нием функции f на множестве E и обозначается через $\omega(f;\ E)$ или просто $\omega(f)$.

Очевидно, что в силу (19.16)

$$\omega(f;\ E)=\sup_{x'\in E,\ x''\in E}[f(x'')-f(x')].$$

Замечание. Из сказанного в этом и предыдущем параграфах, в частности, видно, что в ряде вопросов, относящихся к функциям многих переменных, всю их специфику можно в достаточной мере усмотреть уже в двумерном или трехмерном случае. Благодаря удачно выбранным определениям и обозначениям доказательства теорем автоматически переносятся со случая $n=2$ на произвольный n-мерный случай иногда лишь приводя к некоторому техническому усложнению записи. Случай же $n=2$ имеет преимущество геометрической наглядности и более простой записи, когда в ней участвуют координаты точек. Поэтому для большей ясности и простоты изложения мы, как правило, будем подробно рассматривать лишь случаи $n=2$ или $n=3$, а в случае произвольного n — лишь формулировать соответствующие результаты или даже только отмечать возможность их обобщения на случай произвольного n. Если же при рассмотрении какого-либо вопроса при $n>3$ возникают какие-либо специфические трудности, то этот вопрос будет детально рассматриваться в общем случае.

§ 20. ЧАСТНЫЕ ПРОИЗВОДНЫЕ. ДИФФЕРЕНЦИРУЕМОСТЬ ФУНКЦИЙ МНОГИХ ПЕРЕМЕННЫХ

20.1. ЧАСТНЫЕ ПРОИЗВОДНЫЕ И ЧАСТНЫЕ ДИФФЕРЕНЦИАЛЫ

Рассмотрим сначала случай функций трех переменных.

Определение 1. *Пусть в некоторой окрестности точки $(x_0,\ y_0,\ z_0)$ задана функция $u=u(x,\ y,\ z)$. Фиксируя переменные y и z: $y=y_0$, $z=z_0$, получим функцию одного переменного x: $u=u(x,\ y_0,\ z_0)$. Обычная производная* (см. п. 9.1) *этой функции в точке $x=x_0$ называется частной производной функции $u(x,y,z)$ в точке (x_0,y_0,z_0) по x и обозначается через $\frac{\partial u(x_0,\ y_0,\ z_0)}{\partial x}$.*

Таким образом,

$$\frac{\partial u(x_0,\ y_0,\ z_0)}{\partial x}\overset{\text{def}}{=}\left.\frac{du(x,\ y_0,\ z_0)}{dx}\right|_{x=x_0}$$

Заметим, что обозначение частной производной по переменной x через $\frac{\partial u(x_0,\ y_0,\ z_0)}{\partial x}$ традиционно. Правильнее было бы писать $\frac{\partial u}{\partial x}(x_0,\ y_0,\ z_0)$, так как $\frac{\partial u}{\partial x}$ является единым символом, обозначающим новую функцию, значение которой и рассматривается в точке $(x_0,\ y_0,\ z_0)$.

Если вспомнить определение обычной производной $\frac{du}{dx}$ (см. п. 9.1) то, согласно этому определению, можно написать

$$\frac{\partial u(x_0, y_0, z_0)}{\partial x} = \lim_{\Delta x \to 0} \frac{u(x_0+\Delta x, y_0, z_0) - u(x_0, y_0, z_0)}{\Delta x},$$

или, если ввести обозначение $u(x_0+\Delta x, y_0, z_0) - u(x_0, y_0, z_0) = \Delta_x u$, ($\Delta_x u$ — приращение функции по переменной x),

$$\frac{\partial u}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta_x u}{\Delta x}.$$

Аналогично вводятся частные производные по y и z:

$$\frac{\partial u(x_0, y_0, z_0)}{\partial y} = \left.\frac{du(x_0, y, z_0)}{dy}\right|_{y=y_0},$$
$$\frac{\partial u(x_0, y_0, z_0)}{\partial z} = \left.\frac{du(x_0, y_0, z)}{dz}\right|_{z=z_0},$$

или

$$\frac{\partial u}{\partial y} = \lim_{\Delta y \to 0} \frac{\Delta_y u}{\Delta y}, \quad \frac{\partial u}{\partial z} = \lim_{\Delta z \to 0} \frac{\Delta_z u}{\partial z},$$

где $\Delta_y u$ и $\Delta_z u$ — приращения функции соответственно по переменным y и z.

По аналогии с функциями одной переменной линейные функции $\frac{\partial u}{\partial x} dx$, $\frac{\partial u}{\partial y} dy$, $\frac{\partial u}{\partial z} dz$ переменных dx, dy, dz, называемых *дифференциалами независимых переменных*, называются *частными дифференциалами* функции $u(x, y, z)$ соответственно по переменным x, y, z и обозначаются через

$$d_x u = \frac{\partial u}{\partial x} dx, \quad d_y u = \frac{\partial u}{\partial y} dy, \quad d_z u = \frac{\partial u}{\partial z} dz.$$

Аналогичные определения имеют место для любого числа переменных.

Если функция $y = f(x_1, \ldots, x_n)$ определена в некоторой окрестности точки $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$, то по определению

$$\frac{\partial f(x_1^{(0)}, \ldots, x_n^{(0)})}{\partial x_i} \stackrel{\text{def}}{=} \left.\frac{df(x_1^{(0)}, \ldots, x_{i-1}^{(0)}, x_i, x_{i+1}^{(0)}, \ldots, x_n^{(0)})}{dx_i}\right|_{x_i = x_i^{(0)}}, \quad (20.1)$$

или, что то же, опуская обозначение аргумента, $\frac{\partial y}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{\Delta_{x_i} y}{\Delta x_i}$, где $\Delta_{x_i} y = f(x_1^{(0)}, \ldots, x_{i-1}^{(0)}, x_i^{(0)} + \Delta x_i, x_{i+1}^{(0)}, \ldots, x_n^{(0)}) - f(x_1^{(0)}, \ldots$ $\ldots, x_{i-1}^{(0)}, x_i^{(0)}, x_{i+1}^{(0)}, \ldots, x_n^{(0)})$. Для обозначения частной производной $\frac{\partial y}{\partial x_i}$ применяются также обозначения y_{x_i} или f_{x_i}.

Частный дифференциал $d_{x_i}y$ определяется по формуле

$$d_{x_i}y \overset{\text{def}}{=} \frac{\partial y}{\partial x_i}dx_i, \quad -\infty < dx_i < +\infty, \tag{20.2}$$

и тем самым является линейной функцией переменной dx_i, называемой *дифференциалом независимой переменной* x_i. Здесь везде $i=1, 2 \ldots, n$. В случае $n=1$ частная производная совпадает с обычной производной, а частный дифференциал — с обычным дифференциалом.

Подчеркнем, что $\frac{\partial y}{\partial x_i}$ — единый символ, т. е. в нем числитель и знаменатель не имеют самостоятельного смысла. С другой стороны, частная производная $\frac{\partial y}{\partial x_i}$, конечно, может быть записана и в виде частного двух дифференциалов: $\frac{\partial y}{\partial x_i} = \frac{d_{x_i}y}{dx_i}$.

Из определения частных производных, как обычных производных при условии фиксирования всех переменных, кроме одной, по которой берется производная, следует, что при вычислении частных производных можно пользоваться правилами вычисления обычных производных. Пусть, например, требуется найти производную $\frac{\partial z}{\partial y}$ функции $z = xye^{x/y}$. Для этого, зафиксировав в этой формуле x, получим функцию одной переменной y; вычисляя ее производную, будем иметь:

$$\frac{\partial z}{\partial y} = xe^{x/y} + xye^{x/y}\frac{\partial}{\partial y}\left(\frac{x}{y}\right) = \frac{x(y-x)e^{x/y}}{y}.$$

В заключение этого пункта отметим, что из непрерывности в данной точке функции n переменных не вытекает существование у нее в этой точке частных производных. Соответствующий пример в случае $n=1$ был приведен ранее (см. п. 9.2). Важно заметить, что при $n \geqslant 2$ из существования даже всех частных производных в некоторой точке не следует непрерывность функции в этой точке*). Это естественно, поскольку условие непрерывности функции нескольких переменных в точке накладывает определенное ограничение на ее поведение при приближении к этой точке по в с е м направлениям, в то время как существование частных производных в точке означает, что функция удовлетворяет определенным условиям при приближении к указанной точке лишь в направлении координатных осей.

Чтобы в этом наглядно убедиться, рассмотрим функцию $f(x, y)$, равную 0, если $xy=0$, и 1, если $xy \neq 0$. Очевидно,

---

*) Напомним, что при $n=1$, т. е. для функции одной переменной из существования в точке производной вытекает и непрерывность функции в этой точке (см. п. 9.2).

$f(x, 0) \equiv f(0, y) \equiv 0$, и, следовательно,

$$\frac{\partial f(0, 0)}{\partial x} = \frac{\partial f(0, 0)}{\partial y} = 0.$$

Однако эта функция разрывна в точке (0, 0), так как, например, ее предел вдоль прямой $y=x$ при $(x, y) \to (0, 0)$ равен 1, а $f(0, 0)=0$.

Более того, существуют функции, имеющие частные производные во всех точках и все-таки разрывные. Примером такой функции является функция

$$f(x, y) = \begin{cases} xy/(x^2+y^2) & \text{при} \quad x^2+y^2>0, \\ 0 & \text{при} \quad x=y=0. \end{cases} \tag{20.3}$$

Эта функция имеет частные производные во всей плоскости и разрывна в точке (0, 0) (почему?).

20.2. ДИФФЕРЕНЦИРУЕМОСТЬ ФУНКЦИЙ В ТОЧКЕ

Рассмотрим сначала случай функций двух переменных. Пусть функция $z=f(x, y)$ определена в некоторой δ-окрестности $U = U(M_0; \delta)$ точки $M_0=(x_0, y_0)$ и пусть (рис. 88)

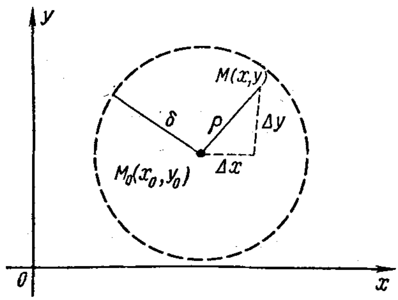

Рис. 88

$$M=(x, y) \in U(M_0; \delta),$$
$$\Delta x = x - x_0, \quad \Delta y = y - y_0$$

и, значит,

$$\rho = \rho(M, M_0) = \sqrt{\Delta x^2 + \Delta y^2} < \delta.$$

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

Обычно Δz называется *полным приращением функции*; это название объясняется тем, что здесь, вообще говоря, все независимые переменные получают приращения, отличные от нуля.

Определение 2. *Функция $z=f(x, y)$ называется дифференцируемой в точке (x_0, y_0), если существуют два такие числа A и B, что*

$$\Delta z = A\,\Delta x + B\,\Delta y + \alpha(\Delta x, \Delta y), \tag{20.4}$$

где при $\rho \neq 0$:

$$\alpha(\Delta x; \Delta y) = \varepsilon(\Delta x, \Delta y)\,\rho, \quad \lim_{\rho \to 0} \varepsilon(\Delta x, \Delta y) = 0\ ^{*)}. \tag{20.5}$$

---

*) Напомним, что, согласно сделанному соглашению, запись $\lim\limits_{\rho \to 0} f$ равносильна записи $\lim\limits_{M \to M_0} f$, где $\rho = \rho(M, M_0)$.

Из (20.4) следует, что $\alpha(0, 0)=0$.

Вместе с тем заметим, что значение функции $\varepsilon(\Delta x, \Delta y)$ в точке $(0, 0)$ не определено формулой (20.5).

Определение 3. *В случае дифференцируемости функции f в точке (x_0, y_0) линейная функция $A\,\Delta x+B\,\Delta y$ переменных Δx и Δy называется полным дифференциалом, или просто дифференциалом, функции f в точке (x_0, y_0) и обозначается через dz.*

Таким образом $dz=A\,\Delta x+B\,\Delta y$.

Вместо Δx и Δy употребляются также равнозначные обозначения dx и dy, т. е. пишут $dz=A\,dx+B\,dy$. Из (20.5) следует, что

$$\lim_{M\to M_0}\frac{\alpha(\Delta x, \Delta y)}{\rho}=0. \tag{20.6}$$

Функции $\alpha(\Delta x, \Delta y)$, обладающие свойством (20.6), будем обозначать по аналогии с функциями одного переменного через $o(\rho)$ при $\rho\to 0$ *). Применяя это обозначение, определение дифференцируемости можно переписать в виде

$$\Delta z=A\,\Delta x+B\,\Delta y+o(\rho),\quad \rho\to 0. \tag{20.7}$$

Лемма 1. *Условие* (20.5) *эквивалентно условию*

$$\alpha(\Delta x, \Delta y)=\varepsilon_1(\Delta x, \Delta y)\,\Delta x+\varepsilon_2(\Delta x, \Delta y)\,\Delta y,\quad \rho\neq 0, \tag{20.8}$$

где $\lim\limits_{\rho\to 0}\varepsilon_1=\lim\limits_{\rho\to 0}\varepsilon_2=0$.

Доказательство. Пусть выполнено условие (20.5), т. е. $\alpha=\varepsilon\rho$, $\rho\neq 0$, где $\varepsilon\to 0$ при $\rho\to 0$; тогда

$$\alpha=\varepsilon\rho=\varepsilon\sqrt{\Delta x^2+\Delta y^2}=$$
$$=\varepsilon\frac{\Delta x}{\sqrt{\Delta x^2+\Delta y^2}}\Delta x+\varepsilon\frac{\Delta y}{\sqrt{\Delta x^2+\Delta y^2}}\Delta y=\varepsilon_1\,\Delta x+\varepsilon_2\,\Delta y,$$

где $\varepsilon_1=\dfrac{\varepsilon\,\Delta x}{\sqrt{\Delta x^2+\Delta y^2}}$, $\varepsilon_2=\dfrac{\varepsilon\,\Delta y}{\sqrt{\Delta x^2+\Delta y^2}}$. Замечая, что $|\Delta x/\sqrt{\Delta x^2+\Delta y^2}|\leqslant 1$, $|\Delta y/\sqrt{\Delta x^2+\Delta y^2}|\leqslant 1$, имеем $|\varepsilon_1|\leqslant|\varepsilon|$, $|\varepsilon_2|\leqslant|\varepsilon|$, откуда $\lim\limits_{\rho\to 0}\varepsilon_1=\lim\limits_{\rho\to 0}\varepsilon_2=0$, т. е. получилось представление функции α в виде (20.8).

Пусть, наоборот, выполнено условие (20.8), т. е. $\alpha=\varepsilon_1\,\Delta x+\varepsilon_2\,\Delta y$, $\rho\neq 0$, где $\varepsilon_1\to 0$ и $\varepsilon_2\to 0$ при $\rho\to 0$; тогда

$$\alpha=\left(\frac{\Delta x}{\sqrt{\Delta x^2+\Delta y^2}}\varepsilon_1+\frac{\Delta y}{\sqrt{\Delta x^2+\Delta y^2}}\varepsilon_2\right)\sqrt{\Delta x^2+\Delta y^2}=\varepsilon\rho,$$

*) Вообще для функций α и β многих переменных $\alpha=o(\beta)$ при $x\to x^{(0)}$, $x\in E\subset R^n$, $x^{(0)}\in R^n$, если $\alpha(x)=\varepsilon(x)\beta(x)$, где $\lim\limits_{x\to x^{(0)},\ x\in E}\varepsilon(x)=0$. В этом случае будем говорить, что функция α является бесконечно малой по сравнению с функцией β при $x\to x^{(0)}$, $x\in E$.

где $\varepsilon = \frac{\Delta x}{\sqrt{\Delta x^2 + \Delta y^2}} \varepsilon_1 + \frac{\Delta y}{\sqrt{\Delta x^2 + \Delta y^2}} \varepsilon_2$ и, значит, $|\varepsilon| \leqslant |\varepsilon_1| + |\varepsilon_2|$; поэтому $\varepsilon \to 0$ при $\rho \to 0$. Таким образом получилось представление функции α в виде (20.5). □

Теорема 1. *Если функция $z = f(x, y)$ дифференцируема в точке (x_0, y_0), то она непрерывна в этой точке.*

Действительно, так как $|\Delta x| \leqslant \rho$ и $|\Delta y| \leqslant \rho$, то из формул (20.4) и (20.5) следует, что $\Delta z \to 0$ при $\rho \to 0$, что и означает непрерывность функции f в точке (x_0, y_0). □

Теорема 2. *Если функция $z = f(x, y)$ дифференцируема в точке (x_0, y_0) и $dz = A\,dx + B\,dy$ — ее дифференциал в этой точке, то в точке (x_0, y_0) у функции f существуют все частные производные и*

$$\frac{\partial f(x_0, y_0)}{\partial x} = A, \quad \frac{\partial f(x_0, y_0)}{\partial y} = B. \tag{20.9}$$

Таким образом,

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy. \tag{20.10}$$

Доказательство. Согласно определению дифференцируемости (см. (20.4) и (20.8)),

$$\Delta z = A\,\Delta x + B\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y,$$

где

$$\lim_{\rho \to 0} \varepsilon_1 = \lim_{\rho \to 0} \varepsilon_2 = 0. \tag{20.11}$$

Полагая $\Delta y = 0$, получим $\Delta z = \Delta_x z = A\,\Delta x + \varepsilon_1 \Delta x$, где $\lim\limits_{\Delta x \to 0} \varepsilon_1 = 0$ (это следует из (20.11), поскольку, полагая $\Delta y = 0$, получим $\rho = |\Delta x|$). Отсюда

$$\frac{\Delta_x z}{\Delta x} = A + \varepsilon_1, \tag{20.12}$$

где при $\Delta x \to 0$ правая часть стремится к пределу, равному A, поэтому и левая часть при $\Delta x \to 0$ имеет тот же предел, а это и означает (см. (20.1)), что в точке (x_0, y_0) существует частная производная $\frac{\partial z}{\partial x} = A$. Аналогично, полагая в (20.4) $\Delta x = 0$ и переходя к пределу, при $\Delta y \to 0$ получим $\frac{\partial z}{\partial y} = B$. □

Следствие. *Если функция $f(x, y)$ дифференцируема в точке (x_0, y_0), то она имеет единственный дифференциал.*

Единственность дифференциала непосредственно вытекает из формул (20.9), так как частные производные в данной точке определяются однозначно.

Вспоминая определения частных дифференциалов (см. (20.2)), формулу (20.10) можно переписать в виде

$$dz = d_x z + d_y z,$$

т. е. полный дифференциал функции (когда он существует) является суммой ее частных дифференциалов.

Заметим, что утверждение, обратное теореме 2, не имеет места: существуют функции, имеющие все частные производные во всех точках плоскости, но не дифференцируемые в некоторой точке. Примером может служить функция (20.3), приведенная в конце предыдущего пункта: в точке (0, 0) эта функция не непрерывна, откуда в силу теоремы 1 вытекает, что в точке (0, 0) она и не дифференцируема.

Из сказанного следует, что не всегда выражение $d_x z + d_y z$, когда оно имеет смысл, является полным дифференциалом функции. Связь между дифференцируемостью функции в точке и существованием в этой точке частных производных сложнее, чем связь между дифференцируемостью и существованием производной у функции одной переменной.

Сформулируем достаточные условия в терминах свойств частных производных для дифференцируемости функции.

Теорема 3. *Пусть функция $z = f(x, y)$ в некоторой окрестности точки (x_0, y_0) имеет частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, которые непрерывны в самой точке (x_0, y_0); тогда функция $z = f(x, y)$ дифференцируема в этой точке.*

Следствие. *Если функция $z = f(x, y)$ имеет в некоторой окрестности точки (x_0, y_0) частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, причем эти частные производные непрерывны в самой точке (x_0, y_0), то и функция $z = f(x, y)$* также *непрерывна в этой точке.*

Доказательство теоремы. Обозначим через $U(\delta)$ δ-окрестность точки (x_0, y_0), в которой определена вместе со своими частными производными f_x и f_y функция f. Выберем Δx и Δy так, чтобы $(x_0 + \Delta x, y_0 + \Delta y) \in U(\delta)$. Замечая, что

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) =$$
$$= [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)] + [f(x_0, y_0 + \Delta y) - f(x_0, y_0)],$$

применим к выражениям, стоящим в квадратных скобках и являющимся приращениями функций только по одной переменной, формулу конечных приращений Лагранжа (см. п. 11.2). Это возможно, поскольку функция $f(x, y_0 + \Delta y)$, рассматриваемая как функция одного переменного x, имеет на отрезке с концами в точках x_0 и $x_0 + \Delta x$ производную (являющуюся частной производной по x функции f), поэтому она и непрерывна на указанном отрезке. Таким образом, функция $f(x, y_0 + \Delta y)$ удовлетворяет всем условиям, при которых была доказана формула конечных приращений Лагранжа. Аналогично проверяется и возможность применения формулы Лагранжа к функции $f(x_0, y)$, рассматриваемой как функция одного переменного y, на отрезке с концами в точках y_0 и $y_0+\Delta y$. Тогда

$$\Delta z = f_x(x_0+\theta_1\Delta x,\ y_0+\Delta y)\,\Delta x + f_y(x_0,\ y_0+\theta_2\Delta y)\,\Delta y,$$
$$0<\theta_1<1,\ 0<\theta_2<1, \qquad (20.13)$$

причем θ_1 и θ_2 зависят, конечно, от выбора точки $(x_0+\Delta x,$ $y_0+\Delta y)$, т. е. от Δx и Δy.

Если

$$f_x(x_0+\theta_1\Delta x,\ y_0+\Delta y) - f_x(x_0,\ y_0) = \varepsilon_1,$$
$$f_y(x_0,\ y_0+\theta_2\Delta y) - f_y(x_0,\ y_0) = \varepsilon_2, \qquad (20.14)$$

то в силу непрерывности частных производных f_x и f_y в точке $(x_0,\ y_0)$ имеем

$$\lim_{\rho\to 0}\varepsilon_1 = \lim_{\rho\to 0}\varepsilon_2 = 0. \qquad (20.15)$$

Подставив (20.14) в (20.13), получим:

$$\Delta z = f_x(x_0,\ y_0)\,\Delta x + f_y(x_0,\ y_0)\,\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y, \qquad (20.16)$$

что в силу выполнения условий (20.15) и означает дифференцируемость функции f в точке $(x_0,\ y_0)$ (см. (20.4) и (20.8)). $\square$

Следствие из теоремы вытекает из того обстоятельства, что функция, дифференцируемая в некоторой точке, является и непрерывной в ней (см. теорему 1).

Теорема 3 имеет важное значение, связанное с тем, что понятие дифференцируемости функции играет первостепенную роль в ряде разделов теории функций многих переменных. Однако непосредственная проверка дифференцируемости функции (например, для выяснения возможности применения тех или иных теорем) часто бывает затруднительна, в то время как проверка непрерывности частных производных, для вычисления которых имеется удобный аналитический аппарат, оказывается проще.

Определение 4. *Функция, имеющая в некоторой точке (или соответственно на некотором множестве) непрерывные частные производные, называется непрерывно дифференцируемой в этой точке (соответственно на этом множестве).*

Сопоставим определение дифференцируемости функции (определение 2) и определение непрерывной дифференцируемости (определение 4). Дифференцируемость функции в точке означает существование в этой точке дифференциала, т. е. справедливость для этой точки формулы (20.4). Непрерывная же дифференцируемость функции в точке означает непрерывность в этой точке ее частных производных. Таким образом, дифференцируемость функции связана с понятием дифференциала, а непрерывная дифференцируемость — с понятием частных производных. Вместе с тем из непрерывной дифференцируемости в точке (на открытом множестве) следует дифференцируемость в этой точке (соответственно на этом множестве); в этом состоит утверждение теоремы 3.

В дальнейшем нам понадобятся некоторые дополнительные свойства функций ε_1 и ε_2 из формулы (20.16).

Определение 5. *Пусть A и B — два плоских множества, $A \subset R^2_{xy}$, $B \subset R^2_{uv}$ и пусть функция $f = f(x, y, u, v)$ определена для $(x, y) \in A$, $(u, v) \in B$.*

Функция f называется равномерно стремящейся к нулю на множестве A переменных x, y при $(u, v) \to (u_0, v_0)$, если для любого $\varepsilon > 0$ существует такое $\delta = \delta(\varepsilon) > 0$, что для всех (u, v), удовлетворяющих условию $\sqrt{(u-u_0)^2+(v-v_0)^2} < \delta$, $(u, v) \neq (u_0, v_0)$ и всех $(x, y) \in A$ выполняется условие $|f(x, y, u, v)| < \varepsilon$.

Общее определение равномерного стремления функции к пределу будет дано в п. 39.4.

Теорема 4. *Пусть функция $z = f(x, y)$ непрерывно дифференцируема на открытом множестве $G \subset R^2$. Тогда*

$$\Delta z = f_x(x, y)\Delta x + f_y(x, y)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y, \qquad (20.17)$$

где функции $\varepsilon_1 = \varepsilon_1(x, y, \Delta x, \Delta y)$ и $\varepsilon_2 = \varepsilon_2(x, y, \Delta x, \Delta y)$ равномерно стремятся к нулю при $\rho = \sqrt{\Delta x^2 + \Delta y^2} \to 0$ на любом компакте $A \subset G$.

Доказательство. Пусть A — компакт, лежащий в G. Тогда замкнутые множества A и $R^2 \setminus G$ не пересекаются, и так как A ограничено (см. п. 18.3, теорему 3), то $d = \rho(A, R^2 \setminus G) > 0$ (см. лемму 7, п. 18.2).

Множество $A_{d/2} = \{(x, y) : \rho((x, y), A) \leqslant d/2\}$ содержится во множестве G и является компактом (см. лемму 11 п. 18.3).

Пусть теперь $\rho = \sqrt{\Delta x^2 + \Delta y^2} < d/2$; тогда при $(x_0, y_0) \in A$ получим (см. (20.13)):

$$(x_0 + \theta_1 \Delta x, y_0 + \Delta y) \in A_{d/2}, \quad (x_0, y_0 + \theta_2 \Delta y) \in A_{d/2},$$

и, следовательно, согласно формулам (20.14), имеем неравенства

$$|\varepsilon_1| \leqslant \omega(\rho; f_x; A_{d/2}), \quad |\varepsilon_2| \leqslant \omega(\rho; f_y; A_{d/2}),$$

где в их правых частях стоят соответственно модули непрерывности функций f_x и f_y. Из непрерывности частных производных f_x и f_y на компакте $A_{d/2}$ следует, что

$$\lim_{\rho \to 0} \omega(\rho; f_x; A_{d/2}) = 0 \quad \text{и} \quad \lim_{\rho \to 0} \omega(\rho; f_y; A_{d/2}) = 0.$$

Поэтому для любого $\varepsilon > 0$ существует $\delta = \delta(\varepsilon) > 0$ такое, что для всех $\rho < \delta$ выполняются неравенства

$$\omega(\rho; f_x; A_{d/2}) < \varepsilon, \quad \omega(\rho; f_y; A_{d/2}) < \varepsilon.$$

Следовательно, для всех $\rho < \delta$ и всех $(x_0, y_0) \in A$ справедливы неравенства

$$|\varepsilon_1| < \varepsilon, \quad |\varepsilon_2| < \varepsilon.$$

Это и означает равномерное стремление к нулю при $\rho \to 0$ функций ε_1 и ε_2 на компакте A. $\square$

Замечание. В предположениях теоремы 2 приращение функции Δz представимо также в виде

$$\Delta z = f_x(x, y)\Delta x + f_y(x, y)\Delta y + \varepsilon\rho, \tag{20.18}$$

где $\varepsilon = \varepsilon(x, y, \Delta x, \Delta y)$ равномерно на каждом компакте $A \subset G$ стремится к нулю, когда $\rho = \sqrt{\Delta x^2 + \Delta y^2} \to 0$. Для доказательства достаточно в формуле (20.18) положить $\varepsilon = \varepsilon_1 \frac{\Delta x}{\rho} + \varepsilon_2 \frac{\Delta y}{\rho}$ (сравните с доказательством леммы в начале этого пункта).

Все определения и утверждения этого пункта переносятся и на случай функции $y = f(x)$, $x = (x_1, \ldots, x_n)$, любого числа n переменных, определенной в некоторой окрестности точки $x^{(0)}$. Например, условие дифференцируемости в данной точке $x^{(0)}$ в общем случае выглядит так:

$$\Delta y = A_1\Delta x_1 + \ldots + A_n\Delta x_n + o(\rho), \quad \rho \to 0, \tag{20.19}$$

где

$$\rho = \sqrt{\sum_{i=1}^{n} \Delta x_i^2}, \quad \Delta y = f(x_1, \ldots, x_n) - f(x_1^{(0)}, \ldots, x_n^{(0)}),$$

$$\Delta x_i = x_i - x_i^{(0)}, \quad i = 1, 2, \ldots, n,$$

причем в этом случае $A_i = \frac{\partial f(x^{(0)})}{\partial x_i}$, $i = 1, 2, \ldots, n$.

Таким образом, если функция f дифференцируема, то

$$f(x) = f(x^{(0)}) + A_1(x_1 - x_1^{(0)}) + \ldots + A_n(x_n - x_n^{(0)}) + o(\rho),$$

$$\rho \to 0, \tag{20.20}$$

т. е. функция f в окрестности данной точки с точностью до бесконечно малых более высокого порядка, чем $\rho = \sqrt{\sum_{i=1}^{n}(x_i - x_i^{(0)})^2}$, равна линейной функции *). Образно говоря, дифференцируемость функции в данной точке означает, что функция f «почти линейна» в окрестности этой точки; точный смысл выражения «почти линейна» заключается в формуле (20.20).

В случае, когда имеет место (20.19), линейная функция $\frac{\partial f(x)}{\partial x}\Delta x_1 + \ldots + \frac{\partial f(x)}{\partial x_n}\Delta x_n$ от переменных $\Delta x_1, \ldots, \Delta x_n$ (здесь вместо $x^{(0)}$ написано x) называется *дифференциалом функции*, или, подробнее, полным дифференциалом функции в данной точке x и обо-

*) Функции вида $y = c_0 + c_1x_1 + \ldots + c_nx_n$, где c_i — постоянные, $i = 0, 1, 2, \ldots, n$, называются *линейными функциями* n переменных, или, что то же самое, *линейными функциями точки* $x = (x_1, \ldots, x_n) \in R^n$.

значается $df(x)$:

$$df(x)=\frac{\partial f(x)}{\partial x_1}\Delta x_1+\ldots+\frac{\partial f(x)}{\partial x_n}\Delta x_n. \tag{20.21}$$

Дифференциал, как и всякая линейная функция n переменных определен на всем n-мерном пространстве R^n. Таким образом, *формула* (20.21) *имеет смысл для всех значений* Δx_i, $i=1, 2, \ldots, n$, в то время как *формула* (20.19) — *только для тех, которые не выводят за область определения функции* f.

Переменные Δx_i называются также *дифференциалами переменных* x_i и обозначаются dx_i, $i=1, 2, \ldots, n$. В этих обозначениях дифференциал функции f записывается в виде

$$df(x)=\frac{\partial f(x)}{\partial x_1}dx_1+\ldots+\frac{\partial f(x)}{\partial x_n}dx_n.$$

Очевидно, что $\Delta f(x)=df(x)+o(\rho)$ при $\rho\to 0$.

Если же рассматривать дифференциал и при изменении точки $x=(x_1, \ldots, x_n)$, то он будет уже являться функцией от $2n$ переменных: $x_1, \ldots, x_n, dx_1, \ldots, dx_n$.

Теоремы 1 — 4 настоящего параграфа очевидным образом обобщаются на функции n переменных, поэтому мы не будем приводить их формулировки.

20.3. ДИФФЕРЕНЦИРОВАНИЕ СЛОЖНОЙ ФУНКЦИИ

Теорема 5. *Пусть функции $x(t)$ и $y(t)$ одного переменного t дифференцируемы в точке t_0 (что, как мы знаем, эквивалентно существованию у них производных в точке t_0,* см. п. 9.2) *и пусть $x_0=x(t_0)$, $y_0=y(t_0)$. Если функция $z=f(x, y)$ дифференцируема в точке (x_0, y_0), то сложная функция $z=f(x(t), y(t))$ определена в некоторой окрестности точки t_0, имеет в t_0 производную и эта производная вычисляется по формуле*

$$\frac{dz}{dt}=\frac{\partial z}{\partial x}\frac{dx}{dt}+\frac{\partial z}{\partial y}\frac{dy}{dt}, \tag{20.22}$$

или, подробнее,

$$\frac{df(x(t_0), y(t_0))}{dt}=\frac{\partial f(x_0, y_0)}{\partial x}\frac{dx(t_0)}{dt}+\frac{\partial f(x_0, y_0)}{\partial y}\frac{dy(t_0)}{dt}.$$

Доказательство. Функция $f(x, y)$, согласно определению дифференцируемости функции, определена в некоторой окрестности точки (x_0, y_0). Из дифференцируемости же функций $x(t)$ и $y(t)$ следует их непрерывность в точке t_0. Поэтому, согласно замечанию к теореме 2 в п. 19.4, в некоторой окрестности точки t_0 определена сложная функция $f(x(t), y(t))$.

Дифференцируемость функции $z=f(x, y)$ в точке (x_0, y_0) означает, что ее полное приращение $\Delta z=f(x_0+\Delta x, y_0+\Delta y)-f(x_0, y_0)$

представимо в виде

$$\Delta z = \frac{\partial z}{\partial x} \Delta x + \frac{\partial z}{\partial y} \Delta y + \varepsilon \sqrt{\Delta x^2 + \Delta y^2}, \qquad (20.23)$$

где функция $\varepsilon = \varepsilon(\Delta x, \Delta y)$ такова, что $\lim\limits_{\rho \to 0} \varepsilon(\Delta x, \Delta y) = 0$. Здесь, как обычно, $\rho = \sqrt{\Delta x^2 + \Delta y^2}$. Доопределим функцию $\varepsilon(\Delta x, \Delta y)$ в точке $(0, 0)$, положив $\varepsilon(0, 0) = 0$ (ср. с доказательством теоремы 4 в п. 9.7). Так, доопределенная функция $\varepsilon(\Delta x, \Delta y)$ является непрерывной в точке $(0, 0)$.

Пусть теперь Δt — приращение переменной t и $\Delta x = x(t_0 + \Delta t) - x(t_0)$, $\Delta y = y(t_0 + \Delta t) - y(t_0)$. Разделим обе части равенства (20.23) на Δt:

$$\frac{\Delta z}{\Delta t} = \frac{\partial z}{\partial x} \frac{\Delta x}{\Delta t} + \frac{\partial z}{\partial y} \frac{\Delta y}{\Delta t} \pm \varepsilon \sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2}. \qquad (20.24)$$

При $\Delta t \to 0$ в силу непрерывности функций $x(t)$ и $y(t)$ в точке t_0 получим $\Delta x \to 0$ и $\Delta y \to 0$, а значит, и $\lim\limits_{\Delta t \to 0} \rho = 0$. Отсюда, по теореме о композиции непрерывных функций (см. п. 19.3), $\lim\limits_{\Delta t \to 0} \varepsilon(\Delta x, \Delta y) = 0$. Заметим, наконец, что существует конечный предел

$$\lim_{\Delta t \to 0} \sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2} = \sqrt{x'^2(t_0) + y'^2(t_0)}.$$

Из всего этого следует, что при $\Delta t \to 0$ правая часть формулы (20.24) стремится к конечному пределу $\frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$ $(t = t_0)$, поэтому и левая часть этой формулы, т. е. $\frac{\Delta z}{\Delta t}$, стремится к тому же пределу, а это и означает, что в точке t_0 существует производная $\frac{dz}{dt}$ и выражается формулой (20.22). □

Отметим, что, хотя в окончательную формулу производной сложной функции (20.22) входят только частные производные $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ функции $z = f(x, y)$, по ходу доказательства существенно использовалось более сильное свойство этой функции, чем существование частных производных, а именно ее дифференцируемость.

Упражнение 1. Показать, что при отказе от требования дифференцируемости функции $z = f(x, y)$ в точке (x_0, y_0), а лишь при предположении существования в этой точке частных производных $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ и существования производных $\frac{dx}{dt}$ и $\frac{dy}{dt}$ в точке t_0 формула (20.22), вообще говоря, несправедлива и, более того, сложная функция $f[x(t), y(t)]$ (предполагается, что она имеет смысл), вообще говоря, не имеет производной в точке t_0.

Следствие. *Пусть функции $x = x(u, v)$ и $y = y(u, v)$ определены в некоторой окрестности точки (u_0, v_0), а функция $z = f(x, y)$ определена в некоторой окрестности точки (x_0, y_0), где $x_0 = x(u_0, v_0)$, $y_0 = y(u_0, v_0)$.*

Если функция $f(x, y)$ дифференцируема в точке (x_0, y_0) и если в точке (u_0, v_0) существуют частные производные $\frac{\partial x}{\partial u}$ и $\frac{\partial y}{\partial u}$, то в этой точке (u_0, v_0) существует и частная производная $\frac{\partial z}{\partial u}$ сложной функции $z = f[x(u, v), y(u, v)]$, причем

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}. \qquad (20.25)$$

Доказательство. Зафиксируем $v = v_0$ и рассмотрим сложную функцию $z = f[x(u, v_0), y(u, v_0)]$ одного переменного u. Согласно теореме 5, эта функция определена в некоторой окрестности точки u_0 и имеет в этой точке производную. Таким образом, производная $\frac{\partial z}{\partial u}$ в точке (u_0, v_0) существует и из формулы (20.22) вытекает формула (20.25). □

Аналогично, если в точке (u_0, v_0) существуют частные производные $\frac{\partial x}{\partial v}$ и $\frac{\partial y}{\partial v}$, то у сложной функции $z = f(x(u, v), y(u, v))$ существует в точке (u_0, v_0) частная производная по v и для нее справедлива формула

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}.$$

Рассмотрим общий n-мерный случай. *Пусть в окрестности точки $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ задана функция $y = y(x_1, \ldots, x_n)$, а на некотором множестве $E_t \subset R^k$ — функции $x_i = x_i(t_1, \ldots, t_k)$, $i = 1, 2, \ldots, n$, такие, что $x_i(t_1^{(0)}, \ldots, t_k^{(0)}) = x_i^{(0)}$. Если функция $y = y(x) = y(x_1, \ldots, x_n)$ дифференцируема в точке $x^{(0)}$ и если в точке $t^{(0)} = (t_1^{(0)}, \ldots, t_k^{(0)})$ существуют частные производные $\frac{\partial x_i}{\partial t_j}$, $j = 1, 2, \ldots, k$, $i = 1, 2, \ldots, n$, то сложная функция $y(x(t))$ имеет в точке $t^{(0)}$ частные производные $\frac{\partial y}{\partial t_j}$, $j = 1, 2, \ldots, k$, причем*

$$\frac{\partial y}{\partial t_j} = \sum_{i=0}^{n} \frac{\partial y}{\partial x_i}\frac{\partial x_i}{\partial t_j}, \quad j = 1, 2, \ldots, k. \qquad (20.26)$$

Заметим, что если при сделанных предположениях частные производные $\frac{\partial y}{\partial x_i}$ и $\frac{\partial x_i}{\partial t_j}$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$, непрерывны ссответственно в точках $x^{(0)}$ и $t^{(0)}$, то в силу формулы (20.26) частные производные сложной функции $y = y(x(t))$ также будут непрерывными в точке $t^{(0)}$, и, следовательно, она будет дифференцируемой в этой точке (см. теорему 3 п. 20.2). В следующем пункте будет доказана дифференцируемость композиций функций при более слабых предположениях.

20.4. ИНВАРИАНТНОСТЬ ФОРМЫ ПЕРВОГО ДИФФЕРЕНЦИАЛА ОТНОСИТЕЛЬНО ВЫБОРА ПЕРЕМЕННЫХ. ПРАВИЛА ВЫЧИСЛЕНИЯ ДИФФЕРЕНЦИАЛОВ

Теорема 6. *Пусть функция $f(x)$, $x=(x_1, \ldots, x_n)$, определена в некоторой окрестности точки $x^{(0)}=(x_1^{(0)}, \ldots, x_n^{(0)})$, а функции $x_i=x_i(t)$, $t=(t_1, \ldots, t_k)$, $i=1, 2, \ldots, n$, определены в некоторой окрестности точки $t^{(0)}=(t_1^{(0)}, \ldots, t_k^{(0)})$ и пусть $x_i^{(0)}=x_i(t^{(0)})$, $i=1, 2, \ldots, n$.*

Тогда, если функция $f(x)$ дифференцируема в точке $x^{(0)}$, а функции $x_i=x_i(t)$, $i=1, 2, \ldots, n$, дифференцируемы в точке $t^{(0)}$, то сложная функция $f(x(t))=f(x_1(t), \ldots, x_n(t))$ определена в некоторой окрестности точки $t^{(0)}$ и дифференцируема в этой точке. При этом дифференциал df функции $f(x(t))$ в точке $t^{(0)}$ может быть записан в следующих двух видах:

$$df=\sum_{j=1}^{k}\frac{\partial f(x(t^{(0)}))}{\partial t_j}dt_j, \tag{20.27}$$

$$df=\sum_{i=1}^{n}\frac{\partial f(x^{(0)})}{\partial x_i}dx_i, \text{ где } dx_i=dx_i(t)\big|_{t=t^{(0)}}. \tag{20.28}$$

Доказательство. Поскольку функции $x_i(t)$, $i=1, 2, \ldots, n$, определены в некоторой окрестности точки $t^{(0)}$ и поскольку из дифференцируемости функций следует их непрерывность, то сложная функция $f(x(t))$ определена в некоторой окрестности точки $t^{(0)}$ (см. замечание к теореме 2 п. 19.4). Зафиксируем какие-либо два числа $\delta>0$ и $\eta>0$ так, чтобы функция $f(x)$ была бы определена на η-окрестности точки $x^{(0)}$, функции $x_i(t)$, $i=1, 2, \ldots, n$, на δ-окрестности точки $t^{(0)}$ и чтобы $(x_1(t), \ldots, x_n(t))\in U(x^{(0)}; \eta)$ при $t\in U(t^{(0)}; \delta)$. Тогда на окрестности $U(t^{(0)}; \delta)$ определена сложная функция $f(x(t))$. Возможность выбора таких чисел δ и η (очевидно δ зависит от выбора η) была показана в п. 19.4. Функция $f(x)$ дифференцируема в точке $x^{(0)}$; поэтому при $r=\sqrt{\sum_{i=1}^{n}\Delta x_i^2}<\eta$ имеем

$$\Delta f=f(x_1^{(0)}+\Delta x_1, \ldots, x_n^{(0)}+\Delta x_n)-f(x_1^{(0)}, \ldots, x_n^{(0)})=$$
$$=\sum_{i=1}^{n}\frac{\partial f(x^{(0)})}{\partial x_i}\Delta x_i+\varepsilon r, \tag{20.29}$$

где $\varepsilon = \varepsilon(\Delta x_1, \ldots, \Delta x_n)$ таково, что $\lim\limits_{r \to 0} \varepsilon = 0$. Положим $\varepsilon(0, \ldots, 0) = 0$. Доопределенная таким образом функция ε является непрерывной в точке $(0, \ldots, 0)$.

В силу дифференцируемости функций $x_i = x_i(t)$, $i = 1, 2, \ldots, n$, в точке $t^{(0)}$ при $\rho = \sqrt{\sum\limits_{j=1}^{k} \Delta t_j^2} < \delta$ получим:

$$\Delta x_i = x_i(t_1^{(0)} + \Delta t_1, \ \ldots, \ t_k^{(0)} + \Delta t_k) - x_i(t_1^{(0)}, \ \ldots, \ t_k^{(0)}) = \\ = \sum_{j=1}^{k} \frac{\partial x_i(t^{(0)})}{\partial t_j} \Delta t_i + \varepsilon_i \rho, \quad i = 1, 2, \ldots, n, \tag{20.30}$$

где $\lim\limits_{\rho \to 0} \varepsilon_i = 0$, $i = 1, 2, \ldots, n$.

Подставив значения Δx_i из (20.30) в (20.29), получим

$$\Delta f = \sum_{i=1}^{n} \frac{\partial f(x^{(0)})}{\partial x_i} \sum_{j=1}^{k} \frac{\partial x_i(t^{(0)})}{\partial t_j} \Delta t_j + \beta, \tag{20.31}$$

где

$$\beta = \sum_{j=1}^{n} \frac{\partial f(x^{(0)})}{\partial x_i} \varepsilon_i \rho + \varepsilon r. \tag{20.32}$$

Переставив порядок суммирования в (20.31), будем иметь

$$\Delta f = \sum_{j=1}^{k} \left(\sum_{i=1}^{n} \frac{\partial f(x^{(0)})}{\partial x_i} \frac{\partial x_i(t^{(0)})}{\partial t_j} \right) \Delta t_j + \beta. \tag{20.33}$$

Теперь, для того чтобы доказать, что сложная функция $f(x(t))$ дифференцируема в точке $t^{(0)}$, надо показать, что $\beta = o(\rho)$ при $\rho \to 0$. В силу непрерывности функций $x_i(t)$, $i = 1, 2, \ldots, n$ в точке $t^{(0)}$ имеем $\lim\limits_{\rho \to 0} \Delta x_i = 0$ и, следовательно, $\lim\limits_{\rho \to 0} r = 0$. Отсюда в силу теоремы о суперпозиции непрерывных функций (см. п. 19.2)

$$\lim_{\rho \to 0} \varepsilon = 0. \tag{20.34}$$

Из (20.32) имеем:

$$\frac{\beta}{\rho} = \sum_{i=1}^{n} \frac{\partial f(x^{(0)})}{\partial x_i} \varepsilon_i + \varepsilon \frac{r}{\rho}. \tag{20.35}$$

Докажем, что отношение r/ρ ограничено. Использовав формулы (20.30), получим

$$\frac{r}{\rho}=\frac{1}{\rho}\sqrt{\sum_{i=1}^{n}\Delta x_i^2}\leqslant\frac{1}{\rho}\sum_{i=1}^{n}|\Delta x_i|^{*)}\leqslant\sum_{i=1}^{n}\sum_{j=1}^{k}\left|\frac{\partial x_i(t^{(0)})}{\partial t_j}\right|\frac{|\Delta t_j|}{\rho}+\varepsilon_i.$$

Поскольку $\lim\limits_{\rho\to 0}\varepsilon_i=0$, то в некоторой окрестности точки $t^{(0)}$ функции ε_i ограничены, и так как $|\Delta t_j|/\rho\leqslant 1$, то функция r/ρ ограничена в некоторой окрестности точки $t^{(0)}$. Поэтому из (20.34) и (20.35) следует, что $\lim\limits_{\rho\to 0}(\beta/\rho)=0$, т. е. что $\beta=o(\rho)$ при $\rho\to 0$. Дифференцируемость сложной функции $f(x(t))$ в точке $t^{(0)}$ доказана.

Из формулы (20.31) имеем

$$df(x^{(0)})=\sum_{i=1}^{n}\frac{\partial f(x^{(0)})}{\partial x_i}\sum_{j=1}^{k}\frac{\partial x_i(t^{(0)})}{\partial t}\Delta t_j.$$

Отсюда, замечая, что $\sum\limits_{j=1}^{k}\frac{\partial x_i(t^{(0)})}{\partial t_j}\Delta t_j=dx_i$, $i=1, 2, \ldots, n$, мы и получаем формулу (20.28). Формула же (20.27) является обычной формулой для дифференциала (см. (20.21)). □

Формально обе записи (20.27) и (20.28) дифференциала функции выглядят одинаково: в обеих формулах дифференциал равен сумме произведений частных производных на соответствующие дифференциалы, однако в случае формулы (20.27) dt_j являются дифференциалами независимых переменных, а в случае формулы (20.28) dx_i суть дифференциалы функций. Это свойство называется *инвариантностью формы первого дифференциала* относительно выбора переменных.

З а м е ч а н и е 1. Из формулы (20.33) следует, что

$$df(x^{(0)})=\sum_{j=1}^{k}\left(\sum_{i=1}^{n}\frac{\partial f(x^{(0)})}{\partial x_i}\frac{\partial x_i(t^{(0)})}{\partial t_j}\right)dt_j.$$

Но коэффициенты дифференциала функции при дифференциалах независимых переменных определяются однозначно и равны соответствующим частным производным, поэтому, сравнивая эту фор-

---

*) Мы воспользовались неравенством $\sqrt{\sum\limits_{i=1}^{n}a_i^2}\leqslant\sum\limits_{i=1}^{n}|a_i|$, которое является следствием очевидного неравенства $\sum\limits_{i=1}^{n}a_i^2\leqslant\left(\sum\limits_{i=1}^{n}|a_i|\right)^2$ (см. (18.11)).

мулу с формулой (20.27), получим

$$\frac{\partial f}{\partial t_j}=\sum_{i=1}^{n}\frac{\partial f(x^{(0)})}{\partial x_i}\frac{\partial x_i(t^{(0)})}{\partial t_j},$$

т. е. снова формулу (20.26). Правда, на этот раз она выведена при более сильных ограничениях, чем раньше; на этот раз предполагалась дифференцируемость функций $x_i(t)$, $i=1, 2, \ldots, n$, в то время как в п. 20.3 — лишь существование у этих функций соответствующих частных производных.

Замечание 2. Если функции $f(x_1, \ldots, x_n)$ и $x_i=x_i(t)$, $t=(t_1, \ldots, t_k)\in R^k$, $i=1, 2, \ldots, n$, имеют непрерывные частные производные соответственно в точке $(x_1^{(0)}, \ldots, x_n^{(0)})\in R^n$ и в точке $t^{(0)}\in R^k$, где $x_i^{(0)}=x_i(t^{(0)})$, то эти функции, согласно теореме 3 п. 20.2 (см. также замечания в конце п. 20.2 об общем случае), дифференцируемы в указанных точках и потому удовлетворяют условиям теоремы 6. Следовательно, для них справедливо утверждение этой теоремы и вытекающая из него формула для вычисления частной производной сложной функции (см. предыдущее замечание).

Инвариантность формы первого дифференциала широко используется при практическом вычислении дифференциалов и частных производных. Если u и v суть функции какого-то числа переменных, то с помощью формулы (20.28) легко получаются следующие:

$$\begin{aligned}&1.\quad d(u+v)=du+dv.\\&2.\quad d(uv)=v\,du+u\,dv.\\&3.\quad d(u/v)=\frac{v\,du-u\,dv}{v^2}.\end{aligned}\tag{20.36}$$

Докажем, например, формулу 3. Пусть $z=u/v$, где $u=u(x_1, \ldots, x_n)$, $v=v(x_1, \ldots, x_n)$. Замечая, что $\frac{\partial z}{\partial u}=\frac{1}{v}$ и $\frac{\partial z}{\partial v}=-\frac{u}{v^2}$, согласно формуле (20.28) имеем

$$dz=\frac{1}{v}\,du-\frac{u}{v^2}\,dv=\frac{v\,du-u\,dv}{v^2}.\ \square$$

При вычислении конкретных дифференциалов функций многих переменных можно широко использовать формулы, полученные нами раньше (см. § 9) для дифференциалов элементарных функций. Заметим для этого следующее: пусть функция $y=y(x_1, \ldots, x_n)$ представлена в виде $y=F(u)$, где $u=u(x_1, \ldots, x_n)$. Тогда при соответствующих предположениях, согласно формуле (20.28),

$$dy=F'(u)\,du,\quad u=u(x_1, \ldots, x_n).$$

Например, если $y=\sin u$, то $dy=\cos u\,du$; если $y=\ln u$, то $dy=\frac{du}{u}$; если $y=\operatorname{arctg} u$, то $dy=\frac{du}{1+u^2}$ и т. д. (подчеркнем, что здесь везде $u=u(x_1, \ldots, x_n)$).

В качестве примера найдем дифференциал функции $z=\operatorname{arctg}\frac{x}{y}$. Вычисления производятся в следующем порядке:

$$dz=d\left(\operatorname{arctg}\frac{y}{x}\right)=\frac{1}{1+\frac{y^2}{x^2}}d\left(\frac{y}{x}\right)=\frac{x^2}{x^2+y^2}\frac{x\,dy-y\,dx}{x^2}=\frac{x\,dy-y\,dx}{x^2+y^2}.$$

Если требуется вычислить частные производные функции многих переменных, особенно если надо вычислить все производные, то целесообразно вычислить дифференциал этой функции, тогда искомыми частными производными будут коэффициенты при соответствующих дифференциалах.

Так, в рассмотренном примере $z=\operatorname{arctg}\frac{y}{x}$, беря коэффициенты при dx и dy из найденного нами выражения для дифференциала, получим

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

Замечание 3. *Всякую функцию $y=f(x_1, \ldots, x_n)$ от n переменных можно рассматривать в определенном смысле и как функцию от любого числа $n+m>n$ переменных $x_1, x_2, \ldots, x_n, \ldots$ $\ldots, x_{n+m}$.* Именно, для всякой функции $f(x_1, \ldots, x_n)$, заданной на множестве $E\subset R^n$, определим функцию $f^*(x_1, \ldots, x_n, \ldots, x_{n+m})$ на множестве точек $(x_1, \ldots, x_n, \ldots, x_{n+m})$ таких, что $(x_1, \ldots, x_n)\in E$, $-\infty<x_j<+\infty$, $j=n+1, \ldots, n+m$, следующим образом:

$$f^*(x_1, \ldots, x_n, \ldots, x_{n+m})=f(x_1, \ldots, x_n). \qquad (20.37)$$

Таким образом, рассмотрение функции n переменных, как функции $n+m$ переменных, означает фактически продолжение по формуле (20.37) функции f с множества ее определения $E\subset R^n$ на множество

$$E^*=\{(x_1, \ldots, x_{n+m}):(x_1, \ldots, x_n)\in E,\ -\infty<x_j<+\infty,\\ j=n+1, \ldots, n+m\},$$

лежащее уже в пространстве R^{n+m}. Для функции f^*, полученной после такого продолжения, имеем

$$\frac{\partial f^*(x_1, \ldots, x_{n+m})}{\partial x_i}=\frac{\partial f(x_1, \ldots, x_n)}{\partial x_i},\quad i=1, 2, \ldots, n,$$

$$\frac{\partial f^*(x_1, \ldots, x_{n+m})}{\partial x_j}=0,\quad j=n+1, \ldots, n+m,$$

поэтому

$$df^*(x_1, \ldots, x_{n+m}) = \sum_{i=1}^{n+m} \frac{\partial f^*(x_1, \ldots, x_{n+m})}{\partial x_i} dx_i =$$

$$= \sum_{i=1}^{n} \frac{\partial f(x_1, \ldots, x_n)}{\partial x_i} dx_i = df(x_1, \ldots, x_n).$$

Например, когда мы говорим, что функцию одного переменного $z = f(x)$, определенную на некотором интервале (a, b), мы рассматриваем как функцию двух переменных $f(x) = F(x, y)$, $x \in (a, b)$, $-\infty < y < +\infty$, это означает, что функция $F(x, y)$ является постоянной, равной $f(x)$ на любой прямой, проходящей через точку x интервала (a, b) оси Ox параллельно оси Oy. При этом

$$\frac{\partial F(x, y)}{\partial x} = f'(x), \quad \frac{\partial F(x, y)}{\partial y} = 0, \quad dF(x, y) = df(x),$$

$$a < x < b, \quad -\infty < y < +\infty.$$

Полезно для дальнейшего отметить в известном смысле обратный факт. Пусть $E \subset R^n$. Если функция $f^*(x_1, \ldots, x_n, x_{n+1})$ определена на множестве

$$E^* = \{(x_1, \ldots, x_n, x_{n+1}) : (x_1, \ldots, x_n) \in E, \quad a < x_{n+1} < b\}$$

и

$$\frac{\partial f^*(x_1, \ldots, x_n, x_{n+1})}{\partial x_{n+1}} = 0 \text{ на } E^*, \tag{20.38}$$

то существует функция $f(x_1, \ldots, x_n)$ от n переменных, определенная на множестве E и такая, что $f^*(x_1, \ldots, x_n, x_{n+1}) = f(x_1, \ldots, x_n)$ для всех $(x_1, \ldots, x_n) \in E$, $x_{n+1} \in (a, b)$. В этом случае говорят, что функция f^* фактически *не зависит от переменной* x_{n+1}. В самом деле, из условия (20.38) следует, что функция f^* постоянна как функция x_{n+1} (см. следствие 1 теоремы 3 из п. 11.2) при фиксированной точке $(x_1, \ldots, x_n)$, т. е. зафиксировав какое-либо $c \in (a, b)$ для любой точки $(x_1, \ldots, x_n) \in E$ и $x_{n+1} \in (a, b)$, имеем $f^*(x_1, \ldots, x_{n+1}) = f^*(x_1, \ldots, x_n, c)$. Искомая функция f, очевидно, определяется равенством $f(x_1, \ldots, x_n) = f^*(x_1, \ldots, x_n, c)$, причем она не зависит от выбора $c \in (a, b)$.

Из вышесказанного, в частности, следует, что формулы (20.36) для дифференциалов остаются справедливыми и в том случае, когда функции u и v зависят от разного числа переменных, так как всегда в силу указанного приема этот случай можно свести к вышеразобранному случаю функций одного числа переменных.

20.5. ГЕОМЕТРИЧЕСКИЙ СМЫСЛ ЧАСТНЫХ ПРОИЗВОДНЫХ И ПОЛНОГО ДИФФЕРЕНЦИАЛА

Для большей геометрической наглядности и для того, чтобы не вводить новых понятий, в этом пункте ограничимся рассмотрением функций двух переменных.

Рассмотрим функцию $z=f(x, y)$, определенную на плоском открытом множестве G, т. е. множестве G, лежащем на плоскости R^2. Пусть $(x_0, y_0)\in G$ и пусть в точке (x_0, y_0) существует частная производная $\frac{\partial z}{\partial x}$. Ее геометрический смысл сразу получается из определения частной производной $\frac{\partial z}{\partial x}$ как обычной производной функции $f(x, y)$ по x при фиксированном y и из геометрического смысла обычной производной (см. п. 9.3). В самом деле, возьмем замкнутый круг Q радиуса r с центром в точке (x_0, y_0) и лежащий в G*). Пусть γ — кривая, заданная представлением

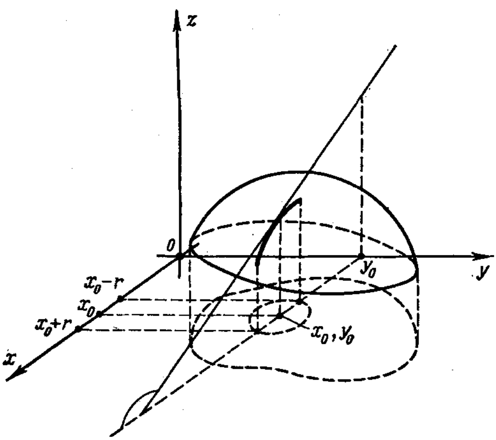

Рис. 89

$$z=f(x, y_0),\quad y=y_0,$$
$$x_0-r\leqslant x\leqslant x_0+r,$$

т. е. кривая, которая получается сечением графика функции $z=f(x, y)$, $(x, y)\in Q$ плоскостью $y=y_0$ (рис. 89). Как известно, $\frac{\partial f(x_0, y_0)}{\partial x}=\left.\frac{df(x, y_0)}{dx}\right|_{x=x_0}=\operatorname{tg}\alpha$, где α — угол, образованный касательной к графику функции $f(x, y_0)$ в точке $(x_0, f(x_0, y_0))$ с осью Ox, т. е. угол, образованный касательной к кривой γ в точке $(x_0, y_0, f(x_0, y_0))$ с осью Ox.

Таким образом,

$$\frac{\partial f(x_0, y_0)}{\partial x}=\operatorname{tg}\alpha$$

— в этом и состоит геометрический смысл частной производной $\frac{\partial f}{\partial x}$.

---

*) Такой круг Q всегда существует. Действительно, в силу определения открытого множества существует такая δ-окрестность U точки (x_0, y_0), что $U\subset G$. Тогда замкнутый круг Q радиуса $\delta/2$ с центром в точке (x_0, y_0) будет заведомо лежать в G.

Аналогично устанавливается и геометрический смысл частной производной $\frac{\partial f(x_0, y_0)}{\partial y}$ как тангенса угла наклона, образованного касательной в точке $(x_0, y_0, f(x_0, y_0))$ к кривой, получающейся сечением графика функции $z=f(x, y)$, $(x, y) \in Q$ плоскостью $x=x_0$, с осью Oy.

Что же касается геометрического смысла дифференциала, то из формул (20.20) и (20.9) для нашего случая, т. е. при $n=2$, получим

$$f(x, y)=z_0+A(x-x_0)+B(y-y_0)+o(\rho),\ \rho \to 0, \qquad (20.39)$$

$$\rho=\sqrt{(x-x_0)^2+(y-y_0)^2}, \qquad z_0=f(x_0, y_0).$$

Уравнение

$$z=z_0+A(x-x_0)+B(y-y_0) \qquad (20.40)$$

является уравнением плоскости, проходящей через точку (x_0, y_0, z_0) и не параллельной оси Oz. Как мы знаем, коэффициенты A и B однозначно определяются из соотношения (20.39), причем

$$A=\frac{\partial f(x_0, y_0)}{\partial x}, \quad B=\frac{\partial f(x_0, y_0)}{\partial y}, \qquad (20.41)$$

и, значит, плоскость (20.40) однозначно определена соотношением (20.39). Эта плоскость называется *касательной плоскостью* к графику функции $z=f(x, y)$ в точке (x_0, y_0, z_0).

Таким образом, мы пришли к следующему определению.

Определение 6. *Касательной плоскостью к графику функции $f(x, y)$ в данной точке называется такая плоскость, что разность ее аппликаты и значения функции $f(x, y)$ является величиной, бесконечно малой по сравнению с ρ при $\rho \to 0$.*

В силу (20.41) уравнение этой касательной плоскости имеет вид

$$z-z_0=\frac{\partial f(x_0, y_0)}{\partial x}(x-x_0)+\frac{\partial f(x_0, y_0)}{\partial y}(y-y_0). \qquad (20.42)$$

В дальнейшем (см. т. 2, п. 50.4) мы познакомимся с другим подходом к понятию касательной плоскости.

Полагая $\Delta x=x-x_0$, $\Delta y=y-y_0$, правую часть уравнения (20.42) запишем в виде

$$\frac{\partial f(x_0, y_0)}{\partial x}\Delta x+\frac{\partial f(x_0, y_0)}{\partial y}\Delta y.$$

Это есть обычная запись дифференциала dz функции $z=f(x, y)$ в точке (x_0, y_0) и поэтому уравнение (20.42) можно переписать:

$$z-z_0=dz.$$

Таким образом, геометрически *полный дифференциал функции в точке* (x_0, y_0) *равен приращению аппликаты плоскости, касательной к графику функции* (рис. 90).

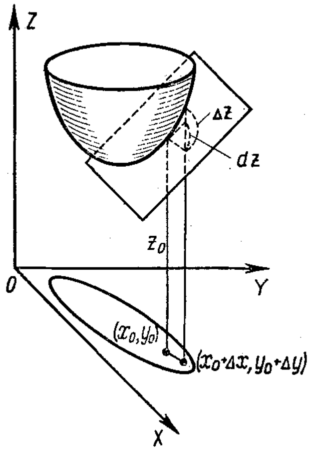

Рис. 90

Более подробно, дифференциал

$$dz = \frac{\partial f(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f(x_0, y_0)}{\partial x}\Delta y,$$

$$\Delta x = x - x_0, \qquad \Delta y = y - y_0$$

совпадает с приращением в точке $x = x_0 + \Delta x$, $y = y_0 + \Delta y$ аппликаты плоскости касательной к графику функции в точке $(x_0, y_0, f(x_0, y_0))$.

20.6. ГРАДИЕНТ ФУНКЦИИ

Пусть функция $F(x, y)$ дифференцируема в точке (x_0, y_0), а кривая γ такова, что функции $x = x(t)$, $y = y(t)$, $a \leqslant t \leqslant b$, с помощью которых она задана в параметрической форме, удовлетворяют уравнению

$$F(x, y) = 0,$$

т. е. посредством его осуществлено неявное задание кривой γ. Пусть $t_0 \in [a, b]$, $x_0 = x(t_0)$, $y_0 = y(t_0)$, а функции $x(t)$ и $y(t)$ дифференцируемы при $t = t_0$.

Дифференцируя при $t = t_0$ тождество $F(x(t), y(t)) = 0$, $a \leqslant t \leqslant b$, получим

$$x'_t \frac{\partial F}{\partial x} + y'_t \frac{\partial F}{\partial y} = 0, \quad t = t_0,$$

т. е. векторы $(x'(t_0), y'(t_0))$ и $\left(\frac{\partial F(x_0, y_0)}{\partial x}, \frac{\partial F(x_0, y_0)}{\partial y}\right)$ ортогональны. Вектор $a = (x'_t, y'_t)$ в случае, когда он не равен нулю, является, как известно, касательным вектором к кривой γ в точке $(x_0, y_0) = (x(t_0), y(t_0))$. Вектор $\left(\frac{\partial F(x_0, y_0)}{\partial x}, \frac{\partial F(x_0, y_0)}{\partial y}\right)$ называется *градиентом функции* F в точке (x_0, y_0) и обозначается через $\operatorname{grad} F(x_0, y_0)$. Из сказанного следует, что градиент функции F ортогонален касательной к кривой, неявно задаваемой уравнением $F(x, y) = 0$. Прямая, перпендикулярная касательной к плоской кривой и лежащая в одной плоскости с ней, называется (см. п. 17.3) *нормалью* к данной кривой.

Таким образом, градиент функции F коллинеарен нормали в соответствующей точке к кривой, задаваемой уравнением $F(x, y) = 0$.

В случае дифференцируемой функции $f(x_1, \ldots, x_n)$ ее градиентом называется вектор $\left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$.

20.7. ПРОИЗВОДНАЯ ПО НАПРАВЛЕНИЮ

Частные производные от функции являются производными «в направлениях координатных осей». Естественно поставить вопрос об определении и вычислении производной по любому фиксированному направлению. Прежде всего определим это понятие. Проведем рассмотрение этого вопроса на примере функций трех переменных.

Пусть функция f определена в δ-окрестности $U(M_0;\ \delta)$ точки $M_0 \in R^3$, пусть $M_1 \in U(M_0;\ \delta)$. Проведем через точки M_0 и M_1 прямую. За положительное направление на этой прямой возьмем направление вектора $\boldsymbol{l} = \overrightarrow{M_0M_1}$, т. е. направление от точки M_0 к точке M_1. Для всякой точки M этой прямой обозначим через M_0M ориентированную длину отрезка с началом в точке M_0 и концом в точке M, т. е. длину этого отрезка со знаком плюс, если вектор $\overrightarrow{M_0M}$ имеет то же направление, что и вектор $\boldsymbol{l}$ и со знаком минус в противном случае.

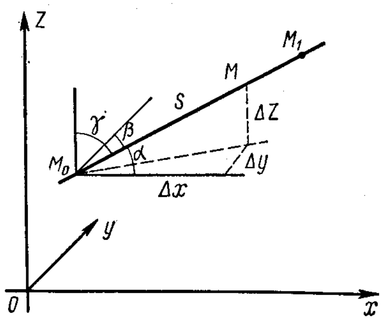

Рис. 91

Определение 7. *Предел* $\lim\limits_{M \to M_0} \dfrac{f(M) - f(M_0)}{M_0M}$, *если он существует, называется производной функции f в точке M_0 по направлению вектора $\boldsymbol{l}$ и обозначается* $\dfrac{\partial f(M_0)}{\partial l}$.

Пусть теперь в пространстве R^3 зафиксирована некоторая система координат $x,\ y,\ z$. Пусть $M_0 = (x_0,\ y_0,\ z_0)$, $M = (x,\ y,\ z)$, $\Delta x = x - x_0$, $\Delta y = y - y_0$, $\Delta z = z - z_0$ и $s = M_0M$. Найдем связь между координатами точки M и ориентированной длиной s отрезка M_0M. Пусть α, β и γ — углы, образованные вектором $\overrightarrow{M_0M_1}$ соответственно с осями Ox, Oy и Oz, тогда (рис. 91)

$$x - x_0 = s\cos\alpha, \quad y - y_0 = s\cos\beta, \quad z - z_0 = s\cos\gamma.$$

Вдоль прямой M_0M функция f является функцией одной переменной s, а именно

$$f(x,\ y,\ z) = f(x_0 + s\cos\alpha,\ y_0 + s\cos\beta,\ z_0 + s\cos\gamma).$$

Производная этой функции по s (если она, конечно, существует) и является *производной функции f в точке M_0 по направлению вектора* $\overrightarrow{M_0M_1}$.

Заметим, что направляющие косинусы $\cos\alpha$, $\cos\beta$ и $\cos\gamma$ вектора $\overrightarrow{M_0M_1}$ через координаты точек $M_0 = (x_0,\ y_0,\ z_0)$ и $M_1 =$

$=(x_1, y_1, z_1)$ определяются следующим образом:

$$\cos\alpha=\frac{x_1-x_0}{\rho},\quad \cos\beta=\frac{y_1-y_0}{\rho},\quad \cos\gamma=\frac{z_1-z_0}{\rho},\tag{20.43}$$
$$\rho=\sqrt{(x_1-x_0)^2+(y_1-y_0)^2+(z_1-z_0)^2}.$$

Вычисляется производная по направлению по правилу дифференцирования сложной функции. Пусть функция $f(x, y, z)$ дифференцируема в точке (x_0, y_0, z_0) и пусть

$$x=x_0+s\cos\alpha,\quad y=y_0+s\cos\beta,\quad z=z_0+s\cos\gamma.\tag{20.44}$$

Согласно определению производной по направлению и формуле производной сложной функции имеем:

$$\frac{\partial F(M_0)}{\partial l}=\lim_{M\to M_0}\frac{f(M)-f(M_0)}{M_0M}=$$
$$=\lim_{s\to 0}\frac{f(x_0+s\cos\alpha,\ y_0+s\cos\beta,\ z_0+s\cos\gamma)-f(x_0, y_0, z_0)}{s}=$$
$$=\left.\frac{df}{ds}\right|_{s=0}=\frac{\partial f(M_0)}{\partial x}\frac{dx}{ds}+\frac{\partial f(M_0)}{\partial y}\frac{dy}{ds}+\frac{\partial f(M_0)}{\partial z}\frac{dz}{ds},$$

но из (20.44) следует, что

$$\frac{dx}{ds}=\cos\alpha,\quad \frac{dy}{ds}=\cos\beta,\quad \frac{dz}{ds}=\cos\gamma,\tag{20.45}$$

поэтому окончательно

$$\frac{\partial f(M_0)}{\partial l}=\frac{\partial f(M_0)}{\partial s}=\frac{\partial f(M_0)}{\partial x}\cos\alpha+\frac{\partial f(M_0)}{\partial y}\cos\beta+\frac{\partial f(M_0)}{\partial z}\cos\gamma.\tag{20.46}$$

Это и есть искомая формула.

Таким образом, доказана следующая теорема.

Теорема 7. *Пусть функция f дифференцируема в точке (x_0, y_0, z_0). Тогда в этой точке функция f имеет производную по любому направлению и эта производная находится по формуле* (20.46).

Любопытно отметить, что из полученной формулы (20.46) для производной по направлению сразу не видно, что эта производная не зависит от выбора системы координат. Эта независимость непосредственно следует из самого определения производной по направлению, откуда в свою очередь вытекает, что правая часть формулы (20.46) не зависит от выбора прямоугольной декартовой системы координат, а определяется только точками M_0 и M_1, или, что то же, точкой M_0 и вектором $\overrightarrow{M_0M_1}$.

Вектор с координатами $\frac{\partial f(M_0)}{\partial x}$, $\frac{\partial f(M_0)}{\partial y}$, $\frac{\partial f(M_0)}{\partial z}$ называется, как мы знаем, *градиентом* функции $f(M)$ в точке M_0 и обозначается $\operatorname{grad} f$. (Мы уже встречались с понятием градиента функций при рассмотрении кривых, заданных неявным образом: см. п. 20.6.)

Таким образом, если $\boldsymbol{i}$, $\boldsymbol{j}$ и $\boldsymbol{k}$ — координатные орты, то

$$\operatorname{grad} f = \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j} + \frac{\partial f}{\partial z}\boldsymbol{k}. \tag{20.47}$$

Часто оказывается удобным использование символического вектора Гамильтона [*]

$$\nabla = \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z},$$

называемого *наблой*. Набла является обозначением определенной операции, которую следует произвести над той или иной функцией.

Для функции f, по определению, полагаем

$$\nabla f = \boldsymbol{i}\frac{\partial f}{\partial x} + \boldsymbol{j}\frac{\partial f}{\partial y} + \boldsymbol{k}\frac{\partial f}{\partial z}.$$

Формально это равенство можно рассматривать как «произведение» вектора ∇ на число f. Итак, $\operatorname{grad} f$ и ∇f являются обозначениями одного и того же выражения.

Пусть теперь вектор $\boldsymbol{l}$ единичный, и, следовательно, $\boldsymbol{l} = (\cos\alpha, \cos\beta, \cos\gamma)$. С помощью градиента формула для производной функции f по направлению вектора $\boldsymbol{l}$ запишется следующим образом:

$$\frac{\partial f}{\partial l} = \cos\alpha\frac{\partial f}{\partial x} + \cos\beta\frac{\partial f}{\partial y} + \cos\gamma\frac{\partial f}{\partial z} = \boldsymbol{l}\operatorname{grad} f, \tag{20.48}$$

где в правой части стоит скалярное произведение вектора $\boldsymbol{l}$ и $\operatorname{grad} f$. Отсюда, поскольку $\boldsymbol{l}$ — единичный вектор,

$$\frac{\partial f}{\partial l} = |\operatorname{grad} f|\cos\varphi,$$

где φ — угол, образованный вектором $\boldsymbol{l}$ и $\operatorname{grad} f$. Из этой формулы видно, что в случае, если в данной точке

$$|\operatorname{grad} f|^2 = \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + \left(\frac{\partial f}{\partial z}\right)^2 \neq 0,$$

то производная дифференцируемой функции по направлению достигает наибольшего значения в единственном направлении, а именно том, при котором $\cos\varphi = 1$, т. е. в направлении градиента. Из этого следует, что для заданной функции точки $f(M)$ градиент в каждой точке однозначно определяется самой функцией, а не зависит от выбора системы координат, как это могло бы сначала показаться из формулы (20.47).

---

[*] У. Гамильтон (1805—1865) — ирландский математик.

Действительно, прежде всего, если градиент равен нулю в одной декартовой системе координат, то он равен нулю и в каждой другой подобной системе координат. В самом деле, равенство нулю градиента в некоторой точке, согласно формуле (20.48), равносильно равенству нулю в этой точке производных по всем направлениям, последнее же не зависит от выбора декартовой системы координат, поскольку от этого выбора не зависит производная по направлению. Если же градиент не равен нулю, то его независимость от выбора декартовой системы координат следует непосредственно из доказанного выше его геометрического смысла: направление градиента показывает направление наибыстрейшего роста функции (оно единственно), а его величина равна производной в этом направлении.

Возьмем теперь любую непрерывно дифференцируемую кривую без особых точек, проходящую через точку (x_0, y_0, z_0), и такую, что вектор $\overrightarrow{M_0M_1}$ является ее касательным вектором. Обозначим через s переменную длину дуги этой кривой, отсчитываемую от точки M_0 в таком направлении, чтобы вектор $\overrightarrow{M_0M_1}$ давал положительное направление на касательной. Если $x=x(s)$, $y=y(s)$, $z=z(s)$ — представление этой кривой, то, как мы знаем (см. п. 16.5), $\frac{dx}{ds}=\cos\alpha$, $\frac{dy}{ds}=\cos\beta$, $\frac{dx}{ds}=\cos\gamma$, т. е. также выполняется (20.45). Поэтому если взять производную в точке (x_0, y_0, z_0) от дифференцируемой функции $f(x, y, z)$ по данной кривой, т. е. при $x=x(s)$, $y=y(s)$, $z=z(s)$, иначе говоря, взять производную от функции $f(x(s), y(s), z(s))$ по s, то для этой производной будет справедлива формула (20.46). Это означает, что производная в некоторой точке от функции вдоль кривой, проходящей через указанную точку, совпадает с производной по направлению касательной к этой кривой в той же точке.

Все сказанное переносится на функции любого числа n переменных ($n \geqslant 2$). Сформулируем лишь определение производной по направлению.

Пусть в некоторой окрестности точки $x^{(0)}=(x_1^{(0)}, \ldots, x_n^{(0)})$ определена функция $f(x)$ и пусть $x^{(1)}=(x_1^{(1)}, \ldots, x_n^{(1)})$ — точка этой окрестности, $x^{(1)} \neq x^{(0)}$.

Проведем прямую через точки $x^{(0)}$ и $x^{(1)}$. Ее уравнение имеет вид (см. (18.44) и (18.45))

$$x_i=x_i^{(0)}+s\cos\alpha_i, \quad i=1, 2, \ldots, n, \quad -\infty<s<+\infty,$$

где $\cos\alpha_i$ — направляющие косинусы вектора

$$l=(x_1^{(1)}-x_1^{(0)}, \ \ldots, \ x_n^{(1)}-x_n^{(0)}).$$

Рассмотрим заданную функцию f только на точках этой прямой, т. е. рассмотрим функцию

$$f(x_1^{(0)}+s\cos\alpha_1, \ \ldots, \ x_n^{(0)}+s\cos\alpha_n).$$

Производная $\frac{\partial f}{\partial l}$ *функции* $f(x_1, \ldots, x_n)$ *в точке* $x^{(0)}$ *в направлении точки* $x^{(1)}$, *или, что то же, в направлении* $(\cos\alpha_1, \ldots, \cos\alpha_n)$, *определяется как производная* $\frac{\partial f}{\partial s}$ *от сложной функции* $f(x_1^{(0)} + s\cos\alpha_1, \ldots, x_n^{(0)} + s\cos\alpha_n)$.

В случае, если функция f дифференцируема в точке $x^{(0)}$, то, согласно формуле для производной сложной функции, имеем в этой точке

$$\frac{\partial f}{\partial l} \overset{\text{def}}{=} \frac{\partial f}{\partial s} = \frac{\partial f}{\partial x_1}\cos\alpha_1 + \ldots + \frac{\partial f}{\partial x_n}\cos\alpha_n.$$

Вспоминая определение градиента функции n переменных (см. п. 20.6), с помощью скалярного произведения n-мерных векторов (см. (18.32)) формулу производной функции f по направлению вектора $\boldsymbol{l}$ для любого n-мерного пространства R^n можно записать в виде (20.48), т. е.

$$\frac{\partial f}{\partial l} = (\operatorname{grad} f, \boldsymbol{l}_0),$$

где $\boldsymbol{l}_0 = (\cos\alpha_1, \ldots, \cos\alpha_n)$.

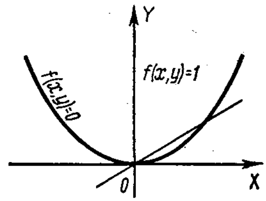

Рис. 92

В заключение отметим, что из того, что функция в некоторой точке имеет производные по всем направлениям, не следует, что функция в этой точке дифференцируема. Например, функция

$$f(x, y) = \begin{cases} 0, & \text{если } y \neq x^2, \text{ или } x = y = 0, \\ 1, & \text{если } y = x^2, \ x^2 + y^2 > 0, \end{cases}$$

имеет в точке (0, 0) по любому направлению производную, равную нулю. Однако, в точке (0, 0) функция f разрывна и, тем более, не дифференцируема (рис. 92).

20.8. ПРИМЕР ИССЛЕДОВАНИЯ ФУНКЦИЙ ДВУХ ПЕРЕМЕННЫХ

С помощью частных производных можно изучать поведение функций многих переменных, подобно тому как исследовалось поведение функции одной переменной с помощью ее производной. Вопросом отыскания наибольших и наименьших значений мы займемся позже в § 40 и § 43, здесь же ограничимся одним примером изучения функции двух переменных, который позволит нам получить одно полезное для дальнейшего неравенство.

Покажем, что для любых $a \geqslant 0$, $b \geqslant 0$, $p > 1$ и числа q, определяемого равенством

$$\frac{1}{p} + \frac{1}{q} = 1 \qquad (20.49)$$

справедливо неравенство

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}. \tag{20.50}$$

Прежде всего отметим, что уравнение (20.49), связывающее числа p и q, равносильно соотношению

$$(p-1)(q-1)=1, \tag{20.51}$$

которое эквивалентно условию

$$q = \frac{p}{p-1}. \tag{20.52}$$

Это устанавливается непосредственной проверкой.

Для доказательства неравенства (20.50) рассмотрим функцию

$$F(x, y) = xy - \frac{x^p}{p} - \frac{y^q}{q}, \quad x \geqslant 0, \ y \geqslant 0. \tag{20.53}$$

Вычислим ее частные производные:

$$\frac{\partial F}{\partial x}(x, y) = y - x^{p-1}, \quad \frac{\partial F}{\partial y}(x, y) = x - y^{q-1}. \tag{20.54}$$

Из (20.51) следует, что при $x \geqslant 0$ и $y \geqslant 0$ уравнения

$$y - x^{p-1} = 0 \tag{20.55}$$

и

$$x - y^{q-1} = 0 \tag{20.56}$$

равносильны. Таким образом, точки (x, y), удовлетворяющие как условию $\frac{\partial F}{\partial x}(x, y) = 0$, так и условию $\frac{\partial F}{\partial y}(x, y) = 0$ лежат на кривой (20.55) или, что то же, на кривой (20.56).

В силу (20.49) и (20.52) вдоль кривой (20.55) имеем:

$$F(x, x^{p-1}) = x^p - \frac{x^p}{p} - \frac{x^{(p-1)q}}{q} =$$
$$= x^p - \frac{x^p}{p} - \frac{x^p}{q} = x^p\left(1 - \frac{1}{p} - \frac{1}{q}\right) = 0. \tag{20.57}$$

Обозначим теперь через G^+ множество всех точек, расположенных выше кривой (20.55), включая саму кривую:

$$G^+ \overset{\text{def}}{=} \{(x, y) : y \geqslant x^{p-1}, \ x \geqslant 0\},$$

а через G^- — множество всех точек первой координатной четверти (включая ось x-ов), лежащих ниже этой кривой:

$$G^- \overset{\text{def}}{=} \{(x, y); 0 \leqslant y \leqslant x^{p-1}, \ x \geqslant 0\}.$$

Согласно формулам (20.54) при $(x, y) \in G^+$, $y \neq x^{p-1}$ имеем $\frac{\partial F}{\partial x}(x, y) > 0$, а при $(x, y) \in G^-$, $y \neq x^{p-1}$, соответственно $\frac{\partial F}{\partial y}(x, y) > 0$ (здесь использована эквивалентность уравнений (20.55) и (20.56)). Поэтому вдоль любого отрезка, лежащего во множестве G^+ и параллельного оси x-ов (рис. 93) функция $F(x, y)$ строго возрастает. Следовательно, если $(x, y) \in G^+$, $y \neq x^{p-1}$, то (см. (20.57))

$$F(x, y) < F(x, x^{p-1}) = 0.$$

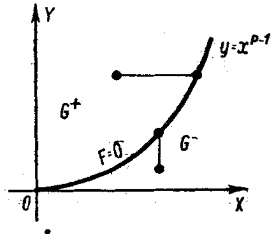

Рис. 93

Аналогично, на любом отрезке, лежащем во множестве G^- и параллельном оси y-ов функция $F(x, y)$ также строго возрастает. Поэтому, если $(x, y) \in G$ и $y \neq x^{p-1}$, то опять

$$F(x, y) < F(x, x^{p-1}) = 0.$$

Таким образом, если $y \neq x^{p-1}$, $x \geqslant 0$, $y \geqslant 0$, то всегда $F(x, y) < 0$.

Итак, вспоминая вид функции F (см. (20.53)), имеем: если $a \geqslant 0$, $b \geqslant 0$, то

$$ab < \frac{a^p}{p} + \frac{b^q}{q} \quad \text{при} \quad b \neq a^{p-1},.$$

$$ab = \frac{a^p}{p} + \frac{b^q}{q} \quad \text{при} \quad b = a^{p-1}.$$

Тем самым неравенство (20.50) доказано.

§ 21. ЧАСТНЫЕ ПРОИЗВОДНЫЕ И ДИФФЕРЕНЦИАЛЫ ВЫСШИХ ПОРЯДКОВ

21.1. ЧАСТНЫЕ ПРОИЗВОДНЫЕ ВЫСШИХ ПОРЯДКОВ

Пусть задана функция $f(x, y)$. Тогда каждая из ее частных производных (если они, конечно, существуют) $\frac{\partial f(x, y)}{\partial x}$ и $\frac{\partial f(x, y)}{\partial y}$, которые называются также *частными производными первого порядка*, снова является функцией независимых переменных x, y и может, следовательно, также иметь частные производные. Частная производная $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$ обозначается через $\frac{\partial^2 f}{\partial x^2}$ или f_{xx}, а $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$ через $\frac{\partial^2 f}{\partial y\, \partial x}$ или f_{xy}. Таким образом,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx}, \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\, \partial x} = f_{xy}$$

и, аналогично,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\, \partial y} = f_{yx}, \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy}.$$