.15}$$

Подставляя эти выражения для производных в исходное дифференциальное уравнение с частными производными, получаем

$$A_1(\xi, \eta)\frac{\partial^2 u}{\partial\xi^2}+2B_1(\xi, \eta)\frac{\partial^2 u}{\partial\xi\partial\eta}+C_1(\xi, \eta)\frac{\partial^2 u}{\partial\eta^2}+ \\ +F_1\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right)=0, \tag{26.16}$$

где

$$\left.\begin{aligned} &A_1(\xi, \eta)=A\left(\frac{\partial\xi}{\partial x}\right)^2+2B\frac{\partial\xi}{\partial x}\frac{\partial\xi}{\partial y}+C\left(\frac{\partial\xi}{\partial y}\right)^2; \\ &B_1(\xi, \eta)=A\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial x}+B\left(\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial y}+\frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial x}\right)+ \\ &\qquad +C\frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial y}; \\ &C_1(\xi, \eta)=A\left(\frac{\partial\eta}{\partial x}\right)^2+2B\frac{\partial\eta}{\partial x}\frac{\partial\eta}{\partial y}+C\left(\frac{\partial\eta}{\partial y}\right)^2. \end{aligned}\right\} \tag{26.17}$$

Непосредственной проверкой можно убедиться в том, что

$$B_1^2 - A_1C_1 = (B^2 - AC)\left(\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial y} - \frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial x}\right)^2 \qquad (26.18)$$

или $B_1^2 - A_1C_1 = (B^2 - AC)I^2$, где якобиан I определен формулой (26.14).

Следовательно, уравнения (26.9) и (26.16) являются уравнениями одного и того же типа, т. е. рассматриваемое преобразование независимых переменных не меняет типа уравнения.

В следующих параграфах будет показано, что функции (26.13) можно выбрать так, чтобы выполнялось только одно из условий: 1) $A_1=0$, $C_1=0$; 2) $A_1=0$, $B_1=0$; 3) $A_1=C_1$, $B_1=0$.

Уравнение (26.16) в каждом из этих случаев принимает наиболее простой вид.

§ 26.5. Приведение уравнения гиперболического типа к каноническому виду

Рассмотрим уравнение с частными производными первого порядка:

$$A\left(\frac{\partial\varphi}{\partial x}\right)^2 + 2B\frac{\partial\varphi}{\partial x}\frac{\partial\varphi}{\partial y} + C\left(\frac{\partial\varphi}{\partial y}\right)^2 = 0, \qquad (26.19)$$

в котором $A=A(x, y)$, $B=B(x, y)$, $C=C(x, y)$ — те же, что и в уравнении (26.9), причем $A\neq 0$.

Уравнение (26.19) можно записать так:

$$\left[A\frac{\partial\varphi}{\partial x} + (B+\sqrt{B^2-AC})\frac{\partial\varphi}{\partial y}\right]\times$$
$$\times\left[A\frac{\partial\varphi}{\partial x} + (B-\sqrt{B^2-AC})\frac{\partial\varphi}{\partial y}\right] = 0, \qquad (26.20)$$

откуда

$$A\frac{\partial\varphi}{\partial x} + (B+\sqrt{B^2-AC})\frac{\partial\varphi}{\partial y} = 0, \qquad (26.21)$$

$$A\frac{\partial\varphi}{\partial x} + (B-\sqrt{B^2-AC})\frac{\partial\varphi}{\partial y} = 0. \qquad (26.22)$$

Для каждого из таких уравнений соответствующая система обыкновенных дифференциальных уравнений (см. § 26.2) сводится к одному уравнению. Эти уравнения принимают вид

$$\frac{dx}{A}=\frac{dy}{B+\sqrt{B^2-AC}}; \tag{26.23}$$

$$\frac{dx}{A}=\frac{dy}{B-\sqrt{B^2-AC}} \tag{26.24}$$

или

$$Ady-(B+\sqrt{B^2-AC})\,dx=0; \tag{26.25}$$

$$Ady-(B-\sqrt{B^2-AC})\,dx=0. \tag{26.26}$$

Последние уравнения можно представить одним уравнением

$$Ady^2-2Bdxdy+Cdx^2=0. \tag{26.27}$$

Пусть

$$\varphi_1(x,\ y)=C_1,\quad \varphi_2(x,\ y)=C_2 \tag{26.28}$$

есть первые интегралы, а $\varphi_1(x,\ y)$, $\varphi_2(x,\ y)$ — независимые интегралы системы, состоящей из уравнений (26.25) и (26.26). В этом случае, как известно (см. § 26.2), функции

$$u=\varphi_1(x,\ y),\quad u=\varphi_2(x,\ y) \tag{26.29}$$

являются соответственно решениями уравнений (26.21) и (26.22), следовательно, и решениями уравнения (26.19).

Линии, определяемые уравнениями (26.28), называются *характеристическими линиями* уравнения (26.9), или *характеристиками* этого уравнения. Уравнение (26.27) называют *уравнением характеристик*, оно распадается на уравнения (26.25) и (26.26).

Если в качестве функций $\xi(x,\ y)$, $\eta(x,\ y)$ в формулах (26.13) взять функции (26.29), то уравнение (26.16) примет более простой вид, так как некоторые его коэффициенты обратятся в нуль (см. формулы (26.17) и уравнение (26.19)).

Пусть уравнение (26.9), т. е. уравнение

$$A(x,\ y)\frac{\partial^2 u}{\partial x^2}+2B(x,\ y)\frac{\partial^2 y}{\partial x\partial y}+C(x,\ y)\frac{\partial^2 u}{\partial y^2}+ \\ +F\left(x,\ y,\ u,\ \frac{\partial u}{\partial x},\ \frac{\partial u}{\partial y}\right)=0,$$

в рассматриваемой области является уравнением гиперболического типа: для него в этой области $B^2-AC>0$.

Предположим, что A и C в этом уравнении не равны одновременно нулю (для определенности будем считать $A\neq 0$).

Рассмотрим уравнение (26.19) и соответствующие уравнения (26.21), (26.22), а также уравнения (26.25) и (26.26). Так как в данном случае $B^2-AC>0$, то система двух последних уравнений имеет вещественные первые интегралы $\varphi_1(x, y)=C_1$, $\varphi_2(x, y)=C_2$.

В формулах (26.13) в качестве функций $\xi(x, y)$ и $\eta(x, y)$ возьмем интегралы указанной системы, т. е. перейдем к новым переменным ξ и η по формулам $\xi=\varphi_1(x, y)$, $\eta=\varphi_2(x, y)$.

Поскольку функции $\varphi_1(x, y)$ и $\varphi_2(x, y)$ являются соответственно решениями уравнений (26.21) и (26.22), а следовательно, и решениями уравнения (26.19), то по формулам (26.17) получаем

$$A_1=A\left(\frac{\partial\varphi_1}{\partial x}\right)^2+2B\frac{\partial\varphi_1}{\partial x}\frac{\partial\varphi_1}{\partial y}+C\left(\frac{\partial\varphi_1}{\partial y}\right)^2\equiv 0,$$

$$C_1=A\left(\frac{\partial\varphi_2}{\partial x}\right)^2+2B\frac{\partial\varphi_2}{\partial x}\frac{\partial\varphi_2}{\partial y}+C\left(\frac{\partial\varphi_2}{\partial y}\right)^2\equiv 0.$$

Уравнение (26.16) принимает вид

$$2B_1(\xi, \eta)\frac{\partial^2 u}{\partial\xi\partial\eta}+F_1\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right)=0.$$

Так как $B_1\neq 0$ (это следует из условия (26.14), неравенства (26.10) и формулы (26.18)), то после деления на $2B_1\neq 0$ получаем

$$\frac{\partial^2 u}{\partial\xi\partial\eta}=\Phi\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right). \qquad (26.30)$$

Уравнение (26.30) называется *каноническим уравнением гиперболического типа.*

Замечание 1. Если $A=C=0$, то уравнение (26.9) имеет канонический вид.

Замечание 2. Уравнение гиперболического типа можно привести и к другому каноническому виду:

$$\frac{\partial^2 u}{\partial\mu^2}-\frac{\partial^2 u}{\partial\nu^2}=\Phi_1\left(\mu, \nu, u, \frac{\partial u}{\partial\mu}, \frac{\partial u}{\partial\nu}\right),$$

если перейти к новым переменным μ и ν по формулам $\xi=\mu+\nu$, $\eta=\mu-\nu$, $\mu=\dfrac{\xi+\eta}{2}$, $\nu=\dfrac{\xi-\eta}{2}$.

Пример. Привести к каноническому виду уравнение

$$x^2\frac{\partial^2 u}{\partial x^2}-y^2\frac{\partial^2 u}{\partial y^2}+3x\frac{\partial u}{\partial x}-3y\frac{\partial u}{\partial y}=0\ (x\neq 0,\ y\neq 0)$$

и проинтегрировать его.

Данное уравнение является уравнением гиперболического типа во всех точках плоскости, кроме точек, лежащих на осях координат, поскольку для него $B^2-AC=x^2y^2>0$, если $x\neq 0$, $y\neq 0$.

Уравнение характеристик (26.27) принимает вид

$$x^2dy^2-y^2dx^2=0,\quad y^2dx^2-x^2dy^2=0$$

и распадается на уравнения $ydx+xdy=0$, $ydx-xdy=0$. Интегрируя эти уравнения, получаем $xy=C_1$, $y/x=C_2$.

Вводим новые переменные ξ, η: $\xi=xy$, $\eta=y/x$ и с помощью формул (26.15) находим частные производные

$$\frac{\partial u}{\partial x}=y\frac{\partial u}{\partial \xi}-\frac{y}{x^2}\frac{\partial u}{\partial \eta},\quad \frac{\partial u}{\partial y}=x\frac{\partial u}{\partial \xi}+\frac{1}{x}\frac{\partial u}{\partial \eta},$$

$$\frac{\partial^2 u}{\partial y^2}=x^2\frac{\partial^2 u}{\partial \xi^2}+2\frac{\partial^2 u}{\partial \xi\partial \eta}+\frac{1}{x^2}\frac{\partial^2 u}{\partial \eta^2},$$

$$\frac{\partial^2 u}{\partial x^2}=y^2\frac{\partial^2 u}{\partial \xi^2}-2\frac{y^2}{x^2}\frac{\partial^2 u}{\partial \xi\partial \eta}+\frac{y^2}{x^4}\frac{\partial^2 u}{\partial \eta^2}+2\frac{y}{x^3}\frac{\partial u}{\partial \eta}.$$

Подставляя эти выражения в исходное уравнение, получаем

$$\frac{\partial^2 u}{\partial \xi\partial \eta}+\frac{1}{xy}\frac{\partial u}{\partial \eta}=0,\quad \frac{\partial^2 u}{\partial \xi\partial \eta}+\frac{1}{\xi}\frac{\partial u}{\partial \eta}=0.$$

Проинтегрируем уравнение $u''_{\xi\eta}+\frac{1}{\xi}u'_{\eta}=0$. Обозначим $u'_{\eta}=w$, тогда $u''_{\eta\xi}=w'_{\xi}$ $(u''_{\eta\xi}=u''_{\xi\eta})$, $w'_{\xi}+\frac{1}{\xi}w=0$, $\frac{w'}{w}+\frac{1}{\xi}=0$, $\ln|w|+$ $+\ln|\xi|=\ln|\varphi_1(\eta)|$, $w\xi=\varphi_1(\eta)$, $u'_{\eta}\xi=\varphi_1(\eta)$, $\xi u=\int\varphi_1(\eta)d\eta+\psi_1(\xi)$.

Следовательно, $u=\frac{1}{\xi}\varphi(\eta)+\psi(\xi)$, где φ и ψ — произвольные дважды дифференцируемые функции своих аргументов.

Возвращаясь к переменным x, y, получаем

$$u=\frac{1}{xy}\varphi\left(\frac{y}{x}\right)+\psi(xy).$$

§ 26.6. Приведение уравнения параболического типа к каноническому виду

Рассмотрим уравнение (26.9), для которого

$$B^2-AC=0. \tag{26.31}$$

Предположим, что A и B одновременно не равны нулю. Пусть $A \neq 0$.

В силу условия (26.31) уравнения (26.21) и (26.22) совпадают. Получаем одно уравнение

$$A \frac{\partial \varphi}{\partial x} + B \frac{\partial \varphi}{\partial y} = 0. \tag{26.32}$$

Покажем, что если функция $\varphi(x, y)$ удовлетворяет уравнению (26.32), то она также удовлетворяет уравнению

$$B \frac{\partial \varphi}{\partial x} + C \frac{\partial \varphi}{\partial y} = 0. \tag{26.33}$$

Действительно, умножая на B обе части уравнения (26.32) и принимая во внимание (26.31), получаем

$$0 = AB \frac{\partial \varphi}{\partial x} + B^2 \frac{\partial \varphi}{\partial y} = AB \frac{\partial \varphi}{\partial x} + AC \frac{\partial \varphi}{\partial y} =$$
$$= A\left(B \frac{\partial \varphi}{\partial x} + C \frac{\partial \varphi}{\partial y}\right),$$

откуда и следует (26.33), так как $A \neq 0$ по предположению.

Первые интегралы (26.28) в этом случае также совпадают. Находим один первый интеграл $\varphi(x, y) = C$.

Как известно (см. § 26.2), функция $u = \varphi(x, y)$ является решением уравнения (26.32). В соответствии с доказанным выше она будет также и решением уравнения (26.33).

В качестве первой из функций (26.13) возьмем $\varphi(x, y)$, т. е. положим

$$\xi = \varphi(x, y). \tag{26.34}$$

При таком выборе функции $\xi = \xi(x, y)$ коэффициенты A_1 и B_1 преобразованного уравнения (26.16) обратятся в нуль. Действительно, преобразовав вторую из формул (26.17), получим

$$B_1 = \left(A \frac{\partial \xi}{\partial x} + B \frac{\partial \xi}{\partial y}\right) \frac{\partial \eta}{\partial x} + \left(B \frac{\partial \xi}{\partial x} + C \frac{\partial \xi}{\partial y}\right) \frac{\partial \eta}{\partial y},$$

откуда $B_1 \equiv 0$, поскольку $\xi = \varphi(x, y)$ — решение уравнений (26.32) и (26.33).

Так как функция $\xi = \varphi(x, y)$ — решение уравнения (26.19), которое свелось в данном случае к одному уравнению (26.32), то $A_1 \equiv 0$ (см. первую из формул (26.17)).

В качестве $\eta=\eta(x, y)$ возьмем любую функцию, которая вместе с функцией (26.34) удовлетворяет условию (26.14). Можно утверждать, что $C_1 \neq 0$ в преобразованном уравнении (26.16).

В самом деле, преобразуем третью из формул (26.17), использовав условие (26.31):

$$C_1 = A\left(\frac{\partial\eta}{\partial x}\right)^2 + 2B\frac{\partial\eta}{\partial x}\frac{\partial\eta}{\partial y} + C\left(\frac{\partial\eta}{\partial y}\right)^2 = \\ = \frac{1}{A}\left[A\frac{\partial\eta}{\partial x} + B\frac{\partial\eta}{\partial y}\right]^2.$$

Отсюда следует, что $C_1 \neq 0$, ибо сумма в квадратной скобке отлична от нуля. Предположив противное, получим $A\frac{\partial\eta}{\partial x} + B\frac{\partial\eta}{\partial y} = 0$. Это уравнение и уравнение (26.32) образуют однородную систему, имеющую ненулевое решение ($A^2 + B^2 \neq 0$ по предположению). Следовательно, определитель этой системы равен нулю (см. § 4.2), т. е.

$$\begin{vmatrix} \frac{\partial\xi}{\partial x} & \frac{\partial\xi}{\partial y} \\ \frac{\partial\eta}{\partial x} & \frac{\partial\eta}{\partial y} \end{vmatrix} = 0,$$

что противоречит условию (26.14).

Итак, уравнение (26.16) принимает вид

$$C_1(\xi, \eta)\frac{\partial^2 u}{\partial\eta^2} + F_1\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right) = 0.$$

Поскольку $C_1 \neq 0$, то уравнение можно записать так:

$$\frac{\partial^2 u}{\partial\eta^2} = \Phi\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right). \tag{26.35}$$

Уравнение (26.35) называется *каноническим уравнением параболического типа.*

Замечание. Если $A=0$, $B=0$, то уравнение (26.9) имеет уже канонический вид.

Пример. Проинтегрировать уравнение

$$x^2\frac{\partial^2 u}{\partial x^2} - 2xy\frac{\partial^2 u}{\partial x\partial y} + y^2\frac{\partial^2 u}{\partial y^2} + x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} = 0.$$

Данное уравнение является уравнением параболического типа на всей плоскости Oxy, так как для него $B^2-AC=(-xy)^2-x^2y^2=0$.

Уравнение характеристик (26.27) принимает вид

$$x^2dy^2+2xydxdy+y^2dx^2=0, \quad (xdy+ydx)^2=0, \quad xdy+ydx=0.$$

Так как $xdy+ydx=d(xy)$, то $d(xy)=0$, откуда $xy=C$, $\varphi(x, y)=xy$.

В формулах (26.13) положим $\xi=xy$. В качестве функции $\eta=\eta(x, y)$ можно взять любую функцию так, чтобы выполнялось условие (26.14), в частности, $\eta=y$.

Преобразуем это уравнение, введя новые переменные ξ и η по формулам $\xi=xy$, $\eta=y$.

Находим выражения для частных производных по x, y через частные производные по ξ, η с помощью формул (26.15):

$$\frac{\partial u}{\partial x}=y\frac{\partial y}{\partial \xi},\quad \frac{\partial u}{\partial y}=x\frac{\partial u}{\partial \xi}+\frac{\partial u}{\partial \eta},\quad \frac{\partial^2 u}{\partial x^2}=y^2\frac{\partial^2 u}{\partial \xi^2},$$

$$\frac{\partial^2 u}{\partial x\partial y}=\frac{\partial u}{\partial \xi}+xy\frac{\partial^2 u}{\partial \xi^2}+y\frac{\partial^2 u}{\partial \xi\partial \eta},$$

$$\frac{\partial^2 u}{\partial y^2}=x^2\frac{\partial^2 u}{\partial \xi^2}+\frac{\partial^2 u}{\partial \eta^2}+2x\frac{\partial^2 u}{\partial \xi\partial \eta}.$$

Подставляя эти выражения в данное уравнение, получаем

$$y\frac{\partial^2 u}{\partial \eta^2}+\frac{\partial u}{\partial \eta}=0,\quad \eta\frac{\partial^2 u}{\partial \eta^2}+\frac{\partial u}{\partial \eta}=0.$$

Введем иовую функцию w по формуле $w=\dfrac{\partial u}{\partial \eta}=u'_\eta$, тогда $\dfrac{\partial w}{\partial \eta}=\dfrac{\partial^2 u}{\partial \eta^2}$, $\eta\dfrac{\partial w}{\partial \eta}+w=0$, $\eta w'_\eta+w=0$, $\dfrac{w'_\eta}{w}+\dfrac{1}{\eta}=0$, $\ln|w|+$ $+\ln|\eta|=\ln|\varphi(\xi)|$, $w\eta=\varphi(\xi)$, $u'_\eta\eta=\varphi(\xi)$, $u'_\eta=\dfrac{1}{\eta}\varphi(\xi)$, $u=\varphi(\xi)\ln|\eta|+\psi(\xi)$.

Следовательно,

$$u=\varphi(xy)\ln|y|+\psi(xy),$$

где $\varphi(xy)$, $\psi(xy)$ — произвольные дважды дифференцируемые функции от произведения xy аргументов x и y.

§ 26.7. Приведение уравнения эллиптического типа к каноническому виду

Пусть уравнение (26.9) является уравнением эллиптического типа, т. е. для него в рассматриваемой области выполняется условие

$$B^2-AC<0. \tag{26.36}$$

В силу этого условия уравнения (26.25) и (26.26) имеют комплексно-сопряженные первые интегралы $\varphi_1(x, y)=C_1$, $\varphi_2(x, y)=C_2$, причем

$$\begin{aligned}\varphi_1(x, y)&=\xi(x, y)+i\eta(x, y),\\ \varphi_2(x, y)&=\xi(x, y)-i\eta(x, y),\end{aligned} \tag{26.37}$$

где $\xi(x, y)$, $\eta(x, y)$ — вещественные функции переменных x и y.

Поскольку функция $\varphi_1(x, y)$ — решение уравнения (26.19), то

$$A\left(\frac{\partial\varphi_1}{\partial x}\right)^2 + 2B\left(\frac{\partial\varphi_1}{\partial x}\right)\left(\frac{\partial\varphi_1}{\partial y}\right) + C\left(\frac{\partial\varphi_1}{\partial y}\right)^2 \equiv 0$$

или

$$A\left(\frac{\partial\xi}{\partial x} + i\frac{\partial\eta}{\partial x}\right)^2 + 2B\left(\frac{\partial\xi}{\partial x} + i\frac{\partial\eta}{\partial x}\right)\left(\frac{\partial\xi}{\partial y} + i\frac{\partial\eta}{\partial y}\right) +$$
$$+ C\left(\frac{\partial\xi}{\partial y} + i\frac{\partial\eta}{\partial y}\right)^2 \equiv 0.$$

Преобразуя левую часть последнего тождества, получаем

$$A\left(\frac{\partial\xi}{\partial x}\right)^2 + 2B\frac{\partial\xi}{\partial x}\frac{\partial\xi}{\partial y} + C\left(\frac{\partial\xi}{\partial y}\right)^2 -$$
$$-\left[A\left(\frac{\partial\eta}{\partial x}\right)^2 + 2B\frac{\partial\eta}{\partial x}\frac{\partial\eta}{\partial y} + C\left(\frac{\partial\eta}{\partial y}\right)^2\right] +$$
$$+ 2i\left[A\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial x} + B\left(\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial y} + \frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial x}\right) + C\frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial y}\right] \equiv 0,$$

откуда

$$A\left(\frac{\partial\xi}{\partial x}\right)^2 + 2B\frac{\partial\xi}{\partial x}\frac{\partial\xi}{\partial y} + C\left(\frac{\partial\xi}{\partial y}\right)^2 -$$
$$-\left[A\left(\frac{\partial\eta}{\partial x}\right)^2 + 2B\frac{\partial\eta}{\partial x}\frac{\partial\eta}{\partial y} + C\left(\frac{\partial\eta}{\partial y}\right)^2\right] \equiv 0,$$

$$A\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial x} + B\left(\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial y} + \frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial x}\right) + C\frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial y} \equiv 0.$$

Принимая во внимание формулы (26.17), заключаем, что

$$A_1 \equiv C_1, \quad B_1 \equiv 0. \tag{26.38}$$

Уравнение (26.16) принимает вид

$$A_1\frac{\partial^2 u}{\partial\xi^2} + A_1\frac{\partial^2 u}{\partial\eta^2} + F_1\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right) = 0.$$

Так как $A_1 \neq 0$ (это следует из (26.36), (26.14) и (26.18)), то последнее уравнение можно переписать так:

$$\frac{\partial^2 u}{\partial\xi^2} + \frac{\partial^2 u}{\partial\eta^2} = \Phi\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right). \tag{26.39}$$

Уравнение (26.39) называется *каноническим уравнением эллиптического типа.*

§ 26.8. Некоторые свойства решений линейных однородных уравнений с частными производными второго порядка

Линейным уравнением с частными производными второго порядка называется уравнение вида

$$A\frac{\partial^2 u}{\partial x^2}+2B\frac{\partial^2 u}{\partial x\partial y}+C\frac{\partial^2 u}{\partial y^2}+D\frac{\partial u}{\partial x}+$$
$$+E\frac{\partial u}{\partial y}+Fu=f(x, y). \qquad (26.40)$$

В это уравнение неизвестная функция $u=u(x, y)$ и ее частные производные могут входить только в первой степени. Уравнение не содержит членов с произведением различных частных производных, с произведением частной производной и функции. Коэффициенты уравнения могут зависеть только от x, y; в частности, они могут быть постоянными, причем не равны нулю одновременно коэффициенты при частных производных второго порядка, т. е. $A^2+B^2+C^2\neq 0$, $f(x, y)$ — заданная функция переменных x и y.

Если $f(x, y)\not\equiv 0$, то уравнение (26.40) называется *линейным неоднородным.* В случае, когда $f(x, y)\equiv 0$, уравнение называется *линейным однородным* и имеет вид

$$A\frac{\partial^2 u}{\partial x^2}+2B\frac{\partial^2 u}{\partial x\partial y}+C\frac{\partial^2 u}{\partial y^2}+D\frac{\partial u}{\partial x}+$$
$$+E\frac{\partial u}{\partial y}+Fu=0. \qquad (26.41)$$

Пусть функции

$$u_1=u_1(x, y),\ u_2=u_2(x, y),\ \ldots,\ u_n=u_n(x, y) \qquad (26.42)$$

являются решениями уравнения (26.41). Нетрудно видеть, что линейная комбинация функций (26.42), т. е. функция

$$u(x, y)=C_1u_1(x, y)+C_2u_2(x, y)+\ldots+C_nu_n(x, y)=$$
$$=\sum_{k=1}^{n}C_ku_k(x, y), \qquad (26.43)$$

где $C_1, C_2, \ldots, C_n$ — произвольные постоянные, также удовлетворяет уравнению (26.41).

Для обыкновенного линейного однородного дифференциального уравнения n-го порядка формулой вида (26.43), где $u_1=u_1(x), u_2=u_2(x), \ldots, u_n=u_n(x)$ — линейно-независимые решения, определяется общее решение (см. формулу (25.31) § 25.2), причем уравнение n-го порядка не может иметь больше n линейно-независимых решений. Как мы увидим ниже, линейное однородное уравнение с частными производными второго порядка может иметь бесконечное множество линейно-независимых решений

$$u_1=u_1(x, y),\ u_2=u_2(x, y),\ \ldots,\ u_n=u_n(x, y),\ \ldots. \quad (26.44)$$

В этом случае рассматриваются ряды

$$\sum_{k=1}^{\infty} C_k u_k(x, y) = C_1 u_1(x, y) + C_2 u_2(x, y) + \ldots$$
$$\ldots + C_n u_n(x, y) + \ldots, \quad (26.45)$$

где $C_1, C_2, \ldots, C_n, \ldots$ — произвольные постоянные; $u_k = u_k(x, y)$ — функции (26.44).

Если ряд (26.45) сходится и его можно дважды дифференцировать почленно, то его сумма, как легко показать, также будет решением уравнения (26.41). Обозначим эту сумму через $u(x, y)$, т. е.

$$u(x, y) = \sum_{k=1}^{\infty} C_k u_k(x, y). \quad (26.46)$$

Итак, функция $u(x, y)$ — решение уравнения (26.41).

Глава 27

ПРОСТЕЙШИЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

В этой главе рассматриваются некоторые уравнения математической физики, т. е. уравнения с частными производными второго порядка, к которым приводят следующие задачи: задача о колебаниях струны, задача о распространении тепла и др.

§ 27.1. Вывод уравнения колебаний струны

Рассмотрим туго натянутую струну, закрепленную на концах. Выведем струну из положения равновесия (оттянув ее или ударив по ней), струна начнет колебаться.

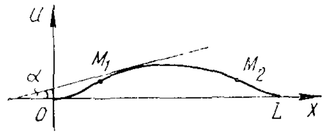

Рис. 27.1

Предположим, что любая точка струны колеблется по прямой, перпендикулярной к исходному положению струны, и струна все время находится в одной и той же плоскости. Выберем в этой плоскости декартову прямоугольную систему координат Oxu. В качестве оси Ox возьмем прямую, на которой находилась струна в положении равновесия, за ось Ou примем прямую, проходящую через левый конец струны перпендикулярно к оси Ox (рис. 27.1).

Отклонение струны от положения равновесия обозначим через u; очевидно, u зависит от абсциссы x точки струны и времени t, т. е. $u=u(x, t)$.

При фиксированном t графиком функции $u=u(x, t)$ в плоскости Oxu является форма струны в данный момент времени t. Угловой коэффициент касательной к графику в точке с абсциссой x равен частной производной по x от функции $u(x, t)$, т. е.

$$\operatorname{tg}\alpha = u'_x(x, t), \quad \operatorname{tg}\alpha = \frac{\partial u}{\partial x}, \tag{27.1}$$

где $\alpha=\alpha(x, t)$ — угол наклона касательной.

Чтобы составить представление о колебаниях струны, необходимо начертить ряд графиков функции $u=u(x, t)$ при различных значениях t.

При фиксированном значении x функция $u=u(x, t)$ определяет закон движения точки с абсциссой x. Эта точка движется по прямой, параллельной оси Ou. Скорость и ускорение указанного движения выражаются соответственно формулами

$$v = \frac{\partial u}{\partial t} = u'_t(x, t), \quad w = \frac{\partial^2 u}{\partial t^2} = u''_{tt}(x, t). \tag{27.2}$$

Будем изучать малые колебания струны, т. е. такие, при которых угол $\alpha=\alpha(x, t)$ (угол наклона касательной к графику функции $u=u(x, t)$ при каждом фиксированном значении t) настолько мал, что его квадратом можно пренебречь, т. е. приближенно считать

$$\alpha^2=0. \tag{27.3}$$

Поскольку

$$\sin\alpha = \alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \ldots, \quad \cos\alpha = 1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} - \ldots,$$

то отсюда следует, что

$$\sin\alpha=\alpha, \quad \cos\alpha=1. \tag{27.4}$$

Далее, так как

$$\operatorname{tg}\alpha-\sin\alpha=\operatorname{tg}\alpha(1-\cos\alpha)=\operatorname{tg}\alpha\cdot 0=0,$$

то

$$\operatorname{tg}\alpha=\sin\alpha. \tag{27.5}$$

Принимая во внимание (27.3)—(27.5), заключаем, что

$$\operatorname{tg}^2\alpha = 0, \quad \text{или} \quad \left(\frac{\partial u}{\partial x}\right)^2 = 0. \tag{27.6}$$

Следовательно, длина дуги струны, ограниченной точками $M_1(x_1, u_1)$, $M_2(x_2, u_2)$, выразится формулой

$$\widehat{M_1M_2} = \int_{x_1}^{x_2} \sqrt{1+\left(\frac{\partial u}{\partial x}\right)^2}\,dx = \int_{x_1}^{x_2} dx = x_2 - x_1,$$

$$\widehat{M_1M_2}=x_2-x_1. \tag{27.7}$$

Соотношение (27.7) означает, что длина любого участка струны остается постоянной.

Будем предполагать струну абсолютно гибкой, что означает следующее: если удалить участки $\widehat{OM_1}$, $\widehat{M_2L}$ (см. рис. 27.1), то их действия на участок $\widehat{M_1M_2}$ заменяются соответственно действием сил натяжения $\mathbf{T}_1$ и $\mathbf{T}_2$, направленных по касательным к графику функции $u=u(x, t)$ в точках M_1 и M_2 (рис. 27.2). Поскольку по предположению точки струны движутся по прямым, параллельным оси Ou, то сумма проекций сил $\mathbf{T}_1$, $\mathbf{T}_2$ на ось Ox равна нулю. Проектируя эти силы на ось Ox, получаем $T_2\cos\alpha_2-T_1\cos\alpha_1=0$, где T_1, T_2 — величины сил $\mathbf{T}_1$, $\mathbf{T}_2$.

На основании второго из равенств (27.4) заключаем, что $T_1=T_2$, т. е. величина силы натяжения остается постоянной. Обозначая ее через T, получаем

$$T=T_1=T_2. \tag{27.8}$$

Проектируя силы $\mathbf{T}_1$, $\mathbf{T}_2$ на ось Ou, находим

$$T_2\sin\alpha_2-T_1\sin\alpha_1=T(\sin\alpha_2-\sin\alpha_1)=T(\operatorname{tg}\alpha_2-\operatorname{tg}\alpha_1).$$

С учетом равенства (27.1) получаем

$$T(\operatorname{tg}\alpha_2-\operatorname{tg}\alpha_1)=$$
$$=T[u'_x(x+\Delta x,\ t)-u'_x(x,\ t)],$$

где x — абсцисса точки M_1; $x+\Delta x$ — абсцисса точки M_2.

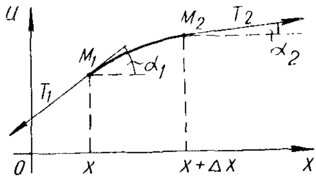

Рис. 27.2

Применяя теорему Лагранжа о конечном приращении дифференцируемой функции, находим, что

$$u'_x(x+\Delta x,\ t)-u'_x(x,\ t)=\frac{\partial^2 u}{\partial x^2}\Delta x,$$

поэтому проекция сил натяжения $\mathbf{T}_1$ и $\mathbf{T}_2$ на ось Ou выразится формулой

$$T(\operatorname{tg}\alpha_2-\operatorname{tg}\alpha_1)=T\frac{\partial^2 u}{\partial x^2}\Delta x. \tag{27.9}$$

Предположим, что на струну действуют также внешние силы, параллельные оси Ou, плотность распределения [1] которых равна $g(x,\ t)$, тогда величина равнодействующей этих сил, приложенных к участку $\overset{\frown}{M_1M_2}$, приближенно равна $g(x,\ t)\Delta x$. Силами сопротивления внешней среды пренебрегаем.

Будем считать струну однородной, обозначим через ρ ее линейную плотность, тогда масса участка $\overset{\frown}{M_1M_2}$ выразится так: $\rho\overset{\frown}{M_1M_2}=\rho\Delta x$, $m=\rho\Delta x$.

В соответствии со вторым законом Ньютона $mw=F$ (произведение массы на ускорение равно действующей силе) получаем

[1] Под плотностью понимают предел средней плотности распределения сил на данном отрезке, когда длина отрезка стремится к нулю; средняя плотность — отношение величины равнодействующей сил к длине отрезка, на котором они приложены.

$$\rho\Delta x \frac{\partial^2 u}{\partial t^2} = T \frac{\partial^2 u}{\partial x^2} \Delta x + g(x, t)\Delta x, \ \rho \frac{\partial^2 u}{\partial t^2} = T \frac{\partial^2 u}{\partial x^2} + g(x, t)$$

или

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + \frac{1}{\rho} g(x, t), \tag{27.10}$$

где

$$a^2 = \frac{T}{\rho}, \ a = \sqrt{\frac{T}{\rho}}. \tag{27.11}$$

Уравнение (27.10) называется *уравнением колебаний струны*, или *одномерным волновым уравнением.*

Если $g(x, t) \equiv 0$ (внешние силы отсутствуют), то уравнение (27.10) принимает вид

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}. \tag{27.12}$$

Уравнение (27.12) называется *уравнением свободных колебаний*; уравнение (27.10) — *уравнением вынужденных колебаний* струны.

§ 27.2. Начальные и краевые условия. Задача Коши

Чтобы из множества решений уравнения с частными производными второго порядка выбрать определенное решение, необходимо задать дополнительные условия.

Так, в случае уравнения (27.10) или (27.12) нужно указать отклонение и скорость движения в начальный момент времени t_0 (будем полагать $t_0 = 0$), т. е.

$$u(x, t)\big|_{t=0} = f(x), \ \left.\frac{\partial u}{\partial t}\right|_{t=0} = F(x), \tag{27.13}$$

где $f(x)$, $F(x)$ — заданные функции, а также зафиксировать отклонения концов струны. Поскольку концы закреплены, то

$$u(x, t)\big|_{x=0} = 0, \quad u(x, t)\big|_{x=l} = 0, \tag{27.14}$$

где l — длина струны.

Условия (27.13) называются *начальными условиями*, а условия (27.14) — *краевыми* (или *граничными*) *условиями.*

Итак, задача о свободных колебаниях струны ставится следующим образом. Найти решение $u = u(x, t)$ линейного однородного уравнения с частными производными второго порядка $\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$, удовлетворяющее начальным условиям $u(x, 0) = f(x)$, $u'_t(x, 0) = F(x)$ и краевым условиям $u(0, t) = 0$, $u(l, t) = 0$.

Функции $f(x)$ и $F(x)$ определены на отрезке $[0, l]$, из краевых условий следует, что $f(0) = 0$, $f(l) = 0$. Можно доказать, что при некоторых предположениях относительно функций $f(x)$ и $F(x)$ поставленная задача имеет единственное решение.

В случае, когда предполагается, что струна является неограниченной, граничные условия не налагаются.

Задача о свободных колебаниях неограниченной струны ставится так. Найти решение $u = u(x, t)$ уравнения с частными производными второго порядка $\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$, удовлетворяющее начальным условиям

$$u(x, t)\big|_{t=0} = f(x), \quad \frac{\partial u}{\partial t}\bigg|_{t=0} = F(x),$$

где $f(x)$ и $F(x)$ — заданные функции, определенные на всей действительной оси. Эта задача называется *задачей Коши*.

§ 27.3. Задача о свободных колебаниях бесконечной струны. Метод Д'Аламбера

Как уже отмечалось, задача о свободных колебаниях бесконечной струны, или задача Коши, состоит в следующем.

Найти решение $u = u(x, t)$ линейного однородного уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} \quad (a \neq 0), \tag{27.15}$$

удовлетворяющее начальным условиям

$$u(x, t)\big|_{t=0} = f(x), \quad \frac{\partial u}{\partial t}\bigg|_{t=0} = F(x), \tag{27.16}$$

где $f(x)$, $F(x)$ — заданные функции, определенные в бесконечном промежутке $]-\infty, +\infty[$.

Уравнение (27.15) перепишем так: $a^2 \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial t^2} = 0$ и (положив $t = y$) сравним его с уравнением (26.9).

Поскольку $B^2-AC=a^2>0$, то уравнение является уравнением гиперболического типа.

Уравнение характеристик $Ady^2-2Bdxdy+Cdx^2=0$ принимает вид $a^2dt^2-dx^2=0$ или $dx^2-a^2dt^2=0$. Оно распадается на два уравнения: $dx-adt=0$, $dx+adt=0$, откуда получаем $x-at=C_1$, $x+at=C_2$.

Введя новые переменные ξ и η по формулам

$$\xi=x-at, \quad \eta=x+at, \tag{27.17}$$

преобразуем уравнение (27.15) к каноническому виду.

Выражаем частные производные по переменным x, t через частные производные по ξ, η:

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial \xi}+\frac{\partial u}{\partial \eta}, \quad \frac{\partial^2 u}{\partial x^2}=\frac{\partial^2 u}{\partial \xi^2}+2\frac{\partial^2 u}{\partial \xi \partial \eta}+\frac{\partial^2 u}{\partial \eta^2},$$

$$\frac{\partial u}{\partial t}=a\left(\frac{\partial u}{\partial \eta}-\frac{\partial u}{\partial \xi}\right), \quad \frac{\partial^2 u}{\partial t^2}=a^2\left(\frac{\partial^2 u}{\partial \xi^2}-2\frac{\partial^2 u}{\partial \xi \partial \eta}+\frac{\partial^2 u}{\partial \eta^2}\right).$$

Подставляя в уравнение (27.15) выражения для частных производных второго порядка, получаем $4a^2\dfrac{\partial^2 u}{\partial \xi \partial \eta}=0$, $\dfrac{\partial^2 u}{\partial \xi \partial \eta}=0$.

Проинтегрируем последнее уравнение. Положим $\dfrac{\partial u}{\partial \xi}=w$, тогда $\dfrac{\partial^2 u}{\partial \xi \partial \eta}=\dfrac{\partial w}{\partial \eta}$, $\dfrac{\partial w}{\partial \eta}=0$.

Следовательно, $w=f(\xi)$, $\dfrac{\partial u}{\partial \xi}=f(\xi)$, $u=\int f(\xi)d\xi+\psi(\eta)$ или $u=\varphi(\xi)+\psi(\eta)$, где $\varphi(\xi)$, $\psi(\eta)$ — произвольные дважды дифференцируемые функции своих аргументов. Принимая во внимание (27.17), последнюю формулу можно записать так:

$$u=\varphi(x-at)+\psi(x+at). \tag{27.18}$$

Формула (27.18) определяет общее решение уравнения (27.15).

Среди всех этих решений найдем то, которое удовлетворяет условиям (27.16). Для функции (27.18) и ее частной производной по t

$$\frac{\partial u}{\partial t}=a[-\varphi'(x-at)+\psi'(x+at)]$$

условия (27.16) принимают вид

$$\varphi(x)+\psi(x)=f(x), \quad a[\psi'(x)-\varphi'(x)]=F(x). \tag{27.19}$$

Второе равенство проинтегрируем по отрезку $[0, x]$. Обозначив переменную интегрирования через z, получим

$$\int_0^x F(z)\,dz = \int_0^x a[\psi'(z) - \varphi'(z)]\,dz = a\psi(z)\Big|_0^x - a\varphi(z)\Big|_0^x =$$

$$= a[\psi(x) - \varphi(x)] + a[\varphi(0) - \psi(0)] = a[\psi(x) - \varphi(x)] + aC,$$

где $C = \varphi(0) - \psi(0)$.

Итак,

$$\int_0^x F(z)\,dz = a[\psi(x) - \varphi(x)] + aC,$$

откуда

$$\psi(x) - \varphi(x) = \frac{1}{a}\int_0^x F(z)\,dz + C.$$

Это уравнение и первое из уравнений (27.19) позволяют определить функции $\varphi(x)$ и $\psi(x)$:

$$\varphi(x) = \frac{1}{2} f(x) - \frac{1}{2a}\int_0^x F(z)\,dz - \frac{C}{2},$$

$$\psi(x) = \frac{1}{2} f(x) + \frac{1}{2a}\int_0^x F(z)\,dz + \frac{C}{2}.$$

Подставляя в эти формулы вместо x соответственно $x - at$ и $x + at$, получаем

$$\varphi(x - at) = \frac{1}{2} f(x - at) - \frac{1}{2a}\int_0^{x-at} F(z)\,dz - \frac{C}{2} =$$

$$= \frac{1}{2} f(x - at) + \frac{1}{2a}\int_{x-at}^0 F(z)\,dz - \frac{C}{2},$$

$$\psi(x + at) = \frac{1}{2} f(x + at) + \frac{1}{2a}\int_0^{x+at} F(z) + \frac{C}{2}.$$

В соответствии с формулой (27.18) находим

$$u = \varphi(x - at) + \psi(x + at) = \frac{f(x - at) + f(x + at)}{2} +$$

$$+ \frac{1}{2a}\int_{x-at}^0 F(z)\,dz + \frac{1}{2a}\int_0^{x+at} F(z)\,dz$$

или

$$u(x, t) = \frac{f(x-at)+f(x+at)}{2} + \frac{1}{2a}\int\limits_{x-at}^{x+at} F(z)\,dz. \quad (27.20)$$

Формула (27.20) представляет *решение Д'Аламбера* рассматриваемой задачи Коши для уравнения колебаний неограниченной струны.

Читателю предлагается непосредственной проверкой убедиться в том, что функция (27.20) удовлетворяет уравнению (27.15) и условиям (27.16).

§ 27.4. Физическая интерпретация решений волнового уравнения

Выясним сначала физический смысл решения (27.18), которое представим так:

$$u=u_1+u_2,\quad u_1=\varphi(x-at),\quad u_2=\psi(x+at). \quad (27.21)$$

Если $\psi(x+at)\equiv 0$, то $u=u_1=\varphi(x-at)$. При фиксированном значении t график функции $u=\varphi(x-at)$ является формой колеблющейся струны в момент времени t. Для точки x_0 при $t=0$ отклонение выразится формулой $u(x_0, 0)=\varphi(x_0)$.

Предположим, что по оси Ox из положения x_0 движется точка в положительном направлении этой оси со скоростью a (a — параметр, входящий в уравнение (27.15) и функцию (27.18)). Закон этого движения выражается формулой $x=x_0+at$. Так как в этом случае $x-at=x_0$, то через момент времени t для точки x получаем отклонение $u=\varphi(x-at)=\varphi(x_0)=u(x_0, 0)$. Это значит, что отклонение для точки x через момент времени t будет тем же, что и для точки x_0 в момент $t=0$.

Следовательно, если мысленно перемещаться вдоль оси Ox в положительном направлении этой оси с постоянной скоростью a, то отклонение струны все время будет казаться постоянным.

Построим графики функции $u_1=\varphi(x-at)$ при различных значениях t: $t_1<t_2<t_3$ (рис. 27.3). Каждый последующий из них получается сдвигом предыдущего вдоль оси Ox на определенную величину. Если эти рисунки по очереди проектировать на неподвижный экран, то первый график «побежит» вправо. Процесс передвижения откло-

нения вдоль прямой, на которой находилась струна в положении равновесия, называется *волной*. Скорость распространения волны равна a, где a определяется формулой (27.11) и входит в уравнение (27.15). Волна распространяется в положительном направлении оси Ox. Явление, описываемое функцией $u_1=\varphi(x-at)$, называется *распространением прямой волны*.

Второе слагаемое формулы (27.18), т. е. функция $u_2=\psi(x+at)$, представляет аналогичный процесс, но только волна будет распространяться влево (в отрицательном направлении оси Ox) с той же скоростью a (рис. 27.3 будет отображать моментальные снимки этой волны, если на него смотреть снизу вверх, т. е. считать $t_3<t_2<t_1$). Явление, описываемое функцией $u_2=\psi(x+at)$, называется *распространением обратной волны*.

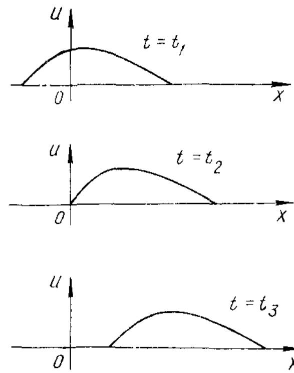

Рис. 27.3

Следовательно, решение (27.18) представляет сумму прямой и обратной волн. Отсюда вытекает следующий графический способ построения формы струны в любой момент времени t. Строим графики функции $u_1=\varphi(x)$, $u_2=\psi(x)$, изображающие прямую и обратную волны в начальный момент времени $t=0$. Не изменяя формы построенных графиков, передвигаем их со скоростью a вдоль оси Ox: первый — вправо, второй — влево. Чтобы получить график струны, достаточно построить алгебраические суммы ординат точек раздвинутых графиков.

§ 27.5. Задача о колебаниях ограниченной струны. Метод Фурье

Задача о колебаниях ограниченной струны, закрепленной на концах, ставится так: найти решение $u=u(x, t)$ уравнения

$$\frac{\partial^2 u}{\partial t^2}=a^2\frac{\partial^2 u}{\partial x^2}, \qquad (27.22)$$

которое удовлетворяет начальным условиям

$$u(x,\ t)\,|_{t=0} = f(x),\quad \left.\frac{\partial u}{\partial t}\right|_{t=0} = F(x) \qquad (27.23)$$

и краевым условиям

$$u(x,\ t)\,|_{x=0} = 0,\quad u(x,\ t)\,|_{x=l} = 0 \qquad (27.24)$$

($f(x)$ и $F(x)$ — заданные функции, определенные на отрезке $[0,\ l]$).

Для решения поставленной задачи применяем *метод Фурье*. Решение $u(x,\ t)$ будем искать в виде

$$u(x,\ t) = X(x)\,T(t), \qquad (27.25)$$

где $X(x)$ — функция только от x; $T(t)$ — функция только от t.

Подставляя выражения для частных производных второго порядка в уравнение (27.22), получаем

$$X(x)\,T''(t) = a^2X''(x)T(t),\quad \frac{1}{a^2}\,\frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)}.$$

Левая часть последнего равенства не зависит от x, она может зависеть только от t; правая часть не зависит от t, она может зависеть только от x. Но поскольку функция $u(x,\ t) = X(x)\,T(t)$ — решение уравнения (27.22), то последнее равенство должно выполняться для всех x и всех t (из рассматриваемых промежутков), что возможно лишь тогда, когда обе части не зависят ни от x, ни от t, т. е. являются постоянными. Обозначим эту постоянную буквой c, тогда

$$\frac{1}{a^2}\,\frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)} = c,$$

откуда получаем два уравнения:

$$T''(t) - ca^2T(t) = 0; \qquad (27.26)$$

$$X''(x) - cX(x) = 0. \qquad (27.27)$$

Эти уравнения являются обыкновенными дифференциальными уравнениями второго порядка с постоянными коэффициентами (они изучались в гл. 25, § 25.7).

Найдем решение уравнения (27.27), удовлетворяющее условиям (27.24). Для функции (27.25) эти условия принимают вид

$$X(0)\,T(t) = 0,\quad X(l)\,T(t) = 0. \qquad (27.28)$$

Поскольку мы ищем нетривиальное решение ($u(x, t) \not\equiv 0$), то из (27.28) следует, что должно быть

$$X(0)=0, \quad X(l)=0. \tag{27.29}$$

Характеристическое уравнение для уравнения (27.27) имеет вид

$$r^2-c=0, \quad r^2=c. \tag{27.30}$$

Рассмотрим случаи, когда $c=\lambda^2$, $c=0$, $c=-\lambda^2$ (т. е. когда $c>0$, $c=0$, $c<0$).

1. Если $c=\lambda^2$, то $r_1=-\lambda$, $r_2=\lambda$, общее решение уравнения (27.27) выражается формулой

$$X(x)=C_1e^{-\lambda x}+C_2e^{\lambda x}. \tag{27.31}$$

Условия (27.29) приводят к системе уравнений

$$\left.\begin{aligned} C_1e^{-0}+C_2e^0&=0,\\ C_1e^{-\lambda l}+C_2e^{\lambda l}&=0,\end{aligned}\right\} \qquad \left.\begin{aligned} C_1+C_2&=0,\\ C_1e^{-\lambda l}+C_2e^{\lambda l}&=0.\end{aligned}\right\}$$

Определитель этой однородной системы относительно неизвестных C_1 и C_2 отличен от нуля:

$$\Delta=\begin{vmatrix} 1 & 1\\ e^{-\lambda l} & e^{\lambda l}\end{vmatrix}=e^{\lambda l}-e^{-\lambda l}\neq 0,$$

поэтому система имеет единственное нулевое решение $C_1=0$, $C_2=0$.

Следовательно, в этом случае решением уравнения (27.27), удовлетворяющим условиям (27.29), является только тривиальное решение $X(x)\equiv 0$.

2. Если $c=0$, то $r_1=0$, $r_2=0$. Общее решение уравнения (27.27) определяется формулой

$$X(x)=C_1+C_2x. \tag{27.32}$$

Среди решений (27.32) только тривиальное решение $X(x)\equiv 0$ удовлетворяет условиям (27.29). Действительно, условия (27.29) для решения (27.32) принимают вид

$$\left.\begin{aligned} C_1+C_2\cdot 0&=0,\\ C_1+C_2l&=0,\end{aligned}\right\}$$

откуда $C_1=0$, $C_2=0$, поэтому $X(x)\equiv 0$.

3. Если $c=-\lambda^2$, то характеристическое уравнение (27.30) имеет мнимые корни $r_1=-i\lambda$, $r_2=i\lambda$, которым соответствует решение

$$X(x)=C_1\cos\lambda x+C_2\sin\lambda x \tag{27.33}$$

уравнения (27.27).

Найдем решение, удовлетворяющее условиям (27.29):

$$\left.\begin{aligned} C_1\cos 0+C_2\sin 0=0,\\ C_1\cos\lambda l+C_2\sin\lambda l=0, \end{aligned}\right\} \qquad \left.\begin{aligned} C_1\cdot 1+C_2\cdot 0=0,\\ C_1\cos\lambda l+C_2\sin\lambda l=0. \end{aligned}\right\}$$

Из этой системы получаем $C_1=0$,

$$\sin\lambda l=0, \tag{27.34}$$

откуда $\lambda l=k\pi$ $(k=\pm 1, \pm 2, \ldots)$, $k\neq 0$, ибо $\lambda\neq 0$,

$$\lambda_k=\frac{k\pi}{l}\quad (k=\pm 1, \pm 2, \ldots). \tag{27.35}$$

Итак, каждая из функций

$$X_k(x)=A_k\sin\frac{k\pi x}{l}\quad (k=1, 2, \ldots), \tag{27.36}$$

где A_k — произвольная постоянная (к ней отнесен и знак; считаем, что k принимает только целые положительные значения), является решением уравнения (27.27), удовлетворяющим условиям (27.29).

Постоянные λ_k, определяемые формулой (27.35), называются *собственными числами*, а функции

$$f_k(x)=\sin\frac{k\pi x}{l}\quad (k=\pm 1, \pm 2, \ldots) \tag{27.37}$$

называются *собственными функциями* уравнения (27.27) с краевыми условиями (27.29).

Отметим, что функции (27.37) ортогональны на отрезке $[0, l]$, так как (см. § 23.4) для них

$$\int_{-l}^{l} f_n(x)\, f_m(x)\, dx=0,\quad n\neq m.$$

Обратимся к уравнению (27.26). Так как $c=-\lambda^2$, $\lambda^2=\frac{k^2\pi^2}{l^2}$, то уравнение принимает вид

$$T''(t)+\frac{a^2k^2\pi^2}{l^2}T(t)=0.$$

Характеристическое уравнение $r^2+\frac{a^2k^2\pi^2}{l^2}=0$ имеет мнимые корни $r_1=-\frac{k\pi a}{l}i$, $r_2=\frac{k\pi a}{l}i$, поэтому общее решение уравнения (27.26) определяется формулой

$$T_k(t) = B_k \cos\frac{k\pi at}{l} + D_k \sin\frac{k\pi at}{l} \quad (k = 1,\ 2,\ \ldots), \tag{27.38}$$

где B_k, D_k — произвольные постоянные.

Подставляя функции (27.36) и (27.38) в формулу (27.25), получаем

$$u_k(x,\ t) = X_k(x)\, T_k(t) =$$
$$= \left(B_k \cos\frac{k\pi at}{l} + D_k \sin\frac{k\pi at}{l}\right) A_k \sin\frac{k\pi x}{l},$$

или

$$u_k(x,\ t) = \left(a_k \cos\frac{k\pi at}{l} + b_k \sin\frac{k\pi at}{l}\right) \sin\frac{k\pi x}{l}, \tag{27.39}$$

где $a_k = B_k A_k$, $b_k = D_k A_k$ $(k = 1,\ 2,\ 3,\ \ldots)$.

Каждая из функций (27.39) является решением уравнения (27.22), удовлетворяющим условиям (27.24), и называется *собственной функцией* этого уравнения, а колебания, определяемые ею, называются *собственными колебаниями.*

Перейдем ко второй части метода Фурье. Составим ряд из функций (27.39), т. е. ряд

$$\sum_{k=1}^{\infty} u_k(x,\ t) = \sum_{k=1}^{\infty} \left(a_k \cos\frac{k\pi at}{l} + b_k \sin\frac{k\pi at}{l}\right) \sin\frac{k\pi x}{l}.$$

Предположим, что этот ряд равномерно сходится, его сумма — непрерывная функция и ряд можно дважды почленно дифференцировать по x и t. Обозначим сумму этого ряда через $u(x,\ t)$, т. е.

$$u(x,\ t) = \sum_{k=1}^{\infty} \left(a_k \cos\frac{k\pi at}{l} + b_k \sin\frac{k\pi at}{l}\right) \sin\frac{k\pi x}{l}. \tag{27.40}$$

Согласно § 26.8 функция (27.40) также является решением уравнения (27.22), причем это решение удовлетворяет условиям (27.24).

Коэффициенты ряда (27.40) a_k и b_k выберем так, чтобы его сумма — функция $u(x,\ t)$ удовлетворяла условиям (27.23). Найдем предварительно частную производную по t функции (27.40):

$$\frac{\partial u}{\partial t} = \sum_{k=1}^{\infty} \frac{k\pi a}{l} \left(-a_k \sin\frac{k\pi at}{l} + b_k \cos\frac{k\pi at}{l}\right) \sin\frac{k\pi x}{l}. \tag{27.41}$$

Подставляя в формулы (27.40), (27.41) значение $t=0$ и принимая во внимание равенства (27.23), получаем

$$\sum_{k=1}^{\infty} a_k \sin\frac{k\pi x}{l} = f(x), \quad \sum_{k=1}^{\infty} b_k \frac{k\pi a}{l} \sin\frac{k\pi x}{l} = F(x).$$

Следовательно, функции $f(x)$ и $F(x)$ представляют собой разложения в ряд Фурье по синусам в промежутке $[0, l]$.

Как известно, коэффициенты таких разложений определяются формулами

$$a_k = \frac{2}{l}\int_0^l f(x)\sin\frac{k\pi x}{l}\,dx,$$

$$b_k\frac{k\pi a}{l} = \frac{2}{l}\int_0^l F(x)\sin\frac{k\pi x}{l}\,dx,$$

откуда

$$\begin{aligned} a_k &= \frac{2}{l}\int_0^l f(x)\sin\frac{k\pi x}{l}\,dx, \\ b_k &= \frac{2}{k\pi a}\int_0^l F(x)\sin\frac{k\pi x}{l}\,dx. \end{aligned} \tag{27.42}$$

Итак, функция (27.40), где a_k, b_k $(k=1, 2, 3, \ldots)$ определены формулами (27.42), даст решение рассматриваемой задачи.

§ 27.6. Понятие о стоячих волнах

Выясним физический смысл собственных функций (27.39):

$$u_k(x, t) = \left(a_k\cos\frac{k\pi at}{l} + b_k\sin\frac{k\pi at}{l}\right)\sin\frac{k\pi x}{l}.$$

Из этой формулы видно, что в моменты времени $t=\frac{2l}{a}$, $t=\frac{4l}{a}$, ... струна возвращается в первоначальное положение. Следовательно, колебания являются незатухающими, периодическими с периодом $T=\frac{2l}{a}$. Это происходит потому, что не учтены силы сопротивления.

Если эти силы учесть, колебания окажутся затухающими.

Преобразуем формулу (27.39), для чего введем обозначения $a_k = F_k \sin\varphi_k$, $b_k = F_k \cos\varphi_k$, откуда $a_k^2 + b_k^2 = F_k^2$,

$$\sqrt{a_k^2 + b_k^2} = F_k, \frac{a_k}{\sqrt{a_k^2 + b_k^2}} = \sin\varphi_k, \quad \frac{b_k}{\sqrt{a_k^2 + b_k^2}} = \cos\varphi_k.$$

Умножив и разделив на $\sqrt{a_k^2 + b_k^2}$ правую часть равен-

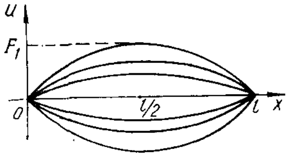

Рис. 27.4

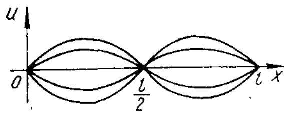

Рис. 27.5

ства (27.39), а также приняв во внимание эти обозначения, получим

$$u_k(x, t) = F_k \sin\left(\frac{k\pi a t}{l} + \varphi_k\right) \sin\frac{k\pi x}{l}. \qquad (27.43)$$

Отсюда следует, что все точки струны совершают колебания с одной и той же частотой $\omega_k = \frac{k\pi a}{l}$ и одной и той же начальной фазой φ_k. Амплитуда колебания зависит от точки x и равна $F_k \sin\frac{k\pi x}{l}$. Все точки струны одновременно проходят через положение равновесия и одновременно достигают максимального отклонения от него в ту или другую сторону. Рассматриваемые колебания струны называются *стоячими волнами*.

На рис. 27.4 показаны различные формы струны в отдельные моменты времени при $k=1$. Отклонение на концах струны равно нулю, наибольшего отклонения достигает точка с абсциссой $x = \frac{l}{2}$. При $k = 2$ неподвижных точек уже будет три, их абсциссы: $x_1 = 0$, $x_2 = \frac{l}{2}$, $x_3 = l$ $\left(\text{это корни уравнения } \sin\frac{2\pi x}{l} = 0\right)$. Наибольшего отклонения достигают две точки струны с абсциссами $x = \frac{l}{4}$, $x = \frac{3l}{4}$ (рис. 27.5).

Стоячая волна (27.43) имеет $k+1$ неподвижных точек, т. е. столько, сколько корней имеет уравнение $\sin\frac{k\pi x}{l}=0$ на отрезке $[0, l]$. Абсциссы этих точек являются корнями указанного уравнения:

$$x_1 = 0, \quad x_2 = \frac{l}{k}, \; x_3 = \frac{2l}{k}, \; \ldots, \; x_k = \frac{k-1}{k}l, \; x_{k+1} = l.$$

Неподвижные точки называются *узлами стоячей волны*. Точки, в которых отклонения достигают максимума, называются *пучностями*.

Каждая струна может иметь собственные колебания только строго определенных частот $\omega_k = \frac{k\pi a}{l}$, которые называются *собственными частотами*. Наименьшей собственной частотой струны является частота

$$\omega_1 = \frac{\pi a}{l} = \frac{\pi}{l}\sqrt{\frac{T}{\rho}},$$

где T — натяжение; ρ — линейная плотность.

При колебаниях струна издает звук, высота которого возрастает вместе с частотой колебаний. Самый низкий тон будет при частоте, равной ω_1. Остальные тона, соответствующие частотам ω_k, называют *обертонами*, или гармониками.

§ 27.7. Уравнение теплопроводности в пространстве

Если некоторое тело неравномерно нагрето, то тепло начнет распространяться от более нагретых участков к менее нагретым. Обозначим температуру тела в точке $M(x, y, z)$ через u. Очевидно, температура будет зависеть от координат x, y, z и времени t, т. е. $u=u(x, y, z, t)$.

Множество точек, в которых функция $u=u(x, y, z, t)$ принимает одно и то же значение C в фиксированный момент времени $t=t_0$, называется *изотермической поверхностью*. Изотермическая поверхность определяется уравнением

$$u(x, y, z, t_0) = C \quad \text{или} \quad u(x, y, z) = C_1. \qquad (27.44)$$

Форма и расположение изотермической поверхности меняется с течением времени t.

Направление наибольшей скорости изменения температуры совпадает с направлением градиента функции $u=u(x, y, z, t)$ при данном значении t (см. § 18.10), т. е. с направлением вектора

$$\operatorname{grad} u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}.$$

В точках изотермической поверхности градиент направлен по нормали $\mathbf{n}$ к этой поверхности в сторону возрастания u, причем $|\operatorname{grad} u| = \frac{\partial u}{\partial n}$.

Считают, что величина теплового потока через малый участок изотермической поверхности площади $\Delta\sigma$ за промежуток времени Δt выражается формулой

$$\Delta Q = -k\frac{\partial u}{\partial n}\Delta\sigma\Delta t, \tag{27.45}$$

где k — коэффициент теплопроводности, который будем считать постоянным.

Замечание 1. Знак минус в формуле (27.45) означает, что тепловой поток считают положительным, если его направление совпадает с выбранным направлением нормали к изотермической поверхности: если $\frac{\partial u}{\partial n} > 0$, то u возрастает в этом направлении, а тепло передается от более нагретых участков к менее нагретым, т. е. в противоположную сторону, получаем $\Delta Q < 0$; если $\frac{\partial u}{\partial n} < 0$, то $\Delta Q > 0$.

Замечание 2. В линейном случае (т. е. при распространении тепла в стержне) изотермическими поверхностями являются поперечные сечения стержня, направление нормали к ним совпадает с направлением оси Ox, поэтому $\frac{\partial u}{\partial n} = \frac{\partial u}{\partial x}$.

В теории теплопроводности доказывается, что формула (27.45) справедлива для любых поверхностей (не только изотермических).

Известно (см. § 18.10), что производная по направлению вектора равна проекции градиента на этот вектор, т. е. $\frac{\partial u}{\partial n} = \operatorname{grad} u \cdot \mathbf{n}$, где $\mathbf{n}$ — единичный вектор нормали. Формула (27.45) принимает вид $\Delta Q = -k \operatorname{grad} u \cdot \mathbf{n}\Delta\sigma\Delta t$.

Введем в рассмотрение вектор

$$\mathbf{a} = -k \operatorname{grad} u \tag{27.46}$$

и назовем его *вектором теплового потока*, тогда

$$\Delta Q = \mathbf{a}\mathbf{n}\Delta\sigma\Delta t, \quad \Delta Q = a_n\Delta\sigma\Delta t, \tag{27.47}$$

где a_n — проекция вектора **a** на направление внешней нормали к поверхности (рис. 27.6).

Выделим в теле некоторый объем V, ограниченный поверхностью σ (см. рис. 27.6). Тепловой поток через всю поверхность σ за промежуток времени Δt выражается формулой

$$Q = \Delta t \iint\limits_{\sigma} a_n d\sigma. \qquad (27.48)$$

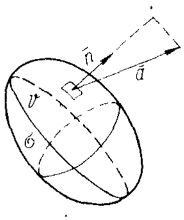

Рис. 27.6

Поток Q будет положительным, если выбранный объем V теряет тепло, и отрицательным, если приобретает его.

Преобразуем этот интеграл с помощью формулы Остроградского — Гаусса

$$\iint\limits_{\sigma} a_n d\sigma = \iiint\limits_{V} \operatorname{div} \mathbf{a} dv.$$

Так как

$$\operatorname{div} \mathbf{a} = \operatorname{div}(-k \operatorname{grad} u) = -k \operatorname{div} \operatorname{grad} u =$$
$$= -k\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right)$$

(см. § 20.14, формула (20.87)) или $\operatorname{div} \mathbf{a} = -k\Delta u$, где

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2},$$

то

$$\iint\limits_{\sigma} a_n d\sigma = -\iiint\limits_{V} k\Delta u dv.$$

Количество тепла, приобретенное данным объемом V за счет прохождения теплового потока, выразится формулой

$$Q_1 = \Delta t \iiint\limits_{V} k\Delta u dv$$

(это выражение отличается от Q только знаком).

Пусть в объеме V имеются источники тепла с плотностью их распределения $F(x, y, z, t)$, тогда за промежуток времени Δt выделится следующее количество тепла:

$$Q_2 = \Delta t \iiint\limits_{V} F(x, y, z, t)\, dv.$$

Итак, в течение указанного промежутка Δt объем V получит Q_1+Q_2 тепла.

Подсчитаем это тепло по-другому. За время Δt температура в точке $M(x, y, z)$ изменилась на величину

$$u(x, y, z, t+\Delta t)-u(x, y, z, t)=\frac{\partial u}{\partial t}\Delta t.$$

Для такого изменения температуры элементарному объему Δv потребовалось количество тепла, равное $c\rho\Delta v\frac{\partial u}{\partial t}\Delta t$, где c — удельная теплоемкость, ρ — плотность, а всему объему —

$$Q_3=\Delta t\iiint_V c\rho\frac{\partial u}{\partial t}\,dv.$$

Поскольку $Q_3=Q_1+Q_2$, то

$$\Delta t\iiint_V c\rho\frac{\partial u}{\partial t}\,dv=\Delta t\iiint_V k\Delta u\,dv+\Delta t\iiint_V F(x, y, z, t)\,dv,$$

откуда

$$\iiint_V\left[c\rho\frac{\partial u}{\partial t}-k\Delta u-F(x, y, z, t)\right]dv=0.$$

Это равенство должно выполняться для любого объема V, выделенного в рассматриваемом теле. Считая все входящие функции непрерывными, из последнего равенства получаем

$$c\rho\frac{\partial u}{\partial t}=k\Delta u+F,\quad \frac{\partial u}{\partial t}=\frac{k}{c\rho}\Delta u+\frac{1}{c\rho}F$$

или

$$\frac{\partial u}{\partial t}=a^2\Delta u+\frac{1}{c\rho}F, \tag{27.49}$$

где

$$a^2=\frac{k}{c\rho},\quad a=\sqrt{\frac{k}{c\rho}}. \tag{27.50}$$

Уравнение (27.49) называется *уравнением теплопроводности в пространстве.*

Если тепловые источники отсутствуют, то уравнение принимает вид

$$\frac{\partial u}{\partial t}=a^2\Delta u,\ \text{или}\ \frac{\partial u}{\partial t}=a^2\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right). \tag{27.51}$$

Отметим частные случаи этого уравнения:

$$\frac{\partial u}{\partial t}=a^2\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right), \qquad (27.52)$$

$$\frac{\partial u}{\partial t}=a^2\frac{\partial^2 u}{\partial x^2}. \qquad (27.53)$$

Уравнение (27.52) является уравнением распространения тепла в пластинке, уравнение (27.53) — уравнением распространения тепла в стержне.

§ 27.8. Начальное и краевые условия для уравнения теплопроводности в пространстве

Начальное условие для уравнения теплопроводности в пространстве определяется равенством

$$u(x,\ y,\ z,\ t)\,|_{t=0}=f(x,\ y,\ z), \qquad (27.54)$$

оно задает температуру каждой точки тела в начальный момент времени $t_0=0$ ($f(x,\ y,\ z)$ — известная функция).

Пусть на поверхности Γ, ограничивающей некоторое тело, происходит теплообмен с окружающей средой. Предположим, что в точке $M(\xi,\ \eta,\ \zeta)$ границы Γ тело имеет температуру $u=u(\xi,\eta,\zeta,t)$, а окружающая среда — $\tilde{u}=\tilde{u}(\xi,\ \eta,\ \zeta,\ t)$, разность $u-\tilde{u}$ называется *перепадом температур*. Теплообмен протекает по закону: поток тепла изнутри тела через любую часть поверхности Γ пропорционален перепаду температур $u-\tilde{u}$, т. е. $Q=h(u-\tilde{u})$, где h — коэффициент теплообмена; h может меняться от точки к точке, но в случае однородности тела и среды $h=\text{const}$.

Выделим часть Γ_1 поверхности Γ. Тепловой поток через Γ_1 за время Δt выразится формулой

$$Q_1=\Delta t\iint\limits_{\Gamma_1} h\,(u-\tilde{u})\,d\sigma.$$

С другой стороны, в соответствии с формулой (27.48) получаем

$$Q_2=\Delta t\iint\limits_{\Gamma_1} a_n d\sigma.$$

Так как тепловой поток, уходящий в окружающее пространство, равен тепловому потоку, подходящему изнутри тела, то $Q_1=Q_2$, т. е.

$$\iint_{\Gamma_1} a_n d\sigma = \iint_{\Gamma_1} h\,(u-\tilde{u})\,d\sigma.$$

Это равенство выполняется для любой части Γ_1 границы Γ. Следовательно, на границе Γ должно выполняться условие $a_n=h(u-\tilde{u})$.

Поскольку

$$a_n = \mathbf{an} = (-k\operatorname{grad} u)\,\mathbf{n} = -k\frac{\partial u}{\partial n},$$

где $\frac{\partial u}{\partial n}$ — производная функции u в точке границы по направлению внешней нормали к поверхности Γ, то последнее равенство можно записать так:

$$-k\frac{\partial u}{\partial n}\bigg|_{\Gamma} = h\,(u|_{\Gamma}-\tilde{u}). \qquad (27.55)$$

Отметим два важных частных случая краевого условия (27.55).

1. Коэффициент теплообмена равен нулю, $h=0$ (на границе тела нет теплообмена с окружающей средой); в этом случае краевое условие принимает вид

$$\frac{\partial u}{\partial n}\bigg|_{\Gamma} = 0. \qquad (27.56)$$

2. Коэффициент теплообмена является достаточно большим. Перепишем условие (27.55) в виде

$$-\frac{k}{h}\,\frac{\partial u}{\partial n} = u|_{\Gamma}-\tilde{u},$$

отсюда при $h\to\infty$ получаем $u|_{\Gamma}-\tilde{u}=0$ или

$$u|_{\Gamma} = \tilde{u}. \qquad (27.57)$$

Краевое условие (27.57) означает, что на границе Γ поддерживается постоянная температура.

Итак, задача о распространении тепла в пространстве (для однородного тела без тепловых источников) формулируется следующим образом.

Найти решение $u=u(x, y, z, t)$ уравнения (27.51), удовлетворяющее начальному условию (27.54) и краевому условию (27.55) (в частном случае условию (27.56) или (27.57)).

В теории дифференциальных уравнений с частными производными доказывается, что при некоторых предположениях относительно соответствующих функций поставленная задача имеет единственное решение.

§ 27.9. Теплопроводность в стержне, концы которого теплоизолированы

Для уравнения (27.53) начальное условие (27.54) принимает вид $u(x, t)\big|_{t=0}=f(x)$, где $f(x)$ — заданная функция ($0\leqslant x\leqslant l$, l — длина стержня), а краевое условие (27.56) запишется так (см. замечание 2 § 27.7):

$$\left.\frac{\partial u}{\partial x}\right|_{x=0}=0, \quad \left.\frac{\partial u}{\partial x}\right|_{x=l}=0.$$

Задача о распространении тепла в стержне, концы которого теплоизолированы, состоит в следующем.

Найти решение уравнения

$$\frac{\partial u}{\partial t}=a^2\frac{\partial^2 u}{\partial x^2}, \tag{27.58}$$

удовлетворяющее начальному условию

$$u(x, t)\big|_{t=0}=f(x) \tag{27.59}$$

и краевым условиям

$$\left.\frac{\partial u}{\partial x}\right|_{x=0}=0, \quad \left.\frac{\partial u}{\partial x}\right|_{x=l}=0. \tag{27.60}$$

Прежде всего отметим, что уравнение принимает более простой вид, если ввести новую переменную τ по формуле

$$\tau=a^2t. \tag{27.61}$$

Действительно,

$$\frac{\partial u}{\partial t}=\frac{\partial u}{\partial \tau}\frac{d\tau}{dt}=\frac{\partial u}{\partial \tau}a^2, \quad a^2\frac{\partial u}{\partial \tau}=a^2\frac{\partial^2 u}{\partial x^2}, \quad \frac{\partial u}{\partial \tau}=\frac{\partial^2 u}{\partial x^2}.$$

Итак, уравнение принимает вид

$$\frac{\partial u}{\partial \tau}=\frac{\partial^2 u}{\partial x^2}. \tag{27.62}$$

Очевидно, начальное условие и краевые условия останутся прежними, так как $\tau=0$ при $t=0$, а условия (27.60) от τ не зависят.

Решение уравнения (27.62), удовлетворяющее начальному условию

$$u(x,\ \tau)\,\big|_{\tau=0}=f(x) \tag{27.63}$$

и краевым условиям (27.60), будем искать с помощью метода Фурье в виде произведения двух функций

$$u(x,\ \tau)=X(x)\,T(\tau), \tag{27.64}$$

где $X(x)$ — функция только от x; $T(\tau)$ — функция только от τ.

Поскольку

$$\frac{\partial u}{\partial \tau}=X(x)\,T'(\tau),\quad \frac{\partial u}{\partial x}=X'(x)\,T(\tau),\quad \frac{\partial^2 u}{\partial x^2}=X''(x)\,T(\tau),$$

то подстановка соответствующих выражений в уравнение (27.62) приводит к соотношению

$$X(x)\,T'(\tau)=X''(x)\,T(\tau)\ \text{ или }\ \frac{X''(x)}{X(x)}=\frac{T'(\tau)}{T(\tau)}.$$

Так как функция (27.64) есть решение уравнения (27.62), то последнее равенство должно выполняться для всех x и τ (из соответствующей области их изменения). Это возможно лишь в случае, когда обе части последнего равенства равны постоянной, ибо левая часть может зависеть только от x, а правая — только от τ. Обозначим эту постоянную через c, тогда

$$\frac{X''(x)}{X(x)}=\frac{T'(\tau)}{T(\tau)}=c,$$

откуда

$$T'(\tau)-cT(\tau)=0; \tag{27.65}$$

$$X''(x)-cX(x)=0. \tag{27.66}$$

Уравнение (27.65) является обыкновенным дифференциальным уравнением первого порядка. Интегрируя это уравнение, получаем

$$\ln T(\tau)=c\tau+\ln C,\quad T(\tau)=e^{c\tau+\ln C},\quad T(\tau)=Ce^{c\tau}.$$

Поскольку температура не может неограниченно возрастать с течением времени (источники тепла отсутствуют), то функция $T(\tau)$ обладает тем же свойством. Следовательно, в последней формуле c может быть только отрицательным, т. е.

$$c=-\lambda^2\quad (\lambda\neq 0). \tag{27.67}$$

Итак, функция $T(\tau)$ выражается формулой

$$T(\tau)=Ce^{-\lambda^2\tau}. \tag{27.68}$$

Уравнение (27.66) с учетом (27.67) принимает вид $X''(x)+\lambda^2 X(x)=0$. Это обыкновенное дифференциальное уравнение второго порядка с постоянными коэффициентами имеет решение

$$X(x)=A\cos\lambda x+B\sin\lambda x, \tag{27.69}$$

где A и B — произвольные постоянные.

Подставляя функции (27.68) и (27.69) в формулу (27.64), получаем

$$u=(\alpha\cos\lambda x+\beta\sin\lambda x)e^{-\lambda^2\tau}, \tag{27.70}$$

где $\alpha=AC$, $\beta=BC$.

Частная производная этой функции по x выражается формулой

$$\frac{\partial u}{\partial x}=\lambda(-\alpha\sin\lambda x+\beta\cos\lambda x)e^{-\lambda^2\tau}.$$

Числа α, β, λ выберем так, чтобы удовлетворялись условия (27.60):

$$\left.\begin{aligned}\lambda(-\alpha\sin 0+\beta\cos 0)e^{-\lambda^2\tau}=0,\\ \lambda(-\alpha\sin\lambda l+\beta\cos\lambda l)e^{-\lambda^2\tau}=0;\end{aligned}\right\}$$

$$\left.\begin{aligned}-\alpha\sin 0+\beta\cos 0=0,\\ -\alpha\sin\lambda l+\beta\cos\lambda l=0.\end{aligned}\right\}$$

Из последних уравнений следует, что $\beta=0$, $\sin\lambda l=0$. Таким образом,

$$\lambda l=n\pi \quad (n=1, 2, 3, \ldots),$$

$$\lambda_n=\frac{n\pi}{l} \quad (n=1,\ 2,\ 3,\ \ldots) \tag{27.71}$$

(взяты только положительные значения n, так как $\cos\lambda x=\cos(-\lambda x)$, $(-\lambda)^2=\lambda^2$; поскольку $\beta=0$, функция (27.70) принимает одинаковые значения при λ и $-\lambda$; далее, $\lambda\neq 0$ в соответствии с условием (27.67)).

Итак, функция (27.70) принимает вид

$$u_n(x,\ \tau)=\alpha_n\cos\lambda_n x e^{-\lambda_n^2\tau},$$

$$u_n(x,\ \tau)=\alpha_n\cos\frac{\pi n x}{l}e^{-\frac{\pi^2 n^2\tau}{l^2}}. \tag{27.72}$$

Решением уравнения (27.62), удовлетворяющим краевым условиям (27.60), будет сумма ряда, составленного из функций (27.72), т. е. ряда

$$u(x, \tau) = \sum_{n=1}^{\infty} u_n(x, \tau) = \sum_{n=1}^{\infty} \alpha_n \cos\frac{n\pi x}{l} e^{-\frac{\pi^2 n^2 \tau}{l^2}}. \quad (27.73)$$

Коэффициенты α_n этого ряда выберем так, чтобы функция $u(x, \tau)$ удовлетворяла также и начальному условию (27.63):

$$u(x, \tau)|_{\tau=0} = f(x), \quad \sum_{n=1}^{\infty} \alpha_n \cos\frac{n\pi x}{l} = f(x).$$

Последнее равенство означает, что функцию $f(x)$ в промежутке $[0, l]$ необходимо разложить в ряд Фурье по косинусам. Как известно, коэффициенты такого разложения определяются формулой

$$\alpha_n = \frac{2}{l}\int_0^l f(x)\cos\frac{n\pi x}{l}\,dx \quad (n = 1, 2, 3, \ldots). \quad (27.74)$$

Следовательно, функция (27.73), для которой α_n определены формулой (27.74), есть решение уравнения (27.62), удовлетворяющее начальному условию (27.63) и краевым условиям (27.60).

Принимая во внимание равенство (27.61), заключаем, что функция

$$u(x, t) = \sum_{n=1}^{\infty} \alpha_n \cos\frac{n\pi x}{l} e^{-\frac{\pi^2 n^2 a^2 t}{l^2}},$$

для которой α_n определены формулой (27.74), является решением уравнения (27.58), удовлетворяющим начальному условию (27.59) и краевым условиям (27.60).

Замечание. Аналогично решается задача о распространении тепла в стержне, на концах которого поддерживается постоянная температура. Эта задача о нахождении решения уравнения (27.58), удовлетворяющего начальному условию (27.59) и краевому условию (27.57), которое принимает вид

$$u|_{x=0} = u_0, \quad u|_{x=l} = u_l. \quad (27.75)$$

Предварительно задачу необходимо свести к задаче с однородными краевыми условиями, т. е. с условиями, которым удовлетворяет тривиальное решение ($u \equiv 0$). Условия (27.75) при $u_0 \neq 0$ и $u_l \neq 0$ не являются однородными. При неоднородных условиях сумма и разность двух решений $u_1(x, t)$ и $u_2(x, t)$ уже не будут решениями, удовлетворяющими этим условиям. В частности, не является решением $u_1(x, t)-u_2(x, t)\equiv 0$. Последнее обстоятельство затрудняет построение общего решения (27.73).

Приведение к задаче с однородными краевыми условиями можно осуществить с помощью новой функции

$$w(x, t)=u(x, t)+(u_0-u_l)\frac{x}{l}-u_0.$$

Для этой функции уравнение (27.58) останется прежним, начальное условие примет вид

$$w(x, t)|_{t=0}=u(x, t)|_{t=0}+(u_0-u_l)\frac{x}{l}-u_0=f_1(x),$$

а краевые условия станут однородными:

$$w(x, t)|_{x=0}=u(x, t)|_{x=0}+(u_0-u_l)\frac{x}{l}\Big|_{x=0}-u_0=u_0-u_0=0,$$

$$w(x, t)|_{x=l}=u(x,t)|_{x=l}+(u_0-u_l)\frac{x}{l}\Big|_{x=l}-u_0=$$

$$=u_l+(u_0-u_l)-u_0=0.$$

§ 27.10. Уравнение диффузии

Если среда заполнена газом с различной концентрацией, то происходит диффузия из мест с большей концентрацией в места с меньшей концентрацией. Аналогичное явление наблюдается и в растворах с переменной концентрацией растворимого вещества для данного объема.

В задачах о диффузии неизвестной функцией является концентрация диффундирующего вещества, которую обозначают через c, $c=c(x, y, z, t)$.

Процесс диффузии во многом схож с процессом распространения тепла, поэтому в предположениях, аналогичных тем, которые были сделаны в § 27.7, получаем, что функция $c=c(x, y, z, t)$ должна удовлетворять уравнению

$$\frac{\partial c}{\partial t}=D\left(\frac{\partial^2 c}{\partial x^2}+\frac{\partial^2 c}{\partial y^2}+\frac{\partial^2 c}{\partial z^2}\right). \qquad (27.76)$$

Постоянная D $(D>0)$ называется *коэффициентом диффузии.*

Начальное условие

$$c(x, y, z, t)|_{t=0}=f(x, y, z), \qquad (27.77)$$

где $f(x, y, z)$ — заданная функция, определяет *начальную концентрацию*.

В качестве краевых условий рассматриваются главным образом следующие условия:

$$\left.\frac{\partial c}{\partial n}\right|_{\Gamma} = 0; \qquad (27.78)$$

$$c(x, y, z, t)\,|_{\Gamma} = c_0, \qquad (27.79)$$

Γ — граница области, в которой происходит диффузия.

Условие (27.78) означает, что граница области для диффундирующего вещества является непроницаемой стенкой. Условие (27.79) определяет концентрацию на границе области.

Линейные задачи диффузии (т. е. задачи о диффузии в тонкой трубке с непроницаемой стенкой) состоят в следующем.

Найти решение $c = c(x, t)$ уравнения $\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$, удовлетворяющее начальному условию $c(x, t)\,|_{t=0} = f(x)$ и краевым условиям вида $\frac{\partial c}{\partial x} = 0$ или $c = c_0$ на конце или на концах трубки (первое условие на одном, второе — на другом конце; первое условие на обоих концах, второе условие на обоих концах).

Эти задачи аналогичны задаче о распространении тепла в стержне и решаются с помощью методов, изложенных в § 27.9.

§ 27.11. Уравнение Лапласа. Задача Дирихле

Уравнением Лапласа называется уравнение

$$\Delta u = 0, \qquad (27.80)$$

где Δu — лапласиан, выражение которого в декартовых, цилиндрических и сферических координатах имеет соответственно вид

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2},$$

$$\Delta u = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\partial^2 u}{\partial z^2},$$

$$\Delta u = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial u}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial \varphi^2}.$$

Уравнение Лапласа часто встречается в приложениях.

Этому уравнению удовлетворяет стационарное распределение температуры в теле. Действительно, если функция $u = u(x, y, z, t)$ не зависит от времени t, то $\frac{\partial u}{\partial t} \equiv 0$. Уравнение (27.51) принимает вид (27.80).

Уравнению Лапласа удовлетворяет потенциал стационарного электрического поля в области, где отсутствуют заряды, и потенциал поля тяготения в области, где отсутствуют массы.

К уравнению Лапласа приводят и другие задачи. Однако при изучении этого уравнения представление функции $u = u(x, y, z)$ как температуры наглядно и удобно. В этой интерпретации функции u наиболее общим граничным условием для уравнения (27.51) является условие (27.55).

Функция, удовлетворяющая уравнению Лапласа, называется *гармонической*. Уравнение Лапласа имеет бесконечное множество решений. Какое-то конкретное решение определяется заданием некоторых дополнительных условий.

Для уравнения Лапласа ставится следующая краевая задача. Найти функцию $u = u(x, y, z)$, гармоническую внутри области, ограниченной замкнутой поверхностью Г, и удовлетворяющую граничному условию

$$H \left. \frac{\partial u}{\partial n} \right|_{\Gamma} = u|_{\Gamma} - \tilde{u},$$

где H и $\tilde{u}$ — функции, заданные на границе Г. В случае задачи о стационарном распределении температуры $H = -\frac{k}{h}$, $\tilde{u}$ — температура окружающей среды на границе тела.

Важный частный случай краевой задачи получается при $H = 0$, соответствующий случаю $h = \infty$, т. е. заданию температуры тела на границе Г: $u|_{\Gamma} = \tilde{u}$. Эта краевая задача называется *задачей Дирихле*.

Задача Дирихле в пространстве ставится так.

Найти функцию $u = u(x, y, z)$, удовлетворяющую уравнению Лапласа внутри области, ограниченной замкнутой поверхностью Г, и принимающую на границе Г заданные значения:

$$u|_{\Gamma} = \tilde{u}. \tag{27.81}$$

В теории уравнений с частными производными доказывается, что при некоторых предположениях относительно Г и $\tilde{u}$ задача Дирихле имеет единственное решение.

Если функция u зависит только от двух переменных (декартовых или полярных координат точки), то уравнение Лапласа имеет соответственно вид

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \quad \text{или} \quad \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2}=0.$$

Задача Дирихле на плоскости состоит в следующем. Найти функцию $u=u(x, y)$, удовлетворяющую уравнению Лапласа внутри области, ограниченной замкнутой кривой Г, и принимающую на границе Г заданные значения $\tilde{u}=\tilde{u}(x, y)$: $u|_\Gamma=\tilde{u}$.

Эта задача также имеет единственное решение. К ней приводят физические задачи двух типов, которые проиллюстрируем на примерах стационарного распределения тепла в теле.

Первый тип задачи возникает при рассмотрении стационарного распределения тепла в тонкой однородной пластинке, параллельной плоскости Oxy, нижняя и верхняя поверхности которой теплоизолированы. Пластинка предполагается настолько тонкой, что можно пренебречь изменением температуры по толщине (температура в этом случае будет функцией только x и y). Край пластинки Г поддерживается при определенной температуре $\tilde{u}$.

Второй тип задачи относится к стационарному распределению температуры в бесконечном однородном цилиндре, у которого образующие параллельны оси Oz, направляющая Г лежит в плоскости Oxy. Температура u остается постоянной на любой прямой, проходящей внутри цилиндра параллельно оси Oz, поэтому $u=u(x, y)$. Боковая поверхность цилиндра поддерживается при определенной температуре $\tilde{u}$.

§ 27.12. Задача Дирихле для круга. Интеграл Пуассона

Задача Дирихле для круга состоит в следующем. Найти функцию, гармоническую внутри круга и принимающую на его границе заданные значения. Введем полярные координаты r и φ, полюс поместим в центре данного круга, радиус которого обозначим через R. Двумерный оператор Лапласа выражаем в полярных координатах. Уравнение Лапласа принимает вид

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u}{\partial\varphi^2}=0. \qquad (27.82)$$

Найдем функцию $u=u(r, \varphi)$, удовлетворяющую уравнению (27.82) при $r<R$ и принимающую на границе Г круга радиуса R заданные значения $u=f(\varphi)$:

$$u(r, \varphi)\,|_{r=R}=f(\varphi). \qquad (27.83)$$

Решения уравнения (27.82) ищем в виде произведения двух функций $U(r)$ и $\Phi(\varphi)$, первая из которых зависит только от r, вторая — только от φ:

$$u(r, \varphi)=U(r)\Phi(\varphi). \qquad (27.84)$$

Подставляя эту функцию в уравнение (27.82), получаем

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dU}{dr}\right)\Phi+\frac{U}{r^2}\frac{d^2\Phi}{d\varphi^2}=0$$

или

$$\frac{r}{U}\frac{d}{dr}\left(r\frac{dU}{dr}\right)=-\frac{1}{\Phi}\frac{d^2\Phi}{d\varphi^2}.$$

Поскольку функция (27.84) — решение уравнения (27.82), то последнее равенство должно выполняться для всех r и φ из данной области ($0\leqslant r<R$, $0\leqslant\varphi<2\pi$). Но это возможно лишь в случае, когда обе части равенства не зависят от r и φ, т. е. являются одной и той же постоянной, так как левая часть его может зависеть только от r, а правая — только от φ.

Обозначая эту постоянную λ, получаем два уравнения:

$$\frac{r^2}{U}\frac{d^2U}{dr^2}+\frac{r}{U}\frac{dU}{dr}=\lambda; \qquad (27.85)$$

$$\frac{d^2\Phi}{d\varphi^2}+\lambda\Phi=0. \qquad (27.86)$$

Это обыкновенные дифференциальные уравнения второго порядка.

Уравнение (27.86) является уравнением с постоянными коэффициентами и имеет решение

$$\Phi(\varphi)=A\cos\sqrt{\lambda}\varphi+B\sin\sqrt{\lambda}\varphi,$$

где A и B — произвольные постоянные. Покажем, что λ не может принимать любые значения. Действительно, прибавление слагаемого 2π к аргументу φ возвращает точку $M(r, \varphi)$ в исходное положение. Это значит, что $\Phi(\varphi+2\pi)=\Phi(\varphi)$, т. е. функция $\Phi(\varphi)$ — периодическая с периодом 2π. Последнее возможно, когда $\sqrt{\lambda}=n$, $\lambda=n^2$ $(n=0, 1, 2, \ldots)$ (отрицательные значения n можно не принимать во внимание, поскольку n влияет только на знак B — произвольной постоянной).

Итак, уравнение (27.86) имеет решения

$$\Phi_n(\varphi)=A_n \cos n\varphi+B_n \sin n\varphi \quad (n=0, 1, 2, \ldots).$$

Уравнение (27.85) при $\lambda=n^2$ принимает вид

$$r^2 \frac{d^2U}{dr^2}+r\frac{dU}{dr}-n^2U=0.$$

Решение этого уравнения находится с помощью подстановки $U=r^{\alpha}$.

Поскольку

$$\frac{dU}{dr}=\alpha r^{\alpha-1}, \quad \frac{d^2U}{dr^2}=\alpha(\alpha-1)r^{\alpha-2},$$

то

$$r^2\alpha(\alpha-1)r^{\alpha-2}+r\alpha r^{\alpha-1}-n^2r^{\alpha}=0, \quad \alpha^2-n^2=0, \quad \alpha=\pm n.$$

В случае $\alpha=-n$ получаем функцию $U(r)=r^{-n}$, которая обращается в бесконечность при $r=0$. Эта функция не может быть использована для построения решения задачи Дирихле (ищется решение непрерывное и конечное в круге радиуса R).

При $\alpha=n$ получаем $U_n(r)=r^n$ $(n=0, 1, 2, \ldots)$. Подставляя выражения для $\Phi_n(\varphi)$ и $U_n(r)$ в формулу (27.84), находим частные решения

$$u_n(r, \varphi)=(A_n \cos n\varphi+B_n \sin n\varphi)r^n$$

уравнения (27.82). Решением этого уравнения является также функция

$$u(r, \varphi)=\sum_{n=0}^{\infty} u_n(r, \varphi)=\sum_{n=0}^{\infty}(A_n \cos n\varphi+B_n \sin n\varphi)r^n.$$

Введя обозначения

$$A_0=\frac{a_0}{2}, \quad A_n=a_n, \quad B_n=b_n \quad (n=1, 2, 3, \ldots),$$

запишем это решение так:

$$u(r,\ \varphi)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos n\varphi+b_n\sin n\varphi)\,r^n. \qquad (27.87)$$

Коэффициенты a_n, b_n определим из условия (27.83):

$$f(\varphi)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos n\varphi+b_n\sin n\varphi)\,R^n,$$

$$f(\varphi)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_nR^n\cos n\varphi+b_nR^n\sin n\varphi).$$

Последнее равенство представляет разложение функции $f(\varphi)$ в ряд Фурье. Как известно,

$$R^na_n=\frac{1}{\pi}\int_0^{2\pi}f(t)\cos nt\,dt,$$

$$R^nb_n=\frac{1}{\pi}\int_0^{2\pi}f(t)\sin nt\,dt,$$

откуда

$$a_n=\frac{1}{\pi R^n}\int_0^{2\pi}f(t)\cos nt\,dt,$$
$$b_n=\frac{1}{\pi R^n}\int_0^{2\pi}f(t)\sin nt\,dt. \qquad (27.88)$$

Подставив выражения (27.88) в формулу (27.87), найдем искомое решение задачи Дирихле для круга.

Преобразуем это решение. Подставляя выражение для коэффициентов a_n, b_n в формулу (27.87), получаем

$$u(r,\ \varphi)=\frac{1}{2\pi}\int_0^{2\pi}\left[1+2\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n\cos n(t-\varphi)\right]f(t)\,dt. \qquad (27.89)$$

Вычислим сумму в квадратной скобке. Принимая во внимание формулы Эйлера (см. § 22.10, формулы (22.76)), находим, что

$$2\cos n(t-\varphi)=e^{in(t-\varphi)}+e^{-in(t-\varphi)}.$$

Следовательно,

$$1+2\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n\cos n\,(t-\varphi)=1+\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n 2\cos n\,(t-\varphi)=$$

$$=1+\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n\left[e^{in\,(t-\varphi)}+e^{-in\,(t-\varphi)}\right]=$$

$$=1+\sum_{n=1}^{\infty}\left[\frac{r}{R}e^{i\,(t-\varphi)}\right]^n+\sum_{n=1}^{\infty}\left[\frac{r}{R}e^{-i\,(t-\varphi)}\right]^n.$$

Комплексные члены каждого ряда образуют геометрические прогрессии, знаменатели которых по модулю меньше единицы. Действительно,

$$\left|\frac{r}{R}\right|\left|e^{i\,(t-\varphi)}\right|=\left|\frac{r}{R}\right|\left|e^{i\,(t-\varphi)}\right|<1,$$

$$\left|\frac{r}{R}\,e^{-i\,(t-\varphi)}\right|=\left|\frac{r}{R}\right|\left|e^{-i\,(t-\varphi)}\right|<1.$$

Сумма членов такой геометрической прогрессии определяется формулой $S=\frac{a}{1-q}$, где a — первый член, q — знаменатель прогрессии.

Находя эти суммы, получаем

$$1+\sum_{n=1}^{\infty}\left[\frac{r}{R}e^{i\,(t-\varphi)}\right]^n+\sum_{n=1}^{\infty}\left[\frac{r}{R}e^{-i\,(t-\varphi)}\right]^n=$$

$$=1+\frac{\frac{r}{R}e^{i\,(t-\varphi)}}{1-\frac{r}{R}e^{i\,(t-\varphi)}}+\frac{\frac{r}{R}e^{-i\,(t-\varphi)}}{1-\frac{r}{R}e^{-i\,(t-\varphi)}}=$$

$$=\frac{R^2-r^2}{R^2-2Rr\cos(t-\varphi)+r^2}.$$

Итак,

$$1+2\sum_{n=1}^{\infty}\left(\frac{r}{R}\right)^n\cos n\,(t-\varphi)=\frac{R^2-r^2}{R^2-2Rr\cos(t-\varphi)+r^2}. \quad (27.90)$$

Подставляя выражение (27.90) в формулу (27.89), находим

$$u(r,\varphi)=\frac{1}{2\pi}\int_0^{2\pi}\frac{R^2-r^2}{R^2-2Rr\cos(t-\varphi)+r^2}f(t)\,dt. \quad (27.91)$$

Функция (27.91) даст решение задачи Дирихле для круга. Интеграл (27.91) называется *интегралом Пуассона.*

РАЗДЕЛ IV

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ

Теория вероятностей изучает объективные закономерности массовых случайных событий. Она является теоретической базой для математической статистики, занимающейся разработкой методов сбора, описания и обработки результатов наблюдений. Путем наблюдений (испытаний, экспериментов), т. е. опыта в широком смысле слова, происходит познание явлений действительного мира.

Глава 28

СЛУЧАЙНЫЕ СОБЫТИЯ И ВЕРОЯТНОСТИ

Здесь вводятся различные определения вероятности события и рассматриваются простейшие свойства вероятностей.

§ 28.1. Классификация событий

Будем называть опытом, или испытанием, всякое осуществление определенного комплекса условий или действий, при которых наблюдается соответствующее явление. Например, опытом является стрельба по мишени, бросание монеты, акт выборки одного изделия из множества готовых изделий и т. д. Возможный результат опыта называют *событием*. Так, при стрельбе по мишени событиями будут попадание и промах, при бросании монеты — герб или цифра на верхней ее стороне; появление бракованного изделия при выборке из множества готовых изделий и т. д.

События будем обозначать заглавными буквами латинского алфавита, например, A — появление цифры, B — попадание в цель и т. п.

Событие называется *достоверным* в данном опыте, если оно обязательно произойдет в этом опыте. Например, если в урне находятся только красные шары, то событие «из урны извлечен красный шар» является достоверным.

Событие называется *невозможным*, если в данном опыте оно произойти не может. Так, если в урне находятся только синие шары, то событие «из урны извлечен белый шар» будет невозможным.

Случайным событием называется событие, которое в данном опыте может произойти, а может и не произойти. Случайными событиями являются, например, попадание и промах при стрельбе, герб или цифра на верхней стороне монеты при ее бросании, выигрыш по билету лотереи и т. д.

Замечание. Говоря о достоверности, невозможности, случайности некоторого события, мы имеем в виду его достоверность, невозможность, случайность по отношению к конкретному опыту, т. е. к наличию определенного комплекса условий. Например, если в урне имеются белые, красные и синие шары, то событие «из урны извлечен белый шар» не будет невозможным, как в приведенном ранее примере.

Два события называются *совместными* в данном опыте, если появление одного из них не исключает появления другого в этом опыте. Так, при бросании двух монет события A (цифра на верхней стороне первой монеты) и B (герб на верхней стороне второй монеты) совместны.

Два события называются *несовместными*, если они не могут произойти вместе при одном и том же испытании. Примеры несовместных событий: попадание и промах при одном выстреле, «герб» и «цифра» — при одном бросании монеты.

Множество событий A_1, A_2, ..., A_n называют *полной группой*, если они попарно-несовместны; появление одного и только одного из них является достоверным событием. Поясним понятие полной группы несовместных событий на следующем примере. Рассмотрим события, появляющиеся при бросании игральной кости (т. е. кубика, на гранях которого написаны цифры 1, 2, 3, 4, 5, 6 или изображены знаки, соответствующие этим цифрам). Если кость упадет, то верхней гранью окажется грань с одной из этих цифр. Событие «верхней гранью оказалась грань с цифрой k» обозначим через A_k ($k=1, 2, 3, 4, 5, 6$). События A_1, A_2, A_3, A_4, A_5, A_6 образуют полную группу: они попарно-несовместны; появление одного и только одного из них будет достоверным событием (только одна из граней окажется верхней, на ней написана только одна из цифр от 1 до 6).

Два события называются *противоположными*, если появление одного из них равносильно непоявлению другого. Примеры противоположных событий: попадание и промах при стрельбе, «герб», «цифра» при бросании монеты. Если одно из противоположных событий обозначено буквой A, то другое обозначают $\bar{A}$. Например, если A — попадание, то $\bar{A}$ — промах при стрельбе.

События считают *равновозможными*, если нет оснований полагать, что одно событие является более возможным, чем другие. Например, при бросании монеты событие A (появление цифры) и событие B (появление герба) равновозможны, так как предполагается, что монета изготовлена из однородного материала, имеет правильную цилиндрическую форму и наличие чеканки не влияет на то, какая сторона монеты (герб или цифра) окажется верхней. При бросании игральной кости события также равновозможны, поскольку предполагается, что кость изготовлена из однородного материала, имеет правильную форму куба и наличие цифр (или очков) на гранях не влияет на то, какая из шести граней куба окажется верхней.

Каждое событие, которое может наступить в итоге опыта, называется *элементарным исходом* (*элементарным событием*, или *шансом*). Например, события A_1, A_2, A_3, A_4, A_5, A_6 — элементарные исходы при бросании игральной кости. Элементарные исходы, при которых данное событие наступает, называются *благоприятствующими* этому событию, или *благоприятными шансами*. Так, при бросании игральной кости элементарные исходы A_3 и A_6 являются благоприятствующими событию «число выпавших очков делится на 3».

§ 28.2. Вероятность события, ее основные свойства. Классическое определение вероятности. Геометрическая вероятность

Рассмотрим конечное множество равновозможных элементарных исходов A_1, A_2, ..., A_n, образующих полную группу событий. Пусть событие A наступает при появлении некоторых из этих элементарных исходов и не наступает при появлении других. Предположим, что среди общего числа n всех равновозможных исходов $A_1, A_2, \ldots, A_n$ m из них благоприятствуют событию A.

Вероятностью события называется отношение числа элементарных исходов, благоприятствующих данному событию, к числу всех равновозможных, образующих полную группу элементарных исходов опыта, в котором может появиться это событие. Вероятность события A обозначим через $P(A)$, тогда по определению

$$P(A) = \frac{m}{n}, \tag{28.1}$$

где m — число элементарных исходов, благоприятствующих событию A; n — число всех равновозможных элементарных исходов опыта, в котором может появляться событие A.

Это определение вероятности называется *классическим*. Оно появилось на начальном этапе развития теории вероятностей. Приведем простейшие примеры вычисления вероятности события на основании ее определения.

Пример 1. Вычислить вероятность появления герба и вероятность появления цифры при бросании монеты.

Пусть A — герб, B — цифра на верхней стороне монеты. События A и B равновозможны и образуют полную группу событий. Для каждого из них число благоприятствующих исходов равно единице, а число всех элементарных исходов равно двум, т. е. $m=1$, $n=2$, поэтому $P(A) = \frac{m}{n} = \frac{1}{2}$, $P(B) = \frac{m}{n} = \frac{1}{2}$.

Пример 2. Найти вероятность появления верхней грани с цифрой 5 при бросании игральной кости.

Пусть A_5 — данное событие, тогда $m=1$, $n=6$, поэтому $P(A_5) = \frac{1}{6}$.

Замечание. Очевидно, ту же вероятность имеют события A_k — появление верхней грани с цифрой k ($k=1, 2, 3, 4, 6$).

Пример 3. Найти вероятность появления верхней грани с числом очков, делящимся на 3, при бросании игральной кости.

В обозначениях предыдущего примера благоприятствующими данному событию A будут A_3 и A_6, а всего элементарных исходов будет шесть. Поскольку $m=2$, $n=6$, то $P(A) = \frac{2}{6} = \frac{1}{3}$.

Из определения вероятности события следуют ее простейшие свойства.

1. Вероятность достоверного события равна единице. Действительно, для достоверного события все элементарные исходы являются благоприятствующими этому собы-

тию, т. е. $m=n$. Обозначим достоверное событие буквой E, тогда

$$P(E)=\frac{n}{n}=1. \qquad (28.2)$$

2. Вероятность невозможного события равна нулю. В самом деле, для невозможного события нет ни одного элементарного исхода, благоприятствующего этому событию, т. е. $m=0$. Обозначим невозможное событие буквой O, тогда

$$P(O)=\frac{0}{n}=0. \qquad (28.3)$$

3. Вероятность случайного события выражается положительным числом, меньшим единицы. Так как для случайного события $0<m<n$, то $0<\frac{m}{n}<1$, т. е.

$$0<P(A)<1. \qquad (28.4)$$

С л е д с т в и е. Вероятность любого события удовлетворяет неравенствам

$$0\leqslant P(A)\leqslant 1. \qquad (28.5)$$

Классическое определение вероятности предполагает, что число всех элементарных исходов конечно. Но на практике часто встречаются опыты, для которых множество таких исходов бесконечно.

Введем понятие геометрической вероятности. Это понятие необходимо, например, при определении вероятности попадания в область g точки, брошенной в область G, которая содержит g. Когда говорят «в некоторую область брошена точка», имеют в виду, что брошено тело, размерами которого можно пренебречь по сравнению с размерами данных областей (например, поперечным сечением пули по сравнению с площадью мишени, поперечным сечением снаряда по сравнению с площадью участка, на котором находятся поражаемые цели).

Общая задача, приведшая к необходимости расширения понятия вероятности, в плоском случае получает следующую формулировку. На плоскости задана квадрируемая область, т. е. область, имеющая площадь. Обозначим эту область буквой G, а ее площадь S_G. В области G содержится область g площади s_g (рис. 28.1). В область G наудачу бросается точка. Будем считать, что брошенная точка может попасть в некоторую часть области G с вероятностью, пропорциональной площади этой части и не зависящей от ее формы и расположения. Требуется определить вероятность попадания данной точки в область g. Пусть A — попадание брошенной точки в область g, тогда геометрическая вероятность этого события определяется формулой $P(A) = \frac{S_g}{S_G}$.

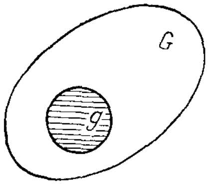

Рис. 28.1

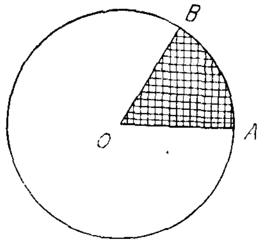

Рис. 28.2

Пример 4. Производится стрельба по мишени, имеющей форму круга и равномерно вращающейся вокруг центра O (рис. 28.2). Попадание в круг — событие достоверное. Сектор OAB, площадь которого равна одной шестой части площади всего круга, окрашен в черный цвет. Найти вероятность попадания в сектор OAB.

В данном случае $S_G = S$, $S_g = \frac{1}{6} S$, где S — площадь рассматриваемого круга, поэтому

$$P(A) = \frac{S_g}{S_G} = \frac{\frac{1}{6} S}{S} = \frac{1}{6}.$$

Аналогично вводится понятие геометрической вероятности при бросании точки в пространственную область G, содержащую область g.

Обозначим меру области g (длину, площадь, объем) через $\operatorname{mes} g$, а меру области G — через $\operatorname{mes} G$, тогда вероятность попадания в область g точки, брошенной в область G, по определению выражается формулой

$$P(A) = \frac{\operatorname{mes} g}{\operatorname{mes} G}, \qquad (28.6)$$

где A (рассматриваемое событие) — попадание точки в область g, которая содержится в области G. Это определение вероятности называется *геометрическим*.

§ 28.3. Действия над событиями. Соотношения между событиями

Суммой, или *объединением*, двух событий называется событие, состоящее в появлении хотя бы одного из них. Сумма двух событий A и B обозначается через $A+B$ или $A\cup B$. Например, если производится два выстрела по мишени и A — попадание при первом выстреле, B — попадание при втором выстреле, то $A+B$ — попадание при одном из выстрелов или при обоих выстрелах. Отметим, что если A и B — несовместные события, то сумма $A+B$ означает появление только одного из них. Аналогично определяется и обозначается сумма n событий — событие, состоящее в появлении одного события, нескольких или всех вместе событий. Рассмотрим n событий A_1, $A_2, \ldots, A_n$, их сумма

$$A_1+A_2+\ldots+A_n, \quad \text{или} \quad A_1\cup A_2\cup\ldots\cup A_n=\bigcup_{i=1}^{n} A_i,$$

означает событие, заключающееся в появлении хотя бы одного из них. В частности, сумма $A+B+C$ трех событий A, B, C означает появление одного (или A, или B, или C), двух (A и B, или A и C, или B и C) или всех трех событий (A, B, C).

Произведением, или *пересечением*, двух событий называется событие, состоящее в одновременном их появлении. Произведение двух событий A и B обозначается через AB, или $A\cap B$. Очевидно, если события A и B несовместны, то их произведение является невозможным событием. Поясним понятие произведения двух событий на следующем примере. Пусть A — выигрыш по лотерее 1, B — выигрыш по лотерее 2, тогда AB — выигрыш по двум лотереям. Аналогично определяется и обозначается произведение в случае большего числа событий. Произведение n событий

$$A_1A_2A_3\ldots A_n, \text{ или } A_1\cap A_2\cap\ldots\cap A_n=\bigcap_{i=1}^{n} A_i,$$

означает событие, состоящее в появлении всех событий A_1, A_2, $\ldots$, A_n. В частности, произведением ABC трех событий A, B, C называется событие, заключающееся в одновременном появлении их всех. Например, если A, B, C — появление «цифры» соответственно при первом, втором и третьем бросании монеты, то ABC — появление «цифры» во всех трех испытаниях.

Введенные понятия можно проиллюстрировать геометрически. На рис. 28.3 показаны сумма и произведение двух событий A (попадание случайной точки в область A) и B (попадание в область B).

Понятия суммы и произведения событий распространяются и на бесконечные последовательности событий.

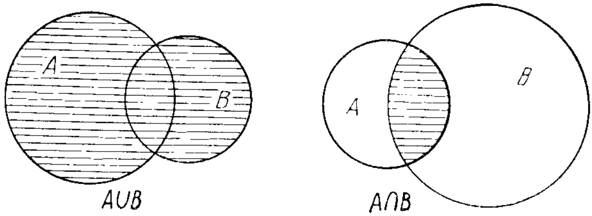

Рис. 28.3

В этих случаях соответственно применяют, например, обозначения:

$$A_1 \cup A_2 \cup \ldots \cup A_n \cup \ldots = \bigcup_{i=1}^{\infty} A_i;$$

$$A_1 \cap A_2 \cap \ldots \cap A_n \cap \ldots = \bigcap_{i=1}^{\infty} A_i.$$

Разностью событий A и B называется событие C, которое означает, что наступает событие A и не происходит событие B. Разность событий A и B обозначается так: $A - B$, или $A \setminus B$. Например, в опыте с бросанием двух игральных костей обозначим события: A — «выпадает четная сумма очков» (на двух верхних гранях), B — «на каждой верхней грани выпадает четное число очков», C — «на каждой верхней грани выпадает нечетное число очков», тогда $A - B = C$ и $A - C = B$.

Операции объединения и пересечения событий обладают некоторыми свойствами, аналогичными свойствам сложения и умножения чисел. Эти операции коммутативны: $A \cup B = B \cup A$, $A \cap B = B \cap A$; ассоциативны: $(A \cup B) \cup C = A \cup (B \cup C) = (A \cup C) \cup B = A \cup B \cup C$, $(A \cap B) \cap C = A \cap (B \cap C) = (A \cap C) \cap B$ и дистрибутивны: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$. Указанные свойства следуют из определения действий объединения и пересечения событий. Так, $(A \cup B) \cap C$ представляет собой совместное появление события C с событием A или с событием B или с A и B вместе. Событие $(A \cap C) \cup (B \cap C)$ тоже состоит в появлении или C вместе с A, или C вместе с B, или C вместе с $A \cap B$. Отметим, что не все законы сложения и умножения чисел справедливы для объединения и пересечения событий. Например, события $A \cup A$ и $A \cap A$ совпадают с A, т. е. $A \cup A = A$, $A \cap A = A$ для любого A.

Если E — достоверное, O — невозможное события, то, очевидно,

$$A\bar{A} = O,\ A \cup \bar{A} = E,\ A \cup O = A,\ AO = O,$$
$$A \cup E = E,\ AE = A,$$

где A — любое событие; $\bar{A}$ — событие, противоположное A (непоявление события A).

Из свойств операций пересечения и объединения следует, что для любых двух событий A и B имеем $A = AE = A(B \cup \bar{B}) = AB \cup A\bar{B}$, т. е.

$$A = AB \cup A\bar{B}. \tag{28.7}$$

Формула (28.7) дает разложение любого события A на два непересекающихся (несовместных) события.

Если событие A обязательно произойдет при появлении некоторого другого события B, то говорят, что событие B представляет собой частный *случай события* A, и пишут $B \subset A$, или $A \supset B$ (говорят также, что B влечет A). Если $B \subset A$ и $A \subset B$, т. е. события A и B в данном опыте могут появиться или не появиться вместе, то их называют *равносильными*, или *эквивалентными*, и пишут $A = B$.

Если $B \subset A$, то $AB = B$ и формула (28.7) принимает вид

$$A = B \cup A\bar{B}. \tag{28.8}$$

Приведем пример, иллюстрирующий действия над событиями и соотношения между ними. Бросается игральная кость; если A — «выпадение шести очков», B — «выпадение трех очков», C — «выпадение четного числа очков», D — «выпадение числа очков, кратного трем», то $A \subset C$, $A \subset D$, $B \subset D$, $A + B = D$, $CD = A$.

§ 28.4. Закон сложения вероятностей

Теорема 28.1. Вероятность суммы двух событий равна сумме вероятностей этих событий без вероятности их совместного появления, т. е.

$$P(A+B)=P(A)+P(B)-P(AB). \qquad (28.9)$$

Доказательство. Пусть n — общее число элементарных исходов опыта, m_1 — число исходов, благоприятствующих событию A, m_2 — число исходов, благоприятствующих событию B, l — число исходов, благоприятствующих произведению AB этих событий, тогда

$$P(A)=\frac{m_1}{n}, \quad P(B)=\frac{m_2}{n}, \quad P(AB)=\frac{l}{n}.$$

Для события $A+B$ будет $m=m_1+m_2-l$ благоприятствующих элементарных исходов. Действительно, складывая число исходов m_1 и m_2, благоприятствующих соответственно событиям A и B, мы дважды считаем исходы, благоприятствующие событию AB.

Следовательно, при подсчете числа исходов, благоприятствующих событию $A+B$, значение l необходимо исключить. Таким образом,

$$P(A+B)=\frac{m_1+m_2-l}{n}=\frac{m_1}{n}+\frac{m_2}{n}-\frac{l}{n}=P(A)+ \\ +P(B)-P(AB).$$

Следствие 1. Поскольку $P(AB)>0$, то из формулы (28.9) получаем, что $P(A+B)\leqslant P(A)+P(B)$, т. е. вероятность суммы двух событий не превосходит суммы вероятностей этих событий.

Отметим, что аналогичное соотношение справедливо для нескольких событий:

$$P(A_1+A_2+\ldots+A_n)\leqslant P(A_1)+P(A_2)+\ldots+P(A_n).$$

Следствие 2. Вероятность суммы двух несовместных событий равна сумме вероятностей этих событий:

$$P(A+B)=P(A)+P(B). \qquad (28.10)$$

Действительно, в этом случае $AB=O$, где O — невозможное событие. Так как $P(O)=0$, то формула (28.9) принимает вид (28.10).

Отметим, что формулу (28.10) можно обобщить на n

слагаемых. Если $A_1, A_2, \ldots, A_n$ — попарно-несовместные события, то

$$P(A_1+A_2+\ldots+A_n)=P(A_1)+P(A_2)+\ldots+P(A_n). \quad (28.11)$$

Действительно, для $n=2$ равенство (28.11) доказано. Покажем, что если формула (28.11) верна при $n=k-1$, то она верна и при $n=k$:

$$P(A_1+A_2+\ldots+A_{k-1}+A_k)=P[(A_1+A_2+\ldots+A_{k-1})+A_k]=$$
$$=P(A_1+A_2+\ldots+A_{k-1})+P(A_k)=$$
$$=P(A_1)+P(A_2)+\ldots+P(A_{k-1})+P(A_k).$$

Сумма вероятностей событий $A_1, A_2, \ldots, A_n$, образующих полную группу, равна единице, т. е.

$$P(A_1)+P(A_2)+\ldots+P(A_n)=1. \quad (28.12)$$

В самом деле, если $A_1, A_2, \ldots, A_n$ образуют полную группу событий (см. § 28.1), то их сумма $A_1+A_2+\ldots$ $\ldots+A_n$ — достоверное событие и

$$P(A_1+A_2+\ldots+A_n)=1. \quad (28.13)$$

События, образующие полную группу, попарно-несовместны, поэтому для них справедливо равенство (28.11). Из равенств (28.11) и (28.13) следует равенство (28.12).

Сумма вероятностей противоположных событий равна единице, т. е.

$$P(A)+P(\bar{A})=1. \quad (28.14)$$

Пусть A и $\bar{A}$ — противоположные события (см. § 28.1). Поскольку A и $\bar{A}$ образуют полную группу, то из формулы (28.12) следует формула (28.14).

Замечание. Если ввести обозначения $P(A)=p$, $P(\bar{A})=q$, то формула (28.14) примет вид

$$p+q=1, \text{ или } p=1-q. \quad (28.15)$$

§ 28.5. Частота события и ее свойства. Статистическое определение вероятности

Классическое определение вероятности предполагает, что все элементарные исходы равновозможны. О равновозможности исходов опыта заключают в силу соображений симметрии (как в случае монеты или игральной кости). Задачи, в которых можно исходить из соображений симметрии, на практике встречаются редко. Во многих случаях трудно указать основания, позволяющие считать, что все элементарные исходы являются равновозможными. В связи с этим появилась необходимость введения еще одного определения вероятности, так называемого статистического. Чтобы дать это определение, предварительно введем понятие относительной частоты события.

Относительной частотой события, или *частотой,* называется отношение числа опытов, в которых появилось это событие, к числу всех произведенных опытов. Обозначим через $W(A)$ относительную частоту события A, тогда по определению

$$W(A) = \frac{m}{n}, \tag{28.16}$$

где m — число опытов, в которых появилось событие A; n — число всех произведенных опытов.

Условной частотой называется частота одного события, вычисленная при условии, что другое событие наступило. Обозначения: $W(A/B)$ — частота события A при условии, что событие B произошло, $W(B/A)$ — частота события B при условии, что событие A наступило. Частота события обладает следующими свойствами: 1) частота случайного события есть число, заключенное между нулем и единицей: $0<W(A)<1$; 2) частота достоверного события E равна единице: $W(E)=1$; 3) частота невозможного события 0 равна нулю: $W(O)=0$; 4) частота суммы двух несовместных событий равна сумме частот этих событий: $W(A+B)=W(A)+W(B)$; 5) частота произведения двух событий равна произведению частоты одного на условную частоту другого: $W(AB)=$ $=W(A)W(B/A)$, $W(AB)=W(B)W(A/B)$.

Первые три свойства очевидны, они следуют из определения частоты (см. формулу (28.16)) и того, что $0<m<n$, т. е. $0<\frac{m}{n}<1$ для случайного события, $m=n$ для достоверного события, $m=0$ для невозможного события. Докажем остальные свойства. Пусть произведено n испытаний, в которых событие A появилось m_1 раз и событие B — m_2 раз, тогда $W(A)=m_1/n$, $W(B)=m_2/n$. Так как A и B несовместны, то событие $A+B$ появилось m_1+m_2 раз, т. е. $W(A+B)=\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n}$.

Следовательно, $W(A+B)=W(A)+W(B)$; четвертое свойство доказано.

Пусть в результате n испытаний событие A появилось m_1 раз, событие B — m_2 раз, причем A и B появились вместе m_3 раз. Поскольку

$$W(A)=\frac{m_1}{n},\quad W(B)=\frac{m_2}{n},\quad W(AB)=\frac{m_3}{n},$$

$$W(B/A)=\frac{m_3}{m_1},\quad W(A/B)=\frac{m_3}{m_2},$$

то

$$W(AB)=\frac{m_3}{n}=\frac{m_3}{m_1}\,\frac{m_1}{n}=W(B/A)\,W(A)=W(A)\,W(B/A),$$

$$W(AB)=\frac{m_3}{n}=\frac{m_3}{m_2}\,\frac{m_2}{n}=W(A/B)\,W(B)=W(B)\,W(A/B).$$

Приведем примеры вычисления относительной частоты некоторых событий.

Пример 1. В результате 25 выстрелов по мишени получено 20 попаданий. Какова относительная частота попаданий?

Так как $m=20$, $n=25$, то $W(A)=\dfrac{20}{25}=0{,}8$.

Пример 2. При многократных бросаниях монеты подсчитывалось число появлений герба, находилась относительная частота этого события (см. [6]):

n	m	$W(A)=\dfrac{m}{n}$
4040	2048	0,5069
12000	6019	0,5016
24000	12012	0,5005

Приведенные данные свидетельствуют о том, что относительная частота появления герба при достаточно больших n мало отличается от его вероятности $P(A)=0{,}5$ (см. § 28.2, пример 1).

Наблюдения позволили установить, что относительная частота события обладает свойством статистической устойчивости: в различных сериях многочисленных испытаний (в каждом из которых может появиться или не появиться это событие) она принимает значения, достаточно близкие к некоторой постоянной. Эту постоянную, являющуюся объективной числовой характеристикой явления, называют вероятностью данного события.

Вероятностью события называется число, около которого группируются значения относительной частоты данного события в различных сериях большого числа испытаний. Это определение вероятности называется *статистическим.* В случае статистического определения вероятность обладает следующими свойствами: 1) вероятность достоверного события равна единице; 2) вероятность невозможного события равна нулю; 3) вероятность суммы двух несовместных событий равна сумме их вероятностей. Как было показано выше, частота события обладает указанными свойствами.

§ 28.6. Аксиоматическое определение вероятности

Теорию вероятностей, как и любую другую математическую науку, можно строить аксиоматическим методом. Аксиомы теории вероятностей вводятся так, чтобы вероятность события обладала основными свойствами частоты.

Как уже можно было заметить (см. § 28.3), операции над событиями тождественны операциям над множествами. Аналогия между событиями и множествами объясняется тем, что каждое событие связано с определенным множеством исходов опыта; оно происходит при появлении одного из исходов, принадлежащих этому множеству, и не происходит при появлении одного из исходов, не принадлежащих этому множеству. Так, при бросании игральной кости событие «выпало четное число очков» связано с множеством исходов A_2, A_4, A_6. В примере, поясняющем понятие геометрической вероятности (см. § 28.2), исходом каждого опыта является попадание в определенную точку, а каждое событие представляет попадание на определенное множество точек (область g).

Пространством элементарных событий называют произвольное множество Ω, а его элементы ω — *элементарными событиями.* Эти понятия являются первоначальными. В реальных опытах элементарным событиям соответствуют взаимно исключающие итоги опыта.

Событиями будем называть подмножества множества Ω и обозначать латинскими буквами A, B, C и т. д. Не исключаются случаи, когда такое подмножество содержит лишь один элемент, совпадает со всем множеством Ω или является пустым. Пустое множество $\varnothing$ называют невозможным событием, а множество Ω — достоверным событием. Случайным событием назовем любое собственное (т. е. отличное от $\varnothing$ и Ω) подмножество множества Ω. Событие $\bar{A}=\Omega-A$ называется *противоположным* событию A; событие $\bar{A}$ означает, что A не произошло. События A и B называются *несовместными*, если $AB=\varnothing$.

Аксиомы теории вероятностей можно ввести следующим образом. Пусть Ω — произвольное пространство элементарных событий, L — некоторая система случайных событий. Система L случайных событий называется *алгеброй событий*, если выполнены условия: 1) $\Omega \in L$; 2) если $A \in L$, $B \in L$, то $AB \in L$, $A+B \in L$, $A \setminus B \in L$. Другими словами, L — алгебра событий, если вместе с любыми двумя событиями она содержит их сумму, произведение и разность, а также множество Ω. Из этих условий следует, что пустое множество $\varnothing$ также принадлежит L. Действительно, поскольку $\varnothing=\Omega-\Omega$ и $\Omega \in L$, то и $\varnothing \in L$. Разъясним введенные понятия на примере опыта с бросанием игральной кости. В этом опыте пространство элементарных событий есть множество $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, где ω_k — элементарный исход опыта, заключающийся в выпадении k очков. В данном опыте можно также говорить, например, о случайном событии, состоящем в том, что выпало нечетное число очков. Такое событие происходит только в том случае, когда осуществилось одно из трех элементарных событий ω_1, ω_3, ω_5; ему соответствует подмножество $A=\{\omega_1, \omega_3, \omega_5\}$ множества Ω. Для этого множества Ω можно выписать все $2^6=64$ события алгебры L, состоящей из подмножеств множества Ω, которые содержат по одному, два, три, четыре, пять, шесть элементарных событий и пустое множество:

$$\varnothing, \{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_6\}; \{\omega_1, \omega_2\}, \{\omega_1, \omega_3\}, \ldots$$
$$\ldots, \{\omega_5, \omega_6\}, \{\omega_1, \omega_2, \omega_3\}, \ldots, \Omega.$$

Приведем еще один пример алгебры событий. Пусть дан единичный квадрат $\Omega=\{(u, v): 0 \leqslant u \leqslant 1, 0 \leqslant v \leqslant 1\}$ в некоторой плоскости. Рассмотрим систему L всех квадрируемых множеств квадрата Ω, т. е. его фигур, имеющих площадь. Можно доказать, что объединение, пересечение и разность квадрируемых фигур является квадрируемой фигурой. Следовательно, система L квадрируемых подмножеств квадрата Ω образует алгебру событий.

Алгебра событий L называется *σ-алгеброй, или борелевской алгеброй*, если из того, что $A_n \in L$ $(n=1, 2, \ldots)$, следует

$$\bigcup_{n=1}^{\infty} A_n \in L, \quad \bigcap_{n=1}^{\infty} A_n \in L.$$

Введем понятие вероятности события. Числовая функция $P(A)$, определенная на алгебре событий L, называется *вероятностью*, если выполнены следующие аксиомы.

1. Каждому событию из L ставится в соответствие неотрицательное число $P(A)$ — его вероятность, т. е. $P(A) \geqslant 0$ для любого $A \in L$.

2. Вероятность достоверного события равна единице: $P(\Omega)=1$.

3. Вероятность суммы двух несовместных событий равна сумме вероятностей этих событий, т. е. $P(A+B) = P(A)+P(B)$, когда A и B несовместны.

Очевидно, эти аксиомы аналогичны соответствующим свойствам частот (см. § 28.5).

Для решения задач, связанных с бесконечными последовательностями событий, аксиомы 1—3 требуется дополнить еще одной аксиомой (аксиомой непрерывности).

4. Для любой убывающей последовательности $A_1 \supset A_2 \supset \ldots \supset A_n \ldots$ событий из L такой, что $\bigcap_{n=1}^{\infty} A_n = \varnothing$, справедливо равенство $\lim_{n\to\infty} P(A_n) = 0$.

Тройка (Ω, L, P), *в которой* L *является σ-алгеброй и* P удовлетворяет аксиомам 1—4, называется *вероятностным пространством*. Таким образом, математической моделью любого случайного явления в современной теории вероятностей служит вероятностное пространство. Приведем примеры вероятностных пространств.

1. Пусть $\Omega=\{\omega_1, \omega_2, \ldots, \omega_n\}$, алгебра событий L состоит из всех подмножеств $A=\{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$, $1 \leqslant i_1 \leqslant i_2 \leqslant \ldots \leqslant i_m (m=1, 2, \ldots, n)$, множества Ω и $P(A) = \frac{m}{n}$. Определенная этим равенством функция $P(A)$ удовлетворяет введенным аксиомам. Получено классическое определение вероятности. Из формулы для $P(A)$ следует, что все элементарные события равновероятны.

2. Пусть Ω — любая квадрируемая область плоскости. Рассмотрим систему L квадрируемых подмножеств A множества Ω. Для любого $A \in L$ положим $P(A) = S_A/S$, где S_A — площадь области A; S — площадь области Ω. Получено (Ω, L, P) — вероятностное пространство и геометрическое определение вероятности.

3. Пусть $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n \ldots\}$ — счетное множество, L — система всех его подмножеств. Очевидно, L является σ-алгеброй. Пусть p_n, $n = 1, 2, \ldots,$ — последовательность неотрицательных чисел, удовлетворяющая условию $\sum_{n=1}^{\infty} p_n = 1$. Для любого $A \in L$ положим $P(A) = \Sigma p_n$, где суммирование распространено на все n, для которых $\omega_n \in A$. Если эта сумма будет конечной, то получим определенное число; если это числовой ряд, то он сходится в силу равенства $\sum_{n=1}^{\infty} p_n = 1$. Можно показать, что функция $P(A)$ удовлетворяет аксиомам 1—4. Полученное вероятностное пространство (Ω, L, P) называется *дискретным вероятностным пространством.*

§ 28.7. Следствия из аксиом вероятности

Выведем простейшие следствия из аксиом вероятности. Поскольку $A + \bar{A} = \Omega$, то из аксиом 2—3 следует, что $P(A) + P(\bar{A}) = 1$, т. е. сумма вероятностей противоположных событий равна 1. Из аксиомы 2 и равенства $P(A) + P(\bar{A}) = 1$, положив в нем $A = \Omega$, получим $P(\varnothing) = 0$, т. е. вероятность невозможного события равна нулю.

Если события $A_1, A_2, \ldots, A_n$ попарно-несовместны (т. е. $A_i A_j = \varnothing$ при любых $i \neq j$, $i, j = 1, 2, \ldots, n$), то

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n). \quad (28.17)$$

Эта формула следует по индукции из аксиомы 3.

Докажем, что для любых событий A и B верна теорема сложения вероятностей:

$$P(A+B) = P(A) + P(B) - P(AB), \quad (28.18)$$

т. е. вероятность суммы двух событий равна сумме вероятностей этих событий без вероятности их совместного появления.

Действительно, представив события $A+B$ и B в виде соответствующих сумм несовместных событий $A+B=$

$=A+B\bar{A}$, $B=B\bar{A}+BA$ (см. формулу (28.7)) и применив аксиому 3, найдем

$$P(A+B)=P(A)+P(B\bar{A}),\ P(B)=P(B\bar{A})+P(AB).$$

Определив $P(\bar{B}A)$ из второго равенства и подставив в первое, получим формулу (28.18).

Из формулы (28.18) и неравенства $P(AB)\geqslant 0$ для любых A и B находим, что

$$P(A+B)\leqslant P(A)+P(B). \qquad (28.19)$$

Отсюда для любых $A_1, A_2, \ldots, A_n$ по индукции следует неравенство

$$P(A_1+A_2+\ldots+A_n)\leqslant P(A_1)+P(A_2)+\ldots+P(A_n).$$

Если событие A влечет за собой событие $B\,(A\subset B)$, то

$$P(A)\leqslant P(B). \qquad (28.20)$$

В этом случае $B=A+\bar{A}B$ (см. формулу (28.8)), поэтому

$$P(B)=P(A)+P(\bar{A}B)\geqslant P(A).$$

Поскольку $\varnothing\subset A\subset\Omega$, $P(\varnothing)=0$, $P(\Omega)=1$, то из (28.20) следует, что $0\leqslant P(A)\leqslant 1$, т. е. вероятность любого события выражается неотрицательным числом, не превосходящим единицы, другими словами, область значений функции $P(A)$ принадлежит отрезку $[0, 1]$.

Замечание. В случае классического понятия вероятности функция $P(A)$ определяется формулой (28.1), в случае геометрического — формулой (28.6), в случае статистического — значение вероятности можно приближенно вычислить по формуле (28.16).

Можно показать, что аксиомы 3 и 4 эквивалентны следующему свойству σ-аддитивности (или счетной аддитивности) вероятности: если события $A_1, A_2, \ldots, A_n, \ldots$ — попарно-несовместны и $A=\bigcup\limits_{n=1}^{\infty}A_n\in L$, то

$$P(A)=\sum_{n=1}^{\infty}P(A_n). \qquad (28.21)$$

Отметим без доказательства еще два свойства, связанных с аксиомой 4 (аксиомой непрерывности).

1. Если $A_1\subset A_2\subset\ldots\subset A_n\subset A_{n+1}\subset\ldots$ и $A=\bigcup\limits_{n=1}^{\infty}A_n$,

то

$$P(A) = \lim_{n \to \infty} P(A_n). \tag{28.22}$$

2. Если $A_1 \supset A_2 \supset \ldots \supset A_n \supset A_{n+1} \supset \ldots$ и $A = \bigcap_{n=1}^{\infty} A_n$, то также выполняется равенство (28.22).

§ 28.8. Условная вероятность. Теорема умножения вероятностей

В ряде случаев приходится рассматривать вероятности событий при дополнительном условии, что произошло некоторое другое событие, имеющее вероятность, отличную от нуля. Такие вероятности называют условными вероятностями.

Вероятность события B при условии, что произошло событие A, называется *условной вероятностью* события B и обозначается так: $P(B/A)$, или $P_A(B)$.

Понятие условной вероятности разъясним на следующем примере. Пусть A — событие, состоящее в извлечении белого шара из урны, содержащей n шаров, в том числе m белых, $n-m$ черных; B — событие, состоящее в извлечении белого шара из той же урны после того, как из нее уже извлечен один шар. Очевидно, если первый извлеченный шар был белым, т. е. если произошло событие A, то в урне после первого извлечения останется $m-1$ белых и $n-m$ черных шаров, поэтому вероятность события B будет равна $\frac{m-1}{n-1}$. Если же первый извлеченный шар был черным (произошло событие $\bar{A}$), то в урне останется m белых и $n-m-1$ черных шаров; искомая вероятность окажется равной $\frac{m}{n-1}$. Следовательно, вероятность события B меняется в зависимости от того, происходит или не происходит событие A, т. е. вероятность события B может принимать два различных значения $\left(\frac{m-1}{n-1}, \frac{m}{n-1}\right)$.

Таким образом,

$$P(B/A) = \frac{m-1}{n-1}, \quad P(B/\bar{A}) = \frac{m}{n-1}.$$

Замечание. Условная вероятность $P(B/A)$ какого-либо события B при определенном условии A может быть или больше, или меньше безусловной вероятности $P(B)$ этого события. Так, в рассмотренном выше примере

$$P(B)=\frac{m}{n},\quad P(B/A)=\frac{m-1}{n-1}<\frac{m}{n}=P(B),$$

$$P(B/\bar{A})=\frac{m}{n-1}>\frac{m}{n}=P(B),$$

где $P(B)$ — вероятность извлечения белого шара из урны, содержащей n шаров, среди которых m белых и $n-m$ черных.

В случае классического определения условные вероятности вычисляются аналогично тому, как вычисляются безусловные вероятности. Пусть среди полной группы элементарных исходов $A_1, A_2, \ldots, A_n$ событию A благоприятствуют m исходов, событию B — k исходов, событию AB — l исходов ($l\leqslant m$, $l\leqslant k$). Если событие B произошло, то это означает, что наступило одно из событий A_i, благоприятствующих B, при этом будет l и только l событий A_i, благоприятствующих AB, поэтому

$$P(A/B)=\frac{l}{k}=\frac{\frac{l}{n}}{\frac{k}{n}}=\frac{P(AB)}{P(B)},$$

$$P(B/A)=\frac{l}{m}=\frac{\frac{l}{n}}{\frac{m}{n}}=\frac{P(AB)}{P(A)}.$$

Итак, получены следующие формулы:

$$P(A/B)=\frac{P(AB)}{P(B)},\quad P(B/A)=\frac{P(AB)}{P(A)}, \tag{28.23}$$

где $P(A)\neq 0$, $P(B)\neq 0$.

При аксиоматическом введении понятия вероятности условные вероятности определяются соответственно формулами

$$P(B/A)=\frac{P(AB)}{P(A)},\quad P(A/B)=\frac{P(AB)}{P(B)}, \tag{28.24}$$

где $P(A)\neq 0$, $P(B)\neq 0$.

Теорема 28.2. Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое событие произошло.

Доказательство. Из формул (28.24) получаем равенства

$$P(AB)=P(A)P(B/A),\quad P(AB)=P(B)P(A/B), \tag{28.25}$$

которые доказывают теорему умножения вероятностей.

Отметим, что теорема верна и в случае, когда одно из событий является невозможным. Если A — невозможное событие, то $P(A)=0$, $P(A/B)=0$, $P(AB)=0$.

§ 28.9. Независимость событий

Введем понятие независимости одного события от другого. Событие B не зависит от события A, если

$$P(B/A)=P(B), \tag{28.26}$$

т. е. вероятность события B не зависит от того, произошло ли событие A.

Из формул (28.25) следует, что $P(A)P(B/A)=$ $=P(B)P(A/B)$. Учитывая (28.26), получаем $P(A)P(B)=$ $=P(B)P(A/B)$. Отсюда следует, что $P(A/B)=P(A)$, так как $P(B)\neq 0$. Это равенство означает, что событие A не зависит от события B. Таким образом, свойство независимости событий является взаимным.

Если события A и B независимы, то независимы также события $\bar{A}$ и B, A и $\bar{B}$, $\bar{A}$ и $\bar{B}$. Докажем независимость событий $\bar{A}$ и B. События A/B и $\bar{A}/B$ противоположны, поэтому $P(A/B)+P(\bar{A}/B)=1$. Поскольку события A и B независимы, т. е. $P(A/B)=P(A)$ и $P(B/A)=P(B)$, то

$$P(A)+P(\bar{A}/B)=1,$$
$$P(\bar{A}/B)=1-P(A)=P(\bar{A}),\ P(\bar{A}/B)=P(\bar{A}).$$

Последнее равенство означает, что $\bar{A}$ не зависит от B. Но в этом случае и B не зависит от $\bar{A}$, т. е. $\bar{A}$ и B независимы. Доказательства независимости событий A и $\bar{B}$, $\bar{A}$ и $\bar{B}$ предоставляются читателю.

Если события A и B независимы, то формулы (28.25) принимают вид

$$P(AB)=P(A)P(B), \tag{28.27}$$

т. е. вероятность произведения независимых событий равна произведению их вероятностей. Формула (28.27) выражает теорему умножения вероятностей двух независимых событий.

Теорему умножения вероятностей можно обобщить на произведение любого числа событий:

$$P(A_1A_2\ldots A_n)=$$
$$=P(A_1)P(A_2/A_1)P(A_3/A_1A_2)\ldots P(A_n/A_1A_2\ldots A_{n-1}). \tag{28.28}$$

Формула (28.28) означает следующее: вероятность произведения n событий равна произведению вероятности одного из них на условные вероятности всех остальных, вычисленные в предположении, что все предыдущие события наступили.

В частности, для трех событий A, B, C получаем

$$P(ABC)=P(A)P(B/A)P(C/AB). \qquad (28.29)$$

Пример. В урне имеется 5 красных, 7 синих и 3 белых шара. Каждое испытание состоит в том, что из урны берут наудачу один шар и не возвращают обратно. Найти вероятность того, что при первом испытании будет взят красный шар (событие A), при втором — синий (событие B), при третьем — белый (событие C).

Поскольку

$$P(A)=\frac{5}{15}=\frac{1}{3},\quad P(B/A)=\frac{7}{14}=\frac{1}{2},\quad P(C/AB)=\frac{3}{13},$$

то по формуле (28.29) находим искомую вероятность:

$$P(ABC)=\frac{1}{3}\,\frac{1}{2}\,\frac{3}{13}=\frac{1}{26}.$$

Введем понятие независимости n событий. События $A_1, A_2, \ldots, A_n$ называются *независимыми в совокупности,* или *независимыми,* если каждое из них и произведение любого числа k остальных ($k=1, 2, \ldots, n-1$) являются независимыми. Например, если события A, B, C независимы в совокупности, то это означает, что независимы A и B, B и C, A и C, B и AC, C и AB, A и BC.

Замечание 1. Из попарной независимости событий не следует их независимость в совокупности. Поясним это на следующем примере. Грани однородного тетраэдра окрашены: первая — в красный цвет, вторая — в синий, третья — в зеленый, четвертая — во все эти три цвета. При бросании тетраэдр упадет на одну грань, окрашенную в некоторый цвет: красный (событие A), синий (событие B), зеленый (событие C), во все три цвета (событие ABC). Являются ли независимыми в совокупности события A, B, C?

Вычислим соответствующие вероятности:

$$P(A)=P(B)=P(C)=\frac{2}{4}=\frac{1}{2},$$

$$P(A/B)=P(A/C)=P(B/A)=P(B/C)=P(C/A)=P(C/B)=\frac{1}{2}.$$

Эти равенства означают, что события A, B, C попарно-независимы. Поскольку $P(A/BC)=1$ и, следовательно, $P(A/BC)\neq P(A)$, то события A, B, C не являются независимыми в совокупности.

Замечание 2. Если события $A_1, A_2, \ldots, A_n$ независимы в совокупности, то противоположные им события $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n$ также независимы в совокупности.

Если события $A_1, A_2, \ldots, A_n$ независимы в совокупности, то

$$P(A_1A_2\ldots A_n)=P(A_1)P(A_2)\ldots P(A_n). \qquad (28.30)$$

Эта формула следует из определения событий, независимых в совокупности, и формулы (28.28). Формула (28.30) означает следующее: вероятность произведения событий, независимых в совокупности, равна произведению вероятностей этих событий. Для трех событий, независимых в совокупности, формула (28.30) принимает вид

$$P(ABC)=P(A)P(B)P(C). \qquad (28.31)$$

Пример. В каждом из трех ящиков имеется по 20 деталей, при этом в первом ящике — 15, во втором — 16, в третьем — 18 стандартных деталей. Из каждого ящика берут по одной детали. Найти вероятность того, что все три извлеченные детали окажутся стандартными.

Введем обозначения: извлечение стандартной детали из первого ящика — событие A, из второго — событие B, из третьего — событие C, тогда

$$P(A)=\frac{15}{20}=\frac{3}{4}, \quad P(B)=\frac{16}{20}=\frac{4}{5}, \quad P(C)=\frac{18}{20}=\frac{9}{10}.$$

Поскольку события A, B, C независимы в совокупности, то по формуле (28.31) находим $P(ABC)=\frac{3}{4}\frac{4}{5}\frac{9}{10}=0{,}54$.

§ 28.10. Вероятность появления хотя бы одного события

Рассмотрим вопрос о том, как найти вероятность того, что наступит хотя бы одно из n независимых событий, которые могут появиться в результате некоторого испытания. Пусть $A_1, A_2, \ldots, A_n$ — независимые события, $\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_n$ — противоположные им события, также независимые. Обозначим через A событие, состоящее в появлении хотя бы одного из событий $A_1, A_2, \ldots, A_n$. Так как события A и $\bar{A}_1\bar{A}_2\ldots\bar{A}_n$ противоположны ($\bar{A}_1\bar{A}_2\ldots\bar{A}_n$ означает, что ни одно из событий $A_1, A_2, \ldots, A_n$ не наступило), то сумма их вероятностей равна единице, т. е. $P(A)+P(\bar{A}_1\bar{A}_2\ldots\bar{A}_n)=1$, откуда $P(A)=1-P(\bar{A}_1\bar{A}_2\ldots\bar{A}_n)$. Применяя формулу (28.30), находим $P(A)=1-P(\bar{A}_1)P(\bar{A}_2)\ldots P(\bar{A}_n)$. Эту формулу можно записать в виде

$$P(A)=1-q_1q_2\ldots q_n, \qquad (28.32)$$

где $q_k=P(\bar{A}_k)$ $(k=1, 2, \ldots, n)$.

Замечание. Если все независимые события A_1, A_2, ..., A_n имеют одну и ту же вероятность p, то вероятность появления хотя бы одного из них определяется формулой $P(A)=1-q^n$.

Пример. Вероятность попадания в цель первым орудием $p_1=0{,}9$, вторым — $p_2=0{,}85$, третьим — $p_3=0{,}8$. Какова вероятность хотя бы одного попадания при одном залпе из трех орудий?

События A_1 (попадание в цель первым орудием), A_2 (вторым орудием), A_3 (третьим орудием) независимы, ибо вероятность попадания в цель при стрельбе любым из трех орудий не зависит от результатов стрельбы другими орудиями. Поскольку $p_1=P(A_1)=0{,}9$, $p_2=P(A_2)=0{,}85$, $p_3=P(A_3)=0{,}8$; $q_1=1-p_1=0{,}1$, $q_2=1-p_2=0{,}15$, $q_3=1-p_3=0{,}2$, то по формуле (28.32) получаем вероятность события A — хотя бы одного попадания при одном залпе из трех орудий $P(A)=1-0{,}1\cdot 0{,}15\cdot 0{,}2=0{,}997$.

§ 28.11. Формула полной вероятности

Пусть A — произвольное событие, события H_1, H_2, ..., H_n попарно-несовместны и $A\subset H_1+H_2+\ldots+H_n$. Вероятности событий H_i $(i=1, 2, \ldots, n)$ известны, причем $P(H_i)\neq 0$; известны также условные вероятности $P(A/H_i)$. Требуется найти вероятность события A.

Событие A можно представить в виде суммы попарно-несовместных событий $A=H_1A+H_2A+\ldots+H_nA$. На основании следствия из аксиомы 3 (см. формулу (28.17)) имеем

$$P(A)=P(H_1A)+P(H_2A)+\ldots+P(H_nA).$$

Применяя теорему умножения вероятностей (см. формулу (28.25)), находим

$$P(H_1A)=P(H_1)P(A/H_1),$$
$$P(H_2A)=P(H_2)P(A/H_2),\ \ldots,\ P(H_nA)=P(H_n)P(A/H_n).$$

Подставив эти выражения в предыдущее равенство, получим

$$P(A)=P(H_1)P(A/H_1)+$$
$$+P(H_2)P(A/H_2)+\ldots+P(H_n)P(A/H_n),$$

или

$$P(A)=\sum_{k=1}^{n} P(H_k)\,P(A/H_k). \qquad (28.33)$$

Формула (28.33) называется *формулой полной вероятности*. Отметим, что события H_1, H_2, ..., H_n иногда называют гипотезами.

Пример. Имеется 5 урн с цветными шарами: две урны состава H_1 — по 3 белых и 4 черных шара, одна урна состава H_2 — 10 белых шаров, две урны состава H_3 — по 2 белых и 5 черных шаров. Наудачу выбирается урна и из нее наудачу извлекается один шар. Какова вероятность извлечь при этом черный шар (событие A)?

Событие A можно представить в виде суммы попарно-несовместных событий $A=H_1A+H_2A+H_3A$ и применить формулу (28.33), которая в данном случае принимает вид

$$P(A)=P(H_1)P(A/H_1)+P(H_2)P(A/H_2)+P(H_3)P(A/H_3).$$

Вычислим вероятности событий, входящих в эту формулу. Так как всего урн 5, две из них состава H_1, то $P(H_1)=2/5$; аналогично находим $P(H_2)=1/5$, $P(H_3)=2/5$. Поскольку в каждой урне состава H_1 всего 7 шаров, из них 4 черных, то $P(A/H_1)=4/7$; аналогично получаем $P(A/H_2)=0$, $P(A/H_3)=5/7$.

Следовательно,

$$P(A)=\frac{2}{5}\,\frac{4}{7}+\frac{1}{5}\,0+\frac{2}{5}\,\frac{5}{7}=\frac{8}{35}+\frac{2}{7}=\frac{18}{35}.$$

§ 28.12. Формулы Байеса

Пусть $H_1, H_2, \ldots, H_n$ — попарно-несовместные события, вероятности которых $P(H_i)\neq 0$ $(i=1, 2, \ldots, n)$, и событие $A\subset H_1+H_2+\ldots+H_n$, для которого известны условные вероятности $P(A/H_i)$ $(i=1, 2, \ldots, n)$. Произведен опыт, в результате которого появилось событие A. Нужно найти условные вероятности событий $H_1, H_2, \ldots, H_n$ относительно события A.

Применяя теорему умножения вероятностей (формула (28.25)), получаем

$$P(AH_k)=P(A)P(H_k/A)=P(H_k)P(A/H_k),$$

откуда

$$P(H_k/A)=\frac{P(H_k)\,P(A/H_k)}{P(A)}.$$

Подставив сюда выражение для $P(A)$ из формулы полной вероятности (28.33), получим

$$P(H_k/A)=\frac{P(H_k)P(A/H_k)}{\sum\limits_{i=1}^{n}P(H_i)\,P(A/H_i)}\quad (k=1, 2, \ldots, n). \qquad (28.34)$$

Эти формулы называют формулами Байеса[1]. Они решают поставленную задачу.

[1] *Томас Байес* (*Thomas Bayes*, 1702—1761) — английский математик.

Замечание. Вероятности $P(H_k)$ ($k=1, 2, \ldots, n$) событий $H_1, H_2, \ldots, H_n$ до опыта называются априорными вероятностями (от латинского a priori, что означает «сперва», т. е. в данном случае до того, как был произведен опыт). Вероятности $P(H_k/A)$ ($k=1, 2, \ldots, n$) тех же событий называются апостериорными (от латинского слова a posteriori, что означает «после», т. е. в данном случае после опыта).

Пример. Имеется 5 урн с белыми и черными шарами различного состава: две урны — по 2 белых и 3 черных шара (состав H_1), две урны — по 1 белому и 4 черных шара (состав H_2), одна урна — 4 белых и 1 черный шар (состав H_3). Из одной наудачу выбранной урны извлечен шар, который оказался черным (событие A). Чему равна апостериорная вероятность того, что шар извлечен из урны второго состава?

Полагая в (28.34) $k=2$, $n=3$, получаем формулу, которой надлежит пользоваться в данном случае:

$$P(H_2/A)=\frac{P(H_2)P(A/H_2)}{P(H_1)P(A/H_1)+P(H_2)P(A/H_2)+P(H_3)P(A/H_3)}.$$

Найдем соответствующие вероятности: $P(H_1)=2/5$, $P(H_2)=2/5$, $P(H_3)=1/5$, $P(A/H_1)=3/5$, $P(A/H_2)=4/5$, $P(A/H_3)=1/5$ и подставим их в данную формулу:

$$P(H_2/A)=\frac{\frac{2}{5}\cdot\frac{4}{5}}{\frac{2}{5}\cdot\frac{3}{5}+\frac{2}{5}\cdot\frac{4}{5}+\frac{1}{5}\cdot\frac{1}{5}}=\frac{\frac{8}{25}}{\frac{15}{25}}=\frac{8}{15}.$$

Аналогично можно найти $P(H_1/A)=6/15$, $P(H_3/A)=1/15$.

Глава 29

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

В этой главе вводятся понятия случайной величины, функции распределения и плотности распределения вероятностей. Под случайной величиной понимают переменную величину, значения которой зависят от случая и для которой соответствующим образом определена функция распределения.

§ 29.1. Случайная величина

Пусть (Ω, L, P) — произвольное вероятностное пространство. Случайной величиной X назовем действительную функцию $X=X(\omega)$, $\omega\in\Omega$, такую, что при любом действительном x

$$\{\omega\colon X(\omega)<x\}\in L. \tag{29.1}$$

Событие, определяемое (29.1), более кратко можно записать в виде $X<x$ (случайная величина X принимает значение, меньшее x).

Так как L является σ-алгеброй, то из (29.1) следует

$$\left.\begin{aligned} &(X \geqslant x) = \overline{(X<x)} \in L; \\ &(x_1 \leqslant X < x_2) = ((X<x_2) - (X<x_1)) \in L; \\ &(X = x) = \bigcap_{n=1}^{\infty} \left(x \leqslant X < x + \frac{1}{n}\right) \in L. \end{aligned}\right\} \qquad (29.2)$$

Соотношения (29.2) означают, что событие, противоположное данному, разность двух данных событий и произведение счетного множества их принадлежат σ-алгебре L. (*Счетным множеством* называется множество, элементы которого можно поставить во взаимно-однозначное соответствие с элементами множества натуральных чисел; другими словами, счетное множество — бесконечное множество, элементы которого можно пронумеровать: x_1, x_2, x_3. . . .) Следовательно, вероятности этих событий определены.

Пример. Пусть в единичный квадрат $\Omega=\{(u, v): 0\leqslant u\leqslant 1, 0\leqslant v\leqslant 1\}$ наудачу бросается точка. Элементарными событиями ω являются точки, событиями — квадрируемые подмножества квадрата Ω; σ-алгебра L представляет систему квадрируемых подмножеств этого квадрата.

Вероятность события равна площади соответствующего квадрируемого подмножества. Случайными величинами будут, например, функция $X(\omega)=\sqrt{u^2+v^2}$ — расстояние брошенной точки $M(u, v)$ от начала координат, функция $X_1(\omega)=u$ — первая координата брошенной точки и др.

§ 29.2. Функция распределения и ее свойства

Функцией распределения случайной величины X называется функция действительной переменной x, определяемая равенством

$$F(x)=P(X<x), \qquad (29.3)$$

где $P(X<x)$ — вероятность того, что случайная величина X примет значение, меньшее x. Геометрически это означает следующее: $F(x)$ — вероятность того, что случайная величина X примет значение, которое изобра-

жается точкой на числовой прямой, расположенной слева от точки x (рис. 29.1).

Значения функции распределения принадлежат отрезку $[0, 1]$, т. е. $0 \leqslant F(x) \leqslant 1$. Последние неравенства следуют из определения $F(x)$ и свойства вероятности (вероятность события заключена между нулем и единицей).

Если все возможные значения случайной величины X

Рис. 29.1

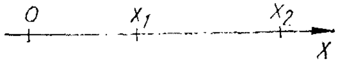

Рис. 29.2

принадлежат интервалу $]a, b[$, то для ее функции распределения $F(x)$

$$F(x)=0 \text{ при } x \leqslant a, \quad F(x)=1 \text{ при } x \geqslant b. \qquad (29.4)$$

Действительно, если $x \leqslant a$, то событие $X<x$ является невозможным (таких значений X не принимает), его вероятность равна нулю; если $x \geqslant b$, то событие $X<x$ будет достоверным, его вероятность равна единице.

Так как $X<-\infty$ — невозможное, а $X<+\infty$ — достоверное события, то

$$F(-\infty)=0, \quad F(+\infty)=1. \qquad (29.5)$$

Поскольку $(X<x_2)=(x_1 \leqslant X<x_2)+(X<x_1)$ и события $(X<x_1)$ и $(x_1 \leqslant X<x_2)$ несовместны (рис. 29.2), то по аксиоме сложения вероятностей несовместных событий имеем

$$P(X<x_2)=P(x_1 \leqslant X<x_2)+P(X<x_1),$$
$$P(x_1 \leqslant X<x_2)=P(X<x_2)-P(X<x_1).$$

С учетом определения (29.3) последнее равенство запишем так:

$$P(x_1 \leqslant X<x_2)=F(x_2)-F(x_1). \qquad (29.6)$$

Следовательно, вероятность того, что случайная величина X примет значение из полуинтервала $[x_1, x_2[$, равна разности значений ее функции распределения в концах этого промежутка.

Поскольку $X<x$ и $X \geqslant x$ — противоположные события, то $P(X<x)+P(X \geqslant x)=1$, откуда $P(X \geqslant x)= 1-P(X<x)$, или $P(X \geqslant x)=1-F(x)$.

Из формул (28.22), (29.2), (29.4) получаем

$$P(X=x)=\lim_{n\to\infty}\left(F\left(x+\frac{1}{n}\right)-F(x)\right)=F(x+0)-F(x). \tag{29.7}$$

Функция распределения $F(x)$ обладает следующими свойствами.

1. $F(x)$ — неубывающая функция; если $x_1<x_2$, то $F(x_1)\leqslant F(x_2)$. Действительно, так как $(X<x_1)\subset(X<x_2)$ при $x_1\leqslant x_2$, то в соответствии с (28.20) $P(X<x_1)\leqslant\leqslant P(X<x_2)$, или $F(x_1)\leqslant F(x_2)$.

2. $\lim\limits_{x\to-\infty}F(x)=F(-\infty)=0,\quad \lim\limits_{x\to+\infty}F(x)=F(+\infty)=1.$

Последовательность событий $(X<-n)$, $n=1, 2, \ldots$, монотонно убывает, т. е. $(X<-1)\supset(X<-2)\supset\ldots$ $\ldots\supset(X<-n)\supset(X<-n-1)\supset\ldots$ и $\bigcap\limits_{n=1}^{\infty}(X<-n)=\varnothing$. По формуле (28.22) получаем

$$\lim_{n\to-\infty}F(-n)=\lim_{n\to-\infty}P(x<-n)=P(\varnothing)=0.$$

Отсюда с учетом монотонности функции $F(x)$ следует, что $\lim\limits_{x\to-\infty}F(x)=F(-\infty)=0$. Второе равенство доказывается аналогично (с использованием монотонной последовательности $(X<n)$, $n=1, 2, \ldots$).

3. $\lim\limits_{x\to x_0-0}F(x)=F(x_0)$; в точке x_0 функция $F(x)$ непрерывна слева. Пусть числовая последовательность (x_n) возрастает и $\lim\limits_{n\to\infty}x_n=x_0$, тогда $(X<x_n)\subset(X<x_{n+1})$, $\bigcup\limits_{n=1}^{\infty}(X<x_n)=(X<x_0)$ и по формуле (28.22) получаем $\lim\limits_{n\to\infty}P(X<x_n)=P(x<x_0)$. С учетом монотонности функции $F(x)$ отсюда следует третье свойство, т. е. $\lim\limits_{x\to x_0-0}F(x)=$ $=F(x_0)$.

На основании свойств функции распределения $F(x)$ можно судить об особенностях ее графика (рис. 29.3).

Замечание. Любая функция $\Phi(x)$, обладающая указанными свойствами, является функцией распределения некоторой случайной величины, т. е. можно построить вероятностное пространство (Ω, L, P) и определить на нем случайную величину X такую, что $F(x)=\Phi(x)$.

Случайная величина X называется *дискретной*, если она может принимать только конечное или счетное множество значений. Для полной вероятностной характеристики дискретной случайной величины X, принимающей значения $x_1, x_2, x_3, \ldots$, достаточно знать вероятности этих значений $P(X=x_k)=p_k$ $(k=1, 2, 3, \ldots)$, где $\sum_{k=1}^{\infty} p_k=1$. Законом распределения дискретной случайной величины X называется соответствие между значениями

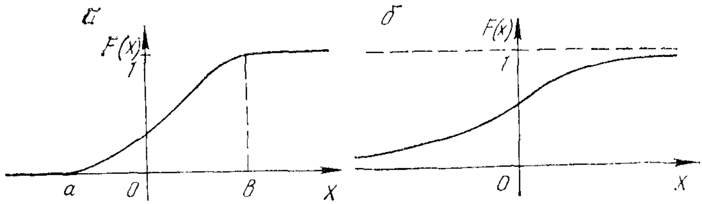

Рис. 29.3

$x_1, x_2, x_3, \ldots$ этой величины и их вероятностями $p_1, p_2, p_3, \ldots$.

Закон распределения дискретной случайной величины X, принимающей значения $x_1, x_2, x_3, \ldots$ с вероятностями $p_1, p_2, p_3, \ldots$, причем $\sum_{k=1}^{\infty} p_k=1$, можно задать таблицей

X	x_1	x_2	x_3	...	x_k	...
P	p_1	p_2	p_3	...	p_k	...

или формулами

$$P(X=x_k)=p_k\,(k=1, 2, 3, \ldots), \quad \sum_{k=1}^{\infty} p_k=1. \qquad (29.8)$$

Аналогично задается закон распределения дискретной случайной величины X, принимающей конечное множество значений $x_1, x_2, \ldots, x_n$ соответственно с вероятностями $p_1, p_2, \ldots, p_n$, где $\sum_{k=1}^{n} p_k=1$.

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

или $P(X=x_k)=p_k\,(k=1, 2, \ldots, n)$, $\sum\limits_{k=1}^{n} p_k=1$.

Функция распределения $F(x)$ дискретной случайной величины имеет вид

$$F(x)=\sum_{x_k<x} P(X=x_k), \tag{29.9}$$

где символ $x_k<x$ означает, что суммируются вероятности тех значений, которые меньше x. Функция $F(x)$ для дискретной случайной величины разрывна.

Случайная величина X называется *непрерывной*, если существует неотрицательная функция $p(x)$ такая, что при любом x

$$F(x)=P(X<x)=\int_{-\infty}^{x} p(u)\,du. \tag{29.10}$$

Как известно, определенный интеграл с переменным верхним пределом является непрерывной функцией (см. § 16.6). Из формулы (29.10) видно, что функция распределения $F(x)$ непрерывна.

§ 29.3. Плотность распределения

Рассмотрим непрерывную случайную величину X, для которой

$$F(x)=P(X<x)=\int_{-\infty}^{x} p(u)\,du.$$

Функция $p(x)$, входящая в это равенство, называется плотностью распределения вероятностей случайной величины X. График функции $p(x)$ называется кривой распределения.

Докажем, что вероятность попадания значений случайной величины X в полуинтервал $[a, b[$ равна определенному интегралу от плотности распределения $p(x)$ по отрезку $[a, b]$, т. е.

$$P(a \leqslant X < b) = \int_a^b p(x)\,dx. \tag{29.11}$$

Действительно, на основании (29.6) имеем $P(a \leqslant X < b) = F(b) - F(a)$. В соответствии с (29.10)

$$F(b) = \int_{-\infty}^{b} p(x)\,dx, \quad F(a) = \int_{-\infty}^{a} p(x)\,dx.$$

Так как

$$\int_{-\infty}^{b} p(x)\,dx = \int_{-\infty}^{a} p(x)\,dx + \int_a^b p(x)\,dx, \text{ или } \int_a^b p(x)\,dx =$$

$$= \int_{-\infty}^{b} p(x)\,dx - \int_{-\infty}^{a} p(x)\,dx = F(b) - F(a),$$

то из двух выражений для $F(b) - F(a)$ следует равенство (29.11). Доказанная формула геометрически означает следующее: вероятность попадания значений непрерывной случайной величины X в полуинтервал $[a, b[$ равна площади криволинейной трапеции, ограниченной кривой распределения, осью Ox и отрезками прямых $x = a$, $x = b$ (рис. 29.4).

Очевидно, что

$$P(X = a) = \lim_{n \to \infty} P\left(a \leqslant X < a + \frac{1}{n}\right) = \lim_{n \to \infty} \int_a^{a+\frac{1}{n}} p(x)dx = 0.$$

Отсюда для непрерывных величин X получаем

$$P(a \leqslant X < b) = P(a < X \leqslant b) = \\ = P(a < X < b) = P(a \leqslant X \leqslant b). \tag{29.12}$$

Из (29.11) и (29.12) следуют, например, формулы

$$P(a \leqslant X \leqslant b) = \int_a^b p(x)\,dx, \quad P(a < X < b) = \int_a^b p(x)\,dx. \tag{29.13}$$

Если x — точка непрерывности функции $p(x)$, то

$$P(x \leqslant X \leqslant x + \Delta x) = p(x)\Delta x + o(\Delta x). \tag{29.14}$$

Действительно, учитывая (29.11), (29.12) и применяя теорему о среднем для определенного интеграла, получаем

$$P(x<X<x+\Delta x)=\int_{x}^{x+\Delta x} p(x)\,dx=p(\xi)\,\Delta x,\ \ \xi=x+\Theta\Delta x,$$
$$0\leqslant\Theta\leqslant 1,$$

или $p(\xi)\Delta x=p(x+\Theta\Delta x)\Delta x=p(x)\Delta x+o(\Delta x)$.

Плотность распределения вероятностей $p(x)$ обладает следующими свойствами.

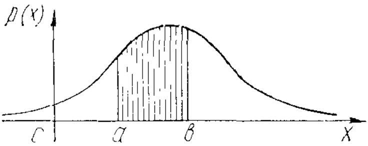

Рис. 29.4

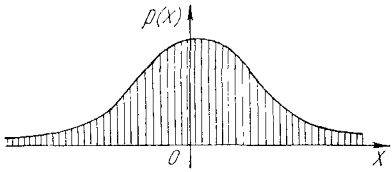

Рис. 29.5

1. Плотность распределения $p(x)$ — неотрицательная функция, т. е. $p(x)\geqslant 0$; это следует из определения функции $p(x)$.

2. В точках дифференцируемости $F(x)$ производная функции распределения равна плотности распределения вероятностей, т. е.

$$F'(x)=p(x). \qquad (29.15)$$

Равенство (29.15) следует из формулы (29.10) и теоремы о производной определенного интеграла с переменным верхним пределом (см. § 16.6).

3. Интервал по бесконечному промежутку $]-\infty, +\infty[$ от плотности распределения вероятностей $p(x)$ равен единице, т. е.

$$\int_{-\infty}^{+\infty} p(x)\,dx=1. \qquad (29.16)$$

В самом деле, исходя из определения несобственного интеграла и свойства 2 функции $F(x)$, получаем

$$\int_{-\infty}^{+\infty} p(x)\,dx=\lim_{x\to+\infty}\int_{-\infty}^{x} p(u)\,du=\lim_{x\to+\infty}F(x)=1.$$

Это свойство имеет следующую геометрическую интерпретацию: площадь фигуры, ограниченной кривой распределения и осью Ox, равна единице (рис. 29.5).

Отметим, что если все возможные значения случайной величины принадлежат отрезку $[a, b]$, то

$$\int_a^b p(x)\,dx = 1, \tag{29.17}$$

так как $p(x) = 0$ вне этого отрезка.

§ 29.4. Совместное распределение двух случайных величин

Пусть на вероятностном пространстве (Ω, L, P) заданы две случайные величины: $X = X(\omega)$, $Y = Y(\omega)$, $\omega \in \Omega$. Каждому элементарному событию ω ставится в соответствие упорядоченная пара значений (x, y) случайных величин X и Y. Упорядоченную пару (X, Y) случайных величин X, Y называют *двумерной случайной величиной*, или *случайным вектором*. В примере § 29.2 двумерной случайной величиной будет (X, Y), где X, Y — координаты точки, в которой оказалась точка, брошенная в данный квадрат. Примером двумерной случайной величины является упорядоченная пара (X, Y), где X, Y — координаты точки попадания снаряда при стрельбе из орудия по некоторой плоской области.

Функцией распределения двумерной случайной величины или случайного вектора (X, Y) называется функция $F(x, y)$ двух действительных переменных x, y, определяемая формулой

$$F(x, y) = P(X < x,\ Y < y), \tag{29.18}$$

где $P(X < x,\ Y < y)$ — вероятность того, что случайная величина X примет значение, меньшее x, а случайная величина Y — значение, меньшее y. Геометрический смысл равенства (29.18): функция $F(x, y)$ есть вероятность того, что случайная точка (X, Y) попадает в бесконечный квадрант с вершиной в точке (x, y); точка (X, Y) будет левее и ниже этой вершины (рис. 29.6).

Из определения $F(x, y)$ и свойств вероятности следует, что значения этой функции принадлежат отрезку $[0, 1]$, т. е. $0 \leqslant F(x, y) \leqslant 1$.

Функция распределения $F(x, y)$ обладает следующими свойствами.

1. $F(x, y)$ является неубывающей функцией по каждой переменной: если $x_1 < x_2$, то $F(x_1, y) \leqslant F(x_2, y)$; если $y_1 < y_2$, то $F(x, y_1) \leqslant F(x, y_2)$. Эти неравенства следуют из определения и свойств функций $F_1(x)$ и $F_2(y)$, получающихся из $F(x, y)$ при фиксировании y и x соответственно: $F_1(x) = F(x, y_0)$, $F_2(y) = F(x_0, y)$.

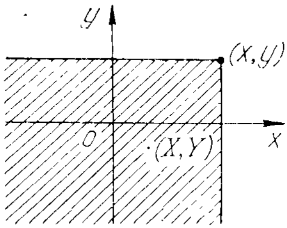

Рис. 29.6

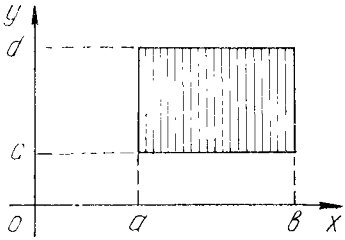

Рис. 29.7

2. Справедливы соотношения:

1) $\lim\limits_{x \to -\infty} F(x, y) = F(-\infty, y) = 0,$

$\lim\limits_{y \to -\infty} F(x, y) = F(x, -\infty) = 0;$

2) $\lim\limits_{\substack{x \to +\infty \\ y \to +\infty}} F(x, y) = F(+\infty, +\infty) = 1.$

Эти соотношения следуют из определения (29.18) и геометрического смысла функции распределения.

Случайный вектор (X, Y) называют вектором дискретного типа, если существует конечное или счетное множество точек (x_i, y_j), $i, j = 1, 2, \ldots$ (без предельных точек), таких, что

$$P(X = x_i,\ Y = y_j) = p_{ij}, \quad \sum_{i, j=1}^{\infty} p_{ij} = 1. \tag{29.19}$$

Случайный вектор (X, Y) называют вектором абсолютно непрерывного типа, если существует функция $p(x, y) \geqslant 0$ такая, что при любых x, y

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} p(u, v)\, du\, dv. \tag{29.20}$$

Функция $p(x, y)$, входящая в равенство (29.20), называется плотностью распределения вероятностей случайного вектора (X, Y) или двумерной случайной величины (X, Y).

Для одномерной случайной величины в § 29.2 была получена формула (29.6), выражающая вероятность попадания значения случайной величины в полуинтервал через функцию распределения. Найдем аналогичную формулу в двумерном случае, т. е. формулу для вероятности попадания значений пары случайных величин в прямоугольник $S = (a \leqslant X < b,\ c \leqslant Y < d)$ (рис. 29.7). Поскольку

$$\begin{gathered}(X<b,\ Y<d) = S + (X<a,\ Y<d) + (X<b,\ Y<c),\\ (X<a,\ Y<d)\,(X<b,\ Y<c) = (X<a,\ Y<c),\end{gathered} \tag{29.21}$$

то по формуле (28.18)

$$\begin{gathered}P\left((X<a,\ Y<d) + (X<b,\ Y<c)\right) =\\ = P(X<a,\ Y<d) + P(X<b,\ Y<c) - P(X<a,\ Y<c) =\\ = F(a, d) + F(b, c) - F(a, c).\end{gathered}$$

Так как в (29.21) S несовместно с другими слагаемыми, то из (29.21) и последнего равенства получаем формулу

$$\begin{gathered}P(S) = P(a \leqslant X < b,\ c \leqslant Y < d) =\\ = F(b, d) + F(a, c) - F(a, d) - F(b, c).\end{gathered} \tag{29.22}$$

Если (X, Y) — абсолютно непрерывный вектор, то из формулы (29.22) следует, что

$$P(a \leqslant X < b,\ \ c \leqslant Y < d) = \int\limits_a^b \int\limits_c^d p(x,\ y)\,dxdy. \tag{29.23}$$

Формула (29.23) является частным случаем более общей формулы

$$P((x,\ y) \in G) = \iint\limits_G p(x,\ y)\,dxdy, \tag{29.24}$$

которая приводится здесь без доказательства.

Формула (29.24) означает следующее: вероятность того, что двумерная случайная величина (X, Y) с плотностью распределения $p(x, y)$ принимает значения из области G, равна двойному интегралу от функции $p(x, y)$ по этой области.

По двумерной функции распределения можно найти одномерные функции распределения. Поскольку $(Y<+\infty)=\Omega$ (достоверное событие), то $(X<x)=(X<x, Y<+\infty)$. Следовательно,

$$F_1(x)=P(X<x)=P(X<x,\ Y<+\infty)=$$
$$=F(x,+\infty)=\lim_{y\to+\infty} F(x, y).$$

Итак,

$$\begin{aligned} F_1(x) &= \lim_{y\to+\infty} F(x, y)=F(x, +\infty), \\ F_2(y) &= \lim_{x\to+\infty} F(x, y)=F(+\infty, y). \end{aligned} \tag{29.25}$$

Если (X, Y) — абсолютно непрерывный вектор, то

$$F_1(x)=\int_{-\infty}^{x}\left[\int_{-\infty}^{+\infty} p(u, v)\, dv\right] du=\int_{-\infty}^{x} p_1(u)\, du,$$

где функция

$$p_1(x)=\int_{-\infty}^{+\infty} p(x,\ v)\, dv \tag{29.26}$$

есть плотность распределения величины X. Аналогично получаем

$$p_2(y)=\int_{-\infty}^{+\infty} p(u,\ y)\, du. \tag{29.27}$$

Пусть вектор (X, Y) является вектором дискретного типа:

$$P(X=x_i,\ Y=y_j)=p_{ij}\ (i, j=1, 2, \ldots),\ \sum_{i, j=1}^{\infty} p_{ij}=1.$$

Найдем соответствующие одномерные распределения. Имеем

$$P(X=x_i)=\sum_{j=1}^{\infty} P(X=x_i,\ Y=y_j)=\sum_{j=1}^{\infty} p_{ij}.$$

Положим

$$p_i=\sum_{j=1}^{\infty} p_{ij},\quad q_j=\sum_{i=1}^{\infty} p_{ij}, \tag{29.28}$$

тогда

$$P(X=x_i)=p_i,\ P(Y=y_j)=q_j. \tag{29.29}$$

Если дискретный вектор принимает конечное множество значений, то

$$P(X=x_i)=\sum_{j=1}^{n} p_{ij}=p_i,\ P(Y=y_j)=\sum_{i=1}^{k} p_{ij}=q_j. \quad (29.30)$$

В этом случае значения можно представить в виде табл. 29.1.

Таблица 29.1

$X \backslash Y$	y_1	y_2	...	y_j	...	y_n	Σ
x_1	p_{11}	p_{12}	...	p_{1j}	...	p_{1n}	p_1
x_2	p_{21}	p_{22}	...	p_{2j}	...	p_{2n}	p_2
...	...	...	...	...	...	...	...
x_i	p_{i1}	p_{i2}	...	p_{ij}	...	p_{in}	p_i
...	...	...	...	...	...	...	...
x_k	p_{k1}	p_{k2}	...	p_{kj}	...	p_{kn}	p_k
Σ	q_1	q_2	...	q_j	...	q_n	1

Из формул (29.30) видно, что для вычисления вероятности $P(X=x_i)=p_i$ в таблице надо просуммировать вероятности в i-й строке, а для вычисления вероятности $P(Y=y_j)=q_j$ — просуммировать вероятности в j-м столбце. Поскольку

$$\sum_{i=1}^{k}\sum_{j=1}^{n} P(X=x_i,\ Y=y_j)=\sum_{i=1}^{k}\sum_{j=1}^{n} p_{ij}=1,$$

то

$$\sum_{i=1}^{k} p_i=\sum_{j=1}^{n} q_j=1.$$

§ 29.5. Совместные распределения нескольких случайных величин

Рассмотрим вероятностное пространство $(\Omega,\ L,\ P)$, на котором заданы случайные величины

$$X_1=X_1(\omega),\ X_2=X_2(\omega),\ \ldots,\ X_n=X_n(\omega),\ \omega\in\Omega.$$

Каждому событию $\omega \in \Omega$ эти величины ставят в соответствие n-мерный вектор $\mathbf{u}(X_1, X_2, \ldots, X_n)$.

Многомерной функцией распределения случайного вектора $\mathbf{u}(X_1, X_2, \ldots, X_n)$ называется числовая функция $F(x_1, x_2, \ldots, x_n)$ n действительных переменных $x_1, x_2, \ldots, x_n$, определяемая равенством

$$F(x_1, x_2, \ldots, x_n) = P(X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n), \quad (29.31)$$

где $P(X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n)$ — вероятность того, что случайная величина X_1 примет значение, меньшее x_1, X_2 — значение, меньшее x_2, $\ldots$, X_n — значение, меньшее x_n.

Из определения и свойств вероятности следует, что $0 \leqslant F(x_1, x_2, \ldots, x_n) \leqslant 1$, т. е. значения функции распределения принадлежат отрезку $[0, 1]$.

Можно показать, что $F(x_1, x_2, \ldots, x_n)$ монотонна по каждому аргументу и выполняются следующие соотношения:

$$\lim_{\substack{x_1 \to +\infty \\ \cdots \\ x_n \to +\infty}} F(x_1, x_2, \ldots, x_n) = 1, \quad \lim_{x_k \to -\infty} F(x_1, x_2, \ldots, x_n) = 0,$$

$$\lim_{x_k \to +\infty} F(x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n) = F_1(x_1, x_2, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n), \quad (29.32)$$

где k — одно из чисел $k = 1, 2, \ldots, n$.

Соотношения (29.32) означают, что предел функции $F(x_1, x_2, \ldots, x_n)$ равен единице, когда все аргументы стремятся к $+\infty$; равен нулю, если один из аргументов стремится к $-\infty$; равен функции $n-1$ переменных, если один из аргументов стремится к $+\infty$.

Случайный вектор $(X_1, X_2, \ldots, X_n)$ называют вектором дискретного типа, если существует конечное или счетное множество точек $(x_{k1}, x_{k2}, \ldots, x_{kn})$, $k = 1, 2, \ldots$ (без предельных точек), таких, что

$$P(X_1 = x_{k1}, X_2 = x_{k2}, \ldots, X_n = x_{kn}) = p_{x_{k_1} x_{k_2} \ldots x_{k_n}},$$

$$\Sigma p_{x_{k_1} x_{k_2} \ldots x_{k_n}} = 1. \quad (29.33)$$

Случайный вектор $(X_1, X_2, \ldots, X_n)$ называют вектором абсолютно непрерывного типа, если существует функция $p(x_1, x_2, \ldots, x_n) \geqslant 0$ такая, что при любых $x_1, x_2, \ldots, x_n$

$$F(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \ldots \int_{-\infty}^{x_n} p(u_1, u_2, \ldots, u_n)\, du_1 du_2 \ldots du_n. \quad (29.34)$$

Функция $p(x_1, x_2, \ldots, x_n)$, входящая в равенство (29.34), называется *плотностью распределения* вероятностей случайного вектора $(X_1, X_2, \ldots, X_n)$.

§ 29.6. Независимость случайных величин

Случайные величины X и Y называются *независимыми*, если для любых x, y выполняется равенство

$$P(X<x, Y<y) = P(X<x)P(Y<y). \quad (29.35)$$

Теорема 29.1. Случайные величины X и Y независимы тогда и только тогда, когда функция распределения двумерной случайной величины (X, Y) равна произведению функций распределений составляющих:

$$F(x, y) = F_1(x)F_2(y). \quad (29.36)$$

Доказательство. *Необходимость.* Пусть случайные величины X, Y независимы, тогда выполнено равенство (29.35). По определению функций распределения $P(X<x,\ Y<y) = F(x,\ y)$, $P(X<x) = F_1(x)$, $P(Y<y) = F_2(y)$. Подставляя эти выражения в равенство (29.35), получаем $F(x, y) = F_1(x)F_2(y)$, т. е. равенство (29.36).

Достаточность. Пусть выполнено равенство (29.36), тогда из определений функций распределения следует равенство (29.35), которое означает, что случайные величины X и Y независимы.

Можно показать, что необходимое и достаточное условие независимости случайных величин X и Y выражается равенством

$$p(x, y) = p_1(x)p_2(y), \quad (29.37)$$

где $p(x, y)$ — плотность распределения вероятностей двумерной случайной величины (X, Y); $p_1(x)$, $p_2(y)$ — плотности распределения вероятностей составляющих случайных величин X и Y.

Пусть (Ω, L, P) — произвольное вероятностное пространство. Случайные величины $X_1, X_2, \ldots, X_n$ называются независимыми в совокупности, или независимыми, если для любой системы $X^{(1)}, X^{(2)}, \ldots, X^{(k)}$ этих величин выполняется равенство

$$P(X^{(1)}<x^{(1)}, X^{(2)}<x^{(2)}, \ldots, X^{(k)}<x^{(k)}) =$$
$$= P(X^{(1)}<x^{(1)})\,P(X^{(2)}<x^{(2)}) \ldots P(X^{(k)}<x^{(k)}),$$

где $x^{(1)}, x^{(2)}, \ldots, x^{(k)}$ — произвольные действительные числа; k — любое из чисел от 1 до n. В частности,

$$P(X_1<x_1, X_2<x_2, \ldots, X_n<x_n) =$$
$$= P(X_1<x_1)\,P(X_2<x_2) \ldots P(X_n<x_n),$$

где $x_1, x_2, \ldots, x_n$ — произвольные действительные числа.

Необходимое и достаточное условие независимости случайных величин $X_1, X_2, \ldots, X_n$ выражается равенством

$$F(x_1, x_2, \ldots, x_n) = F_1(x_1)\,F_2(x_2) \ldots F_n(x_n), \quad (29.38)$$

где $F(x_1, x_2, \ldots, x_n)$ — функция распределения этих величин; $F_k(x_k)$ — функция распределения случайной величины X_k $(k=1, 2, \ldots, n)$.

§ 29.7. Функции случайных величин

Может оказаться, что одна случайная величина Y связана с другой случайной величиной X функциональной зависимостью $Y=f(X)$. Например, если случайная величина X есть радиус обтачиваемого на станке шарика, то его объем $Y=\frac{4}{3}\pi X^3$ — также случайная величина, являющаяся заданной функцией случайной величины X.

Покажем, как найти закон распределения случайной величины Y по известному закону распределения случайной величины X, если $Y=f(X)$, где $f(X)$ — заданная функция.

Рассмотрим сначала случай дискретной случайной величины. Пусть аргумент X имеет следующий закон распределения:

$$P(X=x_i)=p_i \quad (i=1, 2, \ldots), \quad \sum_{i=1}^{\infty} p_i = 1,$$

тогда функция $Y=f(X)$ примет соответствующие значения y_i $(i=1, 2, \ldots)$ с теми же вероятностями p_i.

Действительно, если случайная величина X принимает значение $X=x_i$ с вероятностью p_i, то случайная величина

Y принимает значение $Y=y_i$ с той же вероятностью; этим значением будет $y_i=f(x_i)$, так как функция $f(X)$ известна.

Рассмотрим непрерывную случайную величину X с плотностью распределения $p(x)$ и случайную величину Y, связанную с первой функциональной зависимостью $Y=f(X)$. Предположим, что функция $p(x)$ непрерывна, функция $y=f(x)$ имеет непрерывную производную $y'>0$ в интервале $]a, b[$ возможных значений случайной величины X, а областью ее значений является интервал $]c, d[$. Найдем плотность распределения $\varphi(y)$ случайной величины Y и докажем, что функция $\varphi(y)$ непрерывна в интервале $]c, d[$.

Непрерывная и возрастающая функция $y=f(x)$ взаимно-однозначно отображает интервал $]a, b[$ на интервал $]c, d[$; любому $x\in]a, b[$ соответствует единственное значение $y\in]c, d[$ и обратно. Следовательно, вероятности попаданий случайных величин X и Y в соответствующие интервалы равны между собой, т. е. $P(x<X<x+\Delta x)=P(y<Y<y+\Delta y)$.

Выражая эти вероятности с помощью формулы (29.13) и применяя теорему о среднем, получаем

$$P(x<X<x+\Delta x)=\int_x^{x+\Delta x} p(x)\,dx=p(\xi)\,\Delta x,\quad x<\xi<x+\Delta x;$$

$$P(y<Y<y+\Delta y)=\int_y^{y+\Delta y} \varphi(y)dy=\varphi(\eta)\Delta y,\quad y<\eta<y+\Delta y.$$

Из трех последних равенств следует, что $p(\xi)\Delta x=\varphi(\eta)\Delta y$, откуда

$$\varphi(\eta)=p(\xi)\frac{\Delta x}{\Delta y}. \tag{29.39}$$

По условию функция $y=f(x)$ имеет непрерывную производную $y'>0$ и возрастает, поэтому для нее существует обратная функция $x=g(y)$ с непрерывной производной $x'_y=g'(y)>0$. Переходя к пределу в равенстве (29.39) при $\Delta x\to 0$ $(\Delta y\to 0,\ \xi\to x,\ \eta\to y)$, получаем $\varphi(y)=p(x)x'_y$, или $\varphi(y)dy=p(x)dx$. Равенство $\varphi(y)=p(x)x'_y$ можно записать в виде

$$\varphi(y)=p(g(y))g'(y). \tag{29.40}$$

Формула (29.40) устанавливает связь между плотностями вероятностей $p(x)$ и $\varphi(y)$. Очевидно, функция $\varphi(y)$ непрерывна в интервале $]c, d[$.

Замечание. Если функция $y=f(x)$ убывает и имеет непрерывную производную $y'<0$, то аналогично можно показать, что

$$\varphi(y) = -p(g(y))g'(y). \tag{29.41}$$

Поскольку в данном случае $g'(y)<0$, эту формулу можно записать так:

$$\varphi(y) = p(g(y))|g'(y)|. \tag{29.42}$$

Формула (29.42) справедлива в обоих случаях: для возрастающей и для убывающей функции $y=f(x)$.

В заключение введем понятие функции двух случайных величин. Пусть D — множество возможных пар (x, y) значений двумерной случайной величины (X, Y) и $\{Z\}$ — множество значений случайной величины Z. Если каждой паре значений $(x, y)\in D$ (случайной точке плоскости) по некоторому правилу f ставится в соответствие одно определенное значение $z\in Z$, то говорят, что на множестве D задана функция f двух случайных величин X, Y, и пишут

$$Z=f(X, Y). \tag{29.43}$$

Например, при стрельбе из орудия по некоторой плоской области расстояние Z точки попадания снаряда до начала координат есть функция двух случайных величин X, Y — координат указанной точки, причем $Z=\sqrt{X^2+Y^2}$.

Глава 30

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ СЛУЧАЙНЫХ ВЕЛИЧИН

Функция распределения полностью характеризует случайную величину. Однако для решения многих задач достаточно знать лишь некоторые числа, характеризующие распределения случайных величин. Эти числа называются числовыми характеристиками случайных величин. К ним, в частности, относятся математическое ожидание, дисперсия.

§ 30.1. Математическое ожидание случайной величины

Математическим ожиданием дискретной случайной величины называется сумма произведений всех ее возможных значений на соответствующие вероятности. Математическое ожидание случайной величины обозначается через $M(X)$, или m_x, или a. Если дискретная случайная величина принимает конечное число значений $x_1, x_2, \ldots, x_n$ соответственно с вероятностями $p_1, p_2, \ldots, p_n$, то по определению

$$M(X)=x_1p_1+x_2p_2+\ldots+x_np_n,\ M(X)=\sum_{k=1}^{n}x_kp_k. \quad (30.1)$$

Пример. Найти математическое ожидание дискретной случайной величины по заданному закону ее распределения:

x	2	3	4	5
p	0,1	0,2	0,3	0,4

По формуле (30.1) получаем

$$M(X)=2\cdot 0{,}1+3\cdot 0{,}2+4\cdot 0{,}3+5\cdot 0{,}4=4.$$

Выясним вероятностный смысл математического ожидания дискретной случайной величины. Пусть в результате n испытаний случайная величина X значение x_1 приняла m_1 раз, значение x_2 — m_2 раз, значение x_k — m_k раз, причем $m_1+m_2+\ldots+m_k=n$. Так как сумма принятых значений равна $x_1m_1+x_2m_2+\ldots+x_km_k$, то среднее арифметическое x всех ее значений определяется формулой

$$x=\frac{x_1m_1+x_2m_2+\ \ldots\ +x_km_k}{n},$$

или

$$x=x_1\frac{m_1}{n}+x_2\frac{m_2}{n}+\ \ldots\ +x_k\frac{m_k}{n},\quad x=x_1W_1+x_2W_2+$$
$$+\ \ldots\ +x_kW_k,$$

где $W_i=\frac{m_i}{n}$ — относительная частота значений x_i $(i=1, 2, \ldots, k)$. Если n достаточно велико, то относительная частота события приближенно равна его вероятности, т. е. $W_k\approx p_k$, поэтому получаем приближенную формулу

$$x\approx x_1p_1+x_2p_2+\ldots+x_kp_k,\ x\approx M(X).$$

Итак, математическое ожидание дискретной случайной величины приближенно равно среднему арифметическому ее возможных значений. Вследствие этого математическое ожидание случайной величины называется ее *средним значением*.

Замечание. Математическое ожидание называют также *центром распределения*. Это название заимствовано из механики и объясняется следующим: если в точках $x_1, x_2, \ldots, x_n$ оси Ox находятся соответственно массы $p_1, p_2, \ldots, p_n$, то координата x центра тяжести системы материальных точек вычисляется по формуле $x = \left(\sum_{k=1}^{n} x_k p_k\right) : \left(\sum_{k=1}^{n} p_k\right)$. Так как $\sum_{k=1}^{n} p_k = 1$, то $x = \sum_{k=1}^{n} x_k p_k = M(X)$.

Пусть (X, Y) — система дискретных случайных величин, совместное распределение которых

$$P(X=x_i,\ Y=y_k)=p_{ik},\ P(X=x_i)=p_i,\ P(Y=y_k)=q_k; \quad (30.2)$$

$$\sum_{k=1}^{n} p_{ik}=p_i,\quad \sum_{i=1}^{m} p_{ik}=q_k. \quad (30.3)$$

Математическое ожидание функции $Z=f(X, Y)$, где X, Y — дискретные случайные величины, принимающие соответственно значения x_i $(i=1, 2, \ldots, m)$, y_k $(k=1, 2, \ldots, n)$ с указанными выше вероятностями, определяется формулой

$$\left.\begin{aligned} M(Z) = f(x_1, y_1)\,p_{11} + f(x_1, y_2)\,p_{12} + \ldots +\\ + f(x_1, y_n)\,p_{1n} + f(x_2, y_1)\,p_{21} + f(x_2, y_2)\,p_{22} +\\ + \ldots + f(x_2, y_n)\,p_{2n} +\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ + f(x_m, y_1)\,p_{m1} + f(x_m, y_2)\,p_{m2} + \ldots +\\ + f(x_m, y_n)\,p_{mn}. \end{aligned}\right\} \quad (30.4)$$

Математическое ожидание дискретной случайной величины обладает следующими свойствами.

1. Математическое ожидание случайной величины заключено между ее наименьшим и наибольшим значениями.

2. Математическое ожидание постоянной равно этой постоянной, т. е. $M(C)=C$.

3. Постоянный множитель можно выносить за знак математического ожидания, т. е. $M(CX)=CM(X)$.

4. Математическое ожидание суммы двух случайных

величин равно сумме их математических ожиданий, т. е. $M(X+Y)=M(X)+M(Y)$.

5. Математическое ожидание произведения двух независимых величин равно произведению их математических ожиданий, т. е. $M(XY)=M(X)M(Y)$.

Докажем эти свойства.

1. Обозначим через a и b соответственно наименьшее и наибольшее из всех значений $x_1, x_2, \ldots, x_n$ случайной величины X, которые она принимает с вероятностями $p_1, p_2, \ldots, p_n$, тогда

$$a(p_1+p_2+\ldots+p_n)\leqslant x_1p_1+x_2p_2+\ldots+x_np_n\leqslant$$
$$\leqslant b(p_1+p_2+\ldots+p_n).$$

Так как $\sum\limits_{k=1}^{n} p_k=1$, $\sum\limits_{k=1}^{n} x_kp_k=M(X)$, то $a\leqslant M(X)\leqslant b$, что и требовалось доказать.

2. Постоянную величину C можно рассматривать как случайную величину X, принимающую одно и то же значение C с вероятностью $p=1$, поэтому $M(C)=C\cdot 1=C$.

3. Если случайная величина X принимает значения $x_1, x_2, \ldots, x_n$ соответственно с вероятностями $p_1, p_2, \ldots$ $\ldots, p_n$, то случайная величина CX принимает значения $Cx_1, Cx_2, \ldots, Cx_n$ с теми же вероятностями $p_1, p_2, \ldots, p_n$. Следовательно,

$$M(CX)=\sum_{k=1}^{n}(Cx_k)p_k=C\sum_{k=1}^{n}x_kp_k=CM(X).$$

4. Воспользуемся определением математического ожидания для функции $Z=f(X, Y)$ двух дискретных случайных величин X, Y, совместное распределение которых задано равенствами (30.2). В случае $Z=X+Y$ по формуле (30.4) получаем

$$M(X+Y)=(x_1+y_1)p_{11}+(x_1+y_2)p_{12}+\ldots+$$
$$+(x_1+y_n)p_{1n}+\ldots+(x_m+y_1)p_{m1}+(x_m+y_2)p_{m2}+\ldots+$$
$$+(x_m+y_n)p_{mn}=x_1(p_{11}+p_{12}+\ldots+p_{1n})+x_2(p_{21}+p_{22}+$$
$$+\ldots+p_{2n})+\ldots+x_m(p_{m1}+p_{m2}+\ldots+p_{mn})+y_1(p_{11}+p_{21}+$$
$$+\ldots+p_{m1})+y_2(p_{12}+p_{22}+\ldots+p_{m2})+\ldots+y_n(p_{1n}+p_{2n}+$$
$$+\ldots+p_{mn})=(x_1p_1+x_2p_2+\ldots+x_mp_m)+(y_1q_1+y_2q_2+$$
$$+\ldots+y_nq_n)=M(X)+M(Y).$$

Очевидно,

$$M(X_1+X_2+X_3)=M((X_1+X_2)+X_3)=$$
$$=M(X_1+X_2)+M(X_3)=M(X_1)+M(X_2)+M(X_3).$$

Можно доказать, что

$$M(X_1+X_2+\ldots+X_n)=M(X_1)+M(X_2)+\ldots+M(X_n). \quad (30.5)$$

Следствие. Математическое ожидание разности двух случайных величин равно разности их математических ожиданий:

$$M(X-Y)=M(X)-M(Y). \quad (30.6)$$

Действительно,

$$M(X-Y)=M(X+(-Y))=M(X)+M(-Y)=$$
$$=M(X)-M(Y).$$

5. Рассмотрим дискретную двумерную случайную величину (X, Y) с совместным распределением (30.2). Поскольку величины X и Y независимы, то $P(X=x_i, Y=y_k)=P(X=x_i)P(Y=y_k)$, т. е. $p_{ik}=p_iq_k$ $(i=1, 2, \ldots, m;$ $k=1, 2, \ldots, n)$. По формуле (30.4) получаем

$$M(XY)=x_1y_1p_1q_1+x_1y_2p_1q_2+\ldots+x_1y_np_1q_n+x_2y_1p_2q_1+$$
$$+x_2y_2p_2q_2+\ldots+x_2y_np_2q_n+\ldots+x_my_1p_mq_1+$$
$$+x_my_2p_mq_2+\ldots+x_my_np_mq_n=x_1p_1(y_1q_1+y_2q_2+\ldots+$$
$$+y_nq_n)+x_2p_2(y_1q_1+y_2q_2+\ldots+y_nq_n)+$$
$$+\ldots+x_mp_m(y_1q_1+y_2q_2+\ldots+y_nq_n)=$$
$$=x_1p_1M(Y)+x_2p_2M(Y)+\ldots+x_mp_mM(Y)=$$
$$=(x_1p_1+x_2p_2+\ldots+x_my_m)M(Y)=M(X)M(Y).$$

Если X, Y, Z — независимые случайные величины, то $M(XYZ)=M((XY)Z)=M(XY)M(Z)=M(X)M(Y)M(Z)$.

Можно доказать, что для независимых случайных величин $X_1, X_2, \ldots, X_n$

$$M(X_1, X_2\ldots X_n)=M(X_1)M(X_2)\ldots M(X_n). \quad (30.7)$$

Если дискретная случайная величина X принимает счетное множество значений $x_1, x_2, \ldots, x_n, \ldots$ соответственно с вероятностями $p_1, p_2, \ldots, p_n, \ldots$, то по определению

$$M(X)=\sum_{k=1}^{\infty} x_kp_k \quad (30.8)$$

в предположении, что этот ряд абсолютно сходится.

Можно показать, что свойства 1—5 будут выполнены и в этом случае.

Введем понятие математического ожидания для непрерывной случайной величины. Пусть непрерывная случайная величина X, определяемая плотностью распределения $p(x)$, принимает значения, принадлежащие отрезку $[a, b]$. Отрезок $[a, b]$ разобьем на n элементарных отрезков $[a, x_1]$, $[x_1, x_2]$, ..., $[x_{n-1}, b]$, длины которых выражаются формулой $\Delta x_k = x_k - x_{k-1}$ $(k=1, 2, \ldots, n)$, $x_0 = a$, $x_n = b$. В каждом из элементарных отрезков $[x_{k-1}, x_k]$ выберем произвольно точку ξ_k и составим произведение $p(\xi_k)\Delta x_k$, которое приближенно выражает вероятность p_k попадания значений X в интервал $]x_{k-1}, x_k[$ (см. (29.14)).

Предположив дать определение понятия математического ожидания для непрерывной случайной величины по аналогии с указанным понятием для дискретной случайной величины, образуем сумму произведений значений ξ_k на вероятность $p_k = p(\xi_k)\Delta x_k$ их попадания в соответствующий интервал:

$$\sum_{k=1}^{n} \xi_k p_k = \sum_{k=1}^{n} \xi_k p(\xi_k)\Delta x_k.$$

Переходя к пределу в этой сумме при $\lambda \to 0$ $(\lambda = \max \Delta x_k)$, получаем

$$\lim_{\lambda \to 0} \sum_{k=1}^{n} \xi_k p(\xi_k)\Delta x_k = \int_a^b x p(x)\, dx.$$

Этот определенный интеграл называют математическим ожиданием рассматриваемой непрерывной случайной величины X, т. е. по определению полагают

$$M(X) = \int_a^b x p(x)\, dx. \tag{30.9}$$

Если значения непрерывной случайной величины X принадлежат бесконечному интервалу $]-\infty, +\infty[$, то ее математическое ожидание определяется формулой

$$M(X) = \int_{-\infty}^{+\infty} x p(x)\, dx \tag{30.10}$$

при условии, что интеграл сходится абсолютно.

Математическим ожиданием функции $f(X, Y)$ двух непрерывных случайных величин X и Y с плотностью совместного распределения $p(x, y)$ называют число

$$M(f(X, Y)) = \iint\limits_{\Omega} f(x, y)\, p(x, y)\, dxdy, \qquad (30.11)$$

где Ω — вся плоскость Oxy, если этот интеграл сходится абсолютно.

Математическое ожидание непрерывной случайной величины обладает теми же свойствами, что и математическое ожидание дискретной случайной величины.

Докажем, например, свойство 5. В соответствии с определением (30.11)

$$M(XY) = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} xyp(x, y)\, dxdy,$$

где $p(x, y)$ — плотность распределения двумерной случайной величины (X, Y). Поскольку X, Y — независимые случайные величины и $p(x, y) = p_1(x)p_2(y)$, где $p_1(x)$, $p_2(y)$ — плотности распределения случайных величин X и Y соответственно (см. § 29.6), то

$$M(XY) = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} xyp(x, y)\, dxdy = \int\limits_{-\infty}^{+\infty} \int\limits_{-\infty}^{+\infty} xyp_1(x)p_2(y)dxdy =$$

$$= \int\limits_{-\infty}^{+\infty} xp_1(x)\, dx \cdot \int\limits_{-\infty}^{+\infty} yp_2(y)\, dy = M(X) \cdot M(Y).$$

Отметим, что математическое ожидание случайной величины есть величина постоянная.

Разность $X - M(X)$ называется *отклонением* случайной величины X от ее математического ожидания $M(X)$. Отклонение является случайной величиной. Докажем, что математическое ожидание отклонения равно нулю. Действительно, так как математическое ожидание случайной величины — величина постоянная и математическое ожидание постоянной равно этой постоянной, то

$M(X - M(X)) = M(X) - M(M(X)) = M(X) - M(X) = 0$, т. е.

$$M(X - M(X)) = 0. \qquad (30.12)$$

Это равенство объясняется тем, что отклонения могут быть как положительными, так и отрицательными; в результате их взаимного погашения среднее значение отклонения равно нулю.

§ 30.2. Дисперсия случайной величины

Математическое ожидание характеризует случайную величину не полностью: зная математическое ожидание, нельзя сказать, какие значения принимает случайная величина и как они отклоняются от среднего значения. Чтобы знать, как рассеяны значения случайной величины вокруг ее математического ожидания, вводят другую числовую характеристику, называемую дисперсией. Из равенства (30.12) видно, что за меру рассеяния нельзя принять отклонение случайной величины от ее математического ожидания. Вследствие этого рассматривают модули отклонений или их квадраты, чаще последние.

Дисперсией, или *рассеянием*, случайной величины X называется математическое ожидание квадрата ее отклонения. Дисперсия обозначается через $D(X)$, т. е.

$$D(X)=M((X-M(X))^2). \tag{30.13}$$

Из определения и свойств математического ожидания следует, что дисперсия любой случайной величины X неотрицательна, т. е.

$$D(X)\geqslant 0. \tag{30.14}$$

Для вычисления дисперсии применяется формула

$$D(X)=M(X^2)-(M(X))^2, \tag{30.15}$$

получающаяся из (30.13). Чтобы доказать равенство (30.15), преобразуем (30.13), учтя свойства математического ожидания:

$$\begin{aligned} D(X)&=M((X-M(X))^2)=M(X^2-2XM(X)+(M(X))^2)=\\ &=M(X^2)-2M(XM(X))+M((M(X))^2)=\\ &=M(X^2)-2M(X)M(X)+(M(X))^2=M(X^2)-(M(X))^2. \end{aligned}$$

Итак, дисперсия случайной величины равна разности между математическим ожиданием квадрата этой величины и квадратом ее математического ожидания.

Дисперсия случайной величины обладает следующими свойствами.

1. Дисперсия постоянной величины C равна нулю:

$$D(C)=0 \quad (C=\text{const}). \tag{30.16}$$

2. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат:

$$D(CX)=C^2D(X). \qquad (30.17)$$

3. Дисперсия суммы двух независимых случайных величин равна сумме их дисперсий:

$$D(X+Y)=D(X)+D(Y). \qquad (30.18)$$

4. Дисперсия разности двух независимых случайных величин равна сумме их дисперсий:

$$D(X-Y)=D(X)+D(Y). \qquad (30.19)$$

Докажем эти свойства. Принимая во внимание определение дисперсии и свойства математического ожидания, получаем

$$D(C)=M((C-M(C))^2)=M((C-C)^2)=M(0)=0.$$

Первое свойство доказано, оно означает следующее: постоянная величина не имеет рассеяния, так как принимает одно и то же значение.

Из определения дисперсии и свойств математического ожидания следует, что

$$D(CX)=M((CX-M(CX))^2)=M((CX-CM(X))^2)=$$
$$=M(C^2(X-M(X))^2)=C^2M((X-M(X))^2)=C^2D(X).$$

Для доказательства третьего свойства применяем формулу (30.15):

$$D(X+Y)=M((X+Y)^2)-(M(X+Y))^2=$$
$$=M(X^2+2XY+Y^2)-(M(X)+M(Y))^2=$$
$$=M(X^2)+M(2XY)+M(Y^2)-((M(X))^2+2M(X)M(Y)+$$
$$+(M(Y))^2)=M(X^2)+2M(X)M(Y)+M(Y^2)-$$
$$-(M(X))^2-2M(X)M(Y)-(M(Y))^2=$$
$$=M(X^2)-(M(X))^2+M(Y^2)-(M(Y))^2=D(X)+D(Y).$$

Третье свойство распространяется на любое число попарно-независимых случайных величин. Если $X_1, X_2, \ldots, X_n$ — попарно-независимые случайные величины, то

$$D(X_1+X_2+\ldots+X_n)=D(X_1)+D(X_2)+\ldots+D(X_n). \qquad (30.20)$$

Четвертое свойство следует из формул (30.17) и (30.18):

$$D(X-Y)=D(X+(-Y))=D(X)+D(-Y)=$$
$$=D(X)+(-1)^2D(Y)=D(X)+D(Y).$$

Если случайная величина X является дискретной и задан ее закон распределения $P(X=x_k)=p_k$ $(k=1, 2, \ldots, n)$, то случайная величина $(X-M(X))^2$ имеет следующий закон распределения:

$$P((X=x_k-M(X))^2)=p_k \ (k=1, 2, \ldots, n), \ \sum_{k=1}^{n} p_k=1.$$

Исходя из определения математического ожидания, получаем формулу

$$D(X)=\sum_{k=1}^{n}(x_k-M(X))^2 p_k \qquad (30.21)$$

для дисперсии дискретной случайной величины X, принимающей конечное число значений $x_1, x_2, \ldots, x_n$ соответственно с вероятностями $p_1, p_2, \ldots, p_n$.

Если дискретная случайная величина X принимает счетное множество значений $x_1, x_2, \ldots, x_n, \ldots$ соответственно с вероятностями $p_1, p_2, \ldots, p_n, \ldots$, то ее дисперсия определяется формулой

$$D(X)=\sum_{k=1}^{\infty}(x_k-M(X))^2 p_k \qquad (30.22)$$

при условии, что ряд (30.22) сходится, $\sum_{k=1}^{\infty} p_k=1$.

Дисперсия непрерывной случайной величины X, все возможные значения которой принадлежат отрезку $[a, b]$, определяется формулой

$$D(X)=\int_a^b (x-M(X))^2 p(x)\,dx, \qquad (30.23)$$

где $p(x)$ — плотность распределения этой величины.

Дисперсию можно вычислять по формуле

$$D(X)=\int_a^b x^2 p(x)\,dx-(M(X))^2. \qquad (30.24)$$

Действительно, учитывая определение математического ожидания для непрерывной случайной величины (см. (30.9)) и свойства плотности вероятностей $p(x)$ (см. (29.17)), получаем

$$D(X)=\int_a^b (x-M(X))^2 p(x)\,dx=\int_a^b (x^2-2xM(X)+$$

$$+(M(X))^2)p(x)\,dx=\int_a^b x^2p(x)\,dx-2M(X)\int_a^b xp(x)\,dx+$$

$$+(M(X))^2\int_a^b p(x)\,dx=\int_a^b x^2p(x)\,dx-2M(X)M(X)+$$

$$+(M(X))^2=\int_a^b x^2p(x)\,dx-(M(X))^2.$$

В общем случае дисперсия непрерывной случайной величины X определяется формулой

$$D(X)=\int_{-\infty}^{+\infty}(x-M(X))^2p(x)\,dx, \qquad (30.25)$$

если интеграл сходится.

Можно показать, что

$$D(X)=\int_{-\infty}^{+\infty}x^2p(x)\,dx-(M(X))^2. \qquad (30.26)$$

§ 30.3. Среднее квадратическое отклонение. Числовые характеристики среднего арифметического одинаково распределенных случайных величин

Средним квадратическим отклонением, или *стандартным отклонением*, $\sigma(X)$ случайной величины X называется корень квадратный из ее дисперсии:

$$\sigma(X)=\sqrt{D(X)}. \qquad (30.27)$$

Отметим, что это определение имеет смысл, так как выполнено условие (30.14).

Очевидно, размерность дисперсии равна квадрату размерности случайной величины, поэтому среднее квадратическое отклонение имеет ту же размерность, что и случайная величина X. Среднее квадратическое применяется тогда, когда желательно получить оценку рассеяния случайной величины в тех же единицах, в которых выражены значения самой величины.

Докажем, что среднее квадратическое отклонение суммы независимых случайных величин равно корню квадратному из суммы квадратов средних квадратических этих величин.

Пусть $X_1, X_2, \ldots, X_n$ — независимые случайные величины, $D(X_1), D(X_2), \ldots, D(X_n)$ — их дисперсии, $\sigma(X_1)$, $\sigma(X_2), \ldots, \sigma(X_n)$ — средние квадратические отклонения, т. е. $\sigma^2(X_k) = D(X_k)$ $(k=1, 2, \ldots, n)$. Обозначим через X сумму рассматриваемых случайных величин, т. е. $X = X_1 + X_2 + \ldots + X_n$. На основании свойств дисперсии получаем $D(X) = D(X_1) + D(X_2) + \ldots + D(X_n)$, откуда

$$\sqrt{D(X)} = \sqrt{D(X_1) + D(X_2) + \ldots + D(X_n)},$$

или

$$\sigma(X) = \sqrt{\sigma^2(X_1) + \sigma^2(X_2) + \ldots + \sigma^2(X_n)}. \qquad (30.28)$$

Если $X_1, X_2, \ldots, X_n$ — независимые случайные величины, имеющие одинаковые распределения, то они будут иметь одинаковые соответствующие числовые характеристики, т. е. для $k = 1, 2, \ldots, n$

$$M(X_k) = a, D(X_k) = D, \sigma(X_k) = \sigma. \qquad (30.29)$$

Обозначим через $\widetilde{X}$ среднее арифметическое этих случайных величин, т. е.

$$\widetilde{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}. \qquad (30.30)$$

Числовые характеристики среднего арифметического указанных величин обладают следующими свойствами.

1. Математическое ожидание среднего арифметического n независимых одинаково распределенных случайных величин равно математическому ожиданию каждой из этих величин:

$$M(\widetilde{X}) = a. \qquad (30.31)$$

2. Дисперсия среднего арифметического n независимых одинаково распределенных случайных величин в n раз меньше дисперсии каждой из величин:

$$D(\widetilde{X}) = \frac{D}{n}. \qquad (30.32)$$

3. Среднее квадратическое отклонение среднего арифметического n независимых одинаково распределенных

случайных величин в $\sqrt{n}$ раз меньше среднего квадратического отклонения каждой из величин:

$$\sigma(\tilde{X}) = \frac{\sigma}{\sqrt{n}}. \tag{30.33}$$

Докажем эти свойства. Учитывая свойства математического ожидания и условие $M(X_k)=a$ $(k=1, 2, \ldots, n)$, получаем

$$M(\tilde{X}) = M\left(\frac{X_1+X_2+\ldots+X_n}{n}\right) =$$
$$= \frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n} = \frac{na}{n} = a.$$

На основании свойств дисперсии и с учетом равенств $D(X_k)=D$ находим

$$D(\tilde{X}) = D\left(\frac{X_1+X_2+\ldots+X_n}{n}\right) =$$
$$= \frac{D(X_1)+D(X_2)+\ldots+D(X_n)}{n^2} = \frac{nD}{n^2} = \frac{D}{n},$$
$$\sigma(\tilde{X}) = \sqrt{D(\tilde{X})} = \sqrt{\frac{D}{n}} = \frac{\sqrt{D}}{\sqrt{n}} = \frac{\sigma}{\sqrt{n}}.$$

§ 30.4. Ковариация

Ковариацией двух случайных величин X и Y называется математическое ожидание произведения их отклонений от соответствующих математических ожиданий. Обозначим ковариацию случайных величин X, Y через $\operatorname{cov}(X, Y)$, тогда по определению

$$\operatorname{cov}(X, Y) = M((X-M(X))(Y-M(Y))). \tag{30.34}$$

Используя свойства математического ожидания, преобразуем правую часть формулы (30.34):

$$M((X-M(X))(Y-M(Y))) = M(XY-XM(Y)-YM(X)+$$
$$+M(X)M(Y)) = M(XY)-M(XM(Y))-M(YM(X))+$$
$$+M(M(X)M(Y)) = M(XY)-M(Y)M(X)-M(X)M(Y)+$$
$$+M(X)M(Y) = M(XY)-M(X)M(Y).$$

Следовательно,

$$\operatorname{cov}(X, Y) = M(XY)-M(X)M(Y), \tag{30.35}$$

т. е. ковариация случайных величин равна математическому ожиданию их произведения минус произведение их математических ожиданий.

Очевидно,

$$\operatorname{cov}(X, X) = D(X), \ \operatorname{cov}(X, Y) = \operatorname{cov}(Y, X); \qquad (30.36)$$

эти соотношения следуют из формулы (30.34).

Покажем, что

$$D(X+Y) = D(X) + D(Y) + 2\operatorname{cov}(X, Y). \qquad (30.37)$$

В самом деле,

$$D(X+Y) = M((X+Y) - M(X+Y))^2 = M((X-M(X)) + (Y-M(Y)))^2 = M((X-M(X))^2 + (Y-M(Y))^2 + 2(X-M(X))(Y-M(Y))), \ D(X+Y) = M(X-M(X))^2 + M(Y-M(Y))^2 + 2M(X-M(X))(Y-M(Y)).$$

Из этого равенства и определений (30.13), (30.34) следует (30.37).

Теорема 30.1. Если для случайных величин $X_1, X_2, \ldots, X_n$ существуют $\operatorname{cov}(X_i, X_j) = \sigma_{ij}$ $(i, j = 1, 2, \ldots, n)$, то при любых постоянных $C_1, C_2, \ldots, C_n$

$$D(C_1X_1 + C_2X_2 + \ldots + C_nX_n) = \sum_{i, j=1}^{n} \sigma_{ij} C_i C_j. \qquad (30.38)$$

Доказательство. Положим $Y_n = C_1X_1 + C_2X_2 + \ldots + C_nX_n = \sum_{i=1}^{n} C_iX_i$. Поскольку

$$M(Y_n) = C_1M(X_1) + C_2M(X_2) + \ldots + C_nM(X_n) = \sum_{i=1}^{n} C_iM(X_i),$$

то

$$Y_n - M(Y_n) = \sum_{i=1}^{n} C_i(X_i - M(X_i)),$$

$$(Y_n - M(Y_n))^2 = \sum_{i, j=1}^{n} C_iC_j(X_i - M(X_i))(X_j - M(X_j)).$$

Вычислив математическое ожидание от обеих частей последнего равенства, получим формулу (30.38).

Следствие. Для любых случайных величин $X_1, X_2, \ldots, X_m$ $(m = 1, 2, \ldots)$ определитель из их ковариаций неотрицателен, т. е.

$$\begin{vmatrix} \operatorname{cov}(X_1, X_1) & \operatorname{cov}(X_1, X_2) & \ldots & \operatorname{cov}(X_1, X_m) \\ \operatorname{cov}(X_2, X_1) & \operatorname{cov}(X_2, X_2) & \ldots & \operatorname{cov}(X_2, X_m) \\ \cdot & \cdot & \cdot & \cdot \\ \operatorname{cov}(X_m, X_1) & \operatorname{cov}(X_m, X_2) & \ldots & \operatorname{cov}(X_m, X_m) \end{vmatrix} \geqslant 0. \qquad (30.39)$$

Действительно, правая часть формулы (30.38) есть квадратичная форма от переменных $C_1, C_2, \ldots, C_n$. Так как левая часть равенства неотрицательна в силу (30.14), то указанная квадратичная форма является неотрицательно определенной, поэтому неотрицательны все главные миноры матрицы, составленной из ее коэффициентов. Отсюда следует неравенство (30.39).

Отметим частный случай этого неравенства. При $m=2$ получаем

$$\begin{vmatrix} \operatorname{cov}(X_1, X_1) & \operatorname{cov}(X_1, X_2) \\ \operatorname{cov}(X_2, X_1) & \operatorname{cov}(X_2, X_2) \end{vmatrix} \geqslant 0,$$

$$\begin{vmatrix} D(X_1) & \operatorname{cov}(X_1, X_2) \\ \operatorname{cov}(X_1, X_2) & D(X_2) \end{vmatrix} \geqslant 0,$$

$$D(X_1)D(X_2)-\operatorname{cov}^2(X_1, X_2) \geqslant 0,$$

$$|\operatorname{cov}(X_1, X_2)| \leqslant \sqrt{D(X_1)D(X_2)}. \qquad (30.40)$$

Теорема 30.2. Если случайные величины X, Y независимы, то их ковариация равна нулю:

$$\operatorname{cov}(X, Y)=0. \qquad (30.41)$$

Доказательство. Так как случайные величины X, Y независимы, то математическое ожидание их произведения равно произведению математических ожиданий и равенство (30.41) следует из формулы (30.35).

Следствие. Если $\operatorname{cov}(X, Y) \neq 0$, то случайные величины зависимы. В самом деле, допустив противное (X и Y независимы), по теореме получили бы $\operatorname{cov}(X, Y)=0$, что противоречит условию.

§ 30.5. Коэффициент корреляции

Согласно следствию из теоремы 30.2, если $\operatorname{cov}(X, Y) \neq$ $\neq 0$, то случайные величины X и Y зависимы. В качестве количественной характеристики степени зависимости случайных величин X, Y используется коэффициент корреляции.

Коэффициентом корреляции $\rho(X, Y)$ случайных величин X, Y называется отношение их ковариации к произведению средних квадратических отклонений этих величин:

$$\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sigma(X)\,\sigma(Y)}. \qquad (30.42)$$

Свойства коэффициента корреляции следующие.

1. Коэффициент корреляции по модулю не превосходит единицы: $|\rho(X, Y)| \leqslant 1$.

2. Если величины X, Y независимы, то коэффициент корреляции их равен нулю: $\rho(X, Y) = 0$.

3. Если $Y = AX + B$, где A и B — постоянные, то $|\rho(X, Y)| = 1$ (коэффициент корреляции случайных величин, связанных линейной зависимостью, по модулю равен единице).

Первое свойство следует из неравенства (30.40) и определений (30.27), (30.42):

$$\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sqrt{D(X)}\,\sqrt{D(Y)}}, \quad \left|\frac{\operatorname{cov}(X, Y)}{\sqrt{D(X)\,D(Y)}}\right| \leqslant 1,$$

$$|\rho(X, Y)| \leqslant 1.$$

Второе свойство следует из теоремы 30.2 и определения (30.42). Докажем третье свойство. Положим $M(X) = a$, $D(X) = \sigma^2$, тогда

$$M(Y) = M(AX+B) = M(AX) + M(B) = Aa + B;$$

$$D(Y) = D(AX+B) = D(AX) + D(B) =$$
$$= A^2 D(X) + D(B) = A^2\sigma^2;$$

$$\sigma(Y) = \sqrt{D(Y)} = |A|\,\sigma;$$

$$\operatorname{cov}(X, Y) = M((X-a)(Y-M(Y))) =$$
$$= M((X-a)(AX+B-(Aa+B))) = AD(X) = A\sigma^2,$$

поэтому

$$\rho(X, Y) = \frac{A\sigma^2}{\sigma|A|\sigma} = \frac{A}{|A|}, \quad |\rho(X, Y)| = 1.$$

Отметим, что $\rho(X, Y) > 0$, если $A > 0$; $\rho(X, Y) < 0$, если $A < 0$.

Замечание. Из равенства $\rho(X, Y) = 0$ не следует независимость случайных величин X и Y.

§ 30.6. Моменты случайных величин

Наряду с рассмотренными числовыми характеристиками случайных величин используются и другие, более общие характеристики — начальные и центральные моменты, через которые, в частности, выражаются математическое ожидание и дисперсия.

Начальным моментом k-го порядка, или *моментом k-го порядка*, случайной величины X называется математическое ожидание k-й степени этой величины, т. е.

$$\nu_k = M(X^k). \tag{30.43}$$

Если дискретная случайная величина X принимает значения $x_1, x_2, \ldots, x_n, \ldots$ с вероятностями $p_1, p_2, \ldots, p_n, \ldots$, то в соответствии с определением

$$\nu_k = \sum_{i=1}^{\infty} x_i^k p_i \tag{30.44}$$

при условии, что этот ряд сходится абсолютно.

Для непрерывной случайной величины X с плотностью распределения $p(x)$

$$\nu_k = \int_{-\infty}^{+\infty} x^k p(x)\, dx, \tag{30.45}$$

если этот интеграл сходится абсолютно.

В первом и во втором случае выполняются соответственно следующие соотношения:

$$\sum_{i=1}^{\infty} p_i = 1, \quad \int_{-\infty}^{+\infty} p(x)\, dx = 1. \tag{30.46}$$

Замечание. Если дискретная случайная величина принимает конечное множество значений $x_1, x_2, \ldots, x_n$ с вероятностями $p_1, p_2, \ldots, p_n$, то $\nu_k = \sum_{i=1}^{n} x_i^k p_i$. Если все значения непрерывной случайной величины принадлежат интервалу $]\alpha, \beta[$, то $\nu_k = \int_{\alpha}^{\beta} x^k p(x)\, dx$.

Выпишем формулы, выражающие моменты k-го порядка при $k=0$, 1, 2 соответственно для дискретной и непрерывной случайной величины X:

$$\nu_0 = M(X^0) = \sum_{i=1}^{\infty} p_i = 1, \quad \nu_0 = M(X^0) = \int_{-\infty}^{+\infty} p(x)\,dx = 1;$$

$$\nu_1 = M(X) = \sum_{i=1}^{\infty} x_i p_i, \quad \nu_1 = M(X) = \int_{-\infty}^{+\infty} xp(x)\,dx;$$

$$\nu_2 = M(X^2) = \sum_{i=1}^{\infty} x_i^2 p_i, \quad \nu_2 = M(X^2) = \int_{-\infty}^{+\infty} x^2 p(x)\,dx.$$

Итак, начальный момент нулевого порядка равен единице; начальный момент первого порядка случайной величины X равен ее математическому ожиданию:

$$\nu_1 = M(X). \tag{30.47}$$

Центральным моментом k-го порядка случайной величины X называется математическое ожидание k-й степени отклонения этой величины от ее математического ожидания. Обозначив центральный момент k-го порядка через μ_k и положив $M(X) = a$, по определению получим

$$\mu_k = M((X-a)^k), \text{ или } \mu_k = M(X-a)^k. \tag{30.48}$$

Если дискретная случайная величина принимает значения $x_1, x_2, \ldots, x_n, \ldots$ соответственно с вероятностями $p_1, p_2, \ldots, p_n, \ldots$, то

$$\mu_k = \sum_{i=1}^{\infty} (x_i - a)^k p_i \tag{30.49}$$

при условии, что ряд сходится абсолютно.

Для непрерывной случайной величины X с плотностью распределения $p(x)$

$$\mu_k = \int_{-\infty}^{+\infty} (x-a)^k p(x)\,dx, \tag{30.50}$$

если этот интеграл сходится абсолютно.

Выпишем формулы, выражающие центральные моменты k-го порядка при $k = 0, 1, 2$ для дискретной и непрерывной величины соответственно:

$$\mu_0 = M(X-a)^0 = \sum_{i=1}^{\infty} p_i = 1, \quad \mu_0 = \int_{-\infty}^{+\infty} p(x)\,dx = 1;$$

$$\mu_1 = M(X-a) = 0 \text{ (см. (30.12))};$$

$$\mu_2 = \sum_{i=1}^{\infty} (x_i - a)^2 p_i, \quad \mu_2 = \int_{-\infty}^{+\infty} (x-a)^2 p(x)\, dx.$$

Таким образом, центральный момент нулевого порядка равен единице; центральный момент первого порядка равен нулю; центральный момент второго порядка случайной величины X равен дисперсии этой величины:

$$\mu_2 = M(X-a)^2 = D(X). \tag{30.51}$$

Центральные моменты порядка $k \geqslant 2$ можно выразить через начальные моменты. Учитывая свойства математического ожидания, определения моментов и полагая $M(X) = a$, получаем

$$\mu_2 = M(X-a)^2 = M(X^2 - 2aX + a^2) =$$
$$= M(X^2) - 2aM(X) + M(a^2) = M(X^2) - 2a^2 + a^2 =$$
$$= M(X^2) - a^2 = \nu_2 - \nu_1^2,$$

$$\mu_2 = \nu_2 - \nu_1^2, \text{ или } D(X) = \nu_2 - \nu_1^2; \tag{30.52}$$

$$\mu_3 = M(X-a)^3 = M(X^3 - 3X^2a + 3Xa^2 - a^3) =$$
$$= M(X^3) - 3aM(X^2) + 3a^2M(X) - M(a^3) =$$
$$= M(X^3) - 3aM(X^2) + 3a^2M(X) - a^3 =$$
$$= \nu_3 - 3\nu_1\nu_2 + 3\nu_1^3 - \nu_1^3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3,$$

$$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3; \tag{30.53}$$

$$\mu_4 = M(X-a)^4 = M(X^4 - 4X^3a + 6X^2a^2 - 4Xa^3 + a^4) =$$
$$= M(X^4) - 4aM(X^3) + 6a^2M(X^2) - 4a^3M(X) + M(a^4) =$$
$$= \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 4\nu_1^4 + \nu_1^4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4,$$

$$\mu_4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4. \tag{30.54}$$

Центральный момент третьего порядка характеризует асимметрию распределения случайной величины. За меру отклонения (асимметрии) распределения случайной величины X от симметрии относительно математического ожидания (центра распределения) принимают число

$$\alpha = \frac{\mu_3}{\sigma^3}, \tag{30.55}$$

где σ — среднее квадратическое отклонение.

Эксцессом случайной величины X называется число, определяемое формулой

$$\varepsilon = \frac{\mu_4}{\sigma^4} - 3, \qquad (30.56)$$

где μ_4 — центральный момент четвертого порядка; σ — среднее квадратическое отклонение.

Глава 31

НЕКОТОРЫЕ ЗАКОНЫ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

В этой главе рассматриваются дискретные распределения: биномиальное, геометрическое, распределение Пуассона; непрерывные распределения: равномерное, показательное, гамма-распределение, нормальное, распределение «хи-квадрат». Приводится формула Бернулли, с которой связано биномиальное распределение.

§ 31.1. Формула Бернулли

Пусть производятся испытания, в каждом из которых может появиться событие A. Если вероятность события A в одном испытании не зависит от появления его в любом другом, то испытания называются *независимыми* относительно события A. Будем считать, что испытания происходят в одинаковых условиях и вероятность появления события A в каждом испытании одна и та же. Обозначим эту вероятность через p, а вероятность появления события $\bar{A}$ через q $(q=1-p)$.

Поставим следующую задачу: найти вероятность того, что в серии из n независимых испытаний событие A появится ровно k раз (и не появится $n-k$ раз). Искомую вероятность обозначим через $P_{k,n}$; например, $P_{3,7}$ означает вероятность того, что событие A появится три раза в семи испытаниях.

Отметим, что порядок, в котором появляется событие A при n испытаниях, может быть самым различным. В частности, если при пяти испытаниях событие A появилось четыре раза, то могут быть следующие комбинации:

$$AAAA\bar{A},\ AAA\bar{A}A,\ AA\bar{A}AA,\ A\bar{A}AAA,\ \bar{A}AAAA$$

(запись $AAAA\bar{A}$ означает, что событие A наступило в первых четырех испытаниях и не наступило в пятом испытании и т. д.).

Итак, событие A может появляться k раз при n испытаниях в различных комбинациях, число которых равно C_n^k, т. е. числу сочетаний из n элементов по k. Примером одной из таких комбинаций может служить событие B, состоящее в том, что событие A появляется при k первых испытаниях и не появляется в остальных $n-k$ испытаниях, т. е.

$$B = A_1 A_2 \ldots A_k \bar{A}_{k+1} \bar{A}_{k+2} \ldots \bar{A}_n,$$

где A_i — появление события A при i-м испытании ($i=1, 2, \ldots, k$); $\bar{A}_j$ — непоявление события A в j-м испытании ($j=k+1, k+2, \ldots, n$).

Так как события A_i и $\bar{A}_j$ независимы и

$$P(A_1) = P(A_2) = \ldots = P(A_n) = p,$$
$$P(\bar{A}_1) = P(\bar{A}_2) = \ldots = P(\bar{A}_n) = q,$$

то по теореме умножения вероятностей независимых событий

$$P(B) = P(A_1 A_2 \ldots A_k \bar{A}_{k+1} \ldots \bar{A}_n) =$$
$$= P(A_1) P(A_2) \ldots P(A_k) P(\bar{A}_{k+1}) \ldots P(\bar{A}_n) = p^k q^{n-k}.$$

Поскольку число всех таких событий (событие появляется k раз и не появляется $n-k$ раз) равно C_n^k и эти события несовместны, то по теореме сложения вероятностей несовместных событий находится искомая вероятность

$$P_{k,n} = C_n^k p^k q^{n-k}, \tag{31.1}$$

где

$$C_n^k = \frac{n(n-1) \ldots (n-(k-1))}{k!}, \quad \text{или} \quad C_n^k = \frac{n!}{k!\,(n-k)!}.$$

Формула (31.1) называется формулой Бернулли.

Рассматриваемые события (событие A наступает k раз и не наступает $n-k$ раз) образуют полную группу событий, поэтому

$$\sum_{k=0}^{n} P_{k,n} = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = 1, \tag{31.2}$$

или

$$q^n + C_n^1 p q^{n-1} + C_n^2 p^2 q^{n-2} + \ldots + C_n^k p^k q^{n-k} + \ldots + p^n = 1.$$

Отметим, что члены этой суммы совпадают с членами разложения бинома

$$(q+p)^n = q^n + C_n^1 p q^{n-1} + C_n^2 p^2 q^{n-2} + \dots + \\ + C_n^k p^k q^{n-k} + \dots + p^n.$$

Отсюда видно, что сумма всех вероятностей в правой части этой формулы равна единице, так как $p+q=1$.

§ 31.2. Биномиальное распределение

Предположим, что в одинаковых условиях производится n независимых испытаний, в результате каждого из которых может появиться событие A с вероятностью p или событие $\bar{A}$ с вероятностью $q\,(q=1-p)$. В каждой серии из n испытаний событие A может либо не появиться, либо появиться 1 раз, 2 раза, ..., n раз. Введем в рассмотрение дискретную случайную величину X — число появлений события A при n испытаниях. Найдем закон распределения этой случайной величины. Величина X может принимать следующие значения: $x_0=0$, $x_1=1$, $x_2=2, \dots, x_n=n$.

Вероятность p_k того, что величина X принимает значение x_k, вычисляется по формуле Бернулли

$$p_k = P_{k,n} = C_n^k p^k q^{n-k} \quad (k=0,\ 1,\ 2,\ \dots,\ n).$$

Закон распределения дискретной случайной величины, определяемый формулой Бернулли, называется *биномиальным*. Постоянные n и p, входящие в формулу (31.1), называются параметрами биномиального распределения $(q=1-p)$. Название «биномиальный» объясняется тем, что правую часть равенства (31.1) можно рассматривать как общий член разложения бинома $(q+p)^n$. Первый член этого разложения q^n означает вероятность того, что в n испытаниях событие A не появится ни разу, второй член $C_n^1 p q^{n-1} = npq^{n-1}$ — вероятность того, что событие A появится один раз, наконец, последний член $C_n^n p^n = p^n$ — вероятность того, что событие A появится n раз.

Биномиальный закон распределения дискретной случайной величины можно представить в виде следующей таблицы:

X	0	1	...	k	...	n
p	q^n	$nq^{n-1}p$	...	$C_n^k p^k q^{n-k}$	...	p^n

Рассмотрим вопрос о математическом ожидании числа появлений события A в независимых испытаниях, на который дает ответ следующая

Теорема 31.1. Математическое ожидание числа появлений события A в n независимых испытаниях равно произведению вероятности этого события в каждом испытании на число всех испытаний.

Доказательство. Случайную величину X — число появлений события A при n независимых испытаниях — можно представить в виде суммы $X=X_1+X_2+\ldots+X_n$, где X_k — число появлений события в k-м испытании ($k=1, 2, \ldots, n$). В соответствии с формулой (30.5) имеем $M(X)=M(X_1)+M(X_2)+\ldots+M(X_n)$.

Случайная величина X_k — число появлений события A при одном испытании — может принимать только два значения: $x_1=1$ (событие A наступило) с вероятностью p и $x_2=0$ (событие A не наступило) с вероятностью q ($q=1-p$). Следовательно, $M(X_k)=x_1p_1+x_2p_2=1\cdot p+0\cdot q=p$,

$$M(X_k)=p \quad (k=1, 2, \ldots, n). \qquad (31.3)$$

Используя формулу (31.3), получаем $M(X)=p+p+\ldots+p=np$,

$$M(X)=np. \qquad (31.4)$$

Замечание. Поскольку рассматриваемая случайная величина X распределена по биномиальному закону, полученный результат можно сформулировать следующим образом: математическое ожидание дискретной случайной величины, распределенной по биномиальному закону с параметрами n и p, равна произведению np.

Вопрос о дисперсии числа появлений некоторого события при n независимых испытаниях решает следующая

Теорема 31.2. Дисперсия числа появлений события A при n независимых испытаниях, в каждом из которых вероятность p его появления одна и та же, равна произведению числа всех испытаний на вероятность появления и вероятность непоявления данного события.

Доказательство. Случайную величину X —

число появлений события A при n испытаниях — можно представить как сумму $X=X_1+X_2+\ldots+X_n$, где X_k — число появлений события A при k-м испытании ($k=1, 2, \ldots, n$). В соответствии с формулой (30.20) имеем

$$D(X)=D(X_1)+D(X_2)+\ldots+D(X_n).$$

С помощью формулы (30.21) вычислим $D(X_k)$ для $k=1, 2, \ldots, n$. Случайная величина X_k принимает лишь два значения: $x_1=1$ с вероятностью p и $x_2=0$ с вероятностью $q=1-p$. Случайная величина $(X_k-M(X_k))^2$ также принимает только два значения: $(1-p)^2=q^2$ с вероятностью p и $(0-p)^2=p^2$ с вероятностью q. Принимая во внимание равенства (31.3), по формуле (30.21) получаем

$$D(X_k)=M((X_k-M(X_k))^2)=q^2p+p^2q=pq(q+p)=pq,$$

$$D(X_k)=pq \quad (k=1, 2, \ldots, n). \tag{31.5}$$

Следовательно, $D(X)=D(X_1)+D(X_2)+\ldots+D(X_n)=pq+pq+\ldots+pq=npq$,

$$D(X)=npq. \tag{31.6}$$

Замечание. Поскольку случайная величина X распределена по биномиальному закону, полученный результат означает следующее: дисперсия случайной величины, распределенной по биномиальному закону с параметрами p и n, равна произведению npq.

Среднее квадратическое отклонение случайной величины X, распределенной по биномиальному закону с параметрами n и p, определяется формулой

$$\sigma(X)=\sqrt{npq}, \tag{31.7}$$

которая следует из формул (30.27) и (31.6).

§ 31.3. Геометрическое распределение

Рассмотрим игру с набрасыванием кольца на стержень. Обозначим буквой X число бросаний до первого попадания на стержень при условии, что вероятность попадания при каждом бросании не зависит от результатов предыдущих бросаний и имеет одно и то же значение p $(0<p<1)$. Величина X будет дискретной случайной величиной, возможными значениями которой являются натуральные числа. Найдем закон распределения вероятностей этой случайной величины. Событие $X=1$

означает попадание с первого бросания, его вероятность равна p, т. е. $P(X=1)=p$. Событие $X=2$ означает попадание при втором бросании и, значит, промах при первом бросании. Применяя теорему умножения вероятностей для независимых событий, получаем $P(X=2)=qp$, где $q=1-p$. Событие $X=3$ означает попадание при третьем бросании и, следовательно, промахи при первых двух бросаниях, поэтому $P(X=3)=q\cdot q\cdot p=q^2p$. Продолжая аналогичные рассуждения, находим общую формулу:

$$P(X=m)=q^{m-1}p \quad (m=1, 2, \ldots),$$

выражающую вероятность попадания кольца на стержень при m-м бросании.

Геометрическим распределением называется распределение дискретной случайной величины X, определяемое формулой

$$P(X=m)=(1-p)^{m-1}p,\ 0<p<1 \quad (m=1,\ 2,\ \ldots). \qquad (31.8)$$

Это название связано с тем, что ряд вероятностей (31.8) представляет собой бесконечно убывающую геометрическую прогрессию со знаменателем $q=1-p$; сумма этого ряда равна единице:

$$p+qp+q^2p+\ \ldots\ +q^{m-1}p+\ \ldots\ =p\frac{1}{1-q}=1.$$

Вычислим математическое ожидание и дисперсию дискретной случайной величины, имеющей геометрическое распределение.

Поскольку множество возможных значений величины X бесконечно, на основании формулы (30.8) ее математическое ожидание есть сумма ряда

$$M(X)=\sum_{m=1}^{\infty} mP(X=m)=\sum_{m=1}^{\infty} mq^{m-1}p=p\sum_{m=1}^{\infty} mq^{m-1}$$
$$(q=1-p).$$

Так как члены ряда $\sum\limits_{m=1}^{\infty} mq^{m-1}$ являются производными соответствующих членов ряда $\sum\limits_{m=0}^{\infty} q^m$ и

$$\sum_{m=0}^{\infty} q^m=1+q+q^2+q^3+\ldots=\frac{1}{1-q},$$

то

$$\sum_{m=1}^{\infty} mq^{m-1} = 1+2q+3q^2+\ldots = \left(\frac{1}{1-q}\right)' = \frac{1}{(1-q)^2}.$$

Следовательно, $M(X) = p\,\dfrac{1}{(1-q)^2} = p\,\dfrac{1}{p^2} = \dfrac{1}{p}$,

$$M(X) = \frac{1}{p}\ (p>0). \tag{31.9}$$

Для вычисления дисперсии воспользуемся формулой (30.15) и формулами для сумм степенных рядов:

$$\sum_{m=1}^{\infty} mq^{m-1} = \frac{1}{(1-q)^2}, \quad \sum_{m=2}^{\infty} m(m-1)\,q^{m-2} = \frac{2}{(1-q)^3}.$$

Найдем сначала математическое ожидание квадрата величины X:

$$M(X^2) = \sum_{m=1}^{\infty} m^2q^{m-1}p = p\,\Big(\sum_{m=1}^{\infty} m(m-1)\,q^{m-1} +$$

$$+\sum_{m=1}^{\infty} mq^{m-1}\Big) = p\left(q\,\frac{2}{(1-q)^3} + \frac{1}{(1-q)^2}\right) = p\,\frac{2q+1-q}{(1-q)^3} =$$

$$= p\,\frac{1+q}{p^3} = \frac{1+q}{p^2}.$$

Таким образом,

$$D(X) = M(X^2) - (M(X))^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2} = \frac{1-p}{p^2},$$

$$D(X) = \frac{1-p}{p^2}. \tag{31.10}$$

§ 31.4. Распределение Пуассона

Пусть в одинаковых условиях производится n независимых испытаний, в каждом из которых может появиться событие A с вероятностью p или событие $\bar{A}$ с вероятностью $q\,(q=1-p)$. Вероятность того, что при n испытаниях событие A появится k раз (и не появится $n-k$ раз), определяется формулой Бернулли.

Рассмотрим случай, когда n является достаточно большим, а p — достаточно малым; положим $np=a$, где a — некоторое число. Найдем закон распределения дискретной случайной величины X — числа появлений события A при n испытаниях в указанных предположениях. Так как в данном случае $p = \frac{a}{n}$, $q = 1 - \frac{a}{n}$, то формула (31.1) принимает вид

$$P_{k,n} = C_n^k p^k q^{n-k} = \frac{n(n-1)(n-2)\ \ldots\ (n-(k-1))}{k!} \times$$

$$\times \left(\frac{a}{n}\right)^k \left(1 - \frac{a}{n}\right)^{n-k} = \frac{a^k}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \ldots$$

$$\ldots \left(1 - \frac{k-1}{n}\right)\left(1 - \frac{a}{n}\right)^n \left(1 - \frac{a}{n}\right)^{-k}.$$

Переходя к пределу при $n \to \infty$, находим

$$\lim_{n\to\infty} P_{k,n} = \frac{a^k}{k!} \lim_{n\to\infty}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \ldots \left(1 - \frac{k-1}{n}\right) \times$$

$$\times \lim_{n\to\infty}\left(1 - \frac{a}{n}\right)^n \cdot \lim_{n\to\infty}\left(1 - \frac{a}{n}\right)^{-k} = \frac{a^k}{k!}\cdot 1 \cdot e^{-a} \cdot 1 = \frac{a^k e^{-a}}{k!}.$$

Следовательно, получена приближенная формула $P_{k,n} \approx \approx \frac{a^k e^{-a}}{k!}$. Распределение дискретной случайной величины, определяемое формулой

$$P_{k,n} = \frac{a^k e^{-a}}{k!}, \qquad (31.11)$$

называется *распределением Пуассона*. Постоянную $a = np$, входящую в формулу (31.11), называют параметром этого распределения. Закон распределения Пуассона можно написать в виде следующей таблицы:

x_k	0	1	2	...	k	...
p_k	e^{-a}	ae^{-a}	$\frac{a^2 e^{-a}}{2!}$	...	$\frac{a^k e^{-a}}{k!}$	...

Отметим, что $\sum_{k=0}^{\infty} p_k = 1$. Действительно,

$$\sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} \frac{a^k e^{-a}}{k!} = e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} =$$

$$= e^{-a}\left(1 + \frac{a}{1!} + \frac{a^2}{2!} + \ldots + \frac{a^k}{k!} + \ldots\right) = e^{-a} e^{a} = 1.$$

Теорема 31.3. Математическое ожидание и дисперсия случайной величины X, распределенной по закону Пуассона, одинаковы и каждая из этих числовых характеристик равна параметру a, т. е.

$$M(X)=a, \quad D(X)=a. \tag{31.12}$$

Доказательство. Поскольку закон распределения случайной величины X определяется указанной таблицей, по формуле (30.8) находим

$$M(X)=0\cdot e^{-a}+1\frac{a}{1!}e^{-a}+2\frac{a^2}{2!}e^{-a}+\ldots+k\frac{a^k}{k!}e^{-a}+$$
$$+\ldots=\sum_{k=1}^{\infty}\frac{a^k}{(k-1)!}e^{-a}=ae^{-a}\sum_{k=1}^{\infty}\frac{a^{k-1}}{(k-1)!}=ae^{-a}e^{a}=a.$$

Первое из равенств (31.12) доказано. Оно означает, что центром распределения данной случайной величины X является число a.

Дисперсию $D(X)$ вычислим по формуле (30.15), для чего сначала найдем математическое ожидание квадрата данной величины:

$$M(X^2)=0^2\cdot e^{-a}+1^2\frac{a}{1!}e^{-a}+2^2\frac{a^2}{2!}e^{-a}+\ldots+$$
$$+k^2\frac{a^k}{k!}e^{-a}+\ldots=1\cdot\frac{a}{0!}e^{-a}+2\frac{a^2}{1!}e^{-a}+\ldots+$$
$$+k\frac{a^k}{(k-1)!}e^{-a}+\ldots.$$

Положив $k-1=m$, получим

$$M(X^2)=\sum_{k=1}^{\infty}k\frac{a^k}{(k-1)!}e^{-a}=\sum_{m=0}^{\infty}(m+1)\frac{a^{m+1}}{m!}e^{-a}=$$
$$=\sum_{m=0}^{\infty}m\frac{a^{m+1}}{m!}e^{-a}+\sum_{m=0}^{\infty}\frac{a^{m+1}}{m!}e^{-a}=a^2e^{-a}\sum_{m=1}^{\infty}\frac{a^{m-1}}{(m-1)!}+$$
$$+ae^{-a}\sum_{m=0}^{\infty}\frac{a^m}{m!}=a^2e^{-a}e^{a}+ae^{-a}e^{a}=a^2+a.$$

По формуле (30.15) находим $D(X)=M(X^2)-(M(X))^2=$ $=a^2+a-a^2=a$.

§ 31.5. Равномерное распределение

Случайная величина X называется равномерно распределенной на отрезке $[\alpha, \beta]$, если плотность вероятностей этой величины постоянна на данном отрезке и равна нулю вне этого отрезка. Из определения следует, что

$$p(x)=\begin{cases} c, & \text{если } \alpha\leqslant x\leqslant\beta, \\ 0, & \text{если } x<\alpha \text{ или } x>\beta. \end{cases}$$

Поскольку должно быть выполнено условие (29.16), то

$$\int\limits_{-\infty}^{+\infty} p(x)\,dx=\int\limits_{\alpha}^{\beta} c\,dx=1, \quad c(\beta-\alpha)=1, \quad c=\frac{1}{\beta-\alpha}.$$

Следовательно, плотность вероятности $p(x)$ определяется так:

$$p(x)=\begin{cases} \dfrac{1}{\beta-\alpha}, & \text{если } \alpha\leqslant x\leqslant\beta, \\ 0, & \text{если } x<\alpha \text{ или } x>\beta. \end{cases} \tag{31.13}$$

Для случайной величины X, равномерно распределенной на отрезке $[\alpha, \beta]$, вероятность попадания в интервал $]\gamma, \delta[$, принадлежащий этому отрезку, пропорциональна длине интервала:

$$P(\gamma<X<\delta)=\int\limits_{\gamma}^{\delta} p(x)\,dx=\int\limits_{\gamma}^{\delta}\frac{dx}{\beta-\alpha},$$

$$P(\gamma<X<\delta)=\frac{\delta-\gamma}{\beta-\alpha}. \tag{31.14}$$

Замечание. Выражение «выберем наудачу точку x на отрезке $[\alpha, \beta]$» означает, что координата точки представляет случайную величину с равномерным распределением вероятностей на этом отрезке. Сравните (31.14) и (28.6).

Найдем функцию распределения $F(x)$ по заданной плотности распределения вероятностей (31.13). Применяя формулу (29.10), получаем: при $x\leqslant\alpha$

$$F(x)=\int\limits_{-\infty}^{x} p(u)\,du=0;$$

при $\alpha<x<\beta$

$$F(x)=\int\limits_{-\infty}^{x}p(u)\,du=\int\limits_{-\infty}^{\alpha}p(x)\,dx+\int\limits_{\alpha}^{x}p(u)\,du=$$

$$=\int\limits_{\alpha}^{x}\frac{du}{\beta-\alpha}=\frac{x-\alpha}{\beta-\alpha};$$

при $x\geqslant\beta$

$$F(x)=\int\limits_{-\infty}^{x}p(u)\,du=\int\limits_{-\infty}^{\alpha}p(x)\,dx+\int\limits_{\alpha}^{\beta}p(x)\,dx+$$

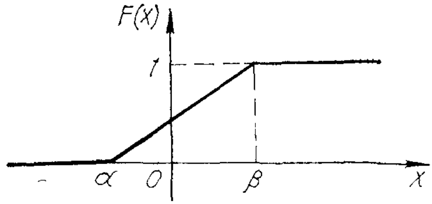

Рис. 31.1

$$+\int\limits_{\beta}^{x}p(u)\,du=\int\limits_{\alpha}^{\beta}\frac{dx}{\beta-\alpha}=1.$$

Итак, функция распределения имеет вид (рис. 31.1)

$$F(x)=\begin{cases}0 & \text{при } x\leqslant\alpha,\\ \dfrac{x-\alpha}{\beta-\alpha} & \text{при } \alpha<x<\beta,\\ 1 & \text{при } x\geqslant\beta.\end{cases}$$

Вычислим математическое ожидание равномерно распределенной случайной величины:

$$M(X)=\int\limits_{-\infty}^{+\infty}xp(x)\,dx=\int\limits_{-\infty}^{\alpha}xp(x)\,dx+\int\limits_{\alpha}^{\beta}xp(x)\,dx+$$

$$+\int\limits_{\beta}^{+\infty}xp(x)\,dx=\int\limits_{\alpha}^{\beta}\frac{xdx}{\beta-\alpha}=\frac{\beta^2-\alpha^2}{2(\beta-\alpha)}=\frac{\beta+\alpha}{2}.$$

Следовательно, математическое ожидание случайной величины, равномерно распределенной на отрезке, есть середина этого отрезка.

Найдем дисперсию равномерно распределенной случайной величины с плотностью распределения (31.13):

$$D(X)=\int_{-\infty}^{+\infty}(x-a)^2p(x)\,dx=\int_{-\infty}^{\alpha}(x-a)^2p(x)\,dx+$$

$$+\int_{\alpha}^{\beta}(x-a)^2p(x)\,dx+\int_{\beta}^{+\infty}(x-a)^2p(x)\,dx=\int_{\alpha}^{\beta}\frac{(x-a)^2}{\beta-\alpha}\,dx=$$

$$=\frac{\left(x-\frac{\alpha+\beta}{2}\right)^3}{3(\beta-\alpha)}\Bigg|_{\alpha}^{\beta}=\frac{\left(\beta-\frac{\alpha+\beta}{2}\right)^3-\left(\alpha-\frac{\alpha+\beta}{2}\right)^3}{3(\beta-\alpha)}=$$

$$=\frac{(\beta-\alpha)^2}{12}.$$

Случайная величина (X, Y) равномерно распределена в области G, если плотность распределения ее вероятностей постоянна в этой области и равна нулю вне области G:

$$p(x, y)=\begin{cases} C \text{ при } (x, y)\in G, \\ 0 \text{ при } (x, y)\notin G. \end{cases} \qquad (31.15)$$

Вероятность попадания точки (X, Y) в любую область g, лежащую внутри G, пропорциональна площади S_g области g:

$$P((X, Y)\in g)=C\cdot S_g. \qquad (31.16)$$

Поскольку попадание в область G — достоверное событие, то

$$P((X, Y)\in G)=C\cdot S_G=1,$$

откуда $C=1/S_G$. Подставляя это значение C в формулу (31.16), получаем

$$P((X, Y)\in g)=\frac{S_g}{S_G};$$

вероятность попадания в область g равна отношению площади области g к площади всей области G.

Замечание. Двумерная случайная величина (X, Y), где X, Y — координаты точки, наудачу брошенной в область G, является равномерно распределенной в этой области.

§ 31.6. Показательное распределение

Если плотность распределения случайной величины X выражается функцией

$$p(x)=\begin{cases}0 & \text{при } x\leqslant 0,\\ \alpha e^{-\alpha x} & \text{при } x>0 \quad (\alpha>0),\end{cases} \tag{31.17}$$

то говорят, что случайная величина X имеет показательное распределение. Кривая распределения вероятностей рассматриваемой величины представлена на рис. 31.2.

Убедимся в том, что эта функция удовлетворяет условию (29.16):

$$\int_{-\infty}^{+\infty} p(x)\,dx=\int_{-\infty}^{0} p(x)\,dx+\int_{0}^{+\infty} p(x)\,dx=\int_{0}^{+\infty}\alpha e^{-\alpha x}\,dx=$$

$$=-e^{-\alpha x}\Big|_0^{+\infty}=-(e^{-\infty}-e^0)=1.$$

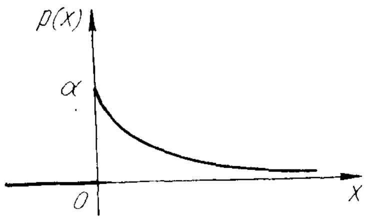

Рис. 31.2

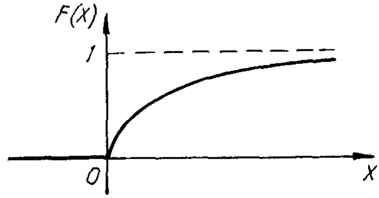

Рис. 31.3

Найдем функцию распределения данной случайной величины с помощью формулы (29.10). При $x\leqslant 0$

$$F(x)=\int_{-\infty}^{x} p(t)\,dt=\int_{-\infty}^{x} 0\,dt=0.$$

Если $x>0$, то

$$F(x)=\int_{-\infty}^{x} p(u)\,du=\int_{-\infty}^{0} p(u)\,du+\int_{0}^{x} p(u)\,du=\int_{-\infty}^{0} 0\,du+$$

$$+\int_{0}^{x}\alpha e^{-\alpha u}\,du=-e^{-\alpha u}\Big|_0^{x}=-(e^{-\alpha x}-e^0)=1-e^{-\alpha x}.$$

Следовательно, функция распределения имеет вид

$$F(x)=\begin{cases}0 & \text{при } x\leqslant 0,\\ 1-e^{-\alpha x} & \text{при } x>0.\end{cases}$$

График функции $F(x)$ изображен на рис. 31.3.

Отметим, что при показательном распределении случайной величины X маловероятны ее большие значения. Действительно, пусть $\frac{N}{\alpha}$ — некоторое положительное число. Найдем вероятность попадания X в интервал $\left]\frac{N}{\alpha}, +\infty\right[$. По второй из формул (29.13) получаем

$$P\left(\frac{N}{\alpha} < X < +\infty\right) = \int\limits_{\frac{N}{\alpha}}^{+\infty} \alpha e^{-\alpha x}\, dx = -e^{-\alpha x}\Big|_{\frac{N}{\alpha}}^{+\infty} =$$

$$= -\left(e^{-\infty} - e^{-\alpha\frac{N}{\alpha}}\right) = e^{-N}.$$

Легко убедиться в том, что уже при $N=5$ вероятность меньше 0,01.

С помощью метода интегрирования по частям можно показать, что

$$M(X) = \frac{1}{\alpha}, \quad D(X) = \frac{1}{\alpha^2}. \tag{31.18}$$

Замечание. Показательный закон распределения вероятностей встречается во многих задачах, связанных с простейшим потоком событий. Под потоком событий понимают последовательность событий, наступающих одно за другим в случайные моменты. Например, поток вызовов на телефонной станции, поток заявок в системе массового обслуживания и др.

§ 31.7. Гамма-распределение

Гамма-распределение является обобщением показательного распределения. Плотность этого распределения задается функцией

$$p(x) = \begin{cases} 0 & \text{при } x \leqslant 0, \\ Cx^{\lambda-1}e^{-\alpha x} & \text{при } x > 0. \end{cases} \tag{31.19}$$

Параметры λ и α могут быть любыми положительными числами. Значение множителя C определяется из условия (29.16), которое в данном случае принимает вид

$$\int\limits_{0}^{+\infty} Cx^{\lambda-1}e^{-\alpha x}\, dx = 1.$$

Сравним интеграл в этом равенстве с гамма-функцией Эйлера (см. § 16.11):

$$\Gamma(\lambda) = \int_0^{+\infty} z^{\lambda-1} e^{-z} dz.$$

Положим $\alpha x = z$, тогда $x = \frac{z}{\alpha}$, $dx = \frac{dz}{\alpha}$,

$$\int_0^{+\infty} C x^{\lambda-1} e^{-\alpha x} dx = \int_0^{+\infty} C \frac{z^{\lambda-1}}{\alpha^{\lambda-1}} e^{-z} \frac{dz}{\alpha} = \frac{C}{\alpha^{\lambda}} \int_0^{+\infty} z^{\lambda-1} e^{-z} dz =$$

$$= \frac{C}{\alpha^{\lambda}} \Gamma(\lambda) = 1, \qquad C = \frac{\alpha^{\lambda}}{\Gamma(\lambda)}. \tag{31.20}$$

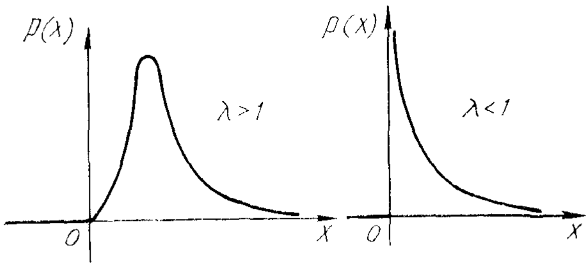

Рис. 31.4

Следовательно,

$$p(x) = \begin{cases} 0 \text{ при } x \leqslant 0, \\ \frac{\alpha^{\lambda}}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x} \text{ при } x > 0. \end{cases} \tag{31.21}$$

На рис. 31.4 показан вид кривых распределения вероятностей при значениях параметра $\lambda > 1$, $\lambda < 1$ (при $\lambda = 1$ получаем показательное распределение, см. рис. 31.2). В случае $\lambda > 1$ кривая распределения имеет один максимум. Точка максимума находится из условия $p'(x) = 0$:

$$p'(x) = Ce^{-\alpha x}((\lambda - 1)x^{\lambda-2} - \alpha x^{\lambda-1}), \quad x = \frac{\lambda - 1}{\alpha} \ (\lambda > 1).$$

В случае $\lambda < 1$ плотность распределения убывает в интервале $]0, +\infty[$.

§ 31.8. Нормальное распределение

Среди всех непрерывных законов распределения вероятностей особое место занимает распределение с плотностью

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}\quad(\sigma>0). \tag{31.22}$$

Это распределение называется *нормальным распределением*, или *распределением Гаусса*, с параметрами a и σ. О случайной величине X с указанным законом распределения вероятностей говорят, что она распределена нормально с параметрами a, σ, и кратко называют нормальной.

Убедимся в том, что функция (31.22) удовлетворяет условию (29.16). В интеграле

$$\int_{-\infty}^{+\infty}p(x)\,dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{(x-a)^2}{2\sigma^2}}dx \tag{31.23}$$

перейдем к новой переменной по формуле

$$t=\frac{x-a}{\sigma}, \tag{31.24}$$

тогда $x=a+\sigma t$, $dx=\sigma dt$. Поскольку новые пределы интегрирования равны старым, то

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}\sigma dt=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=\frac{1}{\sqrt{2\pi}}\cdot\sqrt{2\pi}=1.$$

Здесь принято во внимание равенство

$$\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=\sqrt{2\pi}, \tag{31.25}$$

которое следует из формулы (16.57); интеграл (31.25) с помощью подстановки $t=\sqrt{2}\,x$ сводится к интегралу в формуле (16.57).

Выясним вероятностный смысл параметров a и σ ($\sigma>0$), входящих в функцию (31.22). Покажем, что

$$a=M(X),\ \sigma=\sqrt{D(X)},\ \text{или}\ \sigma^2=D(X). \tag{31.26}$$

В соответствии с формулой (30.10) получаем

$$M(X)=\int_{-\infty}^{+\infty}xp(x)\,dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}xe^{-\frac{(x-a)^2}{2\sigma^2}}dx. \tag{31.27}$$

Введем новую переменную t по формуле (31.24) и преобразуем этот интеграл:

$$M(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(a+\sigma t)\,e^{-\frac{t^2}{2}}\sigma dt=\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt+$$

$$+\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}te^{-\frac{t^2}{2}}dt.$$

Первый из полученных интегралов равен $\sqrt{2\pi}$ (см. (31.25)), второй равен нулю:

$$\int_{-\infty}^{+\infty}te^{-\frac{t^2}{2}}dt=-\int_{-\infty}^{+\infty}d\left(e^{-\frac{t^2}{2}}\right)=-e^{-\frac{t^2}{2}}\Big|_{-\infty}^{+\infty}=0. \quad (31.28)$$

Следовательно, $M(X)=\frac{a}{\sqrt{2\pi}}\sqrt{2\pi}+0=a.$

Учитывая определение (30.23) и равенство $M(X)=a$, находим

$$D(X)=\int_{-\infty}^{+\infty}(x-a)^2p(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}(x-a)^2e^{-\frac{(x-a)^2}{2\sigma^2}}dx. \quad (31.29)$$

Воспользовавшись подстановкой (31.24), получим

$$D(X)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2\sigma^2e^{-\frac{t^2}{2}}\sigma dt=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2e^{-\frac{t^2}{2}}dt.$$

Применяем метод интегрирования по частям. Положим $t=u$, $te^{-\frac{t^2}{2}}dt=dv$, тогда $du=dt$, $v=-e^{-\frac{t^2}{2}}$ (ибо $dv=d\left(-e^{-\frac{t^2}{2}}\right)$),

$$D(X)=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t\cdot te^{-\frac{t^2}{2}}dt=-\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}td\left(e^{-\frac{t^2}{2}}\right)=$$

$$=-\frac{\sigma^2}{\sqrt{2\pi}}\left(te^{-\frac{t^2}{2}}\Big|_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt\right)=-\frac{\sigma^2}{\sqrt{2\pi}}te^{-\frac{t^2}{2}}\Big|_{-\infty}^{+\infty}+$$

$$+\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-\frac{t^2}{2}}dt=0+\frac{\sigma^2}{\sqrt{2\pi}}\sqrt{2\pi}=\sigma^2;\ D(X)=\sigma^2,$$

$$\sqrt{D(X)}=\sigma.$$

Итак, равенства (31.26) доказаны: параметр a — математическое ожидание нормальной случайной величины X, σ — среднее квадратическое отклонение этой величины.

График функции (31.22) называется *нормальной кривой*, или *кривой Гаусса*. Построим нормальную кривую, для чего предварительно исследуем функцию при фиксированных значениях a и σ ($\sigma>0$).

Функция (31.22) определена при всех x, т. е. область ее определения — бесконечный интервал $]-\infty, +\infty[$. Функция принимает только положительные значения при всех $x\in]-\infty, +\infty[$ (это означает, что ее график находится выше оси Ox). Обозначим функцию (31.22) через $f(x)$, найдем ее экстремум.

Так как производная

$$f'(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}\left(-\frac{2(x-a)}{2\sigma^2}\right),\ f'(x)=-\frac{x-a}{\sigma^3\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}$$

обращается в нуль при $x=a$ и при переходе через точку $x=a$ меняет знак с плюса на минус, т. е. $f'(a)=0$, $f'(x)>0$ при $x<a$ и $f'(x)<0$ при $x>a$, то $x=a$ — точка максимума данной функции.

Вычислим значение максимума функции (31.22):

$$\max f(x)=f(a)=\frac{1}{\sigma\sqrt{2\pi}}e^0,\quad \max f(x)=\frac{1}{\sigma\sqrt{2\pi}}.$$

Следовательно, значение максимума зависит от параметра σ: при возрастании значений σ максимум функции убывает, с убыванием σ максимум возрастает.

Поскольку вторая производная

$$f''(x)=-\frac{1}{\sigma^3\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}\left(1-\frac{(x-a)^2}{\sigma^2}\right)$$

обращается в нуль при $x=a-\sigma$, $x=a+\sigma$ и меняет знак при переходе через эти значения, то они являются абсциссами точек перегиба, а точки

$$M_1\left(a-\sigma,\ \frac{1}{\sigma\sqrt{2\pi e}}\right),\quad M_2\left(a+\sigma,\ \frac{1}{\sigma\sqrt{2\pi e}}\right)$$

— точками перегиба нормальной кривой.

График нормальной кривой симметричен относительно прямой $x=a$, так как функция (31.22) принимает одинаковые значения при значениях x, симметричных относительно точки $x=a$, т. е. $f(x_1)=f(x_2)$ при $x_1=a-\varepsilon$ и $x_2=a+\varepsilon$.

Ось Ox является асимптотой графика функции, ибо

$$\lim_{x\to-\infty} f(x)=\lim_{x\to-\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}=0,$$

$$\lim_{x\to+\infty} f(x)=\lim_{x\to+\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma^2}}=0.$$

На рис. 31.5 изображены три нормальные кривые при одном и том же значении a и различных значениях σ.

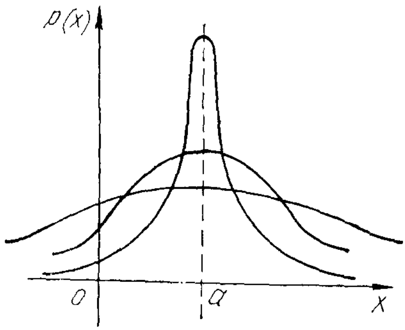

Рис. 31.5

Замечание 1. С изменением значений параметра σ форма нормальной кривой изменяется. Как уже отмечалось, максимальная ордината графика функции (31.22) убывает с возрастанием значения σ (кривая «сжимается» к оси Ox) и возрастает с убыванием значения σ (кривая «растягивается» в положительном направлении оси Oy).

Изменение значений параметра a (при неизменном значении σ) не влияет на форму нормальной кривой; кривая сдвигается вдоль оси Ox вправо, если a возрастает, и влево, если a убывает.

Замечание 2. В случае, когда $a=0$, $\sigma=1$, функция (31.22) принимает вид

$$p(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}. \qquad (31.30)$$

Распределение вероятностей непрерывной случайной величины, определяемое функцией (31.30), называется *нормированным* (в отличие от общего нормального распределения, определяемого функцией (31.22)), или *стандартным*. График функции (31.30) называется нормированной кривой.

§ 31.9. Вероятность попадания в заданный интервал нормальной случайной величины. Правило трех сигм

Как известно, вероятность попадания значений случайной величины в заданный интервал равна определенному интегралу от ее плотности распределения по соответствующему промежутку (см. (29.13)). Если случайная величина имеет нормальный закон распределения вероятностей (31.22), то

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-\frac{(x-a)^2}{2\sigma^2}} dx. \quad (31.31)$$

Преобразуем этот интеграл с помощью новой переменной t:

$$\frac{x-a}{\sigma} = t, \ x = a + \sigma t, \ dx = \sigma dt, \ t_1 = \frac{\alpha - a}{\sigma}, \ t_2 = \frac{\beta - a}{\sigma};$$

$$P(\alpha < X < \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} (\sigma dt) = \frac{1}{\sqrt{2\pi}} \int_{\frac{\alpha-a}{\sigma}}^{0} e^{-\frac{t^2}{2}} dt +$$

$$+ \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\beta-a}{\sigma}} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_{0}^{\frac{\alpha-a}{\sigma}} e^{-\frac{t^2}{2}} dt.$$

Введя в рассмотрение функцию

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} dt, \quad (31.32)$$

предыдущее равенство перепишем так:

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \quad (31.33)$$

Функция (31.32) называется *функцией Лапласа*. Нетрудно убедиться, что функция Лапласа является нечетной, т. е.

$$\Phi(-x) = -\Phi(x). \quad (31.34)$$

Действительно,

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{-x} e^{-\frac{t^2}{2}} dt = -\frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{u^2}{2}} du =$$

$$= -\frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{t^2}{2}} dt = -\Phi(x).$$

Во втором интеграле перешли к новой переменной по формуле $u=-t$, в третьем — переменную интегрирования обозначили буквой t. Таблицы значений функции Лапласа даны в приложениях к учебникам по теории вероятностей (см., например, [5, 6]).

Вычислим вероятность неравенства $|X-a|<\delta$, где X — нормально распределенная случайная величина. Так как в этом случае $-\delta<X-a<\delta$, $a-\delta<X<a+\delta$, то в соответствии с формулой (31.33) имеем

$$P(|X-a|<\delta)=P(a-\delta<X<a+\delta)=$$
$$=\Phi\left(\frac{(a+\delta)-a}{\sigma}\right)-\Phi\left(\frac{(a-\delta)-a}{\sigma}\right)=\Phi\left(\frac{\delta}{\sigma}\right)-\Phi\left(-\frac{\delta}{\sigma}\right).$$

Поскольку функция Лапласа является нечетной, т. е. $\Phi\left(-\frac{\delta}{\sigma}\right)=-\Phi\left(\frac{\delta}{\sigma}\right)$, то

$$P(|X-a|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right). \qquad (31.35)$$

Полагая в этой формуле $\delta=\sigma t$, получаем

$$P(|X-a|<\sigma t)=2\Phi(t).$$

Если $t=3$, т. е. $\sigma t=3\sigma$, то $P(|X-a|<3\sigma)=2\Phi(3)=$ $=2\cdot 0{,}49865=0{,}9973$,

$$P(|X-a|<3\sigma)=0{,}9973. \qquad (31.36)$$

Последнее соотношение означает, что событие, состоящее в осуществлении неравенства $|X-a|<3\sigma$, имеет вероятность, близкую к единице, т. е. является почти достоверным. Формула (31.36) выражает правило трех сигм: если случайная величина распределена по нормальному закону, то модуль ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.

§ 31.10. Нормальное распределение двумерной случайной величины

Распределение двумерной случайной величины (X, Y) называется *нормальным*, если ее плотность распределения определяется функцией

$$p(x, y)=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{(x-a)^2}{2\sigma_1^2}-\frac{(y-b)^2}{2\sigma_2^2}}, \qquad (31.37)$$

где a, b, σ_1, σ_2 — постоянные, причем $\sigma_1>0$, $\sigma_2>0$. Величины σ_1 и σ_2 называются главными средними квадратическими отклонениями. Точка (a, b) называется центром рассеивания двумерной случайной величины (X, Y) с законом распределения (31.37). Функцию (31.37) можно представить в виде

$$p(x, y)=\frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma_1^2}}\cdot\frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{(y-b)^2}{2\sigma_2^2}}$$

и рассматривать как произведение двух плотностей нормальных распределений случайных величин:

$$p_1(x,)=\frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{(x-a)^2}{2\sigma_1^2}}, \quad p_2(y)=\frac{1}{\sigma_2\sqrt{2\pi}}e^{-\frac{(y-b)^2}{2\sigma_2^2}}.$$

Функция (31.37) принимает постоянное значение при x, y, удовлетворяющих условию

$$\frac{(x-a)^2}{\sigma_1^2}+\frac{(y-b)^2}{\sigma_2^2}=c^2 \quad (c=\mathrm{const}). \qquad (31.38)$$

Уравнение (31.38) определяет линии уровня функции (31.37). Эти линии представляют собой эллипсы с полуосями $c\sigma_1$ и $c\sigma_2$ и называются *эллипсами рассеивания*. Оси эллипсов называются осями рассеивания. Указанные эллипсы подобны между собой, имеют один и тот же центр, совпадающий с центром рассеивания (a, b).

В случае, когда $a=0$, $b=0$, функция (31.37) принимает вид

$$p(x,y)=\frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{x^2}{2\sigma_1^2}-\frac{y^2}{2\sigma_2^2}}. \qquad (31.39)$$

Графиком функции (31.39) является поверхность, изображенная на рис. 31.6.

В частном случае $\sigma_1=\sigma_2=\sigma$ нормальное распределение называется круговым и определяется функцией

$$p(x, y)=\frac{1}{2\pi\sigma^2}e^{-\frac{(x-a)^2}{2\sigma^2}-\frac{(y-b)^2}{2\sigma^2}}. \qquad (31.40)$$

Эллипсы рассеивания (31.38) обращаются при этом в окружности

$$\frac{(x-a)^2}{\sigma^2}+\frac{(y-b)^2}{\sigma^2}=c^2,\ (x-a)^2+(y-b)^2=c^2\sigma^2. \quad (31.41)$$

В соответствии с (29.24) вероятность попадания точки (X, Y) в круг D, ограниченный окружностью (31.41), вычисляется по формуле

$$P((X,\ Y)\in D)=\iint\limits_D p(x,\ y)\,dxdy.$$

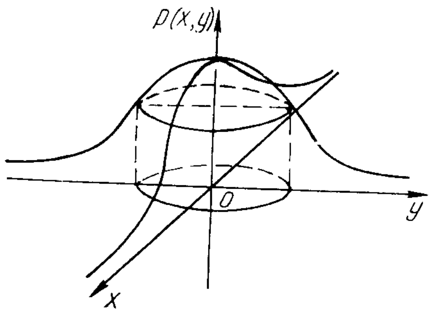

Рис. 31.6

Подставляя сюда выражение (31.40) для $p(x, y)$ и переходя к полярным координатам (r, φ) с полюсом в центре рассеивания (a, b), получаем

$$P((X,Y)\in D)=\frac{1}{2\pi\sigma^2}\int\limits_0^{2\pi}d\varphi\int\limits_0^{c\sigma}e^{-\frac{r^2}{2\sigma^2}}rdr=1-e^{-\frac{c^2}{2}}. \quad (31.42)$$

Отметим, что по этой же формуле находится и вероятность попадания точки в область D_c, ограниченную эллипсом рассеивания (31.39), при общем нормальном двумерном распределении (31.37).

§ 31.11. Распределение «хи-квадрат»

Распределение суммы квадратов независимых случайных величин, подчиненных стандартному нормальному закону распределения, обозначается χ^2 и называется «хи-квадрат». Обозначим исходные величины буквами $X_1, X_2, \ldots, X_k$, а сумму их квадратов — χ_k^2, т. е.

$$\chi_k^2 = X_1^2 + X_2^2 + \ldots + X_k^2.$$

Число слагаемых k в этой формуле называется числом степеней свободы величины χ_k^2. Каждая величина X_i $(i=1, 2, \ldots, k)$ имеет стандартное нормальное распределение, т. е. распределение, определяемое функцией (31.30).

Покажем, что квадрат случайной величины X со стандартным нормальным распределением имеет гамма-распределение с параметрами $\lambda=1/2$, $\alpha=1/2$. Действительно, с помощью формул § 29.7 можно показать, что

$$p_2(y) = p_1(\sqrt{y})\frac{1}{\sqrt{y}}, \tag{31.43}$$

где $p_1(x)$ — плотность распределения случайной величины X; $p_2(y)$ — плотность распределения случайной величины $Y=X^2$ (отметим, что функция $y=x^2$ строго монотонна в каждом из интервалов $]-\infty, 0[$, $]0, +\infty[$).

Если случайная величина X имеет стандартное нормальное распределение с плотностью

$$p_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

то для случайной величины $Y=X^2$ по формуле (31.43) получаем

$$p_2(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} y^{-\frac{1}{2}} \quad (y>0). \tag{31.44}$$

Формула (31.44) означает, что случайная величина $Y=X^2$ имеет гамма-распределение с параметрами $\alpha=1/2$, $\lambda=1/2$.

Можно показать, что если независимые случайные величины X и Y имеют гамма-распределения соответственно с параметрами α, λ и α, μ, то их сумма $X+Y$ — случайная величина, имеющая гамма-распределение с параметрами α, $\lambda+\mu$. Это свойство верно для любого числа слагаемых — независимых случайных величин. Таким образом, случайная величина $U_k = X_1^2 + X_2^2 + \ldots + X_k^2$ имеет гамма-распределение с параметрами $\alpha = \frac{1}{2}$, $\lambda = k \cdot \frac{1}{2}$, т. е. χ_k^2 — распределение с числом степеней свободы k — есть гамма-распределение с параметрами $\alpha = 1/2$, $\lambda = k/2$. Плотность этого распределения выражается функцией

$$p(u) = \begin{cases} C_k e^{-\frac{u}{2}} u^{\frac{k}{2}-1} & \text{при } u \geqslant 0, \\ 0 & \text{при } u < 0, \end{cases} \tag{31.45}$$

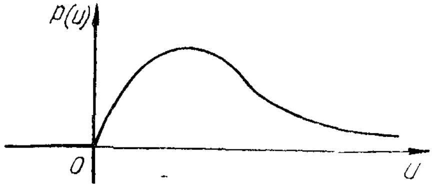

Рис. 31.7

где

$$C_k = \frac{\alpha^\lambda}{\Gamma(\lambda)} = \frac{1}{2^{k/2}\Gamma(k/2)}. \tag{31.46}$$

Например, при $k = 4$ получаем

$$C_4 = \frac{1}{2^2\Gamma(2)} = \frac{1}{4}, \quad p(u) = \frac{u}{4} e^{-\frac{u}{2}} \quad (u > 0).$$

График функции $p(u)$ изображен на рис. 31.7.

Глава 32

ЗАКОН БОЛЬШИХ ЧИСЕЛ. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

Под законом больших чисел понимают общий принцип, в силу которого совместное действие случайных факторов приводит при определенных весьма общих условиях к результату, почти не зависящему от случая. Этот принцип выражается рядом теорем; в числе их теорема Бернулли, теорема Чебышева. Для доказательства указанных теорем используется неравенство Чебышева.

§ 32.1. Неравенства Чебышева

Теорема 32.1. Вероятность того, что модуль отклонения случайной величины X от ее математического ожидания $M(X)$ меньше любого положительного числа ε, больше или равна разности $1-\frac{D(X)}{\varepsilon^2}$:

$$P(|X-M(X)|<\varepsilon)\geqslant 1-\frac{D(X)}{\varepsilon^2}, \qquad (32.1)$$

где $D(X)$ — дисперсия случайной величины X.

Докажем теорему для дискретной случайной величины X, заданной законом распределения $P(x_i)=p_i$ $(i=1, 2, \ldots, n)$.

События $|X-M(X)|<\varepsilon$ и $|X-M(X)|\geqslant\varepsilon$ являются противоположными, поэтому сумма их вероятностей равна единице, т. е.

$$P(|X-M(X)|<\varepsilon)+P(|X-M(X)|\geqslant\varepsilon)=1,$$

откуда

$$P(|X-M(X)|<\varepsilon)=1-P(|X-M(X)|\geqslant\varepsilon). \qquad (32.2)$$

Дисперсия рассматриваемой случайной величины X определяется формулой

$$D(X)=(x_1-M(X))^2p_1+(x_2-M(X))^2p_2+\ldots+ \\ +(x_n-M(X))^2p_n. \qquad (32.3)$$

Предположим, что значения случайной величины занумерованы так, что

$$|x_1-M(X)|<\varepsilon,\ |x_2-M(X)|<\varepsilon, \ldots,\ |x_k-M(X)|<\varepsilon, \\ |x_{k+1}-M(X)|\geqslant\varepsilon,\ |x_{k+2}-M(X)|\geqslant\varepsilon, \ldots,\ |x_n-M(X)|\geqslant\varepsilon.$$

Поскольку обе части каждого из последних неравенств положительны, то

$$(x_{k+1}-M(X))^2\geqslant\varepsilon^2, \\ (x_{k+2}-M(X))^2\geqslant\varepsilon^2, \ldots,\ (x_n-M(X))^2\geqslant\varepsilon^2.$$

Принимая во внимание эти неравенства и тот факт, что каждое слагаемое суммы (32.3) неотрицательно, получаем

$$D(X)\geqslant(x_{k+1}-M(X))^2p_{k+1}+(x_{k+2}-M(X))^2p_{k+2}+\ldots+ \\ +(x_n-M(X))^2p_n\geqslant\varepsilon^2(p_{k+1}+p_{k+2}+\ldots+p_n), \\ D(X)\geqslant\varepsilon^2(p_{k+1}+p_{k+2}+\ldots+p_n). \qquad (32.4)$$

В соответствии с теоремой сложения вероятностей сумма $p_{k+1}+p_{k+2}+\ldots+p_n$ выражает вероятность того, что случайная величина X примет одно из значений x_i $(i=k+1, k+2, \ldots, n)$, для которого $|x_i-M(X)| \geq \varepsilon$, т. е.

$$p_{k+1}+p_{k+2}+\ldots+p_n=P(|X-M(X)| \geq \varepsilon).$$

Подставим это выражение в неравенство (32.4):

$$D(X) \geq \varepsilon^2 P(|X-M(X)| \geq \varepsilon),$$

откуда

$$P(|X-M(X)| \geq \varepsilon) \leq \frac{D(X)}{\varepsilon^2}. \qquad (32.5)$$

Неравенство (32.5) называется *неравенством Чебышева.* Подставив это неравенство в (32.2), получим неравенство (32.1), которое назовем вторым неравенством Чебышева. Отметим, что (32.1) верно и для непрерывных случайных величин.

§ 32.2. Теорема Чебышева

Случайная величина принимает значения, зависящие от многих причин, учесть которые не представляется возможным; вследствие этого невозможно предвидеть, какое значение примет она в результате данного опыта. Поставим вопрос, обладает ли закономерностью поведение суммы достаточно большого числа случайных величин. На первый взгляд может показаться, что такой закономерности нет, как и для каждой случайной величины в отдельности. Но это не так. Поведение суммы достаточно большого числа случайных величин при некоторых условиях утрачивает случайный характер и становится закономерным. Эти условия указаны, например, в теореме Чебышева.

Теорема 32.2. Если случайные величины $X_1, X_2, \ldots, X_n$ независимы, имеют математические ожидания и дисперсии, каждая из которых ограничена одним и тем же числом C, то для любого положительного числа ε выполняется неравенство

$$P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-\right.\right.$$
$$\left.\left.-\frac{M(X_1)+M(X_2)+\ldots+M(X_n)}{n}\right|<\varepsilon\right) \geq 1-\frac{C}{n\varepsilon^2}. \qquad (32.6)$$

Доказательство. Введем в рассмотрение среднее арифметическое данных случайных величин:

$$X = \frac{X_1 + X_2 + \ldots + X_n}{n}, \quad X = \frac{1}{n}\sum_{k=1}^{n} X_k. \qquad (32.7)$$

Так как

$$M(X) = \frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n} = \frac{1}{n}\sum_{k=1}^{n} M(X_k),$$

то неравенство

$$\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{M(X_1) + M(X_2) + \ldots + M(X_n)}{n}\right| < \varepsilon$$

принимает вид $|X - M(X)| < \varepsilon$.

Применяя неравенство (32.1) к случайной величине (32.7), получаем

$$P(|X - M(X)| < \varepsilon) \geqslant 1 - \frac{D(X)}{\varepsilon^2}. \qquad (32.8)$$

Учитывая условие $D(X_k) \leqslant C$ $(k = 1, 2, \ldots, n)$ и свойства дисперсии, находим

$$D(X) = D\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) =$$

$$= \frac{D(X_1) + D(X_2) + \ldots + D(X_n)}{n^2} \leqslant \frac{C + C + \ldots + C}{n^2} =$$

$$= \frac{nC}{n^2} = \frac{C}{n}, \quad D(X) \leqslant \frac{C}{n}.$$

Из последнего неравенства и (32.8) следует неравенство

$$P(|X - M(X)| < \varepsilon) \geqslant 1 - \frac{C}{n\varepsilon^2},$$

или

$$P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} M(X_k)\right| < \varepsilon\right) \geqslant 1 - \frac{C}{n\varepsilon^2}, \qquad (32.9)$$

что равносильно (32.6). Доказанная теорема называется *теоремой Чебышева.*

Введем понятие сходимости по вероятности. Последовательность случайных величин $X_1, X_2, \ldots, X_n$ сходится по вероятности к случайной величине X, если для любого числа $\varepsilon > 0$

$$\lim_{n \to \infty} P(|X_n - X| < \varepsilon) = 1 \qquad (32.10)$$

или $P(|X_n - X| < \varepsilon) \to 1$ при $n \to \infty$; обозначение: $X_n \xrightarrow[p]{} X$.

Замечание 1. Из сходимости по вероятности $x_n \xrightarrow[p]{} a$ не следует, что $\lim\limits_{n\to\infty} x_n = a$ в смысле математического анализа; неравенство $|x_n - a| < \varepsilon$ может не выполняться для отдельных $n > N$.

Следствие. Если случайные величины $X_1, X_2, \ldots, X_n$ удовлетворяют условиям теоремы 32.2, то

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} M(X_k)\right| < \varepsilon\right) = 1. \qquad (32.11)$$

Действительно, переходя к пределу при $n \to \infty$ в неравенстве (32.9), получаем

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} M(X_k)\right| < \varepsilon\right) \geqslant 1.$$

Поскольку вероятность не может быть больше единицы, отсюда следует (32.11).

Теорема Чебышева верна для дискретных и непрерывных случайных величин. Отметим следствие из частного случая теоремы Чебышева. Если все случайные величины имеют одно и то же математическое ожидание a: $M(X_k) = a$ $(k = 1, 2, \ldots, n)$, то $\frac{1}{n}\sum\limits_{k=1}^{n} M(X_k) = \frac{na}{n} = a$ и формула (32.11) принимает вид

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - a\right| < \varepsilon\right) = 1. \qquad (32.12)$$

Замечание 2. Теорема Чебышева имеет важное практическое значение. Этой теоремой пользуются, например, когда при измерении некоторой величины в качестве ее значения берут среднее арифметическое результатов всех проведенных измерений.

В самом деле, результаты n измерений можно рассматривать как случайные величины $X_1, X_2, \ldots, X_n$. Предположим, что выполнены следующие условия: 1) величины $X_1, X_2, \ldots, X_n$ попарно-независимы (результат каждого измерения не зависит от результата любого из остальных измерений); 2) эти величины имеют одно и то же математическое ожидание (измерения проводятся без систематических ошибок — ошибок одного знака; математические ожидания всех величин одинаковы и равны значению измеряемой величины); 3) дисперсии всех величин ограничены одной и той же постоянной (прибор обеспечивает определенную точность измерений; хотя результаты отдельных измерений различны, но рассеяние их ограничено).

Если указанные условия выполнены, можно применить теорему Чебышева (ее частный случай, выражаемый формулой (32.12)), т. е. при достаточно большом числе измерений почти достоверно, что среднее арифметическое полученных результатов как угодно мало отличается от истинного значения измеряемой величины.

Отметим, что среднее арифметическое будет получено лишь с точностью, не превышающей точности прибора (не следует считать, что при увеличении числа измерений можно достичь сколь угодно большой точности).

На теореме Чебышева основан выборочный метод, широко применяемый в статистике. Суть этого метода заключается в том, что по некоторой выборке судят о всей совокупности исследуемых объектов (например, о качестве некоторой массы зерна по небольшой пробе).

§ 32.3. Теорема Бернулли

Теорема 32.3. Если m — число появлений события A в n независимых испытаниях и p — вероятность появления события A в каждом испытании, то при достаточно больших n как угодно близка к единице вероятность того, что модуль отклонения относительной частоты от вероятности меньше любого числа $\varepsilon > 0$, т. е.

$$\lim_{n \to \infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1. \qquad (32.13)$$

Доказательство. Пусть X_k — число появлений события A при k-м испытании ($k=1, 2, \ldots, n$). Каждая из случайных величин X_k может принимать лишь два значения: $x_1=1$ с вероятностью p (событие A наступило) и $x_2=0$ с вероятностью $q=1-p$ (событие A не наступило), поэтому все они имеют одно и то же математическое ожидание a, причем $a=M(X_k)=p$ (см. (31.3)), одну и ту же дисперсию $D=D(X_k)=pq$ (см. (31.5)).

Поскольку $pq \leqslant \frac{1}{4}$ ($p+q=1$, произведение принимает наибольшее значение при $p=q=0{,}5$), то к рассматриваемым величинам $X_1, X_2, \ldots, X_n$ применима теорема Чебышева. Формула (32.12) принимает вид

$$\lim_{n \to \infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - p\right| < \varepsilon\right) = 1. \qquad (32.14)$$

Так как $\sum_{k=1}^{n} X_k = m$, где m — число появлений события A при n испытаниях, то

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1.$$

Доказанная теорема называется *теоремой Бернулли* [1].

Замечание. Теорема Бернулли подтверждает обоснованность статистического определения вероятности.

§ 32.4. Понятие о центральной предельной теореме. Теоремы Лапласа

Многие задачи теории вероятностей связаны с изучением суммы независимых случайных величин, которая при определенных условиях имеет распределение, близкое к нормальному. Эти условия выражаются центральной предельной теоремой, один из вариантов которой принадлежит русскому математику, академику Петербургской академии А. М. Ляпунову (1857—1918). Приведем без доказательства теорему, относящуюся к случаю, когда все случайные величины имеют одинаковые распределения.

Теорема. Если $X_1, X_2, \ldots, X_n$ — независимые случайные величины, имеющие одно и то же распределение с математическим ожиданием a и дисперсией σ^2, то при неограниченном возрастании n закон распределения суммы $X = X_1 + X_2 + \ldots + X_n$ неограниченно приближается к нормальному.

Следствие. Если случайные величины $X_1, X_2, \ldots$ $\ldots, X_n$ удовлетворяют условию теоремы, то их среднее арифметическое при достаточно большом n также имеет распределение, близкое к нормальному.

Действительно, каждое слагаемое суммы

$$\tilde{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{X_1}{n} + \frac{X_2}{n} + \ldots + \frac{X_n}{n}$$

---

[1] Эта теорема опубликована в 1713 г., она положила начало теории вероятностей как науке. Доказательство Я. Бернулли было очень сложным. Более простое доказательство дал П. Л. Чебышев в 1846 г.

имеет математическое ожидание $a:n$ и дисперсию $\sigma^2:n^2$, т. е.

$$M\left(\frac{X_i}{n}\right)=\frac{M(X_i)}{n}=\frac{a}{n},\quad D\left(\frac{X_i}{n}\right)=\frac{D(X_i)}{n^2}=\frac{\sigma^2}{n^2}$$
$$(i=1,\ 2,\ \ldots,\ n),$$

величины $\frac{X_i}{n}$ $(i=1,\ 2,\ \ldots,\ n)$ удовлетворяют условиям теоремы Ляпунова.

Из теоремы Ляпунова следует локальная теорема Лапласа. Обратимся к случайной величине X — числу k появления события A при n независимых испытаниях. Ее можно представить так: $X=X_1+X_2+\ldots+X_n$, где X_i — число появления события A при i-м испытании ($i=1$, $2, \ldots, n$). Поскольку при всех i $M(X_i)=p$, $D(X_i)=pq$ (см. § 31.2), то величины X_i ($i=1, 2, \ldots, n$) удовлетворяют условиям теоремы Ляпунова. Их сумма $X=X_1+$ $+X_2+\ldots+X_n$ имеет распределение, близкое к нормальному: это распределение определяется формулой (31.22). Как известно, $M(X)=np$, $D(X)=npq$ (см. § 31.2). Подставляя в формулу (31.22) значения $x=k$, $a=np$, $\sigma=$ $=\sqrt{npq}$, заключаем: вероятность того, что событие A при n испытаниях появится ровно k раз, приближенно выражается формулой

$$P_{k,n}=\frac{1}{\sqrt{npq}}\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{(k-np)^2}{2npq}}. \tag{32.15}$$

Итак, получена следующая

Локальная теорема Лапласа. Если вероятность появления события A в каждом из n независимых испытаний равна одной и той же постоянной p $(0<p<1)$, то вероятность $P_{k,n}$ того, что во всех этих испытаниях событие A появится ровно k раз, приближенно выражается формулой (32.15), или

$$P_{k,n}=\frac{1}{\sqrt{npq}}\,\varphi(x)\quad\text{при}\quad x=\frac{k-np}{\sqrt{npq}}, \tag{32.16}$$

где

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}}. \tag{32.17}$$

Отметим, что таблицы значений функции (32.17) даны в приложениях к учебникам и учебным пособиям по теории вероятностей (см., например, [5, 6]). Приведем без доказательства еще одну теорему.

Интегральная теорема Лапласа. Если вероятность появления события A в каждом из n независимых испытаний равна одной и той же постоянной p $(0<p<1)$, то вероятность $P_n(k_1, k_2)$ того, что в этих испытаниях событие A появится не менее k_1 раз и не более k_2 раз, приближенно определяется формулой

$$P_n(k_1, k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{x^2}{2}} dx, \quad x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}. \tag{32.18}$$

Эту формулу можно представить в другом виде:

$$P_n(k_1, k_2) = \frac{1}{\sqrt{2\pi}} \int_{x_1}^{0} e^{-\frac{x^2}{2}} dx + \frac{1}{\sqrt{2\pi}} \int_{0}^{x_2} e^{-\frac{x^2}{2}} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{0}^{x_2} e^{-\frac{x^2}{2}} dx - \frac{1}{\sqrt{2\pi}} \int_{0}^{x_1} e^{-\frac{x^2}{2}} dx,$$

или

$$P_n(k_1, k_2) = \Phi(x_2) - \Phi(x_1), \tag{32.19}$$

где $\Phi(x)$ — функция Лапласа (см. (31.32)).

Глава 33

ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Математическая статистика занимается разработкой методов сбора, описания и обработки опытных данных, т. е. результатов наблюдений, с целью получения научных и практических выводов.

§ 33.1. Выборочный метод. Основные понятия

Статистической совокупностью называется множество однородных объектов, объединенных по некоторому общему отличительному признаку. Примеры статистических совокупностей: множество рабочих данного цеха при изучении вопроса о количестве выпускаемой продукции, множество населения данной страны при исследовании ее трудовых ресурсов. Отличительные признаки:

производительность труда в первом случае, возрастной состав населения во втором.

Пусть требуется изучить некоторый признак статистической совокупности. Для этого можно провести сплошное обследование. Но если число объектов достаточно велико, то осуществить указанное обследование не представляется возможным. Если же изучение связано с уничтожением объекта (например, при определении продолжительности времени работы электронного оборудования) или с большими материальными затратами, то сплошное обследование не имеет смысла. Вследствие этого для изучения интересующего признака применяется выборочный метод. Сущность этого метода заключается в том, что обследованию подвергаются не все объекты совокупности, а только некоторая их часть, случайно выбранная из данной совокупности; выводы, полученные при изучении этой части, распространяются на всю совокупность объектов.

Введем основные определения и понятия, связанные с выборочным методом.

Генеральной совокупностью называется совокупность всех однородных объектов, подлежащих изучению. *Выборочной совокупностью*, или *выборкой*, называется совокупность объектов, случайно отобранных из генеральной совокупности. *Объемом* совокупности (генеральной или выборочной) называется число ее объектов. Например, если из 10 000 изготовленных деталей для обследования отобрано 100, то объем генеральной совокупности $N=10\,000$, объем выборки $n=100$. Выборка бывает *повторной* (с возвращением исследуемого объекта в генеральную совокупность) и *бесповторной* (без указанного возвращения). На практике чаще используется бесповторная выборка. Выборка должна быть *репрезентативной* (*представительной*), т. е. такой, по которой можно уверенно судить об интересующем признаке всей генеральной совокупности. Выборка будет репрезентативной, если ее осуществить случайным образом.

При изучении некоторого признака выборочной совокупности производятся испытания (наблюдения). Пусть посредством независимых испытаний, проведенных в одинаковых условиях, получены следующие числовые значения: $x^{(1)}, x^{(2)}, \ldots, x^{(n)}$, где n — объем выборки. Расположим эти значения в порядке их возрастания:

$$x_1, x_2, \ldots, x_n \quad (x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n). \tag{33.1}$$

Последовательность наблюдаемых значений x_i ($i=1, 2, \ldots, n$), записанных в возрастающем порядке (33.1), называется *дискретным вариационным рядом*, а сами эти значения x_i называют *вариантами*. Среди вариант могут оказаться равные, тогда дискретный вариационный ряд можно записать так:

$$\begin{matrix} x_1, & x_2, & \ldots, & x_k, \\ n_1, & n_2, & \ldots, & n_k, \end{matrix} \tag{33.2}$$

где n_i — частота появления значения x_i, причем

$$n_1+n_2+\ldots+n_k=n, \quad \sum_{i=1}^{k} n_i=n. \tag{33.3}$$

Относительной частотой w_i варианты x_i называется отношение ее частоты к объему выборки:

$$w_i = \frac{n_i}{n}. \tag{33.4}$$

Очевидно,

$$\sum_{i=1}^{k} \frac{n_i}{n} = 1, \quad \sum_{i=1}^{k} w_i = 1, \tag{33.5}$$

т. е. сумма относительных частот всех вариант равна единице.

§ 33.2. Статистическое распределение. Полигон и гистограмма

Статистическим распределением выборки называется соответствие между вариантами и их частотами (или относительными частотами). Статистическое распределение может быть задано, например, с помощью таблицы, в которой указаны варианты и соответствующие им частоты.

Пример. Задано распределение частот выборки объема $n=60$:

x_i	4	10	16	20	24	30
n_i	15	18	6	4	5	12

Найти распределение относительных частот.

Применяя формулу (33.4), вычисляем относительные частоты:

$$w_1=\frac{n_1}{n}=\frac{15}{60}=\frac{1}{4},\quad w_2=\frac{n_2}{n}=\frac{18}{60}=\frac{3}{10},\quad w_3=\frac{n_3}{n}=\frac{6}{60}=\frac{1}{10},$$

$$w_4=\frac{n_4}{n}=\frac{4}{60}=\frac{1}{15},\quad w_5=\frac{n_5}{n}=\frac{5}{60}=\frac{1}{12},\quad w_6=\frac{n_6}{n}=\frac{12}{60}=\frac{1}{5}.$$

Следовательно, распределение относительных частот определяется таблицей

x_i	4	10	16	20	24	30
w_i	$\frac{1}{4}$	$\frac{3}{10}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{12}$	$\frac{1}{5}$

Отметим, что $\sum_{i=1}^{6} w_i=\frac{1}{4}+\frac{3}{10}+\frac{1}{10}+\frac{1}{15}+\frac{1}{12}+\frac{1}{5}=1.$

В целях наглядности статистическое распределение изображается графически. Если статистическое распределение задано перечнем вариант x_i и соответствующих частот n_i, то на оси абсцисс откладывают варианты x_i, на оси ординат — частоты n_i, строят точки $M_1(x_1, n_1)$, $M_2(x_2, n_2)$, ..., $M_k(x_k, n_k)$ и последовательно соединяют их отрезками прямых; полученная ломаная называется *полигоном частот*. Аналогично строится полигон относительных частот, т. е. ломаная, отрезки которой последовательно соединяют точки $M_1(x_1, w_1)$, $M_2(x_2, w_2)$, ..., $M_k(x_k, w_k)$, где w_i — относительная частота x_i. На рис. 33.1 изображен полигон частот распределения, указанного в примере.

Кроме дискретных вариационных рядов рассматриваются интервальные вариационные ряды, в которых значения признака могут меняться непрерывно. Примерами непрерывных случайных величин, которые могут принимать любые значения из соответствующих интервалов, являются: урожай какой-либо зерновой культуры и др.

Пусть имеются результаты измерений непрерывной случайной величины X, для которой a и b — соответственно наименьшее и наибольшее значение. Отрезок $[a, b]$ разобьем на k элементарных отрезков $[x_{i-1}, x_i]$ $(i=1, 2, \ldots, k)$ $(x_0=a, x_k=b)$. Обозначим через n_i число значений величины X, принадлежащих интервалу

$]x_{i-1}, x_i[$. Построим табл. 33.1. Эта таблица называется *интервальным вариационным рядом.* Относительной частотой, соответствующей i-му интервалу, называется отношение частоты n_i к объему выборки n, где $n=$

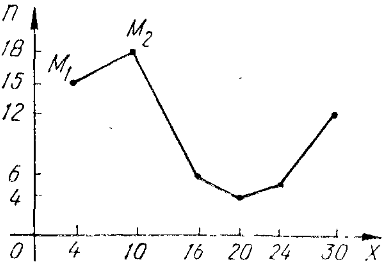

Рис. 33.1

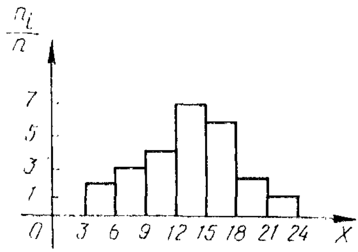

Рис. 33.2

Таблица 33.1

Значения признака	Частота
$]x_0, x_1[$	n_1
$]x_1, x_2[$	n_2
. . . . .	. . .
$]x_{k-1}, x_k[$	n_k
Сумма	n (объем выборки)

$=n_1+n_2+\ldots+n_k$. Очевидно, $\sum\limits_{i=1}^{k}\frac{n_i}{n}=1$. Разности x_1-x_0, x_2-x_1, ..., x_k-x_{k-1} называются *интервальными разностями.* Разность между наибольшим и наименьшим значением признака X, т. е. x_k-x_0 (или $b-a$), называется *размахом вариации.* Плотностью распределения частот на интервале $]x_{i-1}, x_i[$ называется частное

$$\frac{n_i}{x_i-x_{i-1}} \quad (i=1, 2, \ldots, k). \qquad (33.6)$$

Интервальный вариационный ряд будет наиболее простым, когда все интервальные разности равны между собой, т. е. $x_i - x_{i-1} = h$ ($i = 1, 2, \ldots, k$), плотность распределения частот на i-м интервале в этом случае равна n_i/h.

Если статистическое распределение задано перечнем интервалов и соответствующих им частот, то строят гистограмму частот. *Гистограммой частот* называется ступенчатая фигура, состоящая из прямоугольников с основаниями $h_i = x_i - x_{i-1}$ и высотами n_i/h_i. В случае равенства интервальных разностей все прямоугольники имеют одно и то же основание h, их высоты равны n_i/h. На оси абсцисс откладывают частичные интервалы длины h, над i-м интервалом строят прямоугольник высоты n_i/h (плотность частоты). Отметим, что площадь S гистограммы частот равна сумме всех частот, т. е. объему выборки. Действительно, если s_i — площадь i-го прямоугольника, то

$$s_i = h_i \frac{n_i}{h_i} = n_i, \quad S = \sum_{i=1}^{k} s_i = \sum_{i=1}^{k} n_i = n.$$

На рис. 33.2 изображена гистограмма частот распределения объема $n = 75$, указанного в табл. 33.2.

Таблица 33.2

Частичный интервал длины $h=3$	3—6	6—9	9—12	12—15	15—18	18—21	21—24
Сумма частот вариант частичного интервала n_i	6	9	12	21	18	6	3
Плотность частоты $\frac{n_i}{h}$	2	3	4	7	6	2	1

Аналогично строится гистограмма относительных частот, т. е. ступенчатая фигура, состоящая из прямоугольников, основания которых равны h_i (h_i — длина каждого частичного интервала), а высоты — $\frac{w_i}{h_i}$ (это отношение называется плотностью относительной частоты). Очевидно, площадь s гистограммы относительных частот равна сумме всех относительных частот, т. е. единице. В самом деле, если s_i — площадь i-го прямоугольника, то

$$s_i = h_i \frac{w_i}{h_i} = w_i, \quad S = \sum_{i=1}^{k} s_i = \sum_{i=1}^{k} w_i = \sum_{i=1}^{k} \frac{n_i}{n} = 1.$$

§ 33.3. Эмпирическая функция распределения

Эмпирической функцией распределения, или *функцией распределения выборки*, называется функция, определяющая для каждого значения x относительную частоту события $X < x$. Обозначим эмпирическую функцию распределения через $F^*(x)$; если n_x — число вариант, меньших x, n — объем выборки, то по определению

$$F^*(x) = \frac{n_x}{n}. \tag{33.7}$$

Из определения эмпирической функции следует, что $F^*(x)$ обладает следующими свойствами: 1) значения функции $F^*(x)$ принадлежат отрезку $[0, 1]$; 2) $F^*(x)$ — неубывающая функция; 3) если a — наименьшая, b — наибольшая варианта, то $F^*(x) = 0$ при $x \leqslant a$; $F^*(x) = 1$ при $x \geqslant b$. Функцию $F(x)$ распределения генеральной совокупности, в отличие от эмпирической функции $F^*(x)$ распределения выборки, называют *теоретической функцией распределения.* Различие между эмпирической и теоретической функцией распределения состоит в том, что первая определяет относительную частоту события $X < x$, а вторая — вероятность того же события.

Пример. Найти эмпирическую функцию по данному распределению выборки:

варианты x_i	6	8	12	15
частоты n_i	2	3	10	5

Объем выборки $n = \sum_{i=1}^{4} n_i = 2 + 3 + 10 + 5 = 20$. Наименьшая варианта $x_1 = 6$, поэтому $F^*(x) = 0$, если $x \leqslant 6$. Значение $X < 8$, т. е. $x_1 = 6$, наблюдалось 2 раза, поэтому $F^*(x) = \frac{2}{20} = 0,1$, если $6 < x \leqslant 8$.

Значения $X<12$, т. е. $x_1=6$, $x_2=8$, наблюдались $2+3=5$ раз, поэтому $F^*(x)=\frac{5}{20}=0,25$, если $8<x\leqslant 12$.

Значения $X<15$, т. е. $x_1=6$, $x_2=8$, $x_3=12$, наблюдались $2+3+$ $+10=15$ раз, поэтому $F^*(x)=\frac{15}{20}=0,75$, если $12<x\leqslant 15$. Поскольку $x_4=15$ — наибольшая варианта, то $F^*(x)=1$, если $x>15$.

Итак, искомая эмпирическая функция определяется формулами

$$F^*(x)=\begin{cases} 0 & \text{при } x\leqslant 6, \\ 0,1 & \text{при } 6<x\leqslant 8, \\ 0,25 & \text{при } 8<x\leqslant 12, \\ 0,75 & \text{при } 12<x\leqslant 15, \\ 1 & \text{при } x>15. \end{cases}$$

График этой функции изображен на рис. 33.3.

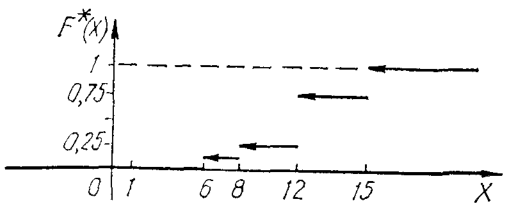

Рис. 33.3

§ 33.4. Оценка параметров по выборке. Понятие несмещенности, состоятельности и эффективности оценки

Пусть случайная величина X имеет распределение $F(x, \alpha)$, содержащее неизвестный параметр. Требуется оценить параметр α, т. е. приближенно определить его значение по некоторой выборке $x_1, x_2, \ldots, x_n$. Оценку параметра α обозначим через $\tilde{\alpha}$; очевидно, $\tilde{\alpha}$ зависит от $x_1, x_2, \ldots, x_n$, т. е.

$$\tilde{\alpha}=\tilde{\alpha}(x_1, x_2, \ldots, x_n). \tag{33.8}$$

Отметим, что $\tilde{\alpha}$ является случайной величиной, так как в i-й серии из n испытаний $\tilde{\alpha}$ принимает некоторое значение $\tilde{\alpha}_i$. Следовательно, можно говорить о распределении этой величины и о числовых характеристиках распределения.

Чтобы оценка $\tilde{\alpha}$ неизвестного параметра α имела практическую ценность, к ней предъявляются некоторые требования.

Оценка параметра α называется *несмещенной*, если математическое ожидание $\tilde{\alpha}$ равно α, т. е.

$$M(\tilde{\alpha}) = \alpha, \tag{33.9}$$

и *смещенной*, если

$$M(\tilde{\alpha}) \neq \alpha. \tag{33.10}$$

Оценка $\tilde{\alpha}$ параметра α называется *состоятельной*, если при любом $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\tilde{\alpha} - \alpha| < \varepsilon) = 1. \tag{33.11}$$

Очевидно, равенство (33.11) выполняется, если

$$\lim_{n\to\infty} D(\tilde{\alpha}) = 0. \tag{33.12}$$

Это следует из неравенства Чебышева (см. (32.1)).

Оценка $\tilde{\alpha}$ называется *эффективной*, если при заданном n она имеет наименьшую дисперсию, т. е.

$$D(\tilde{\alpha}) = D_{\min}. \tag{33.13}$$

Желательно, чтобы полученные из опыта оценки характеристик генеральной совокупности удовлетворяли требованиям несмещенности, состоятельности и эффективности.

§ 33.5. Генеральная средняя. Выборочная средняя. Оценка генеральной средней

Пусть требуется изучить дискретную генеральную совокупность относительно количественного признака X.

Генеральной средней называется среднее арифметическое значений признака X генеральной совокупности. Обозначим генеральную среднюю через $x_г$. Если все значения $x_1, x_2, \ldots, x_N$ признака генеральной совокупности объема N являются различными, то

$$x_г = \frac{x_1 + x_2 + \ldots + x_N}{N}, \quad x_г = \frac{1}{N} \sum_{i=1}^{N} x_i. \tag{33.14}$$

Поскольку признак X можно рассматривать как случайную величину, возможные значения которой $x_1, x_2, \ldots$ $\ldots, x_N$ имеют одну и ту же вероятность $p = 1 : N$, математическое ожидание этой величины определяется формулой

$$M(X) = x_1 \frac{1}{N} + x_2 \frac{1}{N} + \ldots + x_N \frac{1}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i = x_г,$$

$$M(X) = x_г. \tag{33.15}$$

Если значения $x_1, x_2, \ldots, x_m$ имеют соответственно частоты $N_1, N_2, \ldots, N_m$, причем $\sum\limits_{i=1}^{m} N_i = N$, то

$$x_г = \frac{x_1 N_1 + x_2 N_2 + \ldots + x_m N_m}{N}, \quad x_г = \frac{1}{N} \sum_{i=1}^{m} x_i N_i. \tag{33.16}$$

Отметим, что формула (33.15) справедлива и для этого случая. При непрерывном распределении признака X по определению полагают

$$x_г = M(X). \tag{33.17}$$

Предположим, что для изучения генеральной совокупности относительно количественного признака X из нее извлечена выборка объема n.

Выборочной средней $x_в$ называется среднее арифметическое значений признака выборочной совокупности. Если все значения $x_1, x_2, \ldots, x_n$ различны, то

$$x_в = \frac{x_1 + x_2 + \ldots + x_n}{n}, \quad x_в = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{33.18}$$

Если значения $x_1, x_2, \ldots, x_k$ имеют соответственно частоты $n_1, n_2, \ldots, n_k$, причем $n_1 + n_2 + \ldots + n_k = n$, то

$$x_в = \frac{n_1 x_1 + n_2 x_2 + \ldots + n_k x_k}{n}, \quad x_в = \frac{1}{n} \sum_{i=1}^{k} n_i x_i. \tag{33.19}$$

Так как каждой выборке объема n, извлеченной из генеральной совокупности, соответствует некоторое число $\bar{x}_в$, определяемое формулой (33.18) или (33.19), то выборочную среднюю можно рассматривать как случайную величину $\bar{X}_в$.

Выборочную среднюю принимают в качестве оценки генеральной средней. Покажем, что эта оценка является несмещенной и состоятельной, т. е. $M(\bar{X}_в)=\bar{x}_г$, $\lim\limits_{n\to\infty} P(|\bar{X}_в-\bar{x}_г|<\varepsilon)=1$. Не уменьшая общности рассуждений, предполагаем, что все значения $x_1, x_2, \ldots, x_n$ различны. Будем рассматривать эти значения как независимые, одинаково распределенные случайные величины $X_1, X_2, \ldots, X_n$; они имеют одинаковые числовые характеристики, в частности, одно и то же математическое ожидание a. Учитывая свойства математического ожидания, получаем

$$M(\bar{X}_в)=M\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)=a, \quad M(\bar{X}_в)=a. \quad (33.20)$$

Поскольку каждая из величин $X_1, X_2, \ldots, X_n$ имеет то же распределение, что и генеральная совокупность, математическое ожидание признака X генеральной совокупности также равно a, т. е.

$$M(X)=a. \quad (33.21)$$

Из формул (33.15), (33.20) и (33.21) следует, что

$$M(\bar{X}_в)=\bar{x}_г. \quad (33.22)$$

Это равенство означает, что выборочная средняя является несмещенной оценкой генеральной средней. Предполагая, что дисперсии случайных величин $X_1, X_2, \ldots, X_n$ ограничены, на основании следствия из теоремы Чебышева (частный случай) получаем

$$\lim_{n\to\infty} P(|\bar{X}_в-a|<\varepsilon)=1, \quad \text{или} \quad \lim_{n\to\infty} P(|\bar{X}_в-\bar{x}_г|<\varepsilon)=1,$$

так как $a=\bar{x}_г$ в силу равенств (33.15) и (33.21). Следовательно, при неограниченном увеличении объема выборки выборочная средняя стремится по вероятности к генеральной средней. Последнее равенство означает, что выборочная средняя будет состоятельной оценкой генеральной средней.

Из предыдущего следует, что выборочные средние, найденные по нескольким выборкам достаточно большого объема из некоторой генеральной совокупности, приближенно равны между собой. Это утверждение выражает свойство устойчивости выборочных средних.

§ 33.6. Генеральная дисперсия. Выборочная дисперсия. Эмпирическая дисперсия. Эмпирический стандарт

Для характеристики рассеяния значений количественного признака X генеральной совокупности вокруг своего среднего значения служат понятия генеральной дисперсии и генерального среднего квадратического отклонения.

Генеральной дисперсией $D_г$ называется среднее арифметическое квадратов отклонения значений признака X генеральной совокупности от их среднего значения $x_г$. Если все значения $x_1, x_2, \ldots, x_N$ признака генеральной совокупности объема N являются различными, то

$$D_г = \frac{1}{N} \sum_{i=1}^{N} (x_i - x_г)^2. \qquad (33.23)$$

Если значения $x_1, x_2, \ldots, x_k$ имеют соответственно частоты $N_1, N_2, \ldots, N_k$, причем $N_1+N_2+\ldots+N_k=N$, то

$$D_г = \frac{1}{N} \sum_{i=1}^{k} N_i (x_i - x_г)^2. \qquad (33.24)$$

Генеральным средним квадратическим отклонением $\sigma_г$ называется корень квадратный из генеральной дисперсии, т. е.

$$\sigma_г = \sqrt{D_г}. \qquad (33.25)$$

Для характеристики рассеяния значений количественного признака выборки вокруг среднего значения $x_в$ вводят понятия выборочной дисперсии и выборочного среднего квадратического отклонения.

Выборочной дисперсией $D_в$ называют среднее арифметическое квадратов отклонения наблюдаемых значений выборки от их среднего значения $x_в$. Если все значения $x_1, x_2, \ldots, x_n$ признака выборки объема n различны, то

$$D_в = \frac{1}{n} \sum_{i=1}^{n} (x_i - x_в)^2. \qquad (33.26)$$

Если значения $x_1, x_2, \ldots, x_k$ имеют соответственно частоты $n_1, n_2, \ldots, n_k$, причем $n_1+n_2+\ldots+n_k=n$, то

$$D_в = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - x_в)^2. \tag{33.27}$$

Выборочное среднее квадратическое отклонение $\sigma_в$ определяется формулой

$$\sigma_в = \sqrt{D_в}. \tag{33.28}$$

Для вычисления выборочной дисперсии можно пользоваться формулой

$$D_в = x_в^2 - (x_в)^2, \tag{33.29}$$

где

$$x_в = \frac{1}{n}\sum_{i=1}^{k} n_i x_i, \quad x_в^2 = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2. \tag{33.30}$$

Докажем формулу (33.29). Преобразуя формулу (33.27), получаем

$$D_в = \frac{1}{n}\sum_{i=1}^{k} n_i(x_i - x_в)^2 = \frac{1}{n}\sum_{i=1}^{k} n_i(x_i^2 - 2x_i x_в + (x_в)^2) =$$

$$= \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - \frac{2x_в}{n}\sum_{i=1}^{k} n_i x_i + \frac{(x_в)^2}{n}\sum_{i=1}^{k} n_i =$$

$$= x_в^2 - 2x_в x_в + (x_в)^2 = x_в^2 - (x_в)^2.$$

Из формулы (33.24) аналогично находим $D_г = x_г^2 - (x_г)^2$.

Следовательно, для обоих случаев

$$D = x^2 - (x)^2, \tag{33.31}$$

где x^2 — среднее квадратов значений; $(x)^2$ — квадрат общей средней.

Можно доказать, что

$$M(D_в) = \frac{n-1}{n} D_г. \tag{33.32}$$

Так как $M(D_в) \neq D_г$, то выборочная дисперсия $D_в$ является смещенной оценкой генеральной дисперсии $D_г$. Чтобы получить несмещенную оценку генеральной дисперсии $D_г$, вводят понятие так называемой эмпирической (или исправленной) дисперсии s^2.

Эмпирическая, или *исправленная*, дисперсия s^2 определяется формулой

$$s^2 = \frac{n}{n-1} D_в = \frac{n}{n-1} \frac{1}{n} \sum_{i=1}^{k} n_i(x_i - x_в)^2,$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{k} n_i (x_i - x_в)^2. \qquad (33.33)$$

Исправленная дисперсия (33.33) является несмещенной оценкой генеральной дисперсии, так как

$$M(s^2) = M\left(\frac{n}{n-1} D_в\right) = \frac{n}{n-1} M(D_в) =$$
$$= \frac{n}{n-1} \frac{n-1}{n} D_г = D_г.$$

Для оценки среднего квадратического отклонения генеральной совокупности служит «исправленное» среднее квадратическое отклонение, или эмпирический стандарт:

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{k} n_i (x_i - x_в)^2}. \qquad (33.34)$$

В случае, когда все значения $x_1, x_2, \ldots, x_n$ различны, т. е. все $n_i=1$ и $k=n$, формулы (33.33) и (33.34) принимают вид

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - x_в)^2; \qquad (33.35)$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - x_в)^2}. \qquad (33.36)$$

§ 33.7. Доверительная вероятность. Доверительный интервал

Оценка, определяемая одним числом, называется *точечной*. При выборке малого объема точечная оценка может существенно отличаться от истинного значения неизвестного параметра. Вследствие этого пользуются интервальными оценками. Оценка, определяемая двумя числами — концами интервалов, называется *интервальной.*

Пусть $\tilde{\alpha}$ — оценка неизвестного параметра α, полученная по данным выборки. Очевидно, оценка тем точнее, чем меньше модуль разности $\alpha - \tilde{\alpha}$. Если $\delta > 0$ и $|\alpha - \tilde{\alpha}| < \delta$, то чем меньше δ, тем точнее оценка $\tilde{\alpha}$; число δ характеризует точность оценки.

Доверительной вероятностью, или *надежностью,* оценки $\tilde{\alpha}$ параметра α называется вероятность γ, с которой осуществляется неравенство $|\alpha - \tilde{\alpha}| < \delta$, т. е.

$$P(|\alpha - \tilde{\alpha}| < \delta) = \gamma. \qquad (33.37)$$

Обычно надежность γ задается заранее, в качестве γ берут число, близкое к единице. Так как неравенство $|\alpha - \tilde{\alpha}| < \delta$ равносильно неравенствам $-\delta < \alpha - \tilde{\alpha} < \delta$ или $\tilde{\alpha} - \delta < \alpha < \tilde{\alpha} + \delta$, то формулу (33.37) можно записать в виде

$$P(\tilde{\alpha} - \delta < \alpha < \tilde{\alpha} + \delta) = \gamma. \qquad (33.38)$$

Эта формула означает следующее: вероятность того, что интервал $]\tilde{\alpha} - \delta,\ \tilde{\alpha} + \delta[$ заключает в себе (покрывает) неизвестный параметр α, равна γ. Интервал $]\tilde{\alpha} - \delta,\ \tilde{\alpha} + \delta[$, который покрывает неизвестный параметр α с заданной надежностью γ, называется *доверительным интервалом.* Концы доверительного интервала называют *доверительными границами.* Доверительные границы являются случайными величинами (они изменяются от выборки к выборке).

Рассмотрим вопрос о построении доверительного интервала для оценки математического ожидания a нормального распределения при известном значении среднего квадратического отклонения σ.

Пусть количественный признак X генеральной совокупности имеет нормальное распределение с заданным σ и неизвестным a. Оценим неизвестный параметр a выборочной средней $x_в$; найдем доверительный интервал, покрывающий параметр a с надежностью γ. Так как выборочное среднее $x_в$ меняется от выборки к выборке, его можно рассматривать как случайную величину $X_в$. Выборочные значения $x_1, x_2, \ldots, x_n$ также меняются от выборки к выборке. Будем рассматривать их, как одинаково распределенные случайные величины $X_1, X_2, \ldots, X_n$ (математическое ожидание каждой из этих величин равно a, среднее квадратическое отклонение равно σ). В соответствии с формулами § 30.3 (см. (30.32) и (30.33)) имеем

$$M(X_в)=a,\ \sigma(X_в)=\frac{\sigma}{\sqrt{n}}. \qquad (33.39)$$

Потребуем, чтобы

$$P(|X_в-a|<\delta)=\gamma. \qquad (33.40)$$

Поскольку случайная величина $X_в$ также имеет нормальное распределение, то, применяя формулу (см. § 31.9)

$$P(|X-a|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right)$$

к величинам $X_в$ и $\sigma(X_в)=\frac{\sigma}{\sqrt{n}}$, находим

$$P(|X_в-a|<\delta)=2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right)=2\Phi(t), \qquad (33.41)$$

где

$$t=\frac{\delta\sqrt{n}}{\sigma}. \qquad (33.42)$$

Из последнего равенства определим $\delta=t\frac{\sigma}{\sqrt{n}}$ и подставим в формулу (33.41):

$$P\left(|X_в-a|<\frac{t\sigma}{\sqrt{n}}\right)=2\Phi(t), \qquad (33.43)$$

или

$$P\left(x_в-\frac{t\sigma}{\sqrt{n}}<a<x_в+\frac{t\sigma}{\sqrt{n}}\right)=2\Phi(t),$$

где $x_в$ — выборочное среднее.

Поскольку вероятность P задана и равна γ, т. е.

$$P\left(x_в-\frac{t\sigma}{\sqrt{n}}<a<x_в+\frac{t\sigma}{\sqrt{n}}\right)=\gamma, \qquad (33.44)$$

то последнее означает, что доверительный интервал

$$\left]x_{\text{в}} - \frac{t\sigma}{\sqrt{n}}, \quad x_{\text{в}} + \frac{t\sigma}{\sqrt{n}}\right[\tag{33.45}$$

покрывает неизвестный параметр a с надежностью γ. Из формулы (33.42) находим точность оценки

$$\delta = \frac{t\sigma}{\sqrt{n}}. \tag{33.46}$$

Отметим, что число t определяется равенством

$$2\Phi(t) = \gamma, \tag{33.47}$$

получающимся из (33.43) и (33.44); значение t находится с помощью таблиц функции Лапласа.

Глава 34

МАТЕМАТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ

Методы теории вероятностей и математической статистики широко применяются при изучении различных явлений и обработке результатов эксперимента. Изучение явления, как правило, начинается с наблюдений. Под наблюдением имеют в виду регистрацию некоторых (случайных) событий, связанных с изучаемым явлением. Результаты наблюдений делятся на качественные и количественные. К качественным результатам относится появление некоторых событий (зажигание контрольной лампочки, выпадение осадка в растворе и т. п.). Количественные результаты получают при соответствующих подсчетах и измерениях.

§ 34.1. Измерения и их погрешности. Применение методов математической статистики к обработке результатов наблюдений

Под измерением понимают сравнение некоторой величины с другой (однородной) величиной, принятой за единицу. Различают прямые и косвенные измерения. При прямых измерениях исследуемая величина сравнивается с единицей измерения непосредственно или с помощью измерительного прибора (например, измерения длин линейкой, масс на весах и т. п.). При косвенных измерениях величины ее значения определяются по результатам прямых измерений других величин, связанных с рассматриваемой величиной заданной функциональной зависимостью (например, измерения плотности тела по измерениям его массы и объема).

В результате измерений получают приближенные значения величины, а не ее точное значение. При измерениях неизбежны погрешности. *Погрешностью*, или *ошибкой*, измерения называется разность $x-a$ между результатом измерения x и точным значением a измеряемой величины. (Отметим, что погрешности измерений обычно неизвестны, так как неизвестно и точное значение измеряемой величины.)

Различают следующие виды погрешностей (ошибок) измерений: *грубые, систематические, случайные*. К грубым ошибкам (или промахам) относят ошибки, сделанные вследствие неверной записи показаний прибора, неправильно прочитанного отсчета и т. п. Систематические погрешности — погрешности, связанные с ограниченной точностью изготовления прибора, неправильной его установкой или некоторыми другими факторами. Они вызываются вполне определенными причинами, и их величина при всех измерениях остается постоянной (как в случае смещения нуля шкалы прибора) либо изменяется по определенному закону (как в случае неравномерной шкалы). Систематические ошибки можно устранить путем введения соответствующих поправок.

Случайные погрешности вызываются большим числом случайных причин, действие которых на каждое измерение различно и заранее не может быть учтено. Случайные погрешности являются неустранимыми. Хотя их нельзя исключить, но с помощью методов теории вероятностей можно учесть влияние случайных погрешностей на оценку точного значения измеряемой величины. В теоретико-вероятностной модели случайные погрешности $z=x-a$ (и сами результаты измерений $x=a+z$) рассматривают как случайную величину с некоторым законом распределения ее вероятностей. В качестве закона распределения случайных погрешностей измерения чаще всего принимается нормальный закон распределения.

Замечание. Нормальное распределение случайных погрешностей обычно достаточно хорошо согласуется с опытом. В частности, это распределение отражает следующие два свойства случайных погрешностей: 1) симметрии (равные *по модулю и противоположные* по знаку погрешности встречаются почти одинаково часто) и 2) концентрации (малые по модулю случайные погрешности встречаются чаще, чем большие).

При каждом измерении фиксируется один количественный результат. Результаты любой серии из n измерений будут случайным образом колебаться вокруг точного значения измеряемой величины. Следовательно, с этим точным значением связана некоторая случайная величина X. В итоге n независимых измерений получается n ее возможных значений $x_1, x_2, \ldots, x_n$. Если все возможные значения случайной величины X считать генеральной совокупностью, то полученные при n измерениях значения $x_1, x_2, \ldots, x_n$ образуют выборку. По этой выборке и необходимо определить распределение случайной величины X (распределение генеральной совокупности). Таким образом, проведение измерений является частным случаем выборочного метода, когда в качестве генеральной совокупности рассматриваются все возможные значения указанной случайной величины X и исследуется распределение этой величины по выборке (результатам измерений) на основе теории, изложенной в предыдущих главах.

§ 34.2. Оценка точного значения измеряемой величины

Пусть в итоге n независимых измерений некоторой величины X получены следующие результаты:

$$x_1, x_2, \ldots, x_n. \tag{34.1}$$

Будем предполагать, что эти результаты свободны от грубых и систематических ошибок (неверные результаты отброшены, на систематические ошибки введены поправки).

Оценить точное значение a измеряемой величины — это значит:

а) определить функцию $\alpha = \alpha(x_1, x_2, \ldots, x_n)$, которая обеспечивает достаточно близкое приближение к значению a;

б) указать границы интервала $(\alpha - \delta_1, \alpha + \delta_2)$, который с заданной вероятностью γ покрывает истинное значение a (эта оценка называется доверительной оценкой, вероятность γ — доверительной вероятностью, или надежностью оценки, интервал $(\alpha - \delta_1, \alpha + \delta_2)$ — доверительным интервалом, а его границы — доверительными границами). Оценка $\alpha=\alpha(x_1, x_2, \ldots, x_n)$ должна (по возможности) обладать следующими свойствами: несмещенности, состоятельности и эффективности (см. § 33.4).

Введем среднее арифметическое значение (среднее значение) x результатов (34.1), среднее квадратическое отклонение s^* этих результатов от их среднего значения x и эмпирический стандарт s соответственно по формулам

$$x = \frac{x_1 + x_2 + \ldots + x_n}{n}, \quad x = \frac{1}{n}\sum_{i=1}^{n} x_i; \tag{34.2}$$

$$s^* = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2}; \tag{34.3}$$

$$s = s^*\sqrt{\frac{n}{n-1}} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - x)^2}. \tag{34.4}$$

Если все измерения произведены с одинаковой точностью, то в качестве оценки точного значения a измеряемой величины принимают среднее арифметическое значение результатов (34.1):

$$a \approx x = \frac{x_1 + x_2 + \ldots + x_n}{n}, \quad a \approx \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{34.5}$$

Как было показано (см. § 33.5), эта оценка является несмещенной и состоятельной. Введенная оценка оказывается и эффективной при дополнительном предположении о том, что случайные ошибки измерений подчинены нормальному закону распределения. Такое предположение имеется в виду и в дальнейшем. Отметим, что оценка (34.5) относится к числу точечных оценок.

Переходим к *доверительным* оценкам. Будем рассматривать симметрические доверительные оценки, т. е. оценки вида

$$|a-x|<\delta \quad (\delta>0) \tag{34.6}$$

или

$$x-\delta<a<x+\delta, \tag{34.7}$$

где x — среднее значение (см. (34.2)). Величина δ (точность оценки) определяется по заданной доверительной вероятности γ (надежности оценки); γ обычно задается в виде одного из трех значений: $\gamma=0{,}95$, $\gamma=0{,}99$, $\gamma=0{,}999$.

Доверительная оценка при известной точности измерений. Если известно среднее квадратическое отклонение σ, то доверительная оценка (34.6) имеет вид (см. § 33.7, формулы (33.43), (33.46), (33.47))

$$|a-x|<\frac{t\sigma}{\sqrt{n}}, \tag{34.8}$$

где n — число измерений, а значение $t=t(\gamma)$ определяется по заданной доверительной вероятности γ из условия $2\Phi(t)=\gamma$ и находится с помощью таблиц. Точность оценки δ в этом случае выражается формулой

$$\delta=\frac{\sigma t}{\sqrt{n}}. \tag{34.9}$$

Доверительная оценка при неизвестной точности измерений. Если средняя квадратическая погрешность σ заранее неизвестна, то вместо нее применяют эмпирический стандарт s (см. § 33.6), который служит оценкой параметра σ. Доверительная оценка (34.6) принимает вид

$$|a-x|<\frac{ts}{\sqrt{n}}, \tag{34.10}$$

или

$$|a-x|<\frac{ts^*}{\sqrt{k}} \quad (k=n-1), \tag{34.11}$$

где s^* и s определяются соответственно формулами (34.3) и (34.4), а множитель $t=t(\gamma, k)$ зависит не только от доверительной вероятности γ, но и от числа измерений n $(k=n-1)$. Значения этого множителя находят по таблицам.

Замечание. Соответствующие таблицы составлены с помощью распределения Стьюдента, т. е. распределения вероятностей отношения $(x-a)\sqrt{n}/s$; значения $t=t(\gamma, k)$ определены из условия

$$P\left(\left|\frac{X-a}{s/\sqrt{n}}\right|<t\right)=\gamma.$$

Распределение Стьюдента зависит от одного параметра k, называемого числом степеней свободы; в данном случае этот параметр связан с числом измерений формулой $k=n-1$.

Правило трех сигм. Поскольку надежность доверительной оценки выбирается заранее, в практике математической обработки экспериментальных данных широко применяется правило трех сигм (см. § 31.9): отклонение истинного значения измеряемой величины от среднего арифметического результатов измерений не превосходит утроенной средней квадратической погрешности этого среднего значения.

Следовательно, правило трех сигм представляет собой доверительную оценку

$$|a-x|<\frac{3\sigma}{\sqrt{n}} \tag{34.12}$$

при известной величине σ или доверительную оценку

$$|a-x|<\frac{3s}{\sqrt{n}} \tag{34.13}$$

при неизвестной величине σ (s — эмпирический стандарт). Оценка (34.12) имеет надежность $2\Phi(3)=0{,}9973$ (см. § 31.9) независимо от числа измерений. Оценка (34.13) зависит от n — числа измерений (зависимость эта указывается с помощью соответствующих таблиц).

§ 34.3. Оценки точности измерений

Здесь предполагается, что измерения являются независимыми и равноточными (с одной и той же дисперсией), а их погрешности — случайными, причем распределены они по нормальному закону. В качестве показателя точности измерений оценивается дисперсия этого закона σ^2 или средняя квадратическая погрешность $\sigma=\sqrt{\sigma^2}$.

Точечные оценки дисперсии. 1. Если измеряют известную величину a, то в качестве эффективной оценки дисперсии σ^2 применяют квадрат среднего квадратического отклонения s^* результатов измерений (34.1) от значения a:

$$\sigma^2 \approx s^{*2}=\frac{1}{n}\sum_{i=1}^{n}(x_i-a)^2. \tag{34.14}$$

2. При измерениях неизвестной величины в качестве оценки дисперсии σ^2 применяют эмпирическую дисперсию s^2:

$$\sigma^2 \approx s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-x)^2, \tag{34.15}$$

где x — среднее арифметическое значений $x_1, x_2, \ldots, x_n$. Оценка (34.15) является несмещенной и состоятельной, но не является эффективной (она асимптотически эффективна, т. е. ее дисперсия стремится к наименьшему значению при неограниченном увеличении числа измерений n).

3. Если производится m серий измерений некоторой величины и известны количества измерений $n_1, n_2, \ldots, n_m$, а также средние арифметические результаты $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m$ в каждой серии, то в качестве оценки дисперсии применяют эмпирическую дисперсию $\bar{s}^2$ из средних:

$$\sigma^2 \approx \bar{s}^2 = \frac{1}{m-1} \sum_{i=1}^{m} n_i (\bar{x}_i - x)^2, \qquad (34.16)$$

где

$$x = \frac{1}{N} \sum_{i=1}^{m} n_i \bar{x}_i, \quad N = n_1 + n_2 + \ldots + n_m. \qquad (34.17)$$

Эта оценка является несмещенной, состоятельной (и асимптотически эффективной при $m \to \infty$).

Доверительные оценки средней квадратической погрешности. При большом числе измерений доверительную оценку средней квадратической погрешности σ записывают в виде оценки относительного отклонения оцениваемого значения σ от эмпирического стандарта s (или s^*, или $\bar{s}$). Эта оценка имеет вид

$$\left| \frac{\sigma - s}{s} \right| < q, \qquad (34.18)$$

или

$$s(1-q) < \sigma < s(1+q), \qquad (34.19)$$

коэффициент $q = q(\gamma, k)$ находится с помощью соответствующих таблиц в зависимости от доверительной вероятности γ (надежности оценки) и числа степеней свободы k ($k=1$ в случае 1, $k=n-1$ в случае 2, $k=m-1$ в случае 3).

При малом числе измерений симметричная оценка (34.19) приводит к неоправданно большим доверительным интервалам; в этом случае применяют асимметричные доверительные оценки вида

$$s z_1 < \sigma < s z_2, \qquad (34.20)$$

где s — эмпирический стандарт; значения коэффициентов $z_1 = z_1(\gamma, k)$, $z_2 = z_2(\gamma, k)$ находятся по таблицам.

§ 34.4. Метод наименьших квадратов

При обработке опытных данных часто встречаются с задачей об определении параметров функциональной зависимости между переменными величинами x и y посредством формулы $y = f(x)$. Эта задача решается с помощью метода наименьших квадратов, сущность которого состоит в следующем. При измерении двух величин x и y получены следующие данные:

x	x_1	x_2	...	x_n
y	y_1	y_2	...	y_n

Известен также вид функциональной зависимости, т. е.

$$y = f(x, a_0, a_1, \ldots, a_m) = \varphi(x), \qquad (34.21)$$

где f — заданная функция; $a_0, a_1, \ldots, a_m$ — параметры, значения которых требуется определить. Значения y, полученные из формулы (34.21) при заданных значениях x_i $(i = 1, 2, \ldots, n)$, как правило, не совпадают с экспериментальными значениями y_i, приведенными в указанной таблице, т. е. разность $y_i - \varphi(x_i)$ отлична от нуля для всех или некоторых точек x_i $(i = 1, 2, \ldots, n)$. Для каждого i эту разность обозначим через ε_i и назовем *погрешностью*:

$$y_i - \varphi(x_i) = \varepsilon_i \quad (i = 1, 2, \ldots, n). \qquad (34.22)$$

Значения параметров a_k $(k = 0, 1, \ldots, m)$ функции (34.21) требуется выбрать так, чтобы сумма квадратов погрешностей была наименьшей, т. е. так, чтобы функция

$$u = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \varphi(x_i))^2 \qquad (34.23)$$

принимала наименьшее значение. Поскольку эта функция — сумма квадратов некоторых чисел, она принимает неотрицательные значения (каждое слагаемое суммы неотрицательно).

Функция (34.23) является функцией $m+1$ переменных $a_0, a_1, \ldots, a_m$, т. е.

$$u = u(a_0, a_1, \ldots, a_m) = \sum_{i=1}^{n} (y_i - f(x_i, a_0, a_1, \ldots, a_m))^2. \tag{34.24}$$

Если функция $u = u(a_0, a_1, \ldots, a_m)$ имеет непрерывные частные производные по всем переменным, то необходимое условие ее минимума выражается системой уравнений

$$\frac{\partial u}{\partial a_0} = 0, \quad \frac{\partial u}{\partial a_1} = 0, \ldots, \quad \frac{\partial u}{\partial a_m} = 0. \tag{34.25}$$

Из этой системы $m+1$ уравнений находятся искомые значения параметров $a_0, a_1, \ldots, a_m$.

Во многих случаях функция (34.21) определяется формулой

$$y = \sum_{k=0}^{m} a_k f_k(x), \tag{34.26}$$

где $f_0(x), f_1(x), \ldots, f_m(x)$ — известные функции, например, $f_k(x) = x^k$, $f_k(x) = \sin kx$, $f_k(x) = \cos kx$ и т. д.

Функция (34.24) в таких случаях принимает вид

$$u = \sum_{i=1}^{n} \left(y_i - \sum_{k=0}^{m} a_k f_k(x_i) \right)^2, \tag{34.27}$$

а система (34.25) запишется так:

$$\left.\begin{aligned}
&\sum_{i=1}^{n} \Big(y_i - \sum_{k=0}^{m} a_k f_k(x_i) \Big) (-f_0(x_i)) = 0; \\
&\sum_{i=1}^{n} \Big(y_i - \sum_{k=0}^{m} a_k f_k(x_i) \Big) (-f_1(x_i)) = 0; \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
&\sum_{i=1}^{n} \Big(y_i - \sum_{k=0}^{m} a_k f_k(x_i) \Big) (-f_m(x_i)) = 0.
\end{aligned}\right\} \tag{34.28}$$

Решение этой системы может быть получено с помощью метода Гаусса (метод последовательного исключения неизвестных).

Если $f_k(x)=x^k$ $(k=0, 1, 2, \ldots, m)$, то

$$f(x, a_0, a_1, \ldots, a_m)=a_0+a_1x+a_2x^2+\ldots+a_mx^m \quad (34.29)$$

и система (34.28) принимает вид

$$\left.\begin{aligned} &a_0n+a_1\sum_{i=1}^{n}x_i+\ldots+a_m\sum_{i=1}^{n}x_i^m=\sum_{i=1}^{n}y_i;\\ &a_0\sum_{i=1}^{n}x_i+a_1\sum_{i=1}^{n}x_i^2+\ldots+a_m\sum_{i=1}^{n}x_i^{m+1}=\sum_{i=1}^{n}y_ix_i;\\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\ &a_0\sum_{i=1}^{n}x_i^m+a_1\sum_{i=1}^{n}x_i^{m+1}+\ldots+a_m\sum_{i=1}^{n}x_i^{2m}=\sum_{i=1}^{n}y_ix_i^m. \end{aligned}\right\} \quad (34.30)$$

Частные случаи $(m=1,\ m=2)$ последней задачи $(f_k(x)=x^k)$ рассмотрены в § 18.14. При $m=1$, $m=2$ система уравнений (34.30) принимает соответственно вид (18.120), (18.122).

Пример 1. Получены следующие результаты измерений величин x и y:

x	1,00	1,50	2,00	2,50	3,00
y	2,10	2,20	2,70	2,80	2,85

Установить завнсимость между этими величинами и определить параметры эмпирической формулы методом наименьших квадратов.

Будем считать, что соответствующие пары значений (x_i, y_i) $(i=1, 2, 3, 4, 5)$ являются прямоугольными декартовыми координатами точек на плоскости. Построив точки A (1; 2,1), B (1,5; 2,2), C (2; 2,7), D (2,5; 2,8), E (3; 2,85), обнаружим, что они незначительно отклоняются от некоторой прямой. Следовательно, можно предположить, что между величинами x и y существует приближенная линейная зависимость, т. е. $y=ax+b$, где a и b пока неизвестны. Методом наименьших квадратов определим параметры a и b эмпирической формулы $y=ax+b$, использовав нормальную систему уравнений (18.120). Чтобы применить эту систему, необходимо подсчитать входящие в нее коэффициенты. Для подсчета коэффициентов составим таблицу:

i	x_i	y_i	x_iy_i	x_i^2
1	1,00	2,10	2,10	1,00
2	1,50	2,20	3,30	2,25
3	2,00	2,70	5,40	4,00
4	2,50	2,80	7,00	6,25
5	3,00	2,85	8,55	9,00
Σ	10,00	12,65	26,35	22,50

В последней строке таблицы получены коэффициенты системы уравнений (18.120), которая принимает вид $22{,}5a+10b=26{,}35$; $10a+5b=12{,}65$.

Решая эту систему, находим $a=0{,}42$, $b=1{,}69$. Следовательно, зависимость между величинами x и y выражается приближенной формулой $y=0{,}42x+1{,}69$.

Пример 2. Найти параметры a, b, c эмпирической формулы $y=ax^2+bx+c$ по результатам измерений:

x	−3	−2	−1	0	1	2
y	−1,4	−4,3	−5,2	−4,1	−1,1	4,2

Результаты измерений и итоги их обработки представим в таблице:

i	x_i	x_i^2	x_i^3	x_i^4	y_i	y_ix_i	$y_ix_i^2$
1	−3	9	−27	81	−1,4	4,2	−12,6
2	−2	4	−8	16	−4,3	8,6	−17,2
3	−1	1	−1	1	−5,2	5,2	−5,2
4	0	0	0	0	−4,1	0	0
5	1	1	1	1	−1,1	−1,1	−1,1
6	2	4	8	16	4,2	8,4	16,8
Σ	−3	19	−27	115	−11,9	25,3	−19,3

Система уравнений (18.122) в данном случае принимает вид $115a-27b+19c=-19{,}3$; $-27a+19b-3c=25{,}3$; $19a-3b+6c=-11{,}9$.

Решив эту систему, получим $a=1{,}011$, $b=2{,}116$, $c=-4{,}126$. Следовательно, $y=1{,}011x^2+2{,}116x-4{,}126$.

Пример 3. В «Основах химии» Д. И. Менделеев приводит даниые о растворимости азотио-натриевой соли $NaNO_3$ в зависимости от температуры воды. В 100 частях воды растворяется следующее количество условиых частей $NaNO_3$ (y) при соответствующей температуре (x):

x	0	4	10	15	21	29	36	51	68
y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

Д. И. Менделеев указывает, что зависимость между x и y можно выразить формулой $y=0{,}87x+67{,}5$.

Произведем вычислеиия, иеобходимые для проверки этого утверждеиия. Можио показать, что между x и y существует лииейная зависимость. Результаты измерений и итоги их обработки представим в таблице:

i	x_i	y_i	$x_i y_i$	x_i^2
1	0	66,7	0	0
2	4	71,0	284,0	16
3	10	76,3	763,0	100
4	15	80,6	1209,0	225
5	21	85,7	1799,7	441
6	29	92,9	2694,1	841
7	36	99,4	3578,4	1296
8	51	113,6	5793,6	2601
9	68	125,1	8506,8	4624
Σ	234	811,3	24628,6	10144

Нормальная система уравнений (18.120) в этом случае имеет вид $10144a+234b=24628{,}6$; $234a+9b=811{,}3$. Из этой системы находим $a=0{,}87$, $b=67{,}5$, т. е. зависимость между x и y выражается приближенной формулой $y=0{,}87x+67{,}5$.

§ 34.5. Определение параметров эмпирических формул

Выше были рассмотрены примеры нахождения эмпирических формул в случаях линейной и квадратичной зависимостей с помощью метода наименьших квадратов. Существуют и другие способы определения параметров эмпирических формул (более простых и удобных, хотя и менее точных). Рассмотрим некоторые из этих способов, а также вопрос об определении параметров эмпирических формул в случаях показательной, степенной и других зависимостей.

Способы определения параметров эмпирической формулы в случае линейной зависимости. В примерах 1 и 3 § 34.4 значения параметров эмпирической формулы $y=ax+b$ определялись способом наименьших квадратов. Параметры этой эмпирической формулы можно определить и другими способами. К их числу относятся способ выбранных точек и способ средней.

Способ выбранных точек основан на геометрическом подборе прямой. Нанося экспериментальные данные (x_i, y_i), где $i=1, 2, \ldots, n$, на миллиметровку, подбирают графически прямую, ближе всего подходящую к построенным точкам $M_i(x_i, y_i)$, $i=1, 2, \ldots, n$. Выбирая две произвольные точки прямой (которые могут и не совпадать с двумя из построенных точек M_i), определяют их координаты (x_1, y_1), (x_2, y_2). С помощью этих координат находят значения параметров a и b эмпирической формулы $y=ax+b$ из системы двух уравнений $ax_1+b=y_1$, $ax_2+b=y_2$.

Способ средней состоит в следующем. Пусть результаты эксперимента даны в таблице значений (x_i, y_i), $i=1, 2, \ldots, n$. Обозначим через ε_i соответствующие погрешности, т. е. $\varepsilon_i=y_i-ax_i-b$ $(i=1, 2, \ldots, n)$. Если выбирать параметры a и b так, чтобы для всех n наблюдений погрешности уравновешивались, т. е. $\sum\limits_{i=1}^{n}\varepsilon_i=0$, то это привело бы к одному уравнению (относительно двух неизвестных a и b). Потребуем, чтобы уравновешивание происходило не только для всех наблюдений в целом, но и в отдельности для каждой группы, содержащей половину (или почти половину) всех наблюдений. В этом случае приходим к системе уравнений

$$\sum_{i=1}^{m} (y_i - ax_i - b) = 0, \quad \sum_{i=m+1}^{n} (y_i - ax_i - b) = 0,$$

или

$$a \sum_{i=1}^{m} x_i + mb = \sum_{i=1}^{m} y_i, \quad a \sum_{i=m+1}^{n} x_i + (n-m)b = \sum_{i=m+1}^{n} y_i, \qquad (34.31)$$

где m — число наблюдений в первой группе. Обычно m выбирается так, чтобы число наблюдений во второй группе также равнялось m, если n четно, и $m \pm 1$, если n нечетно.

Функциональные шкалы и их применение. Рассмотрим функцию $f(x)$, непрерывную и монотонную на отрезке $[a, b]$. Функциональная шкала для $f(x)$ строится следующим образом. Возьмем ось OM, зафиксировав на ней точку отсчета, и выберем масштаб μ. Отрезок $[a, b]$ разобьем на равные части, вычислим значения функции $f(x)$ в каждой из точек деления, включая концы, и отложим на оси OM отрезок длины $\mu f(x)$, конец этого отрезка обозначим буквой x. Таким образом, в выбранном масштабе откладывается значение функции, а подписывается соответствующее значение аргумента x. Полученная шкала позволяет судить об изменении функции на данном отрезке. Выбрав масштаб μ, тем самым определим длину получающейся шкалы. Чаще поступают наоборот: задав длину шкалы l, определяют масштаб μ из равенства $\mu[f(b) - f(a)] = l$.

Известным примером функциональной шкалы является основная шкала логарифмической линейки, представляющая собой шкалу, построенную для функции $y = \lg x$ на отрезке $[1, 10]$. В данном случае масштаб совпадает с длиной линейки, поскольку $\lg 1 = 0$, $\lg 10 = 1$ ($f(a) = 0$, $f(b) = 1$, $f(b) - f(a) = 1$, $\mu = l$). Чаще всего выбирают $\mu = 25$ см; встречаются также маленькие линейки ($\mu = 15$ см) и большие ($\mu = 50$ см).

Функциональные шкалы находят широкое применение при обработке экспериментальных данных и построении эмпирических формул. Дело в том, что графики многих функций специальным подбором функциональных шкал могут быть преобразованы в графики линейной зависимости (т. е. в прямые линии).

Координатные сетки, построенные с помощью функциональных шкал, называют функциональными сетками. Отметим, что часто используются различные логарифмические сетки, с помощью которых можно «выпрямлять» графики степенных и показательных функций. Если зависимость между x и y определяется уравнением $y=be^{ax}$, то $\lg y = \lg b + (a \lg e)x$. Полагая $\lg y = Y$, $\lg b = B$, $a \lg e = A$, получаем линейную функцию $Y=Ax+B$. В функциональной сетке, для которой на оси Ox оставлена равномерная шкала, а на оси Oy построена логарифмическая шкала, график функции $Y=Ax+B$ является прямой. Указанная сетка называется полулогарифмической. Логарифмической сеткой называют функциональную сетку, у которой на каждой из осей Ox и Oy построена логарифмическая шкала. На такой сетке графики степенных функций изображаются прямыми линиями. В самом деле, если $y=bx^a$, то $\lg y = \lg b + a \lg x$. Это уравнение можно записать в виде $Y=aX+B$, если положить $\lg y = Y$, $\lg b = B$, $\lg x = X$.

Указанные сетки могут быть заготовлены заранее. Бумага с нанесенной на ней полулогарифмической или логарифмической сеткой называется *полулогарифмической или логарифмической бумагой*.

Замечания о выборе вида эмпирической формулы. Если результаты наблюдений (x_i, y_i), изображенные точками $M_i(x_i, y_i)$, $i=1, 2, \ldots, n$, в прямоугольной декартовой системе координат, достаточно хорошо приближаются прямой линией, то коэффициенты эмпирической формулы $y=ax+b$ легко находятся с помощью способов, описанных выше (способ выбранных точек, способ средней, способ наименьших квадратов). Если же эти точки располагаются вблизи некоторой кривой, то по имеющейся весьма ограниченной дуге линии затруднительно судить о том, какого типа функцией следует приближать изучаемую зависимость между x и y (т. е. какого типа эмпирической формулой необходимо воспользоваться). Изобразив результаты наблюдений в той или иной функциональной сетке, получают возможность определить, на какой из них экспериментальные данные ближе всего подходят к прямой, следовательно, выяснить, какую из формул следует выбрать для приближения исследуемой зависимости. После того, как выбран вид эмпирической формулы, необходимо будет найти соответствующие коэффициенты, что можно сделать рассмотренными выше способами.

Предположим, что искомая зависимость между x и y имеет вид

$$y=af(x)+b, \tag{34.32}$$

где $f(x)$ — известная функция, a и b — параметры, значения которых нужно определить по результатам наблюдений. Такой вид зависимости может следовать из некоторых теоретических положений или может быть выбран с помощью функциональной сетки. Положив $f(x)=X$ и построив соответствующую функциональную сетку (равномерная шкала на оси Oy, шкала $f(x)$ на оси Ox), приведем уравнение (34.32) к виду $y=aX+b$. Если при построении опытных данных на указанной функциональной сетке соответствующие точки достаточно близко подходят к некоторой прямой, то можно предположить, что искомая зависимость между x и y выражается формулой (34.32). После этого останется определить значение параметров a и b эмпирической формулы $y=af(x)+b$ (или $y=ax+b$). Для определения искомых значений параметров можно применить способ выбранных точек, способ средней или способ наименьших квадратов.

Определение параметров эмпирической формулы в случае показательной зависимости. Пусть x и y связаны формулой

$$y=be^{ax}. \tag{34.33}$$

Отметим, что к этому виду можно привести любую показательную зависимость $y=c^x$ $(c>0,\ c\neq 1)$, поскольку $c^x=(e^{\ln c})^x=e^{kx}$, где $k=\ln c$. Как уже отмечалось, зависимость $y=be^{ax}$ графически можно изобразить прямой, если воспользоваться полулогарифмической сеткой. Действительно, логарифмируя обе части равенства (34.33), получаем

$$\lg y=\lg b+(a \lg e)x, \text{ или } Y=Ax+B, \tag{34.34}$$

где

$$A=a\lg e,\ B=\lg b,\ Y=\lg y. \tag{34.35}$$

Применяя способ средней, значения A и B определяют из системы уравнений типа (34.31), которая в данном случае принимает вид

$$\sum_{i=1}^{m}(\lg y_i-(Ax_i+B))=0, \quad \sum_{i=m+1}^{n}(\lg y_i-(Ax_i+B))=0,$$

или

$$A\sum_{i=1}^{m}x_i+mB=\sum_{i=1}^{m}\lg y_i,\ A\sum_{i=m+1}^{n}x_i+(n-m)B=\sum_{i=m+1}^{n}\lg y_i, \tag{34.36}$$

где n — число всех наблюдений, m — число наблюдений в первой группе.

Определив A и B из этой системы уравнений, с помощью первых двух равенств (34.35) вычислим искомые значения параметров a и b эмпирической формулы (34.33).

Способ наименьших квадратов в этом случае применяется следующим образом. Составляем сумму квадратов погрешностей $\varepsilon_i = \lg y_i - (Ax_i + B)$:

$$U = \sum_{i=1}^{n} (\lg y_i - (Ax_i + B))^2,$$

получаем функцию двух переменных A, B. Находим частные производные функции $u = u(A, B)$ по A и по B, приравниваем их к нулю, в результате чего получаем систему двух уравнений относительно A и B:

$$A \sum_{i=1}^{n} x_i^2 + B \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i \lg y_i, \quad A \sum_{i=1}^{n} x_i + nB = \sum_{i=1}^{n} \lg y_i. \tag{34.37}$$

Из уравнений (34.37) определяются значения параметров A и B эмпирической формулы (34.34), а из первых двух равенств (34.35) — искомые значения a и b эмпирической формулы (34.33).

Определение параметров эмпирической формулы в случае степенной зависимости. Как уже отмечалось, график функции

$$y = bx^a \tag{34.38}$$

на логарифмической сетке представляет собой прямую линию. Действительно, логарифмируя равенство (34.38), находим

$$\lg y = \lg b + a \lg x. \tag{34.39}$$

Полагая в этом равенстве

$$\lg b = B, \ \lg x = X, \ \lg y = Y, \tag{34.40}$$

получаем

$$Y = aX + B. \tag{34.41}$$

Эмпирическая формула имеет вид (34.38), если соответствующие точки на логарифмической сетке незначительно уклоняются от некоторой прямой.

Для применения метода средней составляется система уравнений типа (34.31):

$$\sum_{i=1}^{m} (\lg y_i - (\lg b + a \lg x_i)) = 0,$$

$$\sum_{i=m+1}^{n} (\lg y_i - (\lg b + a \lg x_i)) = 0,$$

или

$$\left.\begin{aligned} mB + a\sum_{i=1}^{m} \lg x_i &= \sum_{i=1}^{m} \lg y_i, \\ (n-m)B + a\sum_{i=m+1}^{n} \lg x_i &= \sum_{i=m+1}^{n} \lg y_i, \end{aligned}\right\} \tag{34.42}$$

где n — число всех измерений, m — число измерений в первой группе.

Из системы уравнений (34.42) определяются значения a и B. Вычислив значение B по первой из формул (34.40), получим a и b — значения параметров эмпирической формулы (34.38).

Применение способа наименьших квадратов состоит в следующем. Составляем сумму квадратов погрешностей $\varepsilon_i = \lg y_i - (B + a \lg x_i)$:

$$u = u(a, B) = \sum_{i=1}^{n} (\lg y_i - (B + a \lg x_i))^2.$$

Находим частные производные функции $u = u(a, B)$ по a и по B, приравниваем их к нулю, в результате чего получаем систему двух уравнений относительно a и B:

$$\left.\begin{aligned} a\sum_{i=1}^{n} (\lg x_i)^2 + B\sum_{i=1}^{n} \lg x_i &= \sum_{i=1}^{n} \lg x_i \lg y_i, \\ a\sum_{i=1}^{n} \lg x_i + Bn &= \sum_{i=1}^{n} \lg y_i. \end{aligned}\right\} \tag{34.43}$$

Из системы уравнений (34.43) можно найти значения a и B, из равенства $\lg b = B$ (см. (34.40)) — значение b. Таким образом, будут определены искомые значения параметров a и b эмпирической формулы (34.38).

Пример 1. Определить вид эмпирической формулы и найти значения ее параметров по результатам наблюдений:

x	3	6	9	12	15	18	21	24
y	57,6	41,9	31,0	22,7	16,6	12,2	8,9	6,5

В прямоугольной декартовой системе координат построим точки $M_1(3;\ 57{,}6)$, $M_2(6;\ 41{,}9)$, $M_3(9;\ 31{,}0)$, $M_4(12;\ 22{,}7)$, $M_5(15;\ 16{,}6)$, $M_6(18;\ 12{,}2)$, $M_7(21;\ 8{,}9)$, $M_8(24;\ 6{,}5)$. Эти точки располагаются на дуге некоторой кривой.

Вычислим значения десятичных логарифмов чисел y_i:

y	57,6	41,9	31,0	22,7	16,6	12,2	8,9	6,5
$\lg y$	1,7604	1,6222	1,4912	1,3560	1,2201	1,0864	0,9494	0,8129

Построим точки $N_i(x_i,\ \lg y_i)$, $i=1,\ 2,\ \ldots,\ 8$, в полулогарифмической сетке (в системе координат $(x_i,\ \lg y_i)$). Точки N_i оказываются расположенными на некоторой прямой. Следовательно, можно предположить, что между x и y существует зависимость $y=be^{ax}$. Логарифмируя это равенство, получаем $\lg y = \lg b + (a \lg e)x$, $\lg y = \lg b + \frac{a}{\ln 10}x$, $\lg y=\lg b+\frac{a}{2{,}303}x$, $\lg y = \lg b + acx\left(c= \frac{1}{2{,}303}\right)$.

Для определения параметров a и b применяем метод средних. Имеющиеся 8 пар значений $(x_i,\ \lg y_i)$ разбиваем на 2 равные группы и составляем для каждой группы по 4 условных уравнения $\lg y_i= \lg b + \frac{a}{2{,}303}x_i$, а затем суммируем их:

$$\begin{array}{ll}
1{,}7604= \lg b+ 3ca & 1{,}2201= \lg b+15ca \\
1{,}6222= \lg b+ 6ca & 1{,}0864= \lg b+18ca \\
1{,}4192= \lg b+ 9ca & 0{,}9494= \lg b+21ca \\
1{,}3560= \lg b+12ca & 0{,}8129= \lg b+24ca \\
\hline
6{,}2298=4\lg b+30ca & 4{,}0688=4\lg b+78ca
\end{array}$$

Так как $c = \frac{1}{\ln 10}=\frac{1}{2{,}303}$, то из последних двух уравнений находим $a=-0{,}1037$, $\lg b=1{,}8952$, $b=78{,}56$. Следовательно, эмпирическая формула имеет вид $y=78{,}56\, e^{-0{,}1037\, x}$.

ЛИТЕРАТУРА

1. *Боровков А. А.* Теория вероятностей.— М.: Наука, 1976.— 352 с.

2. *Бугров Я. С., Никольский С. М.* Дифференциальное и интегральное исчисление.— М.: Наука, 1980.—432 с.—(Высш. математика).

3. *Бугров Я. С., Никольский С. М.* Дифференциальные уравнения. Кратные интегралы. Ряды. Функции комплексного переменного.— М.: Наука, 1981.—448 с.—(Высш. математика).

4. *Вентцель Е. С., Овчаров Л. А.* Теория вероятностей: Задачи и упражнения.— М.: Высш. шк., 1973.—366 с.

5. *Гмурман В. Е.* Теория вероятностей и математическая статистика.— М.: Высш. шк., 1977.—480 с.

6. *Гнеденко Б. В.* Курс теории вероятностей.— М.: Наука, 1965.— 400 с.

7. *Гусак А. А.* Задачи и упражнения по высшей математике.— Минск: Выш. шк., 1972, ч. 1.—328 с.; 1973, ч. 2.—384 с.

8. *Гусак А. А.* Элементы методов вычислений.— Минск: Изд-во БГУ, 1982.—168 с.

9. *Ильин В. А., Позняк Э. Г.* Основы математического анализа.— М.: Наука, 1971, ч. 1.—600 с.; 1980, ч. 2.—448 с.

10. *Корн Г.* и *Корн Т.* Справочник по математике для научных работников и инженеров.— М.: Наука, 1977.—831 с.

11. *Кудрявцев Л. Д.* Курс математического анализа.— М.: Высш. шк., 1981, т. 1.—287 с.; 1981, т. 2.—584 с.

12. *Кудрявцев В. А., Демидович Б. П.* Краткий курс высшей математики.— М.: Наука, 1975.—624 с.

13. *Матвеев Н. М.* Методы интегрирования обыкновенных дифференциальных уравнений.— Минск: Выш. шк., 1974.—768 с.

14. *Никольский С. М.* Курс математического анализа.— М.: Наука, 1973, т. 1.—432 с.; 1973, т. 2.—392 с.

15. *Смирнов В. И.* Курс высшей математики.— М.: Наука, 1974, т. 1.—479 с.; 1974, т. 2.—655 с.

16. *Смирнов М. М.* Дифференциальные уравнения в частных производных второго порядка.— Минск: Изд-во БГУ, 1974.—232 с.

17. *Смирнов М. М.* Задачи по уравнениям математической физики.— М.: Наука, 1975.—128 с.

18. *Тихонов А. Н., Самарский А. А.* Уравнения математической физики.— М.: Наука, 1972.—736 с.

19. *Чистяков В. П.* Курс теории вероятностей.— М.: Наука, 1981.—255 с.

АЛФАВИТНЫЙ УКАЗАТЕЛЬ

ОГЛАВЛЕНИЕ

РАЗДЕЛ II

МАТЕМАТИЧЕСКИЙ АНАЛИЗ
(продолжение)

Глава 18. Функции нескольких переменных

Глава 19. Кратные интегралы

Глава 20. Криволинейные интегралы. Интегралы по поверхности

Глава 21. Числовые ряды

Глава 22. Функциональные ряды

Глава 23. Ряды Фурье

РАЗДЕЛ III

ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Глава 24. Дифференциальные уравнения первого порядка

Глава 25. Дифференциальные уравнения высших порядков. Системы дифференциальных уравнений

Глава 26. Дифференциальные уравнения с частными производными первого и второго порядка

Глава 27. Простейшие дифференциальные уравнения математической физики

РАЗДЕЛ IV

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ НАБЛЮДЕНИЙ

Глава 28. Случайные события и вероятности

Глава 29. Случайные величины и функции распределения

Глава 30. Числовые характеристики случайных величин

Глава 31. Некоторые законы распределения случайных величин

Глава 32. Закон больших чисел. Предельные теоремы

Глава 33. Элементы математической статистики

Глава 34. Математическая обработка результатов наблюдений