

Capítulo 5

FUNCIONES DE VARIAS VARIABLES

§ 1. Concepto de función de varias variables

1. Sobre las dependencias funcionales entre varias magnitudes variables. Al estudiar muchos problemas de las ciencias naturales se encuentran dependencias entre varias magnitudes variables, en las cuales los valores de una de dichas magnitudes variables se determinan completamente por los valores de las demás variables. Así, por ejemplo, al analizar ciertas características físicas de un cuerpo (por ejemplo, su densidad ρ o temperatura T), hemos de considerar el cambio de estas características cuando pasamos de un punto del cuerpo a otro. Por cuanto cada punto del cuerpo se determina por tres coordenadas cartesianas x , y , z , las características en consideración (densidad ρ o temperatura T) se determinan mediante los valores de *tres* variables x , y , z .

Cuando se analizan los procesos físicos que varían con el tiempo, los valores de las características físicas se determinan por los valores de *cuatro* variables: tres coordenadas de un punto x , y , z y el tiempo t . Por ejemplo, al estudiar las oscilaciones acústicas de un gas la densidad ρ de este gas y la presión p se determinan por los valores de cuatro variables x , y , z y t . Para poder estudiar las dependencias de este género en el presente capítulo se introduce el concepto de función de varias variables y se desarrolla el aparato de investigación de semejantes funciones.

En la teoría de las funciones de varias variables es cómodo usar los términos geométricos. Está claro que de dominio de definición de una función de dos (o de tres) variables sirve cierto conjunto de puntos de un plano (o de un espacio). Con el fin de geometrizar nuestras nociones sobre una función de m variables es conveniente introducir el concepto de espacio m -dimensional que generaliza los conceptos bien conocidos de plano bidimensional y de espacio tridimensional. Empezaremos nuestra exposición ulterior aclarando los conceptos geométricos que nos harán falta.

2. Concepto de plano euclídeo y de espacio euclídeo. Los conceptos de coordenadas de los puntos sobre un plano y en un espacio, que se conocen por el curso de geometría analítica, así como la fórmula para determinar la distancia entre dos puntos, pueden ser empleados para la introducción analítica de los conceptos de plano y de espacio. *Un conjunto de toda clase de pares ordenados (x, y) de números reales x e y lleva el nombre de plano coordenado.*

En este caso cada par (x, y) se llamará punto de dicho plano y se denotará con la letra M . Los números x e y se denominan coordena-

das del punto M . La notación $M(x, y)$ significa que el punto M tiene coordenadas x e y .

Un plano coordenado se llama euclídeo si entre cualesquier dos puntos $M'(x', y')$ y $M''(x'', y'')$ del plano coordenado la distancia $\rho(M', M'')$ se define según la fórmula

$$\rho(M', M'') = \sqrt{(x'' - x')^2 + (y'' - y')^2}.$$

De un modo análogo se introduce el concepto de espacios coordenado y euclídeo. Un conjunto de toda clase de ternas ordenadas (x, y, z) de números x, y, z recibe el nombre de espacio coordenado. En este caso cada terna (x, y, z) se llamará punto de dicho espacio y se denotará con la letra M . Los números x, y, z se llaman coordenadas del punto M . La notación $M(x, y, z)$ significa que el punto M tiene por coordenadas x, y, z .

Un espacio coordenado se denomina euclídeo si entre cualesquier dos puntos $M'(x', y', z')$ y $M''(x'', y'', z'')$ del espacio coordenado la distancia $\rho(M', M'')$ está definida según la fórmula

$$\rho(M', M'') = \sqrt{(x'' - x')^2 + (y'' - y')^2 + (z'' - z')^2}.$$

Los conceptos de plano coordenado y de espacio coordenado, introducidos más arriba, son análogos a la recta numérica, mientras que el plano euclídeo y el espacio euclídeo son análogos a la *recta euclídea*, que puede ser definida como recta numérica, entre cuyos dos puntos x' y x'' cualesquiera está definida la distancia $\rho(x', x'')$ según la fórmula $\rho(x', x'') = \sqrt{(x'' - x')^2} = |x'' - x'|$.

Veamos ciertos conjuntos $\{M\}$ de puntos del plano euclídeo y del espacio euclídeo.

1°. El conjunto $\{M\}$ de puntos del plano euclídeo, cuyas coordenadas x e y satisfacen la inequación $(x - a)^2 + (y - b)^2 \leq R^2$, se denomina, según se sabe, *círculo* de radio R con centro en el punto $M_0(a, b)$. Si las coordenadas x e y satisfacen la desigualdad estricta $(x - a)^2 + (y - b)^2 < R^2$, el conjunto $\{M\}$ se llama *círculo abierto*. En el espacio euclídeo el conjunto $\{M\}$ de puntos, cuyas coordenadas x, y, z satisfacen la desigualdad $(x - a)^2 + (y - b)^2 + (z - c)^2 \leq R^2$, se llama, según se conoce, *esfera* de radio R con centro en el punto $M_0(a, b, c)$. Si las coordenadas x, y, z satisfacen la desigualdad estricta correspondiente, el conjunto $\{M\}$ se denomina *esfera abierta**).

2°. El conjunto $\{M\}$ de puntos del plano euclídeo (del espacio euclídeo), cuyas coordenadas x e y (x, y, z) satisfacen las inequaciones $|x - a| \leq d_1$, $|y - b| \leq d_2$ ($|x - a| \leq d_1$, $|y - b| \leq d_2$ y $|z - c| \leq d_3$), se llama *rectángulo coordenado* (*paralelepípedo* *coor-*

*) Es evidente que el círculo y la esfera son conjuntos $\{M\}$ de puntos de un plano y de un espacio, para los cuales $\rho(M, M_0) \leq R$.

denado) con centro en el punto $M_0(a, b)$ (en el punto $M_0(a, b, c)$).

3. Concepto de función de dos y tres variables. Al hacer uso de los términos geométricos, podemos enunciar del modo siguiente el concepto, ya conocido, de función de una sola variable.

Si a todo punto M de cierto conjunto $\{M\}$ de puntos de la recta euclídea se le pone en correspondencia, de acuerdo con una ley conocida, un cierto número u , se dice que sobre el conjunto $\{M\}$ viene dada una función $u = u(M)$, o bien $u = f(M)$.

Introducamos ahora el concepto de función de dos variables.

Si a todo punto M de cierto conjunto de puntos $\{M\}$ del plano euclídeo se le pone en correspondencia, de acuerdo con una ley conocida, cierto número u , se dice que sobre el conjunto $\{M\}$ viene dada una función $u = u(M)$, o bien $u = f(M)$.

Advirtamos que el concepto de función de dos variables difiere del concepto, enunciado anteriormente, de función de una sola variable solamente en que en lugar de las palabras «recta euclídea» se emplea el término «plano euclídeo». De un modo análogo se introduce el concepto de función de tres variables. Con este fin, en lugar del conjunto $\{M\}$ de puntos del plano euclídeo se debe tomar un conjunto $\{M\}$ de puntos del espacio euclídeo.

Por cuanto el punto M del plano euclídeo se determina mediante dos coordenadas x e y , en tanto que el punto M del espacio euclídeo, mediante tres coordenadas x, y, z , para las funciones de dos y de tres variables se emplearán, respectivamente, las designaciones $u = f(x, y)$ y $u = f(x, y, z)$. Si la función $u = f(M)$ está definida sobre el conjunto $\{M\}$, dicho conjunto se denomina *dominio de definición de la función $u = f(M)$* . El número u , correspondiente al punto dado M del conjunto $\{M\}$, se llamará *valor particular de la función en el punto M* .

La colección $\{u\}$ de todos los valores particulares de la función $u = f(M)$ se llama *conjunto de valores de dicha función*.

Para una función de dos variables puede introducirse el concepto de gráfica: llámase *gráfica de la función $u = f(x, y)$ a una superficie, cuyos puntos tienen coordenadas $(x, y, f(x, y))$* .

He aquí unos ejemplos de funciones de dos y tres variables.

1° $u = \sqrt{4 - x^2 - y^2}$. El dominio de definición de esta función es un círculo de radio 2 con centro en el origen de las coordenadas, mientras que el conjunto de valores es un segmento $0 \leq u \leq 2$.

2° $u = \frac{1}{\sqrt{x^2 + y^2 - 4}}$. El dominio de definición de esta función es un conjunto de puntos dispuestos fuera del círculo de radio 2 con centro en el origen de las coordenadas, en tanto que el conjunto de valores es una semirrecta abierta $u > 0$.

3° $u = \sqrt{\cos(x^2 + y^2)}$. El dominio de definición de esta función es un conjunto $\{M\}$ de puntos, cuyas coordenadas satisfacen la

inecuación $\cos(x^2 + y^2) \geq 0$. Esta inecuación es equivalente a las desigualdades $0 \leq x^2 + y^2 \leq \frac{\pi}{2}$, $2k\pi - \frac{\pi}{2} \leq x^2 + y^2 \leq 2k\pi + \frac{\pi}{2}$, $k = 1, 2, \dots$. De este modo, el conjunto $\{M\}$ se compone de un círculo de radio $\sqrt{\pi/2}$ con centro en el punto $O(0, 0)$ y de los dominios anulares (fig. 5.1).

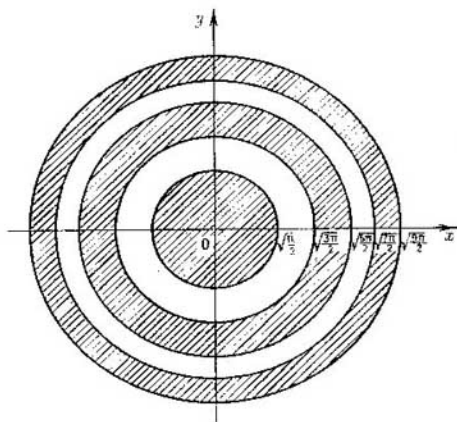


Fig. 5.1

4°. $u = \ln xyz$. El dominio de definición de esta función es un conjunto $\{M\}$ de puntos, cuyas coordenadas satisfacen la desigualdad $xyz > 0$, en tanto que el conjunto de valores es toda la recta numérica $-\infty < u < +\infty$.

5°. $u = x^2 + y^2 + z^2$. El dominio de definición de esta función es todo el espacio euclídeo, en tanto que el conjunto de valores es la semirrecta $u \geq 0$.

4. **Conceptos de espacio coordinado m -dimensional y de espacio euclídeo m -dimensional.** Un conjunto de toda clase de colecciones ordenadas (x_1, x_2, \dots, x_m) de m números x_1, x_2, \dots, x_m se llama *espacio coordinado m -dimensional* A^m .

En este caso cada colección ordenada (x_1, x_2, \dots, x_m) se denominará punto de este espacio y se denotará con la letra M . Los números x_1, x_2, \dots, x_m se llaman coordenadas del punto M . La notación $M(x_1, x_2, \dots, x_m)$ significa que el punto M tiene coordenadas x_1, x_2, \dots, x_m . Introduzcamos el concepto de *espacio euclídeo m -dimensional*. Un espacio coordinado A^m recibe el nombre de *espacio euclídeo m -dimensional* E^m si entre dos puntos M' (x'_1, x'_2, \dots, x'_m)

y $M'' (x_1'', x_2'', \dots, x_m'')$ cualesquiera del espacio coordenado A^m está definida la distancia *) $\rho (M', M'')$ según la fórmula

$$\rho (M', M'') = \sqrt{(x_1' - x_1'')^2 + (x_2' - x_2'')^2 + \dots + (x_m' - x_m'')^2}. \quad (5.1)$$

Los conceptos introducidos de espacio coordenado m -dimensional A^m y de espacio euclídeo m -dimensional E^m son una generalización de los conceptos, mencionados más arriba, de espacio coordenado y de espacio euclídeo.

5. Conjuntos de puntos del espacio euclídeo m -dimensional E^m . Mediante el símbolo $\{M\}$ se denotará cierto conjunto de puntos del espacio euclídeo m -dimensional E^m . Analicemos algunos ejemplos de conjuntos en el espacio euclídeo m -dimensional E^m .

1°. El conjunto $\{M\}$ de toda clase de puntos, cuyas coordenadas x_1, x_2, \dots, x_m satisfacen la inecuación $(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2 + \dots + (x_m - x_m^0)^2 \leq R^2$, se denomina *esfera m -dimensional* de radio R con centro en el punto $M_0 (x_1^0, x_2^0, \dots, x_m^0)$. De este modo,

*) El espacio euclídeo m -dimensional es un *espacio métrico*. Un conjunto arbitrario $\{M\}$, cuyos elementos se llaman puntos, se llama espacio métrico si existe una regla, con cuya ayuda a dos puntos M' y M'' cualesquiera del conjunto $\{M\}$ se les pone en correspondencia cierto número $\rho (M', M'')$, denominada *distancia* entre dichos puntos. En este caso la citada regla ha de ser tal que se cumplan los axiomas (*axiomas del espacio métrico*) siguientes: 1) para M' y M'' cualesquiera $\rho (M', M'') = \rho (M'', M')$ (simetría de la distancia); 2) para M' y M'' cualesquiera $\rho (M', M'') \geq 0$, con la particularidad de que si $\rho (M', M'') = 0$, los puntos M' y M'' coinciden; 3) para tres puntos M', M'', M''' cualesquiera se cumple la desigualdad $\rho (M', M''') \leq \rho (M', M'') + \rho (M'', M''')$ (desigualdad triangular).

Cerciorámonos de que el espacio euclídeo m -dimensional introducido es realmente un espacio métrico. En efecto, la validez de los primeros dos axiomas del espacio métrico es obvia (véase la fórmula (5.1)). Vamos a probar la validez del tercer axioma.

Sean x_1', x_1'', x_1''' las coordenadas de los puntos M', M'', M''' . Tenemos $\rho^2 (M', M''') = \sum_{i=1}^m [(x_i' - x_i'') - (x_i'' - x_i''')]^2 = \sum_{i=1}^m (x_i' - x_i'')^2 + 2 \sum_{i=1}^m (x_i' - x_i'') (x_i'' - x_i''') - x_i'' + \sum_{i=1}^m (x_i'' - x_i''')^2$. Suponiendo $x_i'' - x_i''' = a_i$ y $x_i' - x_i'' = b_i$ y aprovechando la desigualdad de Buniakovski (véase la desigualdad (1.30) en el Complemento 1 al cap. 1) hallaremos que $\left[\sum_{i=1}^m (x_i'' - x_i''') (x_i' - x_i'') \leq \left[\sum_{i=1}^m (x_i'' - x_i''')^2 \times \right. \right.$

$\left. \times \sum_{i=1}^m (x_i' - x_i'')^2 \right]$. De aquí proviene que

$$\rho^2 (M', M''') \leq \left[\sum_{i=1}^m (x_i' - x_i'')^2 + \sum_{i=1}^m (x_i'' - x_i''')^2 \right]^2,$$

es decir, $\rho (M', M''') \leq \rho (M', M'') + \rho (M'', M''')$.

una esfera m -dimensional se define como el conjunto $\{M\}$ de toda clase de puntos M , la distancia ρ de cada uno de los cuales hasta cierto punto M_0 (centro de la esfera) satisface la desigualdad $\rho(M, M_0) \leq R$. Si la distancia $\rho(M, M_0)$ de cada punto del conjunto $\{M\}$ hasta el punto M_0 satisface la desigualdad estricta $\rho(M, M_0) < R$, el conjunto $\{M\}$ se llama *esfera m -dimensional abierta*.

2°. El conjunto $\{M\}$ de puntos, la distancia de cada uno de los cuales hasta cierto punto M_0 satisface la relación $\rho(M, M_0) = R$, se llama *esfera m -dimensional* de radio R con centro en el punto M_0 .

3°. El conjunto $\{M\}$ de puntos, cuyas coordenadas x_1, x_2, \dots, x_m satisfacen las inequaciones $|x_1 - x_1^0| \leq d_1, |x_2 - x_2^0| \leq d_2, \dots, |x_m - x_m^0| \leq d_m$ se llama *paralelepípedo coordenado m -dimensional*. Además, el punto $M_0(x_1^0, x_2^0, \dots, x_m^0)$ se llama centro de este paralelepípedo m -dimensional. Si las coordenadas x_1, x_2, \dots, x_m de los puntos del conjunto $\{M\}$ satisfacen las desigualdades estrictas $|x_1 - x_1^0| < d_1, |x_2 - x_2^0| < d_2, \dots, |x_m - x_m^0| < d_m$, el conjunto $\{M\}$ se llama *paralelepípedo coordenado m -dimensional abierto*.

Introduzcamos el concepto de ε -entorno del punto M_0 de un espacio euclídeo m -dimensional y el de *entorno rectangular del referido punto M_0* . Llamaremos ε -entorno del punto $M_0(x_1^0, x_2^0, \dots, x_m^0)$ del espacio euclídeo m -dimensional E^m una *esfera m -dimensional abierta de radio ε con centro en el punto M_0* . Se denomina *entorno rectangular del punto $M_0(x_1^0, x_2^0, \dots, x_m^0)$ de un espacio euclídeo m -dimensional un paralelepípedo coordenado m -dimensional abierto cualquiera con centro en el punto M_0* .

Es válida la siguiente afirmación evidente.

Cualquier ε -entorno de un punto M_0 del espacio euclídeo m -dimensional E^m contiene cierto entorno rectangular de este punto. Un entorno rectangular cualquiera del punto M_0 contiene cierto ε -entorno del punto M_0 .

Sea $\{M\}$ un conjunto de puntos del espacio euclídeo m -dimensional E^m . Introduzcamos los conceptos siguientes.

El punto M del conjunto $\{M\}$ se llama punto interior de dicho conjunto si existe cierto ε -entorno del punto M , cuyos puntos todos pertenecen al conjunto $\{M\}$.

El punto M^ se denomina punto de frontera del conjunto $\{M\}$ si todo ε -entorno de este punto contiene tanto los puntos pertenecientes al conjunto $\{M\}$, como también los que no pertenecen a $\{M\}$.*

El conjunto $\{M\}$ del espacio E^m se llama conjunto o dominio abierto si cualquier punto de dicho conjunto es interior.

Si todo punto de frontera del conjunto $\{M\}$ es punto de este conjunto, el conjunto $\{M\}$ se denomina cerrado.

* virtamos que en este caso el punto M puede no pertenecer al conjunto $\{M\}$

Si el conjunto $\{M\}$ es un dominio, el conjunto $\{\bar{M}\}$, obtenido sumando a $\{M\}$ todos los puntos de frontera de este conjunto, se denomina *dominio cerrado*.

Advirtamos que si todos los puntos del dominio $\{M\}$ se disponen en el interior de cierta esfera, este dominio se llama *limitado*.

En lo ulterior nos hará falta el concepto de *conjunto conexo*. Introduzcamos preliminarmente el concepto de *curva continua* en el espacio multidimensional E^m .

Llamaremos *curva continua* L en el espacio E^m al conjunto $\{M\}$ de puntos de este espacio, cuyas coordenadas x_1, x_2, \dots, x_m son funciones continuas de parámetro t :

$$x_1 = \varphi_1(t), x_2 = \varphi_2(t), \dots, x_m = \varphi_m(t), \alpha \leq t \leq \beta. \quad (5.2)$$

Diremos que los puntos M' (x'_1, x'_2, \dots, x'_m) y M'' ($x''_1, x''_2, \dots, x''_m$) del espacio E^m pueden unirse mediante una curva continua L si existe una curva continua L tal, definida mediante las ecuaciones paramétricas (14.2), que

$$\begin{aligned} x'_1 &= \varphi_1(\alpha), & x'_2 &= \varphi_2(\alpha), & \dots, & x'_m &= \varphi_m(\alpha), \\ x''_1 &= \varphi_1(\beta), & x''_2 &= \varphi_2(\beta), & \dots, & x''_m &= \varphi_m(\beta). \end{aligned}$$

Enunciamos el concepto de conjunto conexo. *El conjunto $\{M\}$ del espacio E^m se llama conexo si dos puntos cualesquiera del mismo pueden ser unidos con una curva continua, cuyos puntos todos pertenecen a este conjunto.*

OBSERVACION. Advirtamos que a veces llámase dominio a un conjunto abierto y conexo, y no simplemente a un conjunto abierto.

Veamos el ejemplo siguiente.

El conjunto $\{M\}$ de puntos de E^m definido por la ecuación

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \dots + \frac{x_m^2}{a_m^2} = 1, \quad (5.3)$$

se llama *elipsoide m -dimensional*. Los puntos de un elipsoide m -dimensional son puntos de frontera del conjunto $\{M\}$ de puntos M , cuyas coordenadas satisfacen una inequación

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \dots + \frac{x_m^2}{a_m^2} < 1.$$

Este conjunto es un conjunto de puntos interiores de un elipsoide m -dimensional.

El lector mismo puede convencerse con facilidad de que el conjunto de puntos interiores de un elipsoide m -dimensional es un conjunto abierto y conexo. Advirtamos que un elipsoide m -dimensional definido por la relación (14.3) es un conjunto cerrado.

El dominio de definición de la función

$$u = \sqrt{\cos(x^2 + y^2)}$$

es un conjunto *inconexo* (véase ejemplo 3° del p. 3 y fig. 5.1).

Para concluir, convengamos en llamar *entorno del punto M* a todo conjunto abierto conexo en el que está contenido M .

6. Concepto de función de m variables. Introduzcamos el concepto de función de m variables.

Si a todo punto M del conjunto $\{M\}$ de puntos del espacio euclídeo m -dimensional E^m se le pone en correspondencia, de acuerdo con una ley determinada, cierto número u , suele decirse que sobre el conjunto $\{M\}$ está dada la función $u = u(M)$, o $u = f(M)$. En este caso el conjunto $\{M\}$ se llama *dominio de definición de la función $u = f(M)$* .

El número u , correspondiente al punto dado M del conjunto $\{M\}$, se llamará *valor particular de la función en el punto M* . La colección $\{u\}$ de todos los valores particulares de la función $u = f(M)$ se denomina *conjunto de valores* de dicha función. Por cuanto el punto M se define mediante las coordenadas x_1, x_2, \dots, x_m , para la función $u = f(M)$ de m variables se emplea también la designación $u = f(x_1, x_2, \dots, x_m)$.

Veamos unos ejemplos de funciones de m variables.

1°. Sea $u = \sqrt{1 - x_1^2 - x_2^2 - \dots - x_m^2}$. Como dominio de definición de esta función interviene, evidentemente, una esfera m -dimensional de radio 1 con centro en el punto $O(0, 0, \dots, 0)$. Es conjunto de valores de la función en consideración el segmento $[0, 1]$.

2°. Sea $u = \frac{1}{\sqrt{1 - \frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \dots - \frac{x_m^2}{a_m^2}}}$. Como dominio de definición de la función sirve el conjunto $\{M\}$ de puntos interiores de un elipsoide m -dimensional. El conjunto de valores de esta función está representado por una semirrecta $u \geq 1$.

§ 2. Valor límite de una función de varias variables

1. Sucesiones convergentes de puntos en un espacio euclídeo m -dimensional E^m . Criterio de convergencia (de Cauchy) de una sucesión. Examinemos en el espacio euclídeo m -dimensional E^m una sucesión de puntos $\{M_n\}$ *). Enunciemos la definición siguiente.

La sucesión $\{M\}$ de puntos del espacio euclídeo E^m se llama *convergente* si existe un punto A tal que para un número positivo ϵ cualquiera puede indicarse un número N **) tal que para $n \geq N$ se cumple la inecuación $\rho(M_n, A) < \epsilon$. En este caso el punto A se denomina

*) El concepto de sucesión de puntos en el espacio euclídeo E^m se define del modo siguiente. Supongamos que a todo número n de una serie natural de números $1, 2, \dots, n, \dots$ se le pone en correspondencia un punto M_n del espacio euclídeo E^m . La serie de puntos $M_1, M_2, \dots, M_n, \dots$ que aparece en este caso y se examina en el orden indicado, se denomina *sucesión* de puntos del espacio euclídeo E^m . Para mayor brevedad, dicha sucesión se denotará con el símbolo $\{M_n\}$.

**) Por cuanto el número N depende, como regla, de ϵ , a veces se escribe $N = N(\epsilon)$.

límite de la sucesión $\{M_n\}$. Para designar el límite A de la sucesión $\{M_n\}$ se emplea el símbolo siguiente:

$$\lim_{n \rightarrow \infty} M_n = A, \text{ o bien } M_n \rightarrow A \text{ cuando } n \rightarrow \infty.$$

Demostremos el lema siguiente.

Lema 1. *Supongamos que la sucesión $\{M_n\}$ de puntos del espacio euclídeo E^m converge hacia el punto A . Entonces, las sucesiones $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, \dots , $\{x_m^{(n)}\}$ de coordenadas de los puntos M_n convergen hacia las coordenadas correspondientes a_1, a_2, \dots, a_m del punto A , y, viceversa, si las sucesiones $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, \dots , $\{x_m^{(n)}\}$ de coordenadas de los puntos M_n convergen, respectivamente, hacia los números a_1, a_2, \dots, a_m , la sucesión $\{M_n\}$ converge hacia el punto A que tiene coordenadas a_1, a_2, \dots, a_m .*

DEMOSTRACIÓN Demostremos la primera parte del lema. Si la sucesión $\{M_n\}$ converge hacia el punto A , puede indicarse, cualquiera que sea $\varepsilon > 0$, un número N tal que para $n \geq N$ se cumple la desigualdad $\rho(M_n, A) < \varepsilon$. Sean $(x_1^{(n)}, x_2^{(n)}, \dots, x_m^{(n)})$ las coordenadas del punto M_n , y sean (a_1, a_2, \dots, a_m) , las coordenadas del punto A . Entonces, la desigualdad $\rho(M_n, A) < \varepsilon$ puede escribirse del modo siguiente:

$$\sqrt{(x_1^{(n)} - a_1)^2 + (x_2^{(n)} - a_2)^2 + \dots + (x_m^{(n)} - a_m)^2} < \varepsilon. \quad (5.4)$$

De aquí se deduce que, cuando $n \geq N$, se cumplen las inecuaciones

$$|x_1^{(n)} - a_1| < \varepsilon, |x_2^{(n)} - a_2| < \varepsilon, \dots, |x_m^{(n)} - a_m| < \varepsilon.$$

Dicho de otro modo, las sucesiones $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, \dots , $\{x_m^{(n)}\}$ de coordenadas de los puntos M_n convergen, respectivamente, hacia los números a_1, a_2, \dots, a_m . Demostremos ahora la afirmación inversa. Supongamos que las sucesiones citadas de coordenadas de los puntos M_n convergen, respectivamente, hacia los números a_1, a_2, \dots, a_m . Entonces, para cualquier $\varepsilon > 0$ pueden indicarse unos números N_1, N_2, \dots, N_m tales que para $n \geq N_1$, $n \geq N_2, \dots, n \geq N_m$ se cumplan las inecuaciones correspondientes

$$|x_1^{(n)} - a_1| < \frac{\varepsilon}{\sqrt{m}}, |x_2^{(n)} - a_2| < \frac{\varepsilon}{\sqrt{m}}, \dots, |x_m^{(n)} - a_m| < \frac{\varepsilon}{\sqrt{m}}.$$

De aquí se deduce que para $n \geq N = \max\{N_1, N_2, \dots, N_m\}$ se cumple la inecuación (5.4). En otras palabras, cuando $n \geq N$, se cumple la inecuación $\rho(M_n, A) < \varepsilon$, donde A es un punto de E^m con las coordenadas a_1, a_2, \dots, a_m . Así, pues, la sucesión $\{M_n\}$ converge hacia el punto A . El lema queda demostrado.

Formulemos la definición de sucesión fundamental de puntos en un espacio euclídeo m -dimensional. La sucesión $\{M_n\}$ de puntos de un espacio euclídeo m -dimensional se denomina fundamental, o suce-

sión de Cauchy si para cualquier número positivo ε puede indicarse un número N tal que para $n \geq N$ y para cualquier p natural se cumpla la desigualdad $\rho(M_{n+p}, M_n) < \varepsilon$. Es válido el siguiente criterio de convergencia de una sucesión (criterio de Cauchy).

Para que la sucesión $\{M_n\}$ de puntos de un espacio euclídeo m -dimensional sea convergente, es necesario y suficiente que la misma sea fundamental. Con el fin de convencerse de la validez del criterio enunciado, basta notar que el carácter fundamental de la sucesión $\{M_n\}$ determina el carácter fundamental de las sucesiones $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, . . . , $\{x_m^{(n)}\}$ de coordenadas de los puntos M_n y, viceversa, si las sucesiones mencionadas de coordenadas son fundamentales, lo será también la sucesión $\{M_n\}$, y luego basta aplicar el criterio de Cauchy para las sucesiones numéricas a las sucesiones de coordenadas de los puntos $\{M_n\}$ y el lema 1, enunciado más arriba.

2. Algunas propiedades de las sucesiones acotadas de puntos en un espacio euclídeo m -dimensional. Introduzcamos el concepto de sucesión acotada de puntos en un espacio euclídeo m -dimensional. La sucesión $\{M_n\}$ de puntos de un espacio euclídeo m -dimensional se denomina acotada si existe tal número $a > 0$ que para todo n se cumple la desigualdad $\rho(O, M_n) \leq a$, donde O es un punto cuyas coordenadas son todas iguales a cero. Dicho de otro modo, la sucesión $\{M_n\}$ está acotada si todos los puntos M_n de esta sucesión están en el interior o en la frontera de cierta esfera con centro en el origen de las coordenadas.

Es válido el siguiente teorema fundamental.

Teorema 5.1 (teorema de Bolzano—Weierstrass). En sucesión acotada $\{M_n\}$ de puntos de un espacio euclídeo m -dimensional se puede distinguir una subsucesión convergente.

DEMOSTRACIÓN. Primero corciorémonos de que las sucesiones $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, . . . , $\{x_m^{(n)}\}$ de coordenadas de los puntos M_n están acotadas. En efecto, por cuanto la sucesión $\{M_n\}$ está acotada, para todos los n se cumple la desigualdad $\rho(O, M_n) \leq a$. Dado que $\rho(O, M_n) = \sqrt{x_1^{(n)2} + x_2^{(n)2} + \dots + x_m^{(n)2}}$, de aquí se infiere que para todo n se cumplen las inecuaciones $|x_1^{(n)}| \leq a$, $|x_2^{(n)}| \leq a$, . . . , $|x_m^{(n)}| \leq a$. En otras palabras, las sucesiones $\{x_1^{(n)}\}$, $\{x_2^{(n)}\}$, . . . , $\{x_m^{(n)}\}$ de coordenadas de los puntos M_n están acotadas. En virtud del teorema de Bolzano—Weierstrass para las sucesiones numéricas (véase p. 4, § 4, cap. 3, t. 1), en la sucesión $\{x_1^{(n)}\}$ se puede distinguir una sucesión $\{x_1^{(n_{k_1})}\}$ que converge hacia cierto número a_1 . Examinemos la subsucesión correspondiente $\{x_2^{(n_{k_1})}\}$ de la sucesión de segundas coordenadas de los puntos M_n . En vista del mismo teorema, en la subsucesión $\{x_2^{(n_{k_1})}\}$ se puede distinguir una subsucesión $\{x_2^{(n_{k_2})}\}$ que converge hacia cierto número a_2 . Advirtamos que la subsucesión $\{x_1^{(n_{k_2})}\}$ de la sucesión $\{x_1^{(n_{k_1})}\}$

converge hacia el número a_1 . Así, pues, las sucesiones $\{x_2^{(n_{h_1})}\}$ y $\{x_1^{(n_{h_1})}\}$ convergen hacia los números a_1 y a_2 , respectivamente. Es evidente que al distinguir en la subsucesión $\{x_3^{(n_{h_1})}\}$ de la sucesión de terceras coordenadas de los puntos M_n una subsucesión $\{x_3^{(n_{h_2})}\}$ que converge hacia cierto número a_3 , las subsucesiones $\{x_1^{(n_{h_2})}\}$, $\{x_2^{(n_{h_2})}\}$, $\{x_3^{(n_{h_2})}\}$ convergen hacia los números respectivos a_1 , a_2 , a_3 . Continuando estos razonamientos, obtendremos, en fin, una subsucesión $\{x_m^{(n_{h_m})}\}$, convergente hacia cierto número a_m , de la sucesión de m -ésimas coordenadas de los puntos M_n , con la particularidad de que las subsucesiones $\{x_1^{(n_{h_m})}\}$, $\{x_2^{(n_{h_m})}\}$, \dots , $\{x_m^{(n_{h_m})}\}$ convergen hacia los números a_1 , a_2 , \dots , a_m , respectivamente. Mas, en este caso, debido al lema 1, la subsucesión $\{M_{n_{h_m}}\}$ de la sucesión de puntos $\{M_n\}$ converge hacia el punto A que tiene por coordenadas a_1 , a_2 , \dots , a_m . El teorema está demostrado.

OBSERVACIÓN El límite A de la sucesión $\{M_n\}$ de puntos pertenecientes al conjunto cerrado $\{M\}$ pertenece también a dicho conjunto. Para cerciorarnos de ello, basta notar que en cualquier ε -entorno del punto A hay puntos M_n , es decir, puntos del conjunto $\{M\}$, por lo cual el punto A es o bien un punto interior o bien un punto de frontera de $\{M\}$, y, por consiguiente, pertenece a $\{M\}$.

3. Concepto de valor límite de una función de varias variables. Examinemos la función $u = f(M)$, definida sobre el conjunto $\{M\}$ de un espacio euclídeo m -dimensional, y un punto A de dicho conjunto que, quizás, no pertenece al conjunto $\{M\}$, pero posee la propiedad de que en cualquier ε -entorno de este punto está contenido al menos un punto del conjunto $\{M\}$, distinto de A .

Definición 1. El número b se llama valor límite de la función $u = f(M)$ en un punto A (o bien límite de la función para $M \rightarrow A$) si para cualquier sucesión $M_1, M_2, \dots, M_n, \dots$ de puntos del conjunto $\{M\}$, la cual converge hacia A y cuyos elementos M_n son distintos de A ($M_n \neq A$), la sucesión correspondiente $f(M_1), f(M_2), \dots, f(M_n), \dots$ de valores de la función converge hacia b .

La definición aducida se llama definición de valor límite de una función con ayuda de las sucesiones. Formulemos otra definición de valor límite de una función, utilizando la « $\varepsilon - \delta$ » terminología.

Definición 2. El número b se denomina valor límite de la función $u = f(M)$ en el punto A si para cualquier número positivo ε puede indicarse un número positivo δ tal que para todos los puntos M del dominio de definición de la función que satisfacen la condición $0 < \rho(M, A) < \delta$ se verifica la inecuación $|f(M) - b| < \varepsilon$.

OBSERVACIÓN Las definiciones 1 y 2 de valor límite de una función son equivalentes. La validez de esta afirmación puede ser de-

*1 Este requisito se debe, en particular, a que la función $u = f(M)$ puede estar no definida en el punto A .

mostrada del mismo modo que la equivalencia de dos definiciones de valor límite de una función de una sola variable. Para denotar el valor límite b de la función $u = f(M)$ en el punto A se aplica el símbolo siguiente

$$\lim_{M \rightarrow A} f(M) = b, \text{ o bien } \lim_{\substack{x_1 \rightarrow a_1 \\ x_2 \rightarrow a_2 \\ \dots \\ x_m \rightarrow a_m}} f(x_1, x_2, \dots, x_m) = b,$$

donde a_1, a_2, \dots, a_m son las coordenadas del punto A .

Enunciemos la definición de valor límite de una función cuando el punto M tiende al infinito.

Definición 3. El número b se llama valor límite de la función $u = f(M)$ cuando $M \rightarrow \infty$ (o bien límite de la función cuando $M \rightarrow \infty$) si para cualquier número positivo ε puede indicarse el número positivo a tal que para todos los puntos M del dominio de definición de la función que satisfacen la condición $\rho(O, M) > a$, se cumple la inecuación $|f(M) - b| < \varepsilon$.

Las operaciones aritméticas con las funciones de m variables que tienen el valor límite en el punto A conducen a unas funciones que también tienen su valor límite en el mismo punto A . A saber, es válida la afirmación siguiente.

Supongamos que las funciones $f(M)$ y $g(M)$ tienen en el punto A los valores límites b y c . Entonces, las funciones $f(M) + g(M)$, $f(M) - g(M)$, $f(M) \cdot g(M)$ y $\frac{f(M)}{g(M)}$ tienen en A sus valores límites (el cociente, a condición de que $c \neq 0$) iguales a $b + c$, $b - c$, $b \cdot c$, $\frac{b}{c}$, respectivamente.

La demostración de esta afirmación es análoga a la del teorema 4.1.

4. Funciones infinitésimas. La función $u = f(M)$ se llama infinitésima en el punto A (cuando $M \rightarrow A$) si $\lim_{M \rightarrow A} f(M) = 0$.

Es fácil convencerse de que la función $f(M) = (x_1 - a_1)^{n_1} + \dots + (x_m - a_m)^{n_m}$, donde n_1, \dots, n_m son números positivos, es infinitésima en el punto $A(a_1, a_2, \dots, a_m)$ *).

Si la función $u = f(M)$ tiene en A el valor límite igual a b , la función $\alpha(M) = f(M) - b$ es infinitésima en el punto A . En efecto,

$$\lim_{M \rightarrow A} \alpha(M) = \lim_{M \rightarrow A} (f(M) - b) = \lim_{M \rightarrow A} f(M) - \lim_{M \rightarrow A} b = 0.$$

Haciendo uso de este resultado, obtendremos la representación especial para una función que tiene su valor límite igual a b en el punto A :

$$f(M) = b + \alpha(M), \text{ donde } \lim_{M \rightarrow A} \alpha(M) = 0.$$

*) Basta considerar que cada una de las funciones de una sola variable $x_k \rightarrow (x_k - a_k)^{n_k}$ es infinitésima en el punto $x_k = a_k$.

Las funciones infinitésimas de varias variables se comparan de un modo igual al mencionado en el p. 3, § 2, cap. 4, t. 1 para funciones infinitésimas de una sola variable. Advirtamos que, al igual que en el caso de variable única, por símbolo $o(\beta)$ se entenderá cualquier función infinitésima en el punto dado A , cuyo orden de pequeñez es más elevado en comparación con la función $\beta(M)$, infinitésima en el punto dado A .

5. Condición necesaria y suficiente de existencia del valor límite de una función (criterio de Cauchy). Diremos que la función $f(M)$ *satisface la condición de Cauchy en el punto $M = A$* si para cualquier número positivo ε existe un número positivo δ tal, cualesquiera que sean dos puntos M' y M'' del dominio de definición de la función $f(M)$ que satisfacen las inequaciones $0 < \rho(M', A) < \delta$, $0 < \rho(M'', A) < \delta$, para los valores correspondientes de las funciones se verifica la inequación

$$|f(M') - f(M'')| < \varepsilon.$$

Es válido el siguiente teorema *fundamental*.

Teorema 5.2 (criterio de Cauchy). *Para que la función $f(M)$ tenga en el punto $M = A$ su valor límite finito, es necesario y suficiente que la función $f(M)$ satisfaga en este punto la condición de Cauchy.* La demostración de este teorema es plonamente análoga a la del teorema 8.2 y se obtiene de esta última substituyendo las letras x y a por M y A y cambiando las expresiones de la forma $|x - a|$ por el símbolo $\rho(M, A)$.

6. Valores límites reiterados. Para la función de varias variables $u = f(x_1, x_2, \dots, x_n)$ se puede definir el concepto de valor límite respecto de una de las variables x_h , siendo fijos los valores de las demás variables. En esta situación surge el concepto de *valor límite reiterado*. Aclarémoslo con ayuda del ejemplo de función $u = f(x, y)$ de dos variables x e y . Sea que la función $u = f(x, y)$ está definida en cierto entorno rectangular $|x - x_0| < d_1$, $|y - y_0| < d_2$ del punto $M_0(x_0, y_0)$, a excepción, quizás, del propio punto M_0 . Supongamos que para todo y fijo que satisface la condición $0 < y - y_0 < d_2$, existe el valor límite de la función $u = f(x, y)$ de una variable x en el punto $x = x_0$:

$$\lim_{\substack{x \rightarrow x_0 \\ y \text{ fijo}}} f(x, y) = \varphi(y),$$

y que, además, existe el valor límite b de la función $\varphi(y)$ en el punto $y = y_0$:

$$\lim_{y \rightarrow y_0} \varphi(y) = b.$$

En este caso se dice que existe el valor límite *reiterado* b para la función $u = f(x, y)$ en el punto M_0 , el cual se denota del modo siguiente:

$$\lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y) = b.$$

Análogamente se define el valor límite reiterado

$$\lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y).$$

Demos a conocer las condiciones suficientes de igualdad entre dos valores límites reiterados recién introducidos.

Teorema 5.3. Supongamos que la función $u = f(x, y)$ está definida en cierto entorno rectangular $|x - x_0| < d_1$, $|y - y_0| < d_2$ del punto $M_0(x_0, y_0)$ y tiene en dicho punto el valor límite b . Admitamos también que para cualquier x fijo, $0 < |x - x_0| < d_1$, existe el valor límite $\psi(x) = \lim_{y \rightarrow y_0} f(x, y)$ y para cualquier y fijo, $0 < |y - y_0| < d_2$, existe el valor límite $\varphi(y) = \lim_{x \rightarrow x_0} f(x, y)$.

En este caso existen los valores límites reiterados $\lim_{x \rightarrow x_0} \lim_{y \rightarrow y_0} f(x, y)$ y

$\lim_{y \rightarrow y_0} \lim_{x \rightarrow x_0} f(x, y)$, iguales a b .

DEMOSTRACIÓN. Por cuanto la función $u = f(x, y)$ tiene en $M_0(x_0, y_0)$ el valor límite b , podemos indicar para todo $\varepsilon > 0$ tal $\delta > 0$ que cuando $|x - x_0| < \delta$ o $|y - y_0| < \delta$ se cumple la inecuación $|f(x, y) - b| < \varepsilon$. De este modo, en el entorno rectangular $|x - x_0| < \delta$ o $|y - y_0| < \delta$ del punto M_0 los valores de la función $f(x, y)$ se diferencian de b en una magnitud no superior a ε . Mas, en este caso, los valores límites $\psi(x)$ y $\varphi(y)$, aducidos en la formulación del teorema, para x e y que satisfacen las inecuaciones $|x - x_0| < \delta$ o $|y - y_0| < \delta$, también difieren de b en una magnitud no superior a ε . Por consiguiente, los valores límites de estas funciones en los puntos x_0 e y_0 existen y son iguales a b . El teorema está demostrado.

Se puede definir el concepto de límite reiterado para las llamadas sucesiones dobles $\{a_{mn}\}$, cuyos elementos a_{mn} se determinan por dos índices m y n . A saber: el símbolo $\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} a_{mn}$ quiere decir que al principio se define una sucesión $\{b_n\}$, $b_n = \lim_{m \rightarrow \infty} a_{mn}$, y luego se halla el límite de esta sucesión $\{b_n\}$.

Veamos, por ejemplo, la sucesión doble $\{a_{mn}\}$, donde $a_{mn} = \cos^m 2\pi n!x$, x es un número fijo. Demostremos que

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \cos^m 2\pi n!x = \begin{cases} 1 & \text{si } x \text{ es un número racional,} \\ 0 & \text{si } x \text{ es un número irracional.} \end{cases}$$

En efecto, si $x = p/q$, donde p y q son números enteros, para $n \geq q$ tenemos $\cos 2\pi n!x = 1$, y, por eso, $\lim_{m \rightarrow \infty} \cos^m 2\pi n!x = 1$. En otras palabras, si x

es un número racional, se tiene $\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \cos^m 2\pi n!x = 1$. Si, en cambio, x

es irracional, para todo n se cumple la inecuación $|\cos 2\pi n!x| < 1$, por lo que $\lim_{m \rightarrow \infty} \cos^m 2\pi n!x = 0$, es decir, $\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \cos^m 2\pi n!x = 0$.

OBSERVACIÓN. Haciendo uso del resultado obtenido, podemos definir analíticamente la función de Dirichlet (véase p. 1, § 1, cap. 4, t. 1), como límite reiterado $\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \cos^m 2\pi n!x$.

§ 3. Funciones continuas de varias variables

1. Definición de continuidad de una función de varias variables. Supongamos que el punto A pertenece al dominio de definición de la función $u = f(M)$ de varias variables y que cualquier ε -entorno

del punto A contiene puntos, distintos de A , del dominio de definición de dicha función.

Definición 1. La función $u = f(M)$ se llama continua en el punto A si el valor límite de esta función en el punto A existe y es igual al valor particular de $f(A)$. Advertimos que por cuanto $A = \lim_{M \rightarrow A} M$,

la condición de continuidad de la función puede ser escrita en la forma siguiente

$$\lim_{M \rightarrow A} f(M) = f(\lim_{M \rightarrow A} M).$$

Los puntos, donde la función no posee la propiedad de continuidad, se llaman *puntos de discontinuidad* de esta función.

Formulemos la definición de continuidad de una función, empleando la definición de valor límite de una función con ayuda de ε y δ .

Definición 2. La función $u = f(M)$ se llama continua en el punto A si para cualquier número positivo ε puede indicarse tal número positivo δ que para todo M del dominio de definición de la función que satisfice la condición $\rho(M, A) < \delta$, se cumple la inequación $|f(M) - f(A)| < \varepsilon$.

Definición 3. La función $u = f(M)$ se llama continua sobre el conjunto $\{M\}$ si lo es en cada punto de dicho conjunto.

llamemos *incremento* o *incremento total* de la función $u = f(M)$ en el punto A la función Δu que se define mediante la fórmula

$$\Delta u = f(M) - f(A), \quad (5.5)$$

donde M es un punto cualquiera del dominio de definición de la función. Supongamos que A y M tienen coordenadas respectivas a_1, a_2, \dots, a_m y x_1, x_2, \dots, x_m . Denotemos $x_1 - a_1 = \Delta x_1, x_2 - a_2 = \Delta x_2, \dots, x_m - a_m = \Delta x_m$. Utilizando estas denotaciones, obtendremos la siguiente expresión para el incremento de la función Δu , correspondiente a los incrementos de los argumentos $\Delta x_1, \dots, \Delta x_m$:

$$\Delta u = f(a_1 + \Delta x_1, a_2 + \Delta x_2, \dots, a_m + \Delta x_m) - f(a_1, a_2, \dots, a_m). \quad (5.6)$$

Evidentemente, para que la función $u = f(M)$ sea continua en el punto A , es necesario y suficiente que su incremento Δu sea una función infinitésima en el punto A , es decir, es necesario y suficiente que se verifique la igualdad

$$\lim_{M \rightarrow A} \Delta u = \lim_{M \rightarrow A} (f(M) - f(A)) = 0, \quad \text{o bien} \quad \lim_{\substack{\Delta x_1 \rightarrow 0 \\ \Delta x_2 \rightarrow 0 \\ \Delta x_m \rightarrow 0}} \Delta u = 0. \quad (5.7)$$

La condición (5.7) se llamará *forma en diferencias* de la condición de continuidad de la función $u = f(M)$ en el punto A

toda sucesión de puntos $\{M_n\}$ de esta recta convergente hacia el punto M la sucesión correspondiente $\{f(M_n)\}$ de valores de la función tiene por límite el valor particular $f(M)$ de la función en el punto M . Por cuanto sobre una recta la función $u = f(x, y)$ es una función de una sola variable, el concepto de continuidad de la función sobre la recta coincide, obviamente, con el de continuidad de la citada función de una sola variable. En particular, la continuidad de la función en el punto M respecto de cada una de las variables x e y es su continuidad sobre las rectas que pasan por el punto M , paralelas a los ejes coordenados. Demostremos que la función

$$u = \begin{cases} \frac{xy}{x^2+y^2} & \text{para } x^2+y^2 \neq 0, \\ 0 & \text{para } x^2+y^2 = 0 \end{cases}$$

es continua en el punto $O(0, 0)$ respecto de cada una de las variables x e y , es decir, es continua en cada uno de los ejes coordenados, pero no lo es sobre las demás rectas que pasan por este punto, por lo cual no es continua en el punto O . Cada recta, distinta de los ejes coordenados, que pasa por el punto $O(0, 0)$ puede ser representada mediante la ecuación $y = kx$, donde $k \neq 0$. Es obvio que en semejante recta todos los valores de la función son constantes e iguales a $\frac{k}{1+k^2}$. Por eso, si la sucesión $\{M_n\}$ de puntos de semejante recta, distintos del punto O , converge hacia el punto O , la sucesión correspondiente de valores de la función tiene por límite $\frac{k}{1+k^2}$. Por cuanto este límite es distinto de cero cuando $k \neq 0$ y no coincide con el valor particular de la función en el punto O , la función es discontinua en dicho punto de la recta en consideración. La continuidad de una función sobre los ejes coordenados se deduce de que sus valores en los mismos son iguales a cero.

Puede formarse la impresión de que si una función de dos variables es continua sobre una recta cualquiera que pasa por un punto dado, dicha función es continua en el citado punto. El ejemplo que sigue muestra que en el caso general esto no es así.

2°. Veamos la función

$$u = f(M) = \begin{cases} \frac{x^2y}{x^4+y^2} & \text{para } x^4+y^2 \neq 0 \\ 0 & \text{para } x^4+y^2 = 0. \end{cases}$$

Demostremos que, aunque la función mencionada es continua en toda recta que pasa por el punto $O(0, 0)$, no lo es, sin embargo, en este punto. En efecto, los valores de la función en la recta $y = kx$ son iguales a $\frac{kx}{x^2+k^2}$, y por eso, $u \rightarrow 0$ cuando $x \rightarrow 0$. La continuidad de esta función sobre el eje Oy se deduce de que sus valores en este eje son iguales a cero. Por otra parte, los valores de la fun

en el punto $A (a_1, a_2, \dots, a_h)$, mientras que la función $u = f(x_1, x_2, \dots, x_m)$ lo es en el punto $B (b_1, b_2, \dots, b_m)$, donde $b_i = \varphi_i(a_1, a_2, \dots, a_h)$, $i = 1, 2, \dots, m$. Entonces, la función compuesta $u = f(x_1, x_2, \dots, x_m)$, donde x_1, x_2, \dots, x_m son funciones de los argumentos t_1, t_2, \dots, t_h , definidas más arriba, es continua en el punto $A (a_1, a_2, \dots, a_h)$. Esbozemos las etapas principales de demostración de esta afirmación. Sea $\{N_n\}$, $N_n \neq A$, una sucesión arbitraria de puntos, convergente hacia A , del dominio $\{N\}$ de definición de las funciones $\varphi_i (t_1, t_2, \dots, t_h)$, y sea $\{M_n\}$ la sucesión correspondiente de puntos cuyas coordenadas $x_i^{(n)}$ son iguales a $\varphi_i (t_1^{(n)}, t_2^{(n)}, \dots, t_h^{(n)})$. Como las funciones φ_i son continuas en el punto A , la sucesión $\{M_n\}$ converge hacia el punto $B (b_1, b_2, \dots, b_m)$ (no se excluye la posibilidad de que los puntos M_n y B coincidan). En virtud de que las funciones $u = f(x_1, x_2, \dots, x_m)$ son continuas en el punto B , la sucesión $\{f(M_n)\}$ converge hacia $f(B)$. Mas, ésta es una sucesión de valores de una función compuesta que corresponde a la sucesión $\{N_n\}$ de puntos del dominio de su definición que converge hacia A . Ahora, cuando estamos convencidos de que la sucesión $\{f(M_n)\}$ converge hacia el valor particular $f(B)$, la continuidad de la función compuesta queda demostrada.

OBSERVACION La demostración aducida generaliza para el caso de varias variables la demostración del teorema 4.5 de la continuidad de una función compuesta de una sola variable.

3°. **TEOREMA DE LA ESTABILIDAD DEL SIGNO DE UNA FUNCIÓN CONTINUA**

Teorema 5.4. Si la función $u = f(M)$ es continua en el punto A del espacio euclídeo E^m y si $f(A) \neq 0$, existe un δ -entorno del punto A , dentro del cual en todo punto del dominio de definición de la función $f(M)$, ésta no se anula y tiene el signo idéntico al de $f(A)$. La validez de este teorema se deduce directamente de la definición de continuidad de una función en términos de « $\varepsilon - \delta$ ».

4°. **TEOREMA DE QUE UNA FUNCIÓN CONTINUA PASA POR CUALQUIER VALOR INTERMEDIO**

Teorema 5.5. Sea $u = f(M)$ una función continua en todos los puntos del conjunto conexo $\{M\}$ del espacio euclídeo E^m , con la particularidad de que $f(A)$ y $f(B)$ son los valores de esta función en los puntos A y B de dicho conjunto. Supongamos, además, que C es un número cualquiera encerrado entre $f(A)$ y $f(B)$. Entonces, en toda curva continua L que une los puntos A y B y que está íntegramente dispuesta en $\{M\}$ existe un punto N tal que $f(N) = C$.

DEMOSTRACION Sea

$$x_1 = \varphi_1(t), \quad x_2 = \varphi_2(t), \quad \dots, \quad x_m = \varphi_m(t), \quad \alpha \leq t \leq \beta,$$

la ecuación de la curva continua L que une los puntos A y B del conjunto M y que está íntegramente dispuesta en $\{M\}$ (véase p. 5, § 1). Sobre el segmento $[\alpha, \beta]$ está definida la función compuesta $u =$

$= f(x_1, x_2, \dots, x_m)$, donde $x_i = \varphi_i(t)$, $i = 1, 2, \dots, m$, $\alpha \leq t \leq \beta$. Es evidente que los valores de esta función en $[\alpha, \beta]$ coinciden con los de la función $u = f(M)$ en la curva L . La citada función compuesta de una sola variable t es continua sobre el segmento $[\alpha, \beta]$ en virtud de la afirmación en la sección 2° de este punto, y, de acuerdo con el teorema 8.6. en cierto punto ξ del segmento $[\alpha, \beta]$ toma el valor C . Por eso, en el punto N de la curva L con coordenadas $\varphi_1(\xi), \varphi_2(\xi), \dots, \varphi_m(\xi)$ se verifica la igualdad $f(N) = C$. El teorema está demostrado.

5°. CARÁCTER ACOTADO DE UNA FUNCIÓN, CONTINUA SOBRE UN CONJUNTO ACOTADO CERRADO.

Teorema 5.6 (primer teorema de Weierstrass). Si la función $u = f(M)$ es continua en el conjunto acotado cerrado $\{M\}$, es acotada sobre dicho conjunto. Detengámonos en la demostración del carácter acotado de $u = f(M)$ por arriba. Supongamos que $u = f(M)$ no está acotada superiormente en $\{M\}$. Al igual que en la demostración del teorema análogo 8.7, seleccionemos la sucesión $\{M_n\}$ de puntos del conjunto $\{M\}$, para los cuales $f(M_n) > n$. De conformidad con el teorema de Bolzano—Weierstrass (véase p. 2, § 2), se puede seleccionar en $\{M_n\}$ la subsucesión convergente $\{M_{h_n}\}$, cuyo límite M pertenece, en virtud de la observación al teorema de Bolzano—Weierstrass, al conjunto $\{M\}$. Evidentemente, la sucesión $\{f(M_{h_n})\}$ es infinita. Por otra parte, siendo continua la función en el punto M , esta sucesión $\{f(M_{h_n})\}$ ha de converger hacia $f(M)$. La contradicción obtenida demuestra el teorema.

6°. UNA FUNCIÓN, CONTINUA SOBRE UN CONJUNTO ACOTADO CERRADO, ALCANZA SUS COTAS EXACTAS.

Teorema 5.7 (segundo teorema de Weierstrass). Si la función $u = f(M)$ es continua sobre el conjunto acotado cerrado $\{M\}$, la misma alcanza en este conjunto sus cotas exactas superior e inferior. La demostración de este teorema es análoga a la del teorema 8.8 del t. I (segundo teorema de Weierstrass para la función de una sola variable).

7°. CONCEPTO DE CONTINUIDAD UNIFORME DE UNA FUNCIÓN DE VARIAS VARIABLES. La función $u = f(M)$ se llama uniformemente continua sobre el conjunto $\{M\}$ *) del espacio euclídeo E^m , si para todo número positivo ε puede indicarse número positivo δ , dependiente sólo de ε , tal que para cualesquiera dos puntos M' y M'' del conjunto $\{M\}$ que satisfacen la condición $\rho(M', M'') < \delta$ se cumple la inequación $|f(M') - f(M'')| < \varepsilon$. Tiene lugar el teorema siguiente.

Teorema 5.8 (teorema de continuidad uniforme). Una función continua en el conjunto acotado $\{M\}$ es uniformemente continua en

*) Se supone en este caso que el conjunto $\{M\}$ es denso en sí, es decir, en cualquier ε -entorno de cada punto M de dicho conjunto hay puntos del conjunto $\{M\}$ distintos de M .

este conjunto. La demostración de este teorema es análoga a la del teorema 1.2 y se obtiene de esta última sustituyendo el término «segmento $[a, b]$ » por el término «conjunto $\{M\}$ », la letra x , por la letra M , y las expresiones de tipo $|x'' - x'|$, por el símbolo $\rho(M', M'')$.

OBSERVACIÓN Llamaremos *diámetro* del conjunto acotado $\{M\}$ la cota superior exacta de los números $\rho(M', M'')$, donde M' y M'' son toda clase de puntos del conjunto $\{M\}$. Haciendo uso del concepto de diámetro de un conjunto, daremos a conocer la siguiente propiedad de las funciones continuas sobre los conjuntos acotados cerrados. *Supongamos que la función $u = f(M)$ es continua sobre el conjunto acotado cerrado $\{M\}$. Entonces, para cualquier número positivo ε puede indicarse un $\delta > 0$ tal que sobre todo subconjunto cerrado $\{\bar{M}\}$, que pertenece al conjunto $\{M\}$ y cuyo diámetro es inferior a δ , la oscilación ω^* de la función $f(M)$ es inferior a ε .* La demostración de esta propiedad es análoga a la del corolario del teorema 1.2.

§ 4. Derivadas y diferenciales de las funciones de varias variables

1. Derivadas parciales de una función de varias variables. Sea $M(x_1, x_2, \dots, x_m)$ un punto interior del dominio de definición de la función $u = f(x_1, x_2, \dots, x_m)$. Examinemos en el punto fijo dado $M(x_1, x_2, \dots, x_m)$ la razón entre el incremento parcial $\Delta_{x_h} u$ (véase p. 1, § 3, fórmulas (5.8) y (5.9)) y el incremento correspondiente Δx_h del argumento x_h :

$$\frac{\Delta_{x_h} u}{\Delta x_h} = \frac{f(x_1, x_2, \dots, x_{h-1}, x_h + \Delta x_h, x_{h+1}, \dots, x_m) - f(x_1, x_2, \dots, x_m)}{\Delta x_h} \quad (5.12)$$

La razón (5.12) es una función de Δx_h , definida para todos los valores de Δx_h , distintos de cero, para los cuales el punto $M(x_1, x_2, \dots, x_{h-1}, x_h + \Delta x_h, x_{h+1}, \dots, x_m)$ pertenece al dominio de definición de la función u .

Definición. Si existe un límite de la razón (5.12) entre el incremento parcial $\Delta_{x_h} u$ de la función en el punto $M(x_1, x_2, \dots, x_m)$ y el incremento correspondiente Δx_h del argumento x_h para $\Delta x_h \rightarrow 0$, este límite se denomina *derivada parcial* de la función $u = f(x_1, x_2, \dots, x_m)$ en el punto M respecto del argumento x_h y se denota con uno de los símbolos:

$$\frac{\partial u}{\partial x_h}, \quad \frac{\partial f}{\partial x_h}, \quad u'_{x_h}, \quad f'_{x_h}.$$

*) Llámase *oscilación ω* de la función $f(M)$ sobre el conjunto $\{M\}$ a la diferencia entre las cotas exactas superior e inferior de la función $f(M)$ en dicho conjunto.

De este modo,

$$\frac{\partial u}{\partial x_k} = \lim_{\Delta x_k \rightarrow 0} \frac{\Delta x_k u}{\Delta x_k}. \quad (5.13)$$

Advirtamos que la derivada parcial de la función $u = f(x_1, x_2, \dots, x_m)$ respecto del argumento x_k es una derivada ordinaria de la función de una sola variable x_k , siendo fijos los valores de las demás variables. Por eso, el cálculo de las derivadas parciales se realiza de acuerdo con las reglas usuales para el cálculo de las derivadas de las funciones de una sola variable.

EJEMPLOS.

$$1^\circ. u = \operatorname{arctg} \frac{x}{y}, \quad \frac{\partial u}{\partial x} = \frac{y}{x^2 + y^2}, \quad \frac{\partial u}{\partial y} = \frac{-x}{x^2 + y^2}.$$

$$2^\circ. u = xe^{yz} + \ln(x - y + z), \quad \frac{\partial u}{\partial x} = e^{yz} + \frac{1}{x - y + z},$$

$$\frac{\partial u}{\partial y} = xze^{yz} - \frac{1}{x - y + z}, \quad \frac{\partial u}{\partial z} = xye^{yz} + \frac{1}{x - y + z}.$$

$$3^\circ. u = \operatorname{tg} \sqrt{x^2 - yz}, \quad \frac{\partial u}{\partial x} = \frac{x}{\sqrt{x^2 - yz} \cos^2 \sqrt{x^2 - yz}}.$$

$$\frac{\partial u}{\partial y} = -\frac{z}{2\sqrt{x^2 - yz} \cos^2 \sqrt{x^2 - yz}},$$

$$\frac{\partial u}{\partial z} = -\frac{y}{2\sqrt{x^2 - yz} \cos^2 \sqrt{x^2 - yz}}.$$

OBSERVACION 1. Del que una función tiene en el punto dado todas las derivadas parciales no proviene, como regla, la continuidad de la misma en dicho punto. Ya nos hemos convencido de que la función

$$u = \begin{cases} \frac{xy}{x^2 + y^2} & \text{para } x^2 + y^2 \neq 0, \\ 0 & \text{para } x^2 + y^2 = 0 \end{cases}$$

no es continua en el punto $O(0, 0)$ (véase el ejemplo 1°, p. 1, § 3). No obstante, en este punto la función mencionada tiene derivadas parciales respecto de x o y . Esto se deduce de que $f(x, 0) \equiv 0$ y $f(0, y) \equiv 0$, y por eso

$$\left. \frac{\partial f}{\partial x} \right|_{(0, 0)} = 0 \quad \text{y} \quad \left. \frac{\partial f}{\partial y} \right|_{(0, 0)} = 0.$$

OBSERVACION 2. Hemos definido el concepto de derivadas parciales para los puntos interiores del dominio de representación de una función. Para los puntos de frontera del dominio de representación la definición de derivadas parciales aducida, en el caso general, no

es apta. En particular, esto se debe a que en los puntos de frontera del dominio de representación de la función no siempre pueden calcularse los incrementos parciales de dicha función (esto se refiere, por ejemplo, al punto de frontera M_0 del dominio mostrado en la fig. 5.2). Por eso, las derivadas parciales en los puntos de frontera del dominio de representación de la función se hallan, corrientemente, como valores límites de estas derivadas.

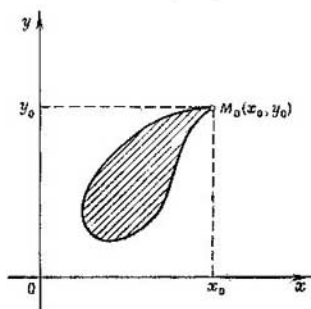


Fig. 5.2

2. Concepto de diferenciabilidad de una función de varias variables. Recordemos que llámase incremento (o incremento total) de la función $u = f(x_1, x_2, \dots, x_m)$ en el punto $M(x_1, x_2, \dots, x_m)$, correspondiente a los incrementos $\Delta x_1, \Delta x_2, \dots, \Delta x_m$ de los argumentos, a la expresión

$$\Delta u = f(x_1 + \Delta x_1, x_2 + \Delta x_2, + \dots, x_m + \Delta x_m) - f(x_1, x_2, \dots, x_m).$$

Definición. La función $u = f(x_1, x_2, \dots, x_m)$ se llama diferenciable en el punto dado $M(x_1, x_2, \dots, x_m)$ si su incremento total en dicho punto puede ser representado en la forma

$$\Delta u = A_1 \Delta x_1 + A_2 \Delta x_2 + \dots + A_m \Delta x_m + \alpha_1 \Delta x_1 + \alpha_2 \Delta x_2 + \dots + \alpha_m \Delta x_m, \quad (5.14)$$

donde A_1, A_2, \dots, A_m son unos números independientes de $\Delta x_1, \Delta x_2, \dots, \Delta x_m$, y $\alpha_1, \alpha_2, \dots, \alpha_m$, unas funciones infinitésimas para $\Delta x_1 \rightarrow 0, \Delta x_2 \rightarrow 0, \dots, \Delta x_m \rightarrow 0$, iguales a cero cuando $\Delta x_1 = \Delta x_2 = \dots = \Delta x_m = 0$. La relación (5.14) se denomina *condición de diferenciabilidad* de la función en el punto dado M . La condición (5.14) de diferenciabilidad de una función puede escribirse también en una forma diferente. Para mostrarlo, examinemos la función $\rho = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_m^2}$ ¹⁾, infinitésima para $\Delta x_1 \rightarrow 0, \Delta x_2 \rightarrow 0, \dots, \Delta x_m \rightarrow 0$, y advirtamos que dicha función se anula sólo cuando $\Delta x_1 = \Delta x_2 = \dots = \Delta x_m = 0$. Cerciorémonos de que la suma $\alpha_1 \Delta x_1 + \alpha_2 \Delta x_2 + \dots + \alpha_m \Delta x_m$ que forma parte del segundo miembro de la relación (5.14), es una función

¹⁾ Geométricamente, esta función es la distancia entre los puntos $M(x_1, x_2, \dots, x_m)$ y $M'(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_m + \Delta x_m)$.

infinitésima de orden más elevado en comparación con ρ . Dicho de otro modo, cerciorémonos de que esta suma es la expresión $o(\rho)$. En efecto, cuando $\rho \neq 0$, se cumple

$$\frac{|\Delta x_i|}{\rho} \leq 1 \text{ y por eso}$$

$$\begin{aligned} |\alpha_1 \Delta x_1 + \alpha_2 \Delta x_2 + \dots + \alpha_m \Delta x_m| &\leq \\ &\leq \left\{ |\alpha_1| \frac{|\Delta x_1|}{\rho} + |\alpha_2| \frac{|\Delta x_2|}{\rho} + \dots + |\alpha_m| \frac{|\Delta x_m|}{\rho} \right\} \rho \leq \\ &\leq \{ |\alpha_1| + |\alpha_2| + \dots + |\alpha_m| \} \rho = o(\rho). \end{aligned}$$

De este modo, la condición (5.14) de diferenciabilidad de una función puede ser escrita en la forma siguiente

$$\Delta u = A_1 \Delta x_1 + A_2 \Delta x_2 + \dots + A_m \Delta x_m + o(\rho). \quad (5.15)$$

La magnitud $o(\rho)$ se considera en este caso igual a cero cuando $\rho = 0$.

Con el fin de demostrar que la condición (5.15) es equivalente a la (5.14), hace falta convencerse de que de la representación (5.15) se deduce, a su vez, la representación (5.14). Para ello representemos $o(\rho)$ en la forma

$$\begin{aligned} o(\rho) &= \frac{o(\rho)}{\rho} \frac{\rho^2}{\rho} = \frac{o(\rho)}{\rho} \frac{\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_m^2}{\rho} \\ &\quad - \left[\frac{o(\rho)}{\rho} \frac{\Delta x_1}{\rho} \right] \Delta x_1 + \left[\frac{o(\rho)}{\rho} \frac{\Delta x_2}{\rho} \right] \Delta x_2 + \dots + \left[\frac{o(\rho)}{\rho} \frac{\Delta x_m}{\rho} \right] \Delta x_m, \end{aligned}$$

considerando que no todos los $\Delta x_1, \Delta x_2, \dots, \Delta x_m$ son nulos *).

Suponiendo $\frac{o(\rho)}{\rho} = \frac{\Delta x_i}{\rho} = \alpha_i$ y considerando que α_i es una función infinitésima para $\rho \rightarrow 0$ y, por lo tanto también para $\Delta x_1 \rightarrow 0, \Delta x_2 \rightarrow 0, \dots, \Delta x_m \rightarrow 0$, llegamos a la representación (5.14).

Así, pues, la condición de diferenciabilidad de una función puede escribirse tanto en la forma (5.14), como en la (5.15).

Si por lo menos uno de los números A_1, A_2, \dots, A_m es distinto de cero, la suma $A_1 \Delta x_1 + A_2 \Delta x_2 + \dots + A_m \Delta x_m$ es la parte *principal, lineal respecto de los incrementos de los argumentos*, del incremento de la función diferenciable. Advertimos que al definir el concepto de función diferenciable, no se excluía la posibilidad de anularse todos los números A_1, A_2, \dots, A_m , por lo cual, si incremento Δu de la función puede representarse en la forma (5.14) o (5.15) para $A_1 = A_2 = \dots = A_m = 0$, la función es diferenciable en el punto dado.

Resulta válido el teorema siguiente.

Teorema 5.9. Si la función $u = f(x_1, x_2, \dots, x_m)$ es diferenciable en el punto $M(x_1, x_2, \dots, x_m)$, en el mismo existen derivadas

*) Si todos los Δx_i son nulos, lo serán también todos los términos en los segundos miembros de las fórmulas (5.14) y (5.15).

parciales respecto de todos los argumentos, con la particularidad de que $\frac{\partial u}{\partial x_i} = A_i$, donde A_i se determinan de la condición (5.14) o (5.15) de diferenciabilidad de una función.

DEMOSTRACION. De la condición (5.14) de diferenciabilidad de la función en el punto $M(x_1, x_2, \dots, x_m)$ se deduce que su incremento parcial $\Delta_{x_i} u$ en este punto es igual a $\Delta_{x_i} u = A_i \Delta x_i + \alpha_i \Delta x_i$. De aquí se desprende que $\frac{\Delta_{x_i} u}{\Delta x_i} = A_i + \alpha_i$ y por eso, dado que $\alpha_i \rightarrow 0$ cuando $\Delta x_i \rightarrow 0$, tenemos $\lim_{\Delta x_i \rightarrow 0} \frac{\Delta_{x_i} u}{\Delta x_i} = \frac{\partial u}{\partial x_i} = A_i$.

Corolario 1. La condición (5.15) de diferenciabilidad de una función en el punto dado M puede escribirse también en la forma siguiente:

$$\Delta u = \frac{\partial u}{\partial x_1} \Delta x_1 + \frac{\partial u}{\partial x_2} \Delta x_2 + \dots + \frac{\partial u}{\partial x_m} \Delta x_m + o(\rho) *). \quad (5.16)$$

Corolario 2. Si la función $u = f(x_1, x_2, \dots, x_m)$ es diferenciable en el punto $M(x_1, x_2, \dots, x_m)$, la representación de su incremento Δu en la forma (5.14) o (5.15) es única. En efecto, los coeficientes A_i de estas representaciones son iguales a las derivadas parciales $\frac{\partial u}{\partial x_i}$ en el punto dado M y, por eso, se definen de un modo único.

Cerciorémonos de la validez de la siguiente propiedad importante de la función diferenciable.

Si la función $u = f(x_1, x_2, \dots, x_m)$ es diferenciable en el punto $M(x_1, x_2, \dots, x_m)$, es también continua en este punto. En efecto, de la condición (5.14) de diferenciabilidad de una función en un punto se deduce que $\lim_{\substack{\Delta x_1 \rightarrow 0 \\ \Delta x_2 \rightarrow 0 \\ \dots \\ \Delta x_m \rightarrow 0}} \Delta u = 0$, lo que precisamente quiere decir

que la función es continua en el punto M (véase p. 1, § 3, fórmula (5.7)).

Para el caso de la función $u = f(x, y)$ de dos variables la condición de diferenciabilidad puede ilustrarse geoméricamente. Introduzcamos el concepto de plano tangente a una superficie en el punto N_0 . El plano π que pasa por el punto N_0 de la superficie se denomina plano tangente en este punto si el ángulo formado por dicho plano y la secante que pasa por el punto N_0 y cualquier punto N_1 de la superficie tiende a cero cuando el punto N_1 tiende a N_0 (fig. 5.3).

Si en el punto N_0 existe un plano tangente, es evidente que una tangente en el punto N_0 a cualquier curva que está dispuesta en la superficie y que pasa por N_0 se dispone en el plano mencionado.

*) Aquí, todas las derivadas parciales $\frac{\partial u}{\partial x_i}$ se toman en el punto dado M .

Cerciorémonos de que de la condición de diferenciabilidad de la función $u = f(x, y)$ en el punto dado $M_0(x_0, y_0)$ se deduce la existencia de un plano tangente a la gráfica S de dicha función en el punto $N_0(x_0, y_0, z_0)$. Pongamos $\Delta x = x - x_0$, $\Delta y = y - y_0$, $\Delta u = u - u_0$, donde $u_0 = f(x_0, y_0)$, $u = f(x, y)$. Obviamente, la

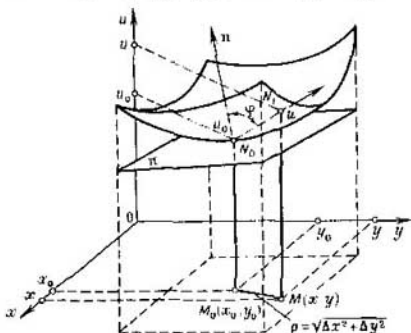


Fig. 5.3

condición (5.14) de diferenciabilidad en el caso que se examina puede ser escrita del modo siguiente:

$$u - u_0 = A(x - x_0) + B(y - y_0) + \alpha \Delta x + \beta \Delta y = \\ = A(x - x_0) + B(y - y_0) + o(\rho),$$

donde A y B son unas constantes iguales a las derivadas parciales $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$ en el punto M_0 , mientras que α y β son funciones infinitésimas para $\Delta x \rightarrow 0$ y $\Delta y \rightarrow 0$; $\rho = \sqrt{\Delta x^2 + \Delta y^2}$.

Veamos la ecuación siguiente:

$$U - u_0 = A(x - x_0) + B(y - y_0).$$

Por el curso de geometría analítica sabemos que esta ecuación define en un sistema cartesiano de coordenadas (x, y, U) un plano π que pasa por el punto $N_0(x_0, y_0, u_0)$ y que cuenta con un vector normal $n = \{A, B, -1\}$ *).

Demostremos que este plano π es tangente en el punto N_0 a la superficie S . Con este fin basta convencerse de que: 1) el plano π pasa por el punto N_0 de la superficie S y 2) el ángulo φ formado por la normal n a dicho plano y cualquier secante N_0N_1 tiende a $\pi/2$,

*) Llámase vector normal de un plano a cualquier vector no nulo n , perpendicular a dicho plano.

cuando el punto N_1 de la superficie S tiende al punto N_0 . La afirmación 1) es evidente. Pasemos a la demostración de la afirmación 2). Calculemos el coseno del ángulo φ utilizando la fórmula conocida para el coseno del ángulo entre dos vectores. Por cuanto las coordenadas del vector \mathbf{n} son iguales a A , B , -1 , mientras que las coordenadas del vector $\overline{N_0N_1}$ de la secante son iguales a $x - x_0$, $y - y_0$, $u - u_0$ (véase fig. 5.3), resulta que

$$\cos \varphi = \frac{A(x - x_0) + B(y - y_0) - (u - u_0)}{\sqrt{A^2 + B^2 + 1} \sqrt{(x - x_0)^2 + (y - y_0)^2 + (u - u_0)^2}}.$$

De la condición de diferenciabilidad de la función $u = f(x, y)$ se deduce que

$$A(x - x_0) + B(y - y_0) - (u - u_0) = o(\rho).$$

Por eso

$$|\cos \varphi| \leq \frac{|o(\rho)|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = \frac{|o(\rho)|}{\rho}.$$

De esta fórmula proviene que $\lim_{\rho \rightarrow 0} \cos \varphi = 0$, es decir, $\lim_{\rho \rightarrow 0} \varphi = \pi/2$.

La afirmación 2) está demostrada.

De este modo, la diferenciabilidad de la función $u = f(x, y)$ en el punto $M_0(x_0, y_0)$ desde el punto de vista geométrico significa la existencia de un plano tangente a la gráfica de la función $u = f(x, y)$ en el punto $N_0(x_0, y_0, u_0)$.

Por cuanto los coeficientes A y B son iguales a las derivadas parciales respectivas calculadas en el punto $M_0(x_0, y_0)$, la ecuación del plano tangente puede ser escrita en la forma

$$U - u_0 = \frac{\partial u}{\partial x}(x - x_0) + \frac{\partial u}{\partial y}(y - y_0). \quad (5.17)$$

El vector normal $\mathbf{n} = \left\{ \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, -1 \right\}$ del plano tangente suele llamarse *normal* a la superficie $u = f(x, y)$ en el punto $N_0(x_0, y_0, u_0)$.

Aclaremos las condiciones suficientes de diferenciabilidad de una función de varias variables.

Teorema 5.10. Si la función $u = f(x_1, x_2, \dots, x_m)$ tiene derivadas parciales respecto de todos los argumentos en cierto entorno del punto $M_0(x_1, x_2, \dots, x_m)$, con la particularidad de que todas estas derivadas parciales son continuas en el propio punto M_0 , la citada función será diferenciable en el punto M_0 .

DEMOSTRACIÓN Con el fin de reducir la anotación, realicemos la demostración para la función de dos variables $u = f(x, y)$. Así, pues, supongamos que ambas derivadas parciales f'_x y f'_y existen en el entorno del punto $M_0(x_0, y_0)$ y son continuas en el mismo. Atri-

buyamos a los argumentos x e y unos incrementos Δx y Δy tan pequeños que el punto $M(x_0 + \Delta x, y_0 + \Delta y)$ no salga de los márgenes del citado entorno del punto M_0 . El incremento total $\Delta u = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ puede escribirse en la forma

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

La expresión $[f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)]$ puede considerarse como incremento de la función $f(x, y_0 + \Delta y)$ de una sola variable x sobre el segmento $[x_0, x_0 + \Delta x]$. Por cuanto la función $u = f(x, y)$ tiene derivadas parciales, la citada función $f(x, y_0 + \Delta y)$ es diferenciable y su derivada respecto de x es una derivada parcial f'_x . Aplicando al incremento mencionado la fórmula de Lagrange, hallaremos tal θ_1 del intervalo $0 < \theta_1 < 1$ que

$$[f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)] = f'_x(x_0 + \theta_1 \Delta x, y_0 + \Delta y) \Delta x.$$

Razonando análogamente, obtenemos que para cierto θ_2 del intervalo $0 < \theta_2 < 1$

$$[f(x_0, y_0 + \Delta y) - f(x_0, y_0)] = f'_y(x_0, y_0 + \theta_2 \Delta y) \Delta y.$$

Por cuanto las derivadas f'_x y f'_y son continuas en el punto M_0 , resulta que

$$\begin{aligned} f'_x(x_0 + \theta_1 \Delta x, y_0 + \Delta y) &= f'_x(x_0, y_0) + \alpha, \\ f'_y(x_0, y_0 + \theta_2 \Delta y) &= f'_y(x_0, y_0) + \beta, \end{aligned}$$

donde α y β son funciones infinitésimas para $\Delta x \rightarrow 0$ y $\Delta y \rightarrow 0$. De aquí, considerando las expresiones aducidas para

$$\begin{aligned} [f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0 + \Delta y)] \text{ y} \\ [f(x_0, y_0 + \Delta y) - f(x_0, y_0)] \end{aligned}$$

y la expresión para Δu , hallaremos

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

Por consiguiente, la función $u = f(x, y)$ es diferenciable en el punto M_0 . Cuando se trata de una función de m variables $u = f(x_1, x_2, \dots, x_m)$, los razonamientos son análogos, mas el incremento total Δu de esta función debe representarse en la forma

$$\begin{aligned} \Delta u &= f(\overset{\circ}{x}_1 + \Delta x_1, \dots, \overset{\circ}{x}_m + \Delta x_m) - f(\overset{\circ}{x}_1, \dots, \overset{\circ}{x}_m) = \\ &= \sum_{k=1}^m [f(\overset{\circ}{x}_1, \dots, \overset{\circ}{x}_{k-1}, \overset{\circ}{x}_k + \Delta x_k, \overset{\circ}{x}_{k+1} + \Delta x_{k+1}, \dots, \\ &\dots, \overset{\circ}{x}_m + \Delta x_m) - f(\overset{\circ}{x}_1, \dots, \overset{\circ}{x}_{k-1}, \overset{\circ}{x}_k, \overset{\circ}{x}_{k+1} + \Delta x_{k+1}, \dots, \overset{\circ}{x}_m + \Delta x_m)]. \end{aligned}$$

El teorema está demostrado.

(14.21), que corresponden a los incrementos elegidos $\Delta t_1, \Delta t_2, \dots, \Delta t_k$ de los argumentos de estas funciones. Por ser las funciones (5.21) diferenciables en el punto $M(\overset{\circ}{t}_1, \overset{\circ}{t}_2, \dots, \overset{\circ}{t}_k)$, los incrementos mencionados Δx_i pueden escribirse en la forma siguiente:

$$\Delta x_i = \frac{\partial x_i}{\partial t_1} \Delta t_1 + \frac{\partial x_i}{\partial t_2} \Delta t_2 + \dots + \frac{\partial x_i}{\partial t_k} \Delta t_k + o(\rho), \quad i = 1, 2, \dots, m, \quad (5.24)$$

donde las derivadas parciales $\frac{\partial x_i}{\partial t_1}, \frac{\partial x_i}{\partial t_2}, \dots, \frac{\partial x_i}{\partial t_k}$ se toman en el punto M , y $\rho = \sqrt{(\Delta t_1)^2 + (\Delta t_2)^2 + \dots + (\Delta t_k)^2}$.

Hemos de convencernos de que al introducir la expresión (5.24) en el segundo miembro de (5.23), el incremento Δu puede ser reducido a la forma

$$\Delta u = A_1 \Delta t_1 + A_2 \Delta t_2 + \dots + A_k \Delta t_k + o(\rho), \quad (5.25)$$

donde

$$A_i = \frac{\partial u}{\partial x_1} \frac{\partial x_1}{\partial t_1} + \frac{\partial u}{\partial x_2} \frac{\partial x_2}{\partial t_1} + \dots + \frac{\partial u}{\partial x_m} \frac{\partial x_m}{\partial t_1}, \quad i = 1, 2, \dots, k. \quad (5.26)$$

Con esto el teorema quedará demostrado, pues la fórmula (5.25) establece el hecho de que una función compuesta es diferenciable y la expresión (5.26) es una derivada parcial de la citada función compuesta (véase el teorema 5.9).

Al introducir en el segundo miembro de (5.23) las expresiones (5.24), obtenemos, además del grupo de sumandos $A_1 \Delta t_1 + A_2 \Delta t_2 + \dots + A_k \Delta t_k$, otros grupos de sumandos. Nos hace falta convencernos de que todos los demás grupos de sumandos representan la magnitud $o(\rho)$. Esto se deduce de los razonamientos siguientes:

1°. Todas las derivadas parciales $\frac{\partial u}{\partial x_i}$ en la fórmula (5.23) se toman en el punto N , es decir, son unas constantes que, siendo multiplicadas por $o(\rho)$, dan de nuevo la magnitud $o(\rho)$.

2°. Todos los Δx_i ($i = 1, 2, \dots, m$) satisfacen la inecuación $|\Delta x_i| \leq \text{const } \rho$. Esto se deduce inmediatamente de las fórmulas (5.24).

3°. Todas las α_i de la fórmula (5.23) son funciones infinitésimas para $\rho \rightarrow 0$. En efecto, todas las α_i son infinitamente pequeñas para $\Delta x_1 \rightarrow 0, \Delta x_2 \rightarrow 0, \dots, \Delta x_m \rightarrow 0$. Pero, todas las funciones (5.21) son diferenciables y, por tanto, continuas en el punto M , razón por la cual $\Delta x_1, \Delta x_2, \dots, \Delta x_m$ tienden a cero cuando $\rho \rightarrow 0$.

4°. Todo producto $\alpha_i \Delta x_i$ representa la magnitud $o(\rho)$. Esto se desprende directamente de los pp. 2° y 3°. El teorema está demostrado.

OBSERVACIÓN. Analicemos un caso particular, de importancia, cuando las funciones (5.21) dependen de un argumento t . En este caso tenemos una función de una sola variable t : $u = f(x_1, x_2, \dots, x_m)$, donde $x_i = \varphi_i(t)$. La derivada $\frac{du}{dt}$ de esta función compuesta se define mediante la fórmula siguiente

$$\frac{du}{dt} = \frac{\partial u}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial u}{\partial x_2} \frac{dx_2}{dt} + \dots + \frac{\partial u}{\partial x_m} \frac{dx_m}{dt}. \quad (5.27)$$

Apliquemos la fórmula (5.27) para demostrar el teorema de Euler de las funciones homogéneas.

La función $u = f(x_1, x_2, \dots, x_m)$ definida sobre el conjunto $\{M\}$ se denomina función homogénea de grado p en este conjunto, si para todo punto $M(x_1, x_2, \dots, x_m)$ del conjunto $\{M\}$ y para todo número t , para el cual un punto $N(tx_1, tx_2, \dots, tx_m)$ pertenece al conjunto $\{M\}$, se verifica la igualdad

$$f(tx_1, tx_2, \dots, tx_m) = t^p f(x_1, x_2, \dots, x_m). \quad (5.28)$$

Teorema 5.12 (teorema de Euler sobre las funciones homogéneas). Si $u = f(x_1, x_2, \dots, x_m)$ es en cierto dominio $\{M\}$ una función homogénea diferenciable de grado p , en cada punto $M(x_1, x_2, \dots, x_m)$ del dominio $\{M\}$ se verifica la igualdad

$$\frac{\partial u}{\partial x_1} x_1 + \frac{\partial u}{\partial x_2} x_2 + \dots + \frac{\partial u}{\partial x_m} x_m = pu. \quad (5.29)$$

DEMOSTRACIÓN. Sea $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ un punto arbitrario del dominio $\{M\}$. Estudiemos una función compuesta $u = f(x_1, x_2, \dots, x_m)$, donde $x_i = tx_i$ ($i = 1, 2, \dots, m$), es decir, una función $u = f(tx_1, tx_2, \dots, tx_m)$. Por cuanto las funciones $x_i = tx_i$ son diferenciables para $t = 1$ y la función $u = f(x_1, x_2, \dots, x_m)$ lo es en el punto correspondiente M_0 , entonces, de acuerdo con el teorema 5.11 y con la observación a este teorema, podemos calcular la derivada $\frac{du}{dt}$ de esta función compuesta en el punto $t = 1$ según la fórmula (5.27). Puesto que $\frac{dx_i}{dt} = \overset{\circ}{x}_i$, se tiene

$$\frac{du}{dt} \Big|_{t=1} = \frac{\partial u}{\partial x_1} \overset{\circ}{x}_1 + \frac{\partial u}{\partial x_2} \overset{\circ}{x}_2 + \dots + \frac{\partial u}{\partial x_m} \overset{\circ}{x}_m, \quad (5.30)$$

donde las derivadas $\frac{\partial u}{\partial x_i}$ se toman en el punto M_0 . Por otra parte en virtud de (5.28), la función compuesta sujeta a la consideración puede ser representada en la forma

$$u = f(tx_1, tx_2, \dots, tx_m) = t^p f(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m). \quad (5.31)$$

De (5.31) se desprende que $\frac{du}{dt} = pt^{p-1}f(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$, es decir,

$$\left. \frac{du}{dt} \right|_{t-1} = pf(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m) = pu. \quad (5.32)$$

Cotejando (5.30) y (5.32), obtendremos la relación (5.29) para el punto M_0 . Por cuanto el punto M_0 es arbitrario en el dominio $\{M\}$, el teorema queda demostrado.

5. Forma invariante de la primera diferencial. En el p. 3 se ha introducido el concepto de primera diferencial du de una función de varias variables y se ha establecido que cuando los argumentos x_1, x_2, \dots, x_m son variables independientes, la diferencial du puede ser representada en la forma

$$du = \frac{\partial u}{\partial x_1} dx_1 + \frac{\partial u}{\partial x_2} dx_2 + \dots + \frac{\partial u}{\partial x_m} dx_m. \quad (5.20)$$

En este punto se demostrará que la fórmula (5.20) es universal y es válida también para el caso en que los argumentos x_1, x_2, \dots, x_m son de por sí funciones diferenciables de las nuevas variables t_1, t_2, \dots, t_k . La propiedad citada de la primera diferencial se suele denominar propiedad *de invariación de su forma*.

Supongamos que los argumentos x_1, x_2, \dots, x_m de la función $u = f(x_1, x_2, \dots, x_m)$ son funciones diferenciables $x_i = \varphi_i(t_1, t_2, \dots, t_k)$ en un punto $A(t_1, t_2, \dots, t_k)$, y la propia función $u = f(x_1, x_2, \dots, x_m)$ lo es en un punto $B(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$, donde $x_i = \varphi_i(t_1, t_2, \dots, t_k)$. En tal caso podemos considerar u como una función compuesta de los argumentos t_1, t_2, \dots, t_k , la cual es, en virtud del teorema 5.11, diferenciable en el punto A . Por eso, la diferencial du de esta función compuesta puede ser representada en la forma

$$du = \frac{\partial u}{\partial t_1} dt_1 + \frac{\partial u}{\partial t_2} dt_2 + \dots + \frac{\partial u}{\partial t_k} dt_k, \quad (5.33)$$

donde $\frac{\partial u}{\partial t_i}$ se definen a partir de las relaciones (5.22). Introduciendo $\frac{\partial u}{\partial t_i}$ de (5.22) en (5.33) y reuniendo los coeficientes de $\frac{\partial u}{\partial x_i}$, obtendremos

$$du = \frac{\partial u}{\partial x_1} \left(\frac{\partial x_1}{\partial t_1} dt_1 + \frac{\partial x_1}{\partial t_2} dt_2 + \dots + \frac{\partial x_1}{\partial t_k} dt_k \right) + \dots + \frac{\partial u}{\partial x_m} \left(\frac{\partial x_m}{\partial t_1} dt_1 + \frac{\partial x_m}{\partial t_2} dt_2 + \dots + \frac{\partial x_m}{\partial t_k} dt_k \right).$$

Nos queda señalar que en la última relación el coeficiente de $\frac{\partial u}{\partial x_i}$ es igual a la diferencial dx_i de la función $x_i = \varphi_i(t_1, t_2, \dots, t_k)$.

Obtendremos para la diferencial du de la función compuesta la fórmula (5.20), en la cual las diferenciales dx_i serán diferenciales de las funciones $x_i = \varphi_i(t_1, t_2, \dots, t_k)$. La invariación de la forma de la primera diferencial queda establecida.

La propiedad de invariación de la forma de la primera diferencial permite establecer las siguientes *reglas de diferenciación*. Sean u y v unas funciones diferenciables de ciertas variables. Entonces

$$d(cu) = c du \quad (c = \text{const}),$$

$$d(u \pm v) = du \pm dv, \quad d(uv) = u dv + v du, \quad d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}.$$

(En la última de las fórmulas escritas v no se anula.)

Demostremos, por ejemplo, la validez de la tercera de las fórmulas citadas. Veamos una función $w = uv$ de dos variables u y v . La diferencial de esta función dw es

$$dw = \frac{\partial w}{\partial u} du + \frac{\partial w}{\partial v} dv.$$

Por cuanto $\frac{\partial w}{\partial u} = v$ y $\frac{\partial w}{\partial v} = u$, tenemos $dw = u dv + v du$. Por ser invariante la forma de la primera diferencial, la expresión $u dv + v du$ expresa la diferencial de la función w también en el caso en que u y v son de por sí funciones diferenciables de ciertas variables.

6. Derivada direccional. Gradiente. Supongamos que una función $u = f(x, y, z)$ de tres variables x, y y z viene dada en cierto entorno del punto $M_0(x_0, y_0, z_0)$. Examinemos una dirección definida por el vector unidad \mathbf{a} , cuyas coordenadas son $\cos \alpha, \cos \beta, \cos \gamma$ (*). Tracemos por el punto M_0 el eje \mathbf{l} , cuyo sentido coincide con el del vector \mathbf{a} , elijamos en dicho eje un punto arbitrario $M(x, y, z)$ y denotemos con l la magnitud del segmento orientado M_0M del mencionado eje (**). Por el curso de geometría analítica se conoce que las coordenadas x, y, z del punto M se definen mediante las igualdades

$$x = x_0 + l \cos \alpha, \quad y = y_0 + l \cos \beta, \quad z = z_0 + l \cos \gamma. \quad (5.34)$$

En el eje \mathbf{l} la función $u = f(x, y, z)$ es, evidentemente, una función compuesta de una sola variable l . Si esta función tiene en el punto $l = 0$ una derivada respecto de la variable l , la citada derivada lleva el nombre de derivada según la dirección \mathbf{l} de la función $u = f(x, y, z)$ en el punto M_0 y se designa por el símbolo $\frac{\partial u}{\partial l}$. De con-

*) Por el curso de geometría analítica se sabe que si el vector unidad \mathbf{a} forma con los ejes coordenados los ángulos α, β, γ , las coordenadas de este vector son iguales a $\cos \alpha, \cos \beta, \cos \gamma$.

**) Llámase magnitud l del segmento orientado M_0M del eje \mathbf{l} a un número, igual a su longitud tomada con el signo más, si la dirección de este segmento coincide con la del eje \mathbf{l} , y tomada con el signo menos, si la dirección de este segmento es opuesta a la del eje \mathbf{l} .

formidad con la observación al teorema 5.11, si la función $u = f(x, y, z)$ es diferenciable en el punto M_0 , la derivada $\frac{\partial u}{\partial l}$ puede calcularse según la fórmula (5.27), en la que el argumento l ha de sustituirse por l . De este modo,

$$\frac{\partial u}{\partial l} = \frac{\partial u}{\partial x} \frac{dx}{dl} + \frac{\partial u}{\partial y} \frac{dy}{dl} + \frac{\partial u}{\partial z} \frac{dz}{dl}.$$

Por cuanto $\frac{dx}{dl} = \cos \alpha$, $\frac{dy}{dl} = \cos \beta$, $\frac{dz}{dl} = \cos \gamma$, de la última fórmula hallamos

$$\frac{\partial u}{\partial l} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma. \quad (5.35)$$

Introduzcamos el concepto de *gradiente* de la función $u = f(x, y, z)$ diferenciable en el punto $M_0(x_0, y_0, z_0)$.

Llámanse *gradiente de la función* $u = f(x, y, z)$ en el punto M_0 a un vector que se denota con el símbolo $\text{grad } u$ y que tiene coordenadas iguales, respectivamente, a las derivadas $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial u}{\partial z}$, tomadas en el punto M_0 . De este modo,

$$\text{grad } u = \left\{ \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z} \right\}. \quad (5.36)$$

Utilizando el concepto de gradiente de una función y teniendo presente que el vector \mathbf{a} que define la dirección del eje \mathbf{l} tiene por coordenadas $\cos \alpha$, $\cos \beta$, $\cos \gamma$, representemos la expresión (5.35) para la derivada $\frac{\partial u}{\partial l}$ según la dirección \mathbf{l} en forma del producto escalar de los vectores $\text{grad } u$ y \mathbf{a} :

$$\frac{\partial u}{\partial l} = \mathbf{a} \text{ grad } u^*). \quad (5.37)$$

Mostremos que el gradiente de la función $u = f(x, y, z)$ en el punto M_0 caracteriza el sentido y la magnitud de crecimiento máximo de esta función en el punto M_0 . A saber: cerciorémonos de que la derivada de la función u en el punto M_0 según la dirección definida por el gradiente de dicha función en el punto mencionado tiene valor máximo en comparación con la derivada según cualquier otra dirección en el punto M_0 , y el valor de la citada derivada es igual a $|\text{grad } u|$, es decir, a la longitud del vector $\text{grad } u$. Reescribamos la fórmula (5.37) en la forma

$$\frac{\partial u}{\partial l} = |\mathbf{a}| |\text{grad } u| \cos \varphi,$$

*) Recordemos que el producto escalar de dos vectores, definido como producto de módulos (longitudes) de los vectores por el coseno del ángulo comprendido entre ellos (cuando los vectores vienen dados por las coordenadas) es igual a la suma de productos de las coordenadas homónimas de estos vectores.

donde φ es el ángulo entre los vectores \mathbf{a} y $\text{grad } u$. Por cuanto

$$|\mathbf{a}| = 1, \text{ tenemos } \frac{\partial u}{\partial l} = |\text{grad } u| \cos \varphi.$$

De la última fórmula se deduce que la derivada direccional tendrá el valor máximo $\left(\frac{\partial u}{\partial l}\right)_{\text{máx}}$ para $\cos \varphi = 1$, es decir, cuando la dirección del vector \mathbf{a} coincide con la del $\text{grad } u$, con la particularidad de que $\left(\frac{\partial u}{\partial l}\right)_{\text{máx}} = |\text{grad } u|$.

Con el fin de aclarar el significado geométrico del vector $\text{grad } u$ introduzcamos el concepto de *superficie de nivel* de una función $u = f(x, y, z)$.

Llamemos superficie de nivel de la función $u = f(x, y, z)$ a toda superficie, sobre la cual la función $u = f(x, y, z)$ conserva constante su valor: $f(x, y, z) = c = \text{const.}$

No es difícil convencerse de que el vector $\text{grad } u$ en un punto dado $M_0(x_0, y_0, z_0)$ es ortogonal con relación a una superficie de nivel de la función $u = f(x, y, z)$, que pasa por el punto dado M_0 .

OBSERVACION. En caso de una función $u = f(x, y)$ de dos variables x y y , el vector unidad \mathbf{a} que define la dirección en el punto M_0 tiene por sus coordenadas $\cos \alpha$ y $\sin \alpha$. Por eso, en este caso la fórmula (5.35) tendrá la forma

$$\frac{\partial u}{\partial l} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \sin \alpha.$$

Notemos que en caso de una función de dos variables el gradiente de la función diferenciable $u(x, y)$ se define como un vector que tiene por coordenadas $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$. La fórmula (5.37) será, evidentemente, válida también en caso de dos variables. Para una función $u = f(x_1, x_2, \dots, x_m)$ de m variables x_1, x_2, \dots, x_m , la derivada direccional y el gradiente se determinan de un modo análogo. A saber: la derivada $\frac{\partial u}{\partial l}$ en el punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ según la dirección \mathbf{l} , que se define mediante el vector unidad $\mathbf{a} = \{\cos \alpha_1, \cos \alpha_2, \dots, \cos \alpha_m\}$ *) se define como derivada respecto de l de la función compuesta $u = f(x_1, x_2, \dots, x_m)$, donde $x_1 = \overset{\circ}{x}_1 + l \cos \alpha_1$, $x_2 = \overset{\circ}{x}_2 + l \cos \alpha_2$, \dots , $x_m = \overset{\circ}{x}_m + l \cos \alpha_m$. Si $u = f(x_1, x_2, \dots, x_m)$ es una función diferenciable, para la derivada direccional tiene lugar la fórmula siguiente

$$\frac{\partial u}{\partial l} = \frac{\partial u}{\partial x_1} \cos \alpha_1 + \frac{\partial u}{\partial x_2} \cos \alpha_2 + \dots + \frac{\partial u}{\partial x_m} \cos \alpha_m.$$

*) En geometría analítica del espacio euclídeo m -dimensional el vector unidad \mathbf{a} se define como vector con las coordenadas $\cos \alpha_1, \cos \alpha_2, \dots, \cos \alpha_m$, donde $\cos^2 \alpha_1 + \cos^2 \alpha_2 + \dots + \cos^2 \alpha_m = 1$.

Llámanse gradiente de una función en el punto dado $M_0 (\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ a un vector que se denota con el símbolo $\text{grad } u$ y que tiene por coordenadas $\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \dots, \frac{\partial u}{\partial x_m}$, con la particularidad de que las derivadas mencionadas se toman en el punto M_0 . Para la derivada direccional de una función diferenciable $u = f(x_1, x_2, \dots, x_m)$ resulta válida la fórmula (5.37).

§ 5. Derivadas parciales y diferenciales de orden superior

1. **Derivadas parciales de orden superior.** Supongamos que una derivada parcial $\frac{\partial u}{\partial x_i}$ respecto del argumento x_i de una función $u = f(x_1, x_2, \dots, x_m)$, definida en un dominio determinado $\{M\}$, existe en cada punto del dominio $\{M\}$. En este caso la citada derivada parcial es una función de las variables x_1, x_2, \dots, x_m , también definida en el dominio $\{M\}$. Puede suceder que esta función $\frac{\partial u}{\partial x_i}$ tenga la derivada parcial respecto del argumento x_k en cierto punto M del dominio $\{M\}$. Entonces, dicha derivada parcial respecto del argumento x_k se denomina segunda derivada parcial o derivada parcial de segundo orden de la función $u = f(x_1, x_2, \dots, x_m)$ en el punto M , primero respecto del argumento x_i , luego, respecto del argumento x_k y se denota con uno de los siguientes símbolos:

$$\frac{\partial^2 u}{\partial x_k \partial x_i}, f_{x_i x_k}^{(2)}, u_{x_i x_k}^{(2)}.$$

Además, si $i \neq k$, la derivada parcial $\frac{\partial^2 u}{\partial x_k \partial x_i}$ se llama *derivada mixta* de segundo orden. Introducido el concepto de segunda derivada parcial, puede introducirse sucesivamente la noción de tercera derivada parcial, luego, de cuarta derivada, etc. Si suponemos que ya tenemos introducida la noción de $(n-1)$ -ésima derivada parcial de la función $u = f(x_1, x_2, \dots, x_m)$ respecto de los argumentos $x_{i_1}, x_{i_2}, \dots, x_{i_{n-1}}$ (algunos o todos los números de la cual pueden coincidir) y que esta $(n-1)$ -ésima derivada tiene en el punto M derivada parcial respecto del argumento x_{i_n} , la citada derivada parcial recibe el nombre de n -ésima derivada parcial (o derivada parcial de n -ésimo orden) de la función $u = f(x_1, x_2, \dots, x_m)$ en el punto M respecto de los argumentos $x_{i_1}, x_{i_2}, \dots, x_{i_{n-1}}, x_{i_n}$. De este modo introducimos el concepto de n -ésima derivada parcial *por el método inductivo*, pasando de la primera derivada parcial a las posteriores. La relación que define la n -ésima derivada parcial respecto de

los argumentos $x_{i_1}, x_{i_2}, \dots, x_{i_{n-1}}, x_{i_n}$, tiene la forma

$$\frac{\partial^n u}{\partial x_{i_n} \partial x_{i_{n-1}} \dots \partial x_{i_2} \partial x_{i_1}} = \frac{\partial}{\partial x_{i_n}} \left(\frac{\partial^{n-1} u}{\partial x_{i_{n-1}} \dots \partial x_{i_2} \partial x_{i_1}} \right).$$

Si no todos los índices i_1, i_2, \dots, i_n coinciden entre sí, la derivada parcial $\frac{\partial^n u}{\partial x_{i_n} \dots \partial x_{i_2} \partial x_{i_1}}$ se llama derivada parcial *mixta*

(*cruzada*) de n -ésimo orden. Por cuanto la derivada parcial respecto del argumento x_i se define como derivada ordinaria de una función de variable única x_i , siendo fijos los valores de las demás variables, la metodología de cálculo de las derivadas parciales de orden superior exige que se sepa calcular sólo las derivadas ordinarias de primer orden. A título de ejemplo, calculemos las derivadas parciales de segundo orden de una función $u = \operatorname{arctg} \frac{x}{y}$. Tenemos

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{y}{x^2 + y^2}, & \frac{\partial u}{\partial y} &= -\frac{x}{x^2 + y^2}, \\ \frac{\partial^2 u}{\partial x^2} &= -\frac{2xy}{(x^2 + y^2)^2}, & \frac{\partial^2 u}{\partial x \partial y} &= \frac{x^2 - y^2}{(x^2 + y^2)^2}, \\ \frac{\partial^2 u}{\partial y \partial x} &= \frac{x^2 - y^2}{(x^2 + y^2)^2}, & \frac{\partial^2 u}{\partial y^2} &= \frac{2xy}{(x^2 + y^2)^2}. \end{aligned}$$

En el ejemplo examinado las derivadas parciales mixtas $\frac{\partial^2 u}{\partial y \partial x}$ y $\frac{\partial^2 u}{\partial x \partial y}$ son iguales. Por regla general, los valores de las derivadas mixtas dependen del orden en que se realiza la diferenciación sucesiva. Cerciorémonos, por ejemplo, de que las derivadas parciales mixtas $\frac{\partial^2 u}{\partial y \partial x}$ y $\frac{\partial^2 u}{\partial x \partial y}$ de la función

$$u = \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2} & \text{para } x^2 + y^2 \neq 0, \\ 0 & \text{para } x^2 + y^2 = 0 \end{cases}$$

existen en el punto $(0, 0)$, pero no son iguales. En efecto,

$$\frac{\partial u}{\partial x} = \begin{cases} \frac{y(x^4 - y^4 + 4x^2y^2)}{(x^2 + y^2)^2} & \text{para } x^2 + y^2 \neq 0, \\ 0 & \text{para } x^2 + y^2 = 0. \end{cases}$$

Por eso

$$\frac{\partial^2 u}{\partial y \partial x} \Big|_{x=0, y=0} = \lim_{y \rightarrow 0} \frac{\frac{\partial u}{\partial x} \Big|_{x=0, y \neq 0} - \frac{\partial u}{\partial x} \Big|_{x=0, y=0}}{y} = -1.$$

Al realizar cálculos análogos, obtendremos $\frac{\partial^2 u}{\partial x \partial y} \Big|_{x=0, y=0} = 1$. De este modo, en el punto $(0, 0)$ $\frac{\partial^2 u}{\partial x \partial y} \neq \frac{\partial^2 u}{\partial y \partial x}$. Aclaremos las condiciones suficientes de independencia de los valores de las derivadas mixtas respecto del orden en que se efectúa la diferenciación sucesiva. Introduzcamos previamente el concepto de función de varias variables n veces diferenciable. Una función $u = f(x_1, x_2, \dots, x_m)$ se denomina n veces diferenciable en un punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ si todas las derivadas parciales de $(n-1)$ -ésimo orden de esta función son funciones diferenciables en el punto M_0 . Hagamos resaltar la afirmación siguiente. Para que una función $u = f(x_1, x_2, \dots, x_m)$ sea n veces diferenciable en un punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$, es suficiente que todas las derivadas parciales cuyas de n -ésimo orden sean continuas en el punto M_0 . La validez de esta afirmación se deduce de la definición de diferenciable de una función y del teorema 5.10 de las condiciones suficientes de diferenciable.

Teorema 5.13. Sea $u = f(x, y)$ dos veces diferenciable en un punto $M_0(x_0, y_0)$. Entonces, en dicho punto las derivadas parciales $f_{xy}^{(2)}$ y $f_{yx}^{(2)}$ son iguales.

DEMOSTRACION. Por cuanto la función $u = f(x, y)$ es dos veces diferenciable en el punto $M_0(x_0, y_0)$, las derivadas parciales f'_x y f'_y están definidas en cierto entorno del punto M_0 y son funciones diferenciables en este punto. Veamos la expresión

$$\Phi = f(x_0 + h, y_0 + h) - f(x_0 + h, y_0) - f(x_0, y_0 + h) + f(x_0, y_0), \quad (5.38)$$

donde h es un número cualquiera tan pequeño que el punto $M(x_0 + h, y_0 + h)$ se encuentra en el entorno mencionado del punto M_0 . La expresión Φ puede considerarse como el incremento $\Delta\varphi = \varphi(x_0 + h) - \varphi(x_0)$ de la función $\varphi(x) = f(x, y_0 + h) - f(x, y_0)$, de una variable x , que es diferenciable sobre el segmento $[x_0, x_0 + h]$. Por eso, de acuerdo con la fórmula de Lagrange, al designar con θ un número del intervalo $0 < \theta < 1$, podemos escribir

$$\begin{aligned} \Phi = \Delta\varphi &= \varphi'(x_0 + \theta h) h = \{f'_x(x_0 + \theta h, y_0 + h) - \\ &- f'_x(x_0 + \theta h, y_0)\} h = \{[f''_{xx}(x_0 + \theta h, y_0 + h) - \\ &- f''_{xx}(x_0, y_0)] - [f''_{xy}(x_0 + \theta h, y_0) - f''_{xy}(x_0, y_0)]\} h. \end{aligned} \quad (5.39)$$

Por cuanto la derivada parcial f''_{xx} es diferenciable en el punto M_0 , resulta que

$$\begin{aligned} [f''_{xx}(x_0 + \theta h, y_0 + h) - f''_{xx}(x_0, y_0)] &= \\ &= f''_{xx}(x_0, y_0) \theta h + f''_{xy}(x_0, y_0) h + \alpha_1 \theta h + \beta_1 h, \\ [f''_{xy}(x_0 + \theta h, y_0) - f''_{xy}(x_0, y_0)] &= f''_{xy}(x_0, y_0) \theta h + \alpha_2 \theta h, \end{aligned}$$

donde α_1 , β_1 y α_2 son funciones infinitésimas para $h \rightarrow 0$. Al introducir las expresiones halladas para $[f'_x(x_0 + \theta h, y_0 + h) - f'_x(x_0, y_0)]$ y $[f'_x(x_0 + \theta h, y_0) - f'_x(x_0, y_0)]$ en la fórmula (14.39), obtendremos

$$\Phi = [f''_{xy}(x_0, y_0) + \alpha] h^2, \quad (5.40)$$

donde $\alpha = \alpha_1\theta + \beta_1 - \alpha_2\theta$ es una función infinitésima para $h \rightarrow 0$. Por otra parte, la expresión Φ que se define mediante (14.38), puede considerarse como el incremento $\Delta\psi = \psi(y_0 + h) - \psi(y_0)$ de la función $\psi(y) = f(x_0 + h, y) - f(x_0, y)$, diferenciable en el segmento $[y_0, y_0 + h]$. Aplicando la fórmula de Lagrange y considerando la diferenciabilidad de la derivada parcial f'_y en el punto M_0 , obtenemos por analogía con el caso antecedente la siguiente expresión para Φ :

$$\Phi = [f''_{yx}(x_0, y_0) + \beta] h^2, \quad (5.41)$$

donde β es una función infinitésima para $h \rightarrow 0$. Al igualar los segundos miembros de las relaciones (5.40) y (5.41) y simplificando ambos miembros de la igualdad obtenida por h^2 , hallaremos que $f''_{xy}(x_0, y_0) + \alpha = f''_{yx}(x_0, y_0) + \beta$. Por cuanto α y β son funciones infinitésimas para $h \rightarrow 0$, de la última igualdad proviene que $f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0)$. El teorema está demostrado.

OBSERVACIÓN El teorema 5.13 afirma que en el punto dado $M_0(x_0, y_0)$ tiene lugar la igualdad $f''_{xy} = f''_{yx}$ si en este punto son diferenciables f'_x y f'_y . Dado que f'_x y f'_y son diferenciables en M_0 se deduce que en este punto existen todas las derivadas parciales de segundo orden. Sin embargo, la igualdad $f''_{xy} = f''_{yx}$ tiene lugar también si existen sólo las derivadas f''_{xy} y f''_{yx} , mas con la exigencia adicional de continuidad de las citadas derivadas en el punto en consideración. Es válida la siguiente afirmación.

Supongamos que en cierto entorno del punto $M_0(x_0, y_0)$ la función $u = f(x, y)$ tiene derivadas parciales f'_x , f'_y , f''_{xy} , f''_{yx} . Además sean las derivadas f''_{xy} y f''_{yx} continuas en el punto M_0 . En este caso $f''_{xy} = f''_{yx}$ en dicho punto.

Para demostrar la afirmación, hagamos uso de la expresión de Φ , definida mediante la relación (5.38). De (5.39) se deduce que Φ es la diferencia de valores de la función $f'_x(x, y)$ en los puntos $(x_0 + \theta h, y_0 + h)$ y $(x_0 + \theta h, y_0)$, multiplicada por h . Al aplicar a dicha diferencia la fórmula de Lagrange de incrementos finitos respecto de la variable y sobre el segmento $[y_0, y_0 + h]$, obtendremos

$$\Phi = f''_{xy}(x_0 + \theta h, y_0 + \theta_1 h) h^2, \text{ donde } 0 < \theta_1 < 1.$$

Por ser f''_{xy} continua en el punto $M_0(x_0, y_0)$, de la última igualdad obtenemos

$$\Phi = [f''_{xy}(x_0, y_0) + \alpha(h)] h^2,$$

donde $\alpha(h) \rightarrow 0$ cuando $h \rightarrow 0$.

Por otro lado, la misma magnitud Φ es la diferencia de valores de la función $f'_y(x, y)$ en los puntos $(x_0 + h, y_0 + \theta_2 h)$ y $(x_0, y_0 + \theta_2 h)$, multiplicada por

h. Al aplicar a dicha diferencia la fórmula de Lagrange de incrementos finitos respecto de la variable x sobre el segmento $[x_0, x_0 + h]$ y al tener en cuenta la continuidad de $f_{yx}^{(2)}$ en el punto $M_0(x_0, y_0)$, obtendremos

$$\Phi = |f_{yx}^{(2)}(x_0, y_0) + \beta(h)| h^2,$$

donde $\beta(h) \rightarrow 0$ cuando $h \rightarrow 0$.

Igualando entre sí las últimas dos expresiones para Φ y razonando del mismo modo que al final de la demostración del teorema 5.13, nos convencemos de que la igualdad requerida

$$f_{xy}^{(2)}(x_0, y_0) = f_{yx}^{(2)}(x_0, y_0)$$

es válida.

Demostremos ahora el teorema de independencia del valor de cualquier n -ésima derivada parcial mixta respecto del orden en que se realiza la diferenciación sucesiva.

Teorema 5.14. *Supongamos que una función $u = f(x_1, x_2, \dots, x_m)$ es n veces diferenciable en el punto $M_0(x_1, x_2, \dots, x_m)$. Entonces, en este punto el valor de cualquier n -ésima derivada parcial mixta no depende del orden en que se realiza la diferenciación sucesiva.*

DEMOSTRACIÓN. Es evidente que basta probar la independencia del valor de cualquier n -ésima derivada mixta respecto del orden en que se realizan dos diferenciaciones sucesivas. Dicho de otro modo, es suficiente probar la igualdad

$$\frac{\partial^n u}{\partial x_{i_1} \dots \partial x_{i_{h+1}} \partial x_{i_h} \dots \partial x_{i_1}} = \frac{\partial^n u}{\partial x_{i_1} \dots \partial x_{i_h} \partial x_{i_{h+1}} \dots \partial x_{i_1}}. \quad (5.42)$$

Examinemos una función $\frac{\partial^{h-1} u}{\partial x_{i_{h-1}} \dots \partial x_{i_1}}$. Esta es una función dos veces diferenciable de las variables x_{i_h} y $x_{i_{h+1}}$. Por eso, en virtud del teorema 5.13,

$$\frac{\partial^{h+1} u}{\partial x_{i_{h+1}} \partial x_{i_h} \partial x_{i_{h-1}} \dots \partial x_{i_1}} = \frac{\partial^{h+1} u}{\partial x_{i_h} \partial x_{i_{h+1}} \partial x_{i_{h-1}} \dots \partial x_{i_1}}.$$

De aquí precisamente se desprende la validez de la igualdad (5.42). El teorema está demostrado.

Advirtamos que en el caso de la función $u = f(x_1, x_2, \dots, x_m)$, n veces diferenciable en el punto M_0 , cualquier derivada parcial suya de n -ésimo orden puede escribirse en la forma

$$\frac{\partial^n u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_m^{\alpha_m}},$$

donde $\alpha_1, \alpha_2, \dots, \alpha_m$ son números enteros que satisfacen las condiciones $0 \leq \alpha_i \leq n$, $\alpha_1 + \alpha_2 + \dots + \alpha_m = n$.

2. Diferenciales de orden superior. Para denotar las diferenciales de los argumentos de la función $u = f(x_1, x_2, \dots, x_m)$ y la diferencial de la misma función más arriba hemos empleado los símbolos dx_1, dx_2, \dots, dx_m y du , respectivamente.

Ahora hemos de utilizar también otros símbolos para denotar las diferenciales de los argumentos de la citada función y la diferencial

de la propia función. En particular, las diferenciales de los argumentos de la función $u = f(x_1, x_2, \dots, x_m)$ y la diferencial de la propia función se denotarán con los símbolos $\delta x_1, \delta x_2, \dots, \delta x_m$ y δu , respectivamente. Utilizando estas designaciones la expresión para la primera diferencial de dicha función (5.20), invariante por su forma (véase p. 5, § 4) se representa de la manera siguiente:

$$\delta u = \frac{\partial u}{\partial x_1} \delta x_1 + \frac{\partial u}{\partial x_2} \delta x_2 + \dots + \frac{\partial u}{\partial x_m} \delta x_m.$$

Volviendo a las designaciones anteriores, veamos la expresión (5.20) para la primera diferencial de la función $u = f(x_1, x_2, \dots, x_m)$ diferenciable en el punto dado $M(x_1, x_2, \dots, x_m)$:

$$du = \frac{\partial u}{\partial x_1} dx_1 + \frac{\partial u}{\partial x_2} dx_2 + \dots + \frac{\partial u}{\partial x_m} dx_m. \quad (5.20)$$

Supongamos que la magnitud que forma parte del segundo miembro de (5.20), es una función de los argumentos x_1, x_2, \dots, x_m , diferenciable en un punto dado $M(x_1, x_2, \dots, x_m)$. Para esto es suficiente exigir que la función $u = f(x_1, x_2, \dots, x_m)$ sea dos veces diferenciable en el punto $M(x_1, x_2, \dots, x_m)$, y los argumentos x_1, x_2, \dots, x_m sean o bien variables independientes, o bien funciones dos veces diferenciables de ciertas variables independientes.

Para estas suposiciones podemos estudiar la diferencial

$$\delta(du) = \delta \left[\sum_{h=1}^m \frac{\partial u}{\partial x_h} dx_h \right]$$

de la magnitud (5.20).

Definición 1. El valor $\delta(du)$ de la diferencial de la primera diferencial (5.20), tomado para $\delta x_1 = dx_1, \delta x_2 = dx_2, \dots, \delta x_m = dx_m$, se llama segunda diferencial de la función $u = f(x_1, x_2, \dots, x_m)$ (en un punto dado $M(x_1, x_2, \dots, x_m)$) y se denota con el símbolo d^2u . Así, pues, por definición *)

$$d^2u = \delta(du) \left| \begin{array}{l} \delta x_1 = dx_1, \\ \delta x_2 = dx_2, \\ \dots \\ \delta x_m = dx_m \end{array} \right. = \left\{ \delta \left[\sum_{h=1}^m \frac{\partial u}{\partial x_h} dx_h \right] \right\} \left| \begin{array}{l} \delta x_1 = dx_1, \\ \delta x_2 = dx_2, \\ \dots \\ \delta x_m = dx_m \end{array} \right.$$

*) El símbolo $\left\{ \begin{array}{l} \delta x_1 = dx_1, \\ \delta x_2 = dx_2, \\ \dots \\ \delta x_m = dx_m \end{array} \right\}$ significa que en la expresión encerrada entre llaves se debe adoptar $\delta x_1 = dx_1, \delta x_2 = dx_2, \dots, \delta x_m = dx_m$.

La diferencial $d^n u$ de cualquier orden n se introduce *por inducción*.

Supongamos que ya está introducida la diferencial $d^{n-1}u$ de orden $n-1$ y que $u = f(x_1, x_2, \dots, x_m)$ es n veces diferenciable en un punto dado $M(x_1, x_2, \dots, x_m)$, mientras que sus argumentos x_1, x_2, \dots, x_m son variables independientes, o bien funciones n veces diferenciables de ciertas variables independientes t_1, t_2, \dots, t_k .

Definición 2. El valor $\delta(d^{n-1}u)$ de la diferencial de la $(n-1)$ -ésima diferencial $d^{n-1}u$, tomada para $\delta x_1 = dx_1, \delta x_2 = dx_2, \dots, \delta x_m = dx_m$, se denomina *n-ésima diferencial de la función* $u = f(x_1, x_2, \dots, x_m)$ (en un punto dado $M(x_1, x_2, \dots, x_m)$) y se denota con el símbolo $d^n u$. Así, por definición,

$$d^n u = \delta(d^{n-1}u) \Big|_{\substack{\delta x_1 = dx_1, \\ \delta x_2 = dx_2, \\ \dots \\ \delta x_m = dx_m.}}$$

Al calcular las diferenciales segunda y posteriores hay que distinguir rigurosamente dos casos: 1) el caso en que los argumentos x_1, x_2, \dots, x_m son variables independientes, 2) el caso en que los argumentos x_1, x_2, \dots, x_m son funciones de ciertas variables independientes t_1, t_2, \dots, t_k , diferenciables el número correspondiente de veces.

Analicemos, al principio, el primer caso. Si x_1, x_2, \dots, x_m son variables *independientes*, tenemos el derecho de considerar que dx_1, dx_2, \dots, dx_m no dependen de x_1, x_2, \dots, x_m .

Cada diferencial dx_h podemos tomarlo igual a un mismo incremento Δx_h para todos los puntos $M(x_1, x_2, \dots, x_m)$. En este caso tendremos

$$\delta(dx_h) = \sum_{i=1}^m \frac{\partial(dx_h)}{\partial x_i} \delta x_i = 0.$$

La última relación y las reglas de diferenciación establecidas al final del p. 5, § 4, permiten escribir para una función $u = f(x_1, x_2, \dots, x_m)$, dos veces diferenciable en un punto dado M , la siguiente cadena de igualdades:

$$\begin{aligned} d^2 u &= \delta(du) \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \delta \left[\sum_{h=1}^m \frac{\partial u}{\partial x_h} dx_h \right] \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \\ &= \sum_{h=1}^m \delta \left[\frac{\partial u}{\partial x_h} dx_h \right] \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \sum_{h=1}^m \left\{ dx_h \cdot \delta \left(\frac{\partial u}{\partial x_h} \right) + \right. \end{aligned}$$

$$\begin{aligned}
& + \frac{\partial u}{\partial x_h} \delta(dx_h) \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \sum_{h=1}^m \left\{ dx_h \sum_{i=1}^m \frac{\partial}{\partial x_i} \left(\frac{\partial u}{\partial x_h} \right) \delta x_i \right\} \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \\
& = \left[\sum_{h=1}^m \sum_{i=1}^m \frac{\partial^2 u}{\partial x_i \partial x_h} \delta x_i \cdot dx_h \right] \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \sum_{h=1}^m \sum_{i=1}^m \frac{\partial^2 u}{\partial x_i \partial x_h} dx_i dx_h. \quad (5.43)
\end{aligned}$$

(También hemos aprovechado el hecho de que para una función dos veces diferenciable las derivadas mixtas de segundo orden no depende de la sucesión en que se realiza la diferenciación.)

Hemos llegado, pues, a que cuando los argumentos x_1, x_2, \dots, x_m son variables independientes, para la segunda diferencial de la función dos veces diferenciable en un punto dado $u = f(x_1, x_2, \dots, x_m)$ se verifica la representación

$$d^2u = \sum_{i=1}^m \sum_{h=1}^m \frac{\partial^2 u}{\partial x_i \partial x_h} dx_i dx_h. \quad (5.44)$$

OBSERVACION 1. Una función de m variables t_1, t_2, \dots, t_m del tipo $\Phi = \sum_{i=1}^m \sum_{h=1}^m a_{ih} t_i t_h$, donde a_{ih} son unos números reales constantes, se denomina *forma cuadrática* de las variables t_1, t_2, \dots, t_m , y los números a_{ih} , sus coeficientes.

Una forma cuadrática se llama simétrica, si sus coeficientes satisfacen la condición $a_{ih} = a_{hi}$ (para todo $i = 1, 2, \dots, m$, y todo $h = 1, 2, \dots, m$).

La expresión obtenida (5.44) permite afirmar que para el caso en que los argumentos x_1, x_2, \dots, x_m son variables independientes, la segunda diferencial de la función $u = f(x_1, x_2, \dots, x_m)$, dos veces diferenciable en el punto dado M , es una forma cuadrática simétrica *) de las variables dx_1, dx_2, \dots, dx_m , cuyos coeficientes son iguales a las correspondientes derivadas parciales de segundo orden de la función $u = f(x_1, x_2, \dots, x_m)$ tomadas en el punto dado M .

Advirtamos que la expresión obtenida para la diferencial de segundo orden (5.44) puede escribirse también en una forma diferente,

*) La simetría de esta forma cuadrática deriva de la igualdad $\frac{\partial^2 u}{\partial x_i \partial x_h}(M) = \frac{\partial^2 u}{\partial x_h \partial x_i}(M)$

utilizando el símbolo formal

$$d = dx_1 \frac{\partial}{\partial x_1} + dx_2 \frac{\partial}{\partial x_2} + \dots + dx_m \frac{\partial}{\partial x_m}. \quad (5.45)$$

Con ayuda de este símbolo la expresión (5.44) puede escribirse en la forma

$$d^2u = \left(dx_1 \frac{\partial}{\partial x_1} + dx_2 \frac{\partial}{\partial x_2} + \dots + dx_m \frac{\partial}{\partial x_m} \right)^2 u. \quad (5.46)$$

Es fácil convencerse por inducción de que cuando los argumentos x_1, x_2, \dots, x_m de la función $u = f(x_1, x_2, \dots, x_m)$, n veces diferenciable en el punto dado $M(x_1, x_2, \dots, x_m)$, son variables independientes, para la n -ésima diferencial de esta función se verifica la representación

$$d^n u = \sum_{i_1=1}^m \sum_{i_2=1}^m \dots \sum_{i_n=1}^m \frac{\partial^n u}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_n}} dx_{i_1} dx_{i_2} \dots dx_{i_n}.$$

Se puede escribir esta representación con ayuda del símbolo formal (5.45) en la forma siguiente:

$$d^n u = \left(dx_1 \frac{\partial}{\partial x_1} + dx_2 \frac{\partial}{\partial x_2} + \dots + dx_m \frac{\partial}{\partial x_m} \right)^n u. \quad (5.47)$$

Las representaciones para la segunda diferencial y las ulteriores tienen una forma absolutamente diferente cuando los argumentos x_1, x_2, \dots, x_m de la función $u = f(x_1, x_2, \dots, x_m)$ son funciones de ciertas variables independientes t_1, t_2, \dots, t_h , diferenciables un número correspondiente de veces.

Volviendo a este caso, establezcamos una expresión para la segunda diferencial de la función $u = f(x_1, x_2, \dots, x_m)$, dos veces diferenciable en un punto dado $M(x_1, x_2, \dots, x_m)$, cuyos argumentos x_1, x_2, \dots, x_m son funciones dos veces diferenciables de ciertas variables independientes t_1, t_2, \dots, t_h .

Al repetir los razonamientos de la cadena (5.43), obtendremos

$$\begin{aligned} d^2u &= \delta(du) \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \sum_{h=1}^m \left\{ dx_h \delta \left(\frac{\partial u}{\partial x_h} \right) + \frac{\partial u}{\partial x_h} \delta(dx_h) \right\} \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} = \\ &= \sum_{k=1}^m \left\{ dx_k \sum_{i=1}^m \frac{\partial}{\partial x_i} \left(\frac{\partial u}{\partial x_k} \right) \delta x_i \right\} \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} + \\ &\quad + \sum_{h=1}^m \left\{ \frac{\partial u}{\partial x_h} \delta(dx_h) \right\} \Big|_{\substack{\delta x_1 = dx_1, \\ \dots \\ \delta x_m = dx_m}} \end{aligned}$$

Advirtamos que, en virtud de la definición de segunda diferencial de la función $u = x_h$ (donde k es cualquiera de los números $1, 2, \dots, \dots, m$)

$$[\delta(dx_h)] \Big|_{\substack{\delta x_1 = dx_1, \\ \vdots \\ \delta x_m = dx_m}} \cdot d^2x_h.$$

Considerando esta relación, llegamos a la siguiente representación para la segunda diferencial:

$$d^2u = \sum_{k=1}^m \sum_{l=1}^m \frac{\partial^2 u}{\partial x_l \partial x_k} dx_l dx_k + \sum_{h=1}^m \frac{\partial u}{\partial x_h} d^2x_h$$

o bien, empleando el símbolo (5.45),

$$d^2u = \left(dx_1 \frac{\partial}{\partial x_1} + dx_2 \frac{\partial}{\partial x_2} + \dots + dx_m \frac{\partial}{\partial x_m} \right)^2 u + \\ + \left(\frac{\partial u}{\partial x_1} d^2x_1 + \frac{\partial u}{\partial x_2} d^2x_2 + \dots + \frac{\partial u}{\partial x_m} d^2x_m \right). \quad (5.48)$$

Comparando la representación obtenida (5.48) con la (5.46), nos convencemos de que la segunda diferencial (en diferencia de la primera) ya no posee la propiedad de invariación de la forma.

Con mayor razón no poseen esta propiedad todas las diferenciales ulteriores.

OBSERVACION 2. Demos a conocer un caso particular de importancia cuando la segunda diferencial y las posteriores de una función $u = f(x_1, x_2, \dots, x_m)$ de m variables poseen, no obstante, la invariación de la forma y se definen mediante la misma fórmula (5.47) que en caso de las variables independientes x_1, x_2, \dots, x_m .

Diremos que las variables x_1, x_2, \dots, x_m son *funciones lineales* de las variables independientes t_1, t_2, \dots, t_h si se definen mediante las igualdades

$$x_i = a_{i0} + a_{i1}t_1 + a_{i2}t_2 + \dots + a_{ih}t_h \quad (i = 1, 2, \dots, m),$$

en las cuales con $a_{i0}, a_{i1}, \dots, a_{ih}$ están designadas ciertas constantes.

Advirtamos que si la función $u = f(x_1, x_2, \dots, x_m)$ es n veces diferenciable en un punto dado $M(x_1, x_2, \dots, x_m)$ y si sus argumentos x_1, x_2, \dots, x_m son funciones lineales de las variables independientes t_1, t_2, \dots, t_h , la n -ésima diferencial de la función $u = f(x_1, x_2, \dots, \dots, x_m)$ se define mediante la misma fórmula (5.47) que en caso de las variables independientes x_1, x_2, \dots, x_m .

Con el fin de cerciorarnos de ello, advirtamos que por cuanto t_1, t_2, \dots, t_h son variables independientes, la n -ésima diferencial de x_i , siendo diferencial de las funciones de los argumentos $t_1, t_2, \dots, \dots, t_h$, se define mediante una igualdad del tipo (5.47), y, con mayor

precisión, mediante la igualdad

$$d^n x_i = \left(dt_1 \frac{\partial}{\partial t_1} + dt_2 \frac{\partial}{\partial t_2} + \dots + dt_n \frac{\partial}{\partial t_n} \right)^n x_i.$$

Mas, cualquier derivada parcial de orden superior al primero de una función lineal x_i es igual a cero.

Por consiguiente, $d^2 x_i = 0$, $d^3 x_i = 0$, ..., $d^n x_i = 0$.

La igualdad $d^2 x_i = 0$ (para todos los $i = 1, 2, \dots, m$) y la representación (5.48) permiten concluir que $d^2 u$ se define mediante la igualdad (5.46). De un modo análogo, aprovechando las relaciones $d^3 x_i = 0, \dots, d^n x_i = 0$, demostraremos por inducción que $d^3 u, d^4 u, \dots, d^n u$ se definen mediante la igualdad (5.47).

OBSERVACION 3. Al realizar los cálculos se necesita, a veces, descubrir la igualdad (5.47) y, teniendo presente que ésta contiene términos que coinciden, escribir todos los términos diferentes de esta igualdad junto con sus coeficientes.

Con este fin puede emplearse la fórmula del polinomio de Newton que tiene la forma

$$(a_1 + a_2 + \dots + a_m)^n = \sum_{\substack{\alpha_1 + \alpha_2 + \dots + \alpha_m = n \\ 0 \leq \alpha_i \leq n}} \frac{(\alpha_1 + \alpha_2 + \dots + \alpha_m)!}{\alpha_1! \alpha_2! \dots \alpha_m!} a_1^{\alpha_1} a_2^{\alpha_2} \dots a_m^{\alpha_m}$$

(la sumación en el segundo miembro de esta fórmula se realiza respecto de todos los índices de números enteros $\alpha_1, \alpha_2, \dots, \alpha_m$, cada uno de los cuales satisface las inequaciones $0 \leq \alpha_i \leq n$, a condición de que la suma de todos estos índices $\alpha_1 + \alpha_2 + \dots + \alpha_m$ sea igual a n).

La fórmula (5.49) se establece con facilidad por inducción. En efecto, para $m = 2$ y para cualquier n natural esta fórmula es a ciencia cierta válida, pues se transforma en la fórmula conocida del binomio de Newton.

Supongamos que esta fórmula es válida para cierto número $m \geq 2$ y todo n natural y comprobemos que en este caso es válida también para cualquier número $m + 1$ y todo n natural.

Al representar $(a_1 + a_2 + \dots + a_m + a_{m+1})^n$ en la forma

$$(a_1 + a_2 + \dots + a_{m+1})^n = [(a_1 + a_2 + \dots + a_m) + a_{m+1}]^n,$$

calculemos, con ayuda del binomio de Newton, el coeficiente de $a_1^{\alpha_1} a_2^{\alpha_2} \dots a_m^{\alpha_m} a_{m+1}^{\alpha_{m+1}}$. En virtud de la igualdad $\alpha_1 + \alpha_2 + \dots + \alpha_{m+1} = n$, de la fórmula del binomio de Newton y la suposición sobre la validez de la fórmula (5.49), para el número m y todo n natural, este coeficiente será: *)

$$\begin{aligned} C_n^{\alpha_{m+1}} \cdot \frac{(\alpha_1 + \alpha_2 + \dots + \alpha_m)!}{(\alpha_1)! (\alpha_2)! \dots (\alpha_m)!} &= \\ &= \frac{(\alpha_1 + \alpha_2 + \dots + \alpha_m + \alpha_{m+1})!}{(\alpha_{m+1})! (\alpha_1 + \alpha_2 + \dots + \alpha_m)!} \cdot \frac{(\alpha_1 + \alpha_2 + \dots + \alpha_m)!}{(\alpha_1)! (\alpha_2)! \dots (\alpha_m)!} = \\ &= \frac{(\alpha_1 + \alpha_2 + \dots + \alpha_m + \alpha_{m+1})!}{(\alpha_1)! (\alpha_2)! \dots (\alpha_m)! (\alpha_{m+1})!}. \end{aligned}$$

*) Tenemos en cuenta que $C_n^k = \frac{n!}{k! (n-k)!}$, y, por eso,

$$C_n^{\alpha_{m+1}} = C_n^{\alpha_{m+1} + \alpha_2 + \dots + \alpha_{m+1}} = \frac{(\alpha_1 + \alpha_2 + \dots + \alpha_m + \alpha_{m+1})!}{(\alpha_{m+1})! (\alpha_1 + \alpha_2 + \dots + \alpha_m)!}.$$

La expresión obtenida para el coeficiente de $a_1^{\alpha_1} \cdot a_2^{\alpha_2} \dots a_m^{\alpha_m} \cdot a_{m+1}^{\alpha_{m+1}}$ coincide plenamente con la expresión que se obtiene a partir de la fórmula (5.49) sustituyendo en ella m por $m+1$.

La inducción queda terminada y la fórmula (5.49) está demostrada.

La fórmula (5.49) nos permite escribir la expresión (5.47) para la n -ésima diferencial en la forma siguiente

$$d^n u(M) = \sum_{\substack{\alpha_1 + \alpha_2 + \dots + \alpha_m = n \\ 0 \leq \alpha_i \leq n}} \frac{\partial^n u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_m^{\alpha_m}}(M) \cdot (dx_1)^{\alpha_1} \cdot (dx_2)^{\alpha_2} \dots (dx_m)^{\alpha_m}.$$

3. Fórmula de Taylor para una función de m variables con el término residual en forma de Lagrange. Denotaremos la diferencial de k -ésimo orden de la función $u = f(x_1, x_2, \dots, x_m)$ en el punto M con el símbolo $d^k u|_M$. Demostremos el teorema siguiente.

Teorema 5.15. *Supongamos que una función $u = f(M) = f(x_1, x_2, \dots, x_m)$ viene dada en cierto ε -entorno *) del punto $M_0(x_1, x_2, \dots, x_m)$ y que es $n+1$ veces diferenciable en el ε -entorno citado. Entonces, el incremento total $\Delta u = f(M) - f(M_0)$ de dicha función en el punto M_0 para cualquier punto de M del ε -entorno indicado puede ser representado en la forma siguiente:*

$$\Delta u = du|_{M_0} + \frac{1}{2!} d^2 u|_{M_0} + \dots + \frac{1}{n!} d^n u|_{M_0} + \frac{1}{(n+1)!} d^{n+1} u|_N, \quad (5.50)$$

donde N es cierto punto del ε -entorno mencionado que en el caso general sólo depende de $M(x_1, x_2, \dots, x_m)$, y las diferenciales dx_i de las variables x_i que integran las expresiones $d^k u|_{M_0}$ y $d^{n+1} u|_N$ son iguales a $\Delta x_i = x_i - x_{i0}$. La fórmula (5.50) se llama fórmula de Taylor para la función $u = f(M)$ con centro del desarrollo en el punto M_0 .

DEMOSTRACION. Para reducir las notaciones, analicemos el caso de la función $u = f(x, y)$ de dos variables x e y . Previamente escribamos en una forma especial la fórmula de Taylor para la función $u = F(t)$ de una sola variable t , $n+1$ veces diferenciable en cierto entorno de un punto t_0 . Recordemos que la fórmula de Taylor con centro del desarrollo en t_0 para la función $u = F(t)$ de una sola variable tiene la forma (el término residual se toma en forma de Lagrange):

$$F(t) = F(t_0) + F'(t_0)(t-t_0) + \frac{1}{2!} F^{(2)}(t_0)(t-t_0)^2 + \dots \\ \dots + \frac{1}{n!} F^{(n)}(t_0)(t-t_0)^n + \frac{1}{(n+1)!} F^{(n+1)} \times \\ \times (t_0 + \theta(t-t_0))(t-t_0)^{n+1}, \quad 0 < \theta < 1. \quad (5.51)$$

*) En lugar del ε -entorno del punto M_0 se puede tomar el llamado *entorno estelar* de dicho punto que se define como un entorno del punto M_0 , el cual, junto con cada uno de sus puntos M , contiene íntegramente el segmento M_0M .

Por cuanto el argumento t es una variable independiente, el incremento $\Delta t = t - t_0$ es la diferencial dt de la variable independiente t . Por eso

$$F^{(k)}(t_0)(t - t_0)^k = F^{(k)}(t_0) dt^k = d^k F(t_0) = d^k u|_{t_0}$$

y

$$F^{(n+1)}(t_0 + \theta(t - t_0))(t - t_0)^{n+1} = d^{n+1} u|_{t_0 + \theta(t - t_0)}. \quad (5.52)$$

Al denotar la diferencia $F(t) - F(t_0)$ con Δu , la fórmula de Taylor (5.51) puede ser escrita, de acuerdo con (5.52), en la siguiente forma especial:

$$\Delta u = du|_{t_0} + \frac{1}{2!} d^2 u|_{t_0} + \dots \\ \dots + \frac{1}{n!} d^n u|_{t_0} + \frac{1}{(n+1)!} d^{(n+1)} u|_{t_0 + \theta(t - t_0)}. \quad (5.53)$$

Veamos ahora en el ε -entorno del punto $M_0(x_0, y_0)$ un punto arbitrario $M(x_0 + \Delta x, y_0 + \Delta y)$ y unamos los puntos M_0 y M mediante una recta. Evidentemente, las coordenadas x e y de los puntos de la citada recta son las siguientes *funciones lineales* de la nueva variable t :

$$x = x_0 + t \Delta x, \quad y = y_0 + t \Delta y, \quad (5.54)$$

y, además, las coordenadas de los puntos del segmento M_0M corresponden a los valores de la variable t del segmento $[0, 1]$. Notemos que al valor de $t = 0$ le corresponde el punto M_0 , y al valor de $t = 1$, el punto M . Dado que, por hipótesis, la función $u = f(x, y)$ de dos variables x e y es $n + 1$ veces diferenciable en el entorno que se considera, de las fórmulas (5.54) se deduce que en la recta M_0M esta función viene a ser la función compuesta de la variable t , $(n + 1)$ veces diferenciable al menos para todos los valores de t del segmento $[0, 1]$. Denotemos esta función compuesta con $F(t)$ y escribamos para ella la fórmula de Taylor con centro de desarrollo en el punto $t_0 = 0$ en la forma especial (5.53) para

$$\Delta u = F(1) - F(0) = f(M) - f(M_0).$$

Las diferenciales de diferentes órdenes que figuran en la fórmula (5.53) son diferenciales de la función compuesta $u = f(x, y)$, donde x e y son funciones lineales (5.54). De conformidad con la observación del punto anterior, en estas condiciones las diferenciales de cualquier orden de la función $u = f(x, y)$ pueden ser escritas en la forma (5.47). Por eso,

$$\left. \begin{aligned} d^k u|_{t_0=0} &= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy \right)^k u|_{M_0(x_0, y_0)} = d^k u|_{M_0}, \\ d^{n+1} u|_{t_0+\theta(t-t_0)} &= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy \right)^{n+1} \times \\ &\quad \times u|_{N(x_0+\theta \Delta x, y_0+\theta \Delta y)} = d^{n+1} u|_N, \end{aligned} \right\} \quad (5.55)$$

con la particularidad de que en las fórmulas (5.55) dx y dy se hallan a partir de las relaciones (5.54) para $dt = \Delta t = 1 - 0 = 1$. De este modo, en las fórmulas (5.55)

$$dx = dt \Delta x = \Delta x \quad \text{y} \quad dy = dt \Delta y = \Delta y. \quad (5.56)$$

Introduciendo $d^h u|_{t_0}$ y $d^{n+1}u|_{t_0+\theta(t-t_0)}$ de (5.55) en la fórmula (5.53) y considerando las relaciones (5.56), obtendremos la fórmula de Taylor (5.50). El teorema queda demostrado.

He aquí la expresión desarrollada de la fórmula de Taylor (5.50) para la función $u = f(x_1, x_2, \dots, x_m)$:

$$\begin{aligned} f(x_1, x_2, \dots, x_m) &= f(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m) + \\ &+ \sum_{h=1}^n \frac{1}{h!} \left[\frac{\partial}{\partial x_1} (x_1 - \overset{\circ}{x}_1) + \frac{\partial}{\partial x_2} (x_2 - \overset{\circ}{x}_2) + \dots \right. \\ &\quad \left. \dots + \frac{\partial}{\partial x_m} (x_m - \overset{\circ}{x}_m) \right]^h f(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m) + \\ &+ \frac{1}{(n+1)!} \left[\frac{\partial}{\partial x_1} (x_1 - \overset{\circ}{x}_1) + \frac{\partial}{\partial x_2} (x_2 - \overset{\circ}{x}_2) + \dots \right. \\ &\quad \left. \dots + \frac{\partial}{\partial x_m} (x_m - \overset{\circ}{x}_m) \right]^{n+1} f[\overset{\circ}{x}_1 + \theta(x_1 - \overset{\circ}{x}_1), \\ &\quad \overset{\circ}{x}_2 + \theta(x_2 - \overset{\circ}{x}_2), \dots, \overset{\circ}{x}_m + \theta(x_m - \overset{\circ}{x}_m)]. \end{aligned} \quad (5.57)$$

4. Fórmula de Taylor con el término residual en forma de Peano.

Teorema 5.15*. Supongamos que $n \geq 1$ es un número entero y que está dada una función $u = f(M) = f(x_1, x_2, \dots, x_m)$, $(n-1)$ veces diferenciable en el ε -entorno del punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ y n veces diferenciable en el propio punto M_0 (*). Entonces, para todo punto M del mencionado ε -entorno de M_0 resulta ser válida la fórmula siguiente:

$$\begin{aligned} f(M) &= f(M_0) + \frac{1}{1!} du|_{M_0} + \frac{1}{2!} d^2u|_{M_0} + \dots \\ &\quad \dots + \frac{1}{n!} d^n u|_{M_0} + o(\rho^n), \end{aligned} \quad (5.58)$$

en la cual con ρ está designada la distancia $\rho(M_0, M)$, y el símbolo $o(\rho^n)$ denota una función infinitésima para $\rho \rightarrow 0$ (o para $M \rightarrow M_0$), cuyo orden de infinitud es más elevado que ρ^n .

La fórmula (5.58) se denomina *fórmula de Taylor* (con centro en el punto M_0), cuyo término residual se expresa en la forma de Peano.

*) Cuando $n = 1$ se debe exigir que la función $u = f(x_1, x_2, \dots, x_m)$ esté definida sólo en el ε -entorno del punto M_0 y sea diferenciable en el propio punto M_0 .

OBSERVACIÓN. En la notación más detallada la fórmula de Taylor (5.58) tiene la forma

$$f(x_1, x_2, \dots, x_m) = f(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m) + \sum_{k=1}^n \frac{1}{k!} \left[(x_1 - \overset{\circ}{x}_1) \frac{\partial}{\partial x_1} + \dots \right. \\ \left. \dots + (x_m - \overset{\circ}{x}_m) \frac{\partial}{\partial x_m} \right]^k f(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m) + o(\rho^n). \quad (5.59)$$

Notemos que el segundo miembro de (5.59) contiene la suma del polinomio de n -ésimo grado de m variables x_1, x_2, \dots, x_m y del término residual $o(\rho^n)$.

Denotemos con $R_{n+1}(M)$ la diferencia entre $f(M)$ y el polinomio mencionado, es decir, pongamos

$$R_{n+1}(M) = f(M) - f(M_0) - \sum_{k=1}^n \frac{1}{k!} \left[(x_1 - \overset{\circ}{x}_1) \frac{\partial}{\partial x_1} + \dots \right. \\ \left. \dots + (x_m - \overset{\circ}{x}_m) \frac{\partial}{\partial x_m} \right]^k f(M_0). \quad (5.60)$$

El teorema 5.15* quedará demostrado si establecemos que, cumplidas las condiciones de este teorema, $R_{n+1}(M) = o(\rho^n)$.

Antes de demostrar el teorema 14.15* aduzcamos dos lemas.

Lema 1. Si una función $f(M) = f(x_1, x_2, \dots, x_m)$ es n veces diferenciable en un punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$, tanto la propia función $R_{n+1}(M)$, definida mediante la igualdad (5.60), como todas las derivadas parciales suyas respecto de cualesquiera variables x_1, x_2, \dots, x_m de orden hasta n inclusive se anulan en el punto M_0 .

DEMOSTRACIÓN. Cuando $n = 1$, la función (5.60) toma la forma

$$R_2(M) = f(M) - f(M_0) - (x_1 - \overset{\circ}{x}_1) \frac{\partial f}{\partial x_1}(M_0) - \dots \\ \dots - (x_m - \overset{\circ}{x}_m) \frac{\partial f}{\partial x_m}(M_0),$$

y las igualdades $R_2(M_0) = 0$, $\frac{\partial R_2}{\partial x_i}(M_0) = 0$ (para todos $i = 1, 2, \dots, m$) se comprueban de un modo elemental.

Para realizar la inducción, supongamos que el lema es válido con cierto número $n \geq 1$ y demostremos que en tal caso lo es también para el número $n + 1$.

Sea $f(M)$ una función $(n + 1)$ veces diferenciable en el punto M_0 y

$$R_{n+2}(M) = f(M) - f(M_0) - \sum_{k=1}^{n+1} \frac{1}{k!} \left[(x_1 - \overset{\circ}{x}_1) \frac{\partial}{\partial x_1} + \dots \right. \\ \left. \dots + (x_m - \overset{\circ}{x}_m) \frac{\partial}{\partial x_m} \right]^k f(M_0). \quad (5.61)$$

La igualdad $R_{n+2}(M_0) = 0$ se comprueba de un modo elemental (basta considerar que cada diferencia entre paréntesis $(x_i - \dot{x}_i)$ en (5.61) se anula en el punto M_0).

Queda por demostrar que para todo $i = 1, 2, \dots, m$ la propia función $\frac{\partial R_{n+2}}{\partial x_i}(M)$ y todas las derivadas parciales de esta función de orden hasta n inclusive se reducen a cero en el punto M_0 , y con este fin, en virtud de la suposición admitida sobre la validez del lema para el número n , es suficiente señalar que la función $\frac{\partial R_{n+2}}{\partial x_i}(M)$ se define mediante la igualdad (4.60), y con mayor precisión, mediante la igualdad

$$\begin{aligned} \frac{\partial R_{n+2}}{\partial x_i}(M) &= \frac{\partial f}{\partial x_i}(M) - \frac{\partial f}{\partial x_i}(M_0) - \\ &- \sum_{k=1}^n \frac{1}{k!} \left[(x_1 - \dot{x}_1) \frac{\partial}{\partial x_1} + \dots + (x_m - \dot{x}_m) \frac{\partial}{\partial x_m} \right]^k \frac{\partial f}{\partial x_i}(M_0). \end{aligned} \quad (5.62)$$

Por cuanto todas las variables x_i ($i = 1, 2, \dots, m$) son equitativas e intervienen en la expresión para $R_{n+2}(M)$ de un modo simétrico, es suficiente demostrar la igualdad (5.62) para $i = 1$, es decir, demostrar la igualdad

$$\begin{aligned} \frac{\partial R_{n+2}}{\partial x_1}(M) &= \frac{\partial f}{\partial x_1}(M) - \frac{\partial f}{\partial x_1}(M_0) - \sum_{k=1}^n \frac{1}{k!} \left[(x_1 - \dot{x}_1) \frac{\partial}{\partial x_1} + \dots \right. \\ &\quad \left. \dots + (x_m - \dot{x}_m) \frac{\partial}{\partial x_m} \right]^k \frac{\partial f}{\partial x_1}(M_0). \end{aligned} \quad (5.63)$$

La fórmula (5.61) evidencia que para demostrar (5.63) basta convencerse de que para cada número $k = 1, 2, \dots, n+1$ con x_2, x_3, \dots, x_m fijos

$$\begin{aligned} \frac{d}{dx_1} \left[(x_1 - \dot{x}_1) \frac{\partial}{\partial x_1} + (x_2 - \dot{x}_2) \frac{\partial}{\partial x_2} + \dots \right. \\ \left. \dots + (x_m - \dot{x}_m) \frac{\partial}{\partial x_m} \right]^k f(M_0) &= k \left[(x_1 - \dot{x}_1) \frac{\partial}{\partial x_1} + \right. \\ &\left. + (x_2 - \dot{x}_2) \frac{\partial}{\partial x_2} + \dots + (x_m - \dot{x}_m) \frac{\partial}{\partial x_m} \right]^{k-1} \frac{\partial f}{\partial x_1}(M_0). \end{aligned} \quad (5.64)$$

Al diferenciar respecto de x_1 , las variables x_2, x_3, \dots, x_m son fijas, razón por la cual la magnitud

$$D = (x_2 - \dot{x}_2) \frac{\partial}{\partial x_2} + \dots + (x_m - \dot{x}_m) \frac{\partial}{\partial x_m}$$

puede considerarse, al diferenciar respecto de x_1 , como una magnitud constante. Conviene añadir que por cuanto los símbolos $\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots$

... , $\frac{\partial}{\partial x_m}$ se emplean para formar las derivadas parciales de la función f en un punto fijo M_0 , entonces, al diferenciar respecto de x_1 , los citados símbolos también deben tomarse por magnitudes constantes.

En virtud de lo dicho, para demostrar la igualdad (14.64), basta convencerse de que es válida la igualdad

$$\frac{\partial}{\partial x_1} \left[(x_1 - \overset{\circ}{x}_1) \frac{\partial}{\partial x_1} + D \right]^k = k \left[(x_1 - \overset{\circ}{x}_1) \frac{\partial}{\partial x_1} + D \right]^{k-1} \frac{\partial}{\partial x_1}. \quad (5.65)$$

Al diferenciar la función $\left[(x_1 - \overset{\circ}{x}_1) \frac{\partial}{\partial x_1} + D \right]^k$ respecto de x_1 como función compuesta y considerando la independencia, señalada más arriba, de los símbolos D y $\frac{\partial}{\partial x_1}$ respecto de x_1 , obtendremos la igualdad (5.65). La inducción se da por terminada.

El lema 1 está demostrado.

Lema 2. Sea $R(M) = R(x_1, x_2, \dots, x_m)$ una función arbitraria que satisface dos exigencias:

- 1) $R(M)$ es n veces diferenciable en un punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$,
- 2) la propia función $R(M)$ y todas las derivadas parciales suyas respecto de cada una de las variables x_1, x_2, \dots, x_m de orden hasta n inclusive se reducen a cero en el punto mencionado M_0 . Entonces, para la función $R(M)$ es válida la estimación

$$R(M) = o(\rho^n), \quad (5.66)$$

donde la letra ρ designa la distancia $\rho(M_0, M)$ entre los puntos M_0 y M .

DEMOSTRACIÓN. Cuando $n = 1$, la afirmación del lema se desprende de la condición de diferenciabilidad *) de la función $R(M)$ en el punto M_0 , la cual tiene la forma

$$R(M) - R(M_0) = \sum_{k=1}^m \frac{\partial R}{\partial x_k}(M_0)(x_k - \overset{\circ}{x}_k) + o(\rho).$$

Considerando que $R(M_0) = 0$, $\frac{\partial R}{\partial x_k}(M_0) = 0$ para todo $k = 1, 2, \dots, m$, llegamos a que $R(M) = o(\rho)$.

Al realizar la inducción, supongamos que el lema 2 es válido para cierto número $n \geq 1$ y demostremos que en tal caso lo es también para el número $n + 1$.

Supongamos que la función $R(M)$ satisface dos requisitos del lema 2 para el número $n + 1$. Entonces, evidentemente, cualquier derivada parcial de esta función de primer orden $\frac{\partial R}{\partial x_k}(M)$ ($k =$

*) Véase la relación (5.16) del p. 2 § 4 de este capítulo.

$= 1, 2, \dots, m$) satisfará los dos requisitos del lema 2 para el número n , por lo cual (en virtud de la suposición admitida sobre la validez del lema 2 para el número n) será válida la estimación

$$\frac{\partial R}{\partial x_k}(M) = o(\rho^n). \quad (5.66^*)$$

Advirtamos, ahora, que por cuanto $n \geq 1$, tenemos $n + 1 \geq 2$, y la función $R(M)$, que satisface dos requisitos del lema 2 para el número $n + 1$, es en todo caso diferenciable una vez en un entorno del punto M_0 . Por eso, para esta función quedan cumplidas las condiciones del teorema 5.15 para el número $n = 0$. De acuerdo con el teorema mencionado, para todo punto M de un ε -entorno suficientemente pequeño del punto M_0 en el segmento M_0M existe un punto N tal que es válida la fórmula

$$R(M) = R(M_0) + \frac{1}{1!} \sum_{k=1}^m (x_k - \dot{x}_k) \frac{\partial R}{\partial x_k}(N). \quad (5.67)$$

Diremos ahora que por cuanto el punto N se encuentra entre los puntos M_0 y M , y ρ es la distancia entre los puntos M_0 y M , entonces $\rho(N, M_0) \leq \rho$, y, por eso, de (5.66*) se desprende que

$$\frac{\partial R}{\partial x_k}(N) = o(\rho^n).$$

Introduciendo la última estimación en (5.67) y considerando que $R(M_0) = 0$, obtendremos

$$R(M) = o(\rho^n) \sum_{k=1}^m |x_k - \dot{x}_k|.$$

Dado que $|x_k - \dot{x}_k| \leq \sqrt{\sum_{i=1}^m (x_i - \dot{x}_i)^2} = \rho$, obtenemos en definitiva que $R(M) = o(\rho^n)$.

La inducción se da por terminada. El lema 2 está demostrado.

LA DEMOSTRACION DEL TEOREMA 5.15* se lleva a cabo fácilmente con ayuda de los lemas 1 y 2.

En efecto, ya se ha hecho constar más arriba que para demostrar el teorema 5.15*, basta establecer la validez de la estimación

$$R_{n+1}(M) = o(\rho^n)$$

a condición de que se cumplan las condiciones del teorema para la función (5.60).

En virtud del lema 1, la propia función (5.60) y todas las derivadas parciales suyas respecto de cualquiera de las variables x_1, x_2, \dots, x_m de orden hasta n inclusive se anulan en el punto M_0 . Pero, en este caso, en virtud del lema 2, para la función (14.60) es válida la estimación $R_{n+1}(M) = o(\rho^m)$.

El teorema 5.15* está demostrado.

§ 6. Extremo local de una función de m variables

1. Concepto de extremo de una función de m variables. Condiciones necesarias de un extremo local. Supongamos que la función de m variables $u = f(M) = f(x_1, x_2, \dots, x_m)$ está definida en cierto entorno del punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ de un espacio E^m .

Definición 1. Diremos que la función $u = f(M)$ tiene en el punto M_0 un máximo local (mínimo local) si existe un δ -entorno del punto M_0 , en cuyos márgenes el valor $f(M_0)$ es el máximo (mínimo) entre todos los valores $f(M)$ de esta función.

Definición 2. Diremos que la función $u = f(M)$ tiene en el punto M_0 un extremo local si tiene en dicho punto o bien el máximo local, o bien el mínimo local.

Establezcamos las condiciones necesarias de un extremo local de la función $u = f(M)$ que posee en el punto dado M_0 derivadas parciales de primer orden respecto de todas las variables.

Demostremos la afirmación siguiente: si una función $u = f(M) = f(x_1, x_2, \dots, x_m)$ posee en un punto $M_0(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ derivadas parciales de primer orden respecto de todas las variables x_1, x_2, \dots, x_m y tiene en este punto un extremo local, todas las derivadas parciales de primer orden se reducen a cero en el punto M_0 , es decir, se verifican las igualdades

$$\frac{\partial u}{\partial x_1}(M_0) = 0, \quad \frac{\partial u}{\partial x_2}(M_0) = 0, \quad \dots, \quad \frac{\partial u}{\partial x_m}(M_0) = 0. \quad (5.68)$$

DEMOSTRACION Establezcamos la validez de la primera igualdad (5.68). Fijemos en la función $u = f(x_1, x_2, \dots, x_m)$ los argumentos x_2, x_3, \dots, x_m , adoptándolos iguales a las coordenadas correspondientes del punto M_0 , es decir, haciendo $x_2 = \overset{\circ}{x}_2, x_3 = \overset{\circ}{x}_3, \dots, x_m = \overset{\circ}{x}_m$. Obtenemos, pues, la función $u = f(x_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ de una sola variable x_1 . La derivada de esta función de una sola variable en el punto $x_1 = \overset{\circ}{x}_1$ coincide con la derivada parcial $\frac{\partial u}{\partial x_1}(M_0)$.

Por cuanto la función de m variables $u = f(M)$ posee su extremo local en el punto M_0 , la citada función de una sola variable $u = f(x_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ tiene el extremo local en el punto $x_1 = \overset{\circ}{x}_1$, y, por eso (en virtud de los resultados del p. 2, § 7, cap. 8, t. 1) la derivada de esta función de una sola variable en el punto $x_1 = \overset{\circ}{x}_1$, que coincide con la derivada parcial $\frac{\partial u}{\partial x_1}(M_0)$, es igual a cero.

La primera igualdad de (5.68) está demostrada. Las demás igualdades de (5.68) se demuestran de un modo análogo.

Recalquemos que las igualdades (5.68) (es decir, la anulación en el punto dado M_0 de todas las derivadas parciales de primer

orden) son sólo condiciones necesarias, pero no suficientes para que haya extremo local de la función $u = f(M)$ en el punto M_0 .

Por ejemplo, ambas derivadas parciales de una función de dos variables $u = xy$, $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$, se anulan en el punto $M_0(0, 0)$, no obstante, la citada función no tiene ningún extremo en el punto $M_0(0, 0)$, puesto que la función $u = xy$ es nula en el mismo punto $M_0(0, 0)$, mientras que toma valores tanto positivos, como negativos en un δ -entorno tan pequeño como se quiera.

Los puntos, en los que se anulan todas las derivadas parciales de primer orden de la función $y = f(M)$, se denominan *puntos de extremo eventual* de dicha función.

En cada punto de extremo eventual la función $u = f(M)$ puede tener un extremo local, pero, la existencia de este extremo puede establecerse sólo con ayuda de las condiciones suficientes de un extremo local que se analizan en el punto siguiente.

De la afirmación demostrada más arriba se desprende también otra forma de las condiciones necesarias de un extremo local:

si una función $u = f(M)$ es diferenciable en un punto M_0 y tiene en el mismo un extremo local, la diferencial $du|_{M_0}$ de esta función en el punto M_0 es idénticamente igual a cero respecto de las diferenciales de las variables independientes dx_1, dx_2, \dots, dx_m .

En efecto, por cuanto

$$du|_{M_0} = \frac{\partial u}{\partial x_1}(M_0) dx_1 + \frac{\partial u}{\partial x_2}(M_0) dx_2 + \dots + \frac{\partial u}{\partial x_m}(M_0) dx_m,$$

de las igualdades (14.68) se deduce que, cualesquiera que sean dx_1, dx_2, \dots, dx_m , queda válida la igualdad $du|_{M_0} = 0$.

2. Condiciones suficientes de un extremo local. En la formulación de las condiciones suficientes de un extremo local de la función de m variables $u = f(M)$ un papel importante lo desempeña la segunda diferencial de esta función en el punto M_0 sujeto al examen.

En el p. 2, § 5 de este capítulo nos hemos convencido de que cuando los argumentos x_1, x_2, \dots, x_m de una función dos veces diferenciable $u = f(x_1, x_2, \dots, x_m)$ son o bien variables independientes, o bien funciones lineales de ciertas variables independientes, la segunda diferencial de esta función en el punto dado M_0 es una forma cuadrática respecto de las diferenciales de los argumentos dx_1, dx_2, \dots, dx_m del tipo siguiente:

$$d^2u|_{M_0} = \sum_{i=1}^m \sum_{h=1}^m a_{ih} dx_i dx_h, \quad (5.69)$$

donde

$$a_{ih} = a_{hi} = \frac{\partial^2 u}{\partial x_i \partial x_h}(M_0). \quad (5.70)$$

Para formular las condiciones suficientes de un extremo local nos harán falta algunos datos de la teoría de las formas cuadráticas que daremos a conocer, para mayor comodidad del lector, más abajo *).

La forma cuadrática respecto de las variables h_1, h_2, \dots, h_m

$$\Phi(h_1, h_2, \dots, h_m) = \sum_{i=1}^m \sum_{h=1}^m a_{ih} h_i h_h \quad (5.71)$$

se denomina *definida positiva* (*definida negativa*) si para cualesquiera valores h_1, h_2, \dots, h_m , que no son simultáneamente nulos, esta forma toma valores estrictamente positivos (estrictamente negativos).

La forma cuadrática (5.71) se llama forma *de signo definido* si es o bien definida positiva, o bien definida negativa.

La forma cuadrática (5.71) se llama forma *de signo variable* si toma valores tanto estrictamente positivos, como estrictamente negativos.

La forma cuadrática (5.71) se llama forma *de signo casi definido* si toma o bien sólo valores no negativos, o bien sólo valores no positivos, anulándose para los valores h_1, h_2, \dots, h_m que no son simultáneamente nulos.

Enunciamos el llamado *criterio de Sylvester* que caracteriza la definición del signo de una forma cuadrática **).

Llamemos *matriz de forma cuadrática* (5.71) la siguiente expresión:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix}. \quad (15.72)$$

Si todos los elementos de la matriz A satisfacen la condición $a_{ik} = a_{ki}$ ($s = 1, 2, \dots, m; k = 1, 2, \dots, m$), esta matriz se llama *simétrica*.

Llamemos *menores principales* de la matriz simétrica (14.72) los siguientes determinantes:

$$A_1 = a_{11}, \quad A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad \dots, \quad A_m = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{vmatrix}.$$

*). Todas las definiciones y afirmaciones aquí aducidas pueden encontrarse, por ejemplo, en *Algebra lineal* de V. A. Ilín y E. G. Pozniak (en ruso).

**). J. Sylvester (1814-1897), matemático inglés.

El criterio de Sylvester se enuncia en forma de las dos afirmaciones siguientes:

1°. Para que la forma cuadrática (5.71) con la matriz simétrica (5.72) sea definida positiva, es necesario y suficiente que todos los menores principales de la matriz (5.72) sean positivos, es decir, que se cumplan las desigualdades

$$A_1 > 0, A_2 > 0, \dots, A_m > 0.$$

2°. Para que la forma cuadrática (5.71) con la matriz simétrica (5.72) sea definida negativa, es necesario y suficiente que los signos de los menores principales de la matriz (5.72) se alternen, con la particularidad de que el signo de A_1 sea negativo, es decir, que se verifiquen las desigualdades

$$A_1 < 0, A_2 > 0, A_3 < 0, A_4 > 0, \dots$$

Ahora ya estamos listos para poder enunciar y demostrar un teorema que establece las condiciones de un extremo local.

Teorema 5.16. Supongamos que una función de m variables $u = f(M) = f(x_1, x_2, \dots, x_m)$ es una vez diferenciable en cierto entorno del punto $M_0(x_1^0, x_2^0, \dots, x_m^0)$ y dos veces diferenciable en el propio punto M_0 . Admitamos, además, que el punto M_0 es el punto de extremo eventual de la función $u = f(M)$, es decir, que $du|_{M_0} = 0$. Entonces, si la segunda diferencial (5.69)–(5.70) es una forma cuadrática definida positiva (definida negativa) de las variables dx_1, dx_2, \dots, dx_m , la función $u = f(M)$ tiene en el punto M_0 un mínimo local (máximo local). En cambio, si la segunda diferencial (5.69)–(5.70) es una forma cuadrática de signo variable, la función $u = f(M)$ no tiene extremo local en el punto M_0 .

DEMOSTRACION. Demostremos, al principio, la primera parte del teorema, suponiendo, para concretar, que la segunda diferencial (5.69)–(5.70) es una forma cuadrática definida positiva de las variables dx_1, dx_2, \dots, dx_m . Demostremos que en este caso la función $u = f(M)$ tiene en el punto M_0 un mínimo local.

Desarrollemos la función $u = f(M)$ en un entorno del punto M_0 de conformidad con la fórmula de Taylor con término residual en forma de Peano, adoptando en la citada fórmula $n = 2^*$.

Obtendremos que

$$f(M) - f(M_0) = du|_{M_0} + \frac{1}{2!} d^2u|_{M_0} + o(\rho^2), \quad (14.73)$$

con la particularidad de que en la igualdad (5.73) las diferenciales dx_h de las variables x_h , que figuran en las expresiones para $du|_{M_0}$ y $d^2u|_{M_0}$, son iguales a los incrementos correspondientes $(x_h - x_h^0)$

* Para la función $u = f(M)$ quedan cumplidas, con $n = 2$, todas las condiciones del teorema 5.15* (véase p. 4, § 5 de este capítulo).

de estas variables, mientras que la magnitud ρ tiene la forma

$$\rho = \sqrt{(dx_1)^2 + (dx_2)^2 + \dots + (dx_n)^2} = \sqrt{(x_1 - \dot{x}_1)^2 + (x_2 - \dot{x}_2)^2 + \dots + (x_m - \dot{x}_m)^2}. \quad (5.74)$$

Por la hipótesis del teorema, el punto M_0 es un punto de extremo eventual. Por eso, en vista de los resultados del punto anterior, $du_{M_0} = 0$. Teniendo en cuenta esta igualdad y suponiendo en las expresiones (5.69)–(5.70) para la segunda diferencial $dx_k = x_k - \dot{x}_k$, demos a la fórmula de Taylor (5.73) la forma siguiente:

$$f(M) - f(M_0) = \frac{1}{2} \sum_{i=1}^m \sum_{k=1}^m a_{ik} (x_i - \dot{x}_i) (x_k - \dot{x}_k) + o(\rho^2). \quad (5.75)$$

Es suficiente probar que para todos los ρ suficientemente pequeños el segundo miembro de (5.75) es positivo. (Esto significará precisamente que en un entorno suficientemente pequeño del punto M_0 la diferencia $f(M) - f(M_0)$ es positiva, es decir, la función $u = f(M)$ tiene en el punto M_0 un mínimo local).

Adoptemos $h_i = \frac{x_i - \dot{x}_i}{\rho}$, donde $i = 1, 2, \dots, m$. Entonces, de la expresión (5.74) para ρ se deducen las relaciones siguientes:

$$|h_i| \leq 1, \quad h_1^2 + h_2^2 + \dots + h_m^2 = 1. \quad (5.76)$$

Con ayuda de las designaciones introducidas la igualdad (5.75) puede ser escrita en la forma

$$f(M) - f(M_0) = \frac{\rho^2}{2} \sum_{i=1}^m \sum_{k=1}^m a_{ik} h_i h_k + o(\rho^2). \quad (5.75^*)$$

La relación $\frac{o(\rho^2)}{\rho^2}$ es una función infinitésima para $\rho \rightarrow 0$ (o bien para $M \rightarrow M_0$), la cual se denotará con $\alpha(\rho)$. Introduciendo esta función se puede escribir la igualdad $o(\rho^2) = \rho^2 \alpha(\rho)$, con cuya ayuda daremos a la relación (5.75*) la forma siguiente:

$$f(M) - f(M_0) = \rho^2 \left[\frac{1}{2} \sum_{i=1}^m \sum_{k=1}^m a_{ik} h_i h_k + \alpha(\rho) \right]. \quad (5.75^{**})$$

Ahora ya no es difícil demostrar que el segundo miembro de (5.75**) es positivo para cualquier ρ suficientemente pequeño. La

forma cuadrática $\Phi = \sum_{i=1}^m \sum_{h=1}^m a_{ih} h_i h_h$ es una función definida y continua en la superficie de la esfera unidad (5.76) que constituye un conjunto cerrado y acotado. De acuerdo con el segundo teorema de Weierstrass (véase teorema 5.7 del p. 2, § 3, cap. 5), esta función alcanza en este conjunto su cota inferior exacta μ ; además, dado que la forma cuadrática (5.71) es definida positiva y que h_1, h_2, \dots, h_m , que satisfacen la relación (5.76), no son iguales a cero simultáneamente, la citada cota inferior exacta μ es estrictamente positiva.

Por cuanto la función $\alpha(\rho)$, infinitésima para $\rho \rightarrow 0$, satisface, para todos los ρ suficientemente pequeños, la desigualdad $|\alpha(\rho)| < \mu$, todo el miembro derecho de (5.75) es positivo para todos los ρ suficientemente pequeños, es decir, para todos los M suficientemente próximos a M_0 .

Esto es precisamente un indicio de que la función $u = f(M)$ tiene un mínimo local en el punto M_0 .

De un modo análogo se demuestra que cuando la segunda diferencial (5.69)–(5.70) es una forma cuadrática definida negativa, la función $u = f(M)$ tiene en el punto M_0 un máximo local.

Demostremos ahora la segunda parte del teorema, es decir, probemos que si la segunda diferencial (5.69)–(5.70) es una forma cuadrática de signo variable, la función $u = f(M)$ no tiene extremo local en el punto M_0 .

Establezcamos, ante todo, la siguiente propiedad auxiliar de la forma cuadrática de signo variable (5.71).

Si una forma cuadrática $\Phi(h_1, h_2, \dots, h_m)$ es de signo variable, existen dos colecciones de variables $(h'_1, h'_2, \dots, h'_m)$ y $(h''_1, h''_2, \dots, h''_m)$ tales que

$$(h'_1)^2 + (h'_2)^2 + \dots + (h'_m)^2 = 1, \quad (h''_1)^2 + (h''_2)^2 + \dots + (h''_m)^2 = 1, \quad (5.77)$$

con la particularidad de que

$$\Phi(h'_1, h'_2, \dots, h'_m) > 0, \quad \Phi(h''_1, h''_2, \dots, h''_m) < 0. \quad (5.78)$$

En efecto, en virtud de la definición de la forma cuadrática de signo variable se encontrarán dos colecciones de argumentos $(t'_1, t'_2, \dots, t'_m)$ y $(t''_1, t''_2, \dots, t''_m)$, compuestas de números no nulos simultáneamente, de tal género que

$$\Phi(t'_1, t'_2, \dots, t'_m) > 0, \quad \Phi(t''_1, t''_2, \dots, t''_m) < 0. \quad (5.79)$$

Adoptando

$$h'_i = \frac{t'_i}{\sqrt{(t'_1)^2 + (t'_2)^2 + \dots + (t'_m)^2}}, \quad (5.80)$$

$$h''_i = \frac{t''_i}{\sqrt{(t''_1)^2 + (t''_2)^2 + \dots + (t''_m)^2}}$$

y considerando que de la definición (5.71) de la forma cuadrática se deduce inmediatamente que

$$\Phi(h'_1, h'_2, \dots, h'_m) = \frac{1}{(t'_1)^2 + (t'_2)^2 + \dots + (t'_m)^2} \cdot \Phi(t'_1, t'_2, \dots, t'_m),$$

$$\Phi(h''_1, h''_2, \dots, h''_m) = \frac{1}{(t''_1)^2 + (t''_2)^2 + \dots + (t''_m)^2} \cdot \Phi(t''_1, t''_2, \dots, t''_m),$$

obtendremos (en virtud de (5.79)) las desigualdades (5.78), con la particularidad de que de las relaciones (5.80) proviene directamente las igualdades (5.77).

La propiedad auxiliar de la forma cuadrática de signo variable está demostrada.

Volvamos ahora a la demostración de la segunda parte del teorema.

Fijemos dos colecciones de variables $(h'_1, h'_2, \dots, h'_m)$ y $(h''_1, h''_2, \dots, h''_m)$ que satisfacen las relaciones (5.77) y (5.78) y demostremos que para todo $\rho > 0$ existen dos puntos M' (x'_1, x'_2, \dots, x'_m) y M'' (x''_1, \dots, x''_m) del espacio E^m de tal género que $\rho(M', M_0) = \rho(M'', M_0) = \rho$, con la particularidad de que

$$\frac{x'_i - \overset{\circ}{x}_i}{\rho} = h'_i, \quad \frac{x''_i - \overset{\circ}{x}_i}{\rho} = h''_i \quad \text{para todo } i = 1, 2, \dots, m, \quad (5.81)$$

En efecto, adoptando para cualquier $\rho > 0$ y para todo número i ($i = 1, 2, \dots, m$)

$$x'_i = \overset{\circ}{x}_i + \rho h'_i, \quad x''_i = \overset{\circ}{x}_i + \rho h''_i,$$

satisfaremos las relaciones (5.81); además, en virtud de las igualdades (5.77), serán válidas las igualdades

$$\begin{aligned} \rho(M', M_0) &= \sqrt{(x'_1 - \overset{\circ}{x}_1)^2 + (x'_2 - \overset{\circ}{x}_2)^2 + \dots + (x'_m - \overset{\circ}{x}_m)^2} = \\ &= \rho \sqrt{(h'_1)^2 + (h'_2)^2 + \dots + (h'_m)^2} = \rho, \end{aligned}$$

$$\begin{aligned} \rho(M'', M_0) &= \sqrt{(x''_1 - \overset{\circ}{x}_1)^2 + (x''_2 - \overset{\circ}{x}_2)^2 + \dots + (x''_m - \overset{\circ}{x}_m)^2} = \\ &= \rho \sqrt{(h''_1)^2 + (h''_2)^2 + \dots + (h''_m)^2} = \rho. \end{aligned}$$

Ahora ya no es difícil convencerse de que cuando la segunda diferencial (5.69)–(5.70) es una forma cuadrática de signo variable, la función $u = f(M)$ no tiene extremo en el punto M_0 .

Al escribir para la función $u = f(M)$ el desarrollo en un entorno del punto M_0 según la fórmula de Taylor con término residual en forma de Peano y al tomar este desarrollo en los puntos M' y M'' mencionados más arriba, obtendremos en lugar de (5.75) los dos desa-

rollos siguientes:

$$f(M') - f(M_0) = \frac{1}{2} \sum_{i=1}^m \sum_{h=1}^m a_{ih} (x'_i - \overset{\circ}{x}_i) (x'_h - \overset{\circ}{x}_h) + o(\rho^2). \quad (5.82)$$

$$f(M'') - f(M_0) = \frac{1}{2} \sum_{i=1}^m \sum_{h=1}^m a_{ih} (x''_i - \overset{\circ}{x}_i) (x''_h - \overset{\circ}{x}_h) + o(\rho^2), \quad (5.83)$$

válidos para todos los $\rho > 0$ suficientemente pequeños.

Introduciendo en estos desarrollos los valores $(x'_i - \overset{\circ}{x}_i)$ y $(x''_i - \overset{\circ}{x}_i)$ de las igualdades (5.81) y considerando que $o(\rho^2) = \rho^2 \cdot \alpha(\rho)$, donde $\alpha(\rho) \rightarrow 0$ cuando $\rho \rightarrow 0$, daremos a los desarrollos (5.82) y (5.83) la forma siguiente

$$f(M') - f(M_0) = \rho^2 \left[\frac{1}{2} \sum_{i=1}^m \sum_{h=1}^m a_{ih} h'_i h'_h + \alpha(\rho) \right],$$

$$f(M'') - f(M_0) = \rho^2 \left[\frac{1}{2} \sum_{i=1}^m \sum_{h=1}^m a_{ih} h''_i h''_h + \alpha(\rho) \right].$$

Las últimas dos relaciones pueden reescribirse en la forma

$$f(M') - f(M_0) = \rho^2 \left[\frac{1}{2} \Phi(h'_1, h'_2, \dots, h'_m) + \alpha(\rho) \right], \quad (5.82^*)$$

$$f(M'') - f(M_0) = \rho^2 \left[\frac{1}{2} \Phi(h''_1, h''_2, \dots, h''_m) + \alpha(\rho) \right]. \quad (5.83^*)$$

Teniendo en cuenta las relaciones (5.78) y el hecho de que las magnitudes $\Phi(h'_1, h'_2, \dots, h'_m) > 0$ y $\Phi(h''_1, h''_2, \dots, h''_m) < 0$ no dependen de ρ , y recordando que $\rho = \rho(M', M_0) = \rho(M'', M_0)$, obtenemos de las relaciones (5.82*) y (5.83*) que para $\rho > 0$, tan pequeño como se quiera, se cumplen las desigualdades $f(M') > f(M_0)$ y $f(M'') < f(M_0)$ que demuestran precisamente la ausencia de extremo en el punto M_0 .

El teorema 5.16 está plenamente demostrado.

OBSERVACION 1. Si la segunda diferencial de una función $u = f(M)$, dos veces diferenciable en un punto dado de extremo eventual M_0 , es en dicho punto una forma cuadrática de signo casi definido, no puede decirse nada concreto sobre la existencia o ausencia en este punto de un extremo local.

Así, por ejemplo, en cada una de las dos funciones $u_1 = x^3 + y^3$ y $u_2 = x^4 + y^4$ la segunda diferencial en el punto de extremo eventual $M_0(0, 0)$ es idénticamente igual a cero (es decir, es una forma cuadrática de signo casi definido), mas, sólo la segunda de las funciones mencionadas tienen en dicho punto un extremo local.

Con el fin de resolver la cuestión de extremo local para el caso en que la segunda diferencial es una forma cuadrática de signo casi definido, se deben examinar las diferenciales de órdenes más elevados, pero esto sale de los márgenes del curso dado.

OBSERVACION 2 *El requisito $d^2u | M \geq 0$ ($d^2u | M_0 \leq 0$, respectivamente) es la condición necesaria para que exista un mínimo (máximo) local en el punto M_0 de una función $u = f(M)$ dos veces diferenciable en este punto.*

En efecto, supongamos, para concretar, que $u = f(M)$ tiene en el punto M_0 un mínimo local, pero la condición $d^2u | M_0 \geq 0$ no está cumplida. Entonces, se encontrarán tales h_1, h_2, \dots, h_m que

$$d^2u | M_0 = \sum_{i=1}^m \sum_{k=1}^m \frac{\partial^2 f}{\partial x_i \partial x_k} (M_0) h_i h_k < 0.$$

Veamos la función $F(t) = f(\overset{\circ}{x}_1 + th_1, \overset{\circ}{x}_2 + th_2, \dots, \overset{\circ}{x}_m + th_m)$, definida a ciencia cierta para cualesquiera t suficientemente pequeños en módulo. La función $F(t)$ ha de tener un mínimo local en el punto $t = 0$, lo que está en contradicción con la condición

$$F''(0) = d^2u | M_0 < 0.$$

3. Caso de una función de dos variables. En la práctica se encuentra frecuentemente el problema de extremo de una función de dos variables $u = f(x, y)$. En este punto daremos a conocer los resultados referentes al caso dado.

Denotemos las derivadas parciales $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x \partial y}$, $\frac{\partial^2 u}{\partial y^2}$ en cierto punto $M_0(\overset{\circ}{x}, \overset{\circ}{y})$ con los símbolos a_{11} , a_{12} , a_{22} , respectivamente. Resulta ser válida la afirmación siguiente.

Supongamos que una función de dos variables $u = f(x, y)$, una vez diferenciable en un entorno del punto $M_0(\overset{\circ}{x}, \overset{\circ}{y})$ y dos veces diferenciable en el propio punto M_0 , y admitamos que M_0 es un punto de extremo eventual. Entonces, si en el punto M_0 se cumple la condición $a_{11}a_{22} - a_{12}^2 > 0$, la función $u = f(x, y)$ tiene en este punto extremo local (máximo, cuando $a_{11} < 0$, y mínimo, cuando $a_{11} > 0$). En cambio, si en dicho punto M_0 se tiene $a_{11}a_{22} - a_{12}^2 < 0$, la función $u = f(x, y)$ no tiene en este punto extremo local.*

DEMOSTRACION La validez de la primera parte de la afirmación enunciada se deduce directamente del teorema 14.16 y del criterio de Sylvester sobre la forma cuadrática de signo definido, puesto que

$$A_1 = a_{11}, \quad A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}^2.$$

* El caso de $a_{11}a_{22} - a_{12}^2 = 0$ requiere un análisis adicional.

Demostremos la segunda parte de la afirmación. Así, pues, supongamos que en el punto M_0 se cumple la desigualdad $a_{11}a_{22} - a_{12}^2 < 0$. Demostremos que en este caso la segunda diferencial d^2u en el punto M_0 es una forma de signo variable. Examinemos, al principio, el caso de $a_{11} \neq 0$. Al hacer uso de las designaciones introducidas anteriormente

$$\rho = \sqrt{(x-\bar{x})^2 + (y-\bar{y})^2}, \quad h_1 = \frac{x-\bar{x}}{\rho}, \quad h_2 = \frac{y-\bar{y}}{\rho},$$

obtendremos la expresión siguiente para la segunda diferencial:

$$\begin{aligned} d^2u|_{M_0} &= \rho^2 \sum_{i=1}^2 \sum_{k=1}^2 a_{ik} h_i h_k = \rho^2 (a_{11}h_1^2 + 2a_{12}h_1h_2 + a_{22}h_2^2) = \\ &= \frac{\rho^2}{a_{11}} [(a_{11}h_1 + a_{12}h_2)^2 + (a_{11}a_{22} - a_{12}^2)h_2^2]. \end{aligned}$$

Es fácil comprobar que cuando $h_1 = 1$, $h_2 = 0$, y para $h_1 = \frac{a_{12}}{\sqrt{a_{11}^2 + a_{12}^2}}$, $h_2 = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{12}^2}}$ (*), la diferencial $d^2u|_{M_0}$ tiene signos diferentes, es decir, es una forma de signo variable, por lo cual, de acuerdo con el teorema 14.16, la función no tiene en M_0 un extremo local.

Examinemos ahora el caso de $a_{11} = 0$. De la condición $a_{11}a_{22} - a_{12}^2 < 0$ se desprende que $a_{12} \neq 0$. Por consiguiente, de la expresión para $d^2u|_{M_0}$, escrita más arriba, se obtiene

$$d^2u|_{M_0} = \rho^2 h_2 (2a_{12}h_1 + a_{22}h_2). \quad (5.84)$$

Supongamos que $h_1 \neq 0$ y que la magnitud h_2 es tan pequeña (de la condición $h_1^2 + h_2^2 = 1$ se deduce que semejante elección de h_1 y h_2 es posible) que la expresión $(2a_{12}h_1 + a_{22}h_2)$ conserva invariable el signo de la magnitud $2a_{12}h_1$. Entonces, de la fórmula (5.84) se deduce que $d^2u|_{M_0}$ tiene signos diferentes para $h_2 > 0$ y $h_2 < 0$, es decir, la función $u = f(x, y)$ no tiene un extremo local en el punto M_0 . La afirmación está completamente demostrada.

4. Ejemplos de análisis del extremo de una función. 1) Hállense los puntos del extremo local de una función de m variables

$$u = \lambda x_1^2 + x_2^2 + \dots + x_m^2 + 2x_2 + \dots + 2x_m, \quad (5.85)$$

donde λ es un número real distinto de cero.

Para hallar los puntos de extremo eventual obtenemos las siguientes ecuaciones:

$$\frac{\partial u}{\partial x_1} = 2\lambda x_1 = 0, \quad \frac{\partial u}{\partial x_2} = 2x_2 + 2 = 0, \quad \dots, \quad \frac{\partial u}{\partial x_m} = 2x_m + 2 = 0. \quad (5.86)$$

* En este caso ρ puede ser una magnitud tan pequeña como se quiera. La condición $h_1^2 + h_2^2 = 1$ está cumplida.

A partir de las ecuaciones (5.86) concluimos que el único punto de extremo eventual es $M_0(0, -1, \dots, -1)$.

Con el fin de analizar la función (5.85) en dicho punto M_0 con ayuda de las condiciones suficientes de extremo, calculemos la segunda diferencial

$$d^2u|_{M_0} = 2\lambda(dx_1)^2 + 2(dx_2)^2 + \dots + 2(dx_m)^2. \quad (5.87)$$

Es obvio que cuando $\lambda > 0$, todos los valores de la segunda diferencial (5.87) son estrictamente positivos para dx_1, dx_2, \dots, dx_m simultáneamente distintos de cero, es decir, cuando $\lambda > 0$, la segunda diferencial (5.87) es una forma cuadrática definida positiva. Por eso, cuando $\lambda > 0$, la función (5.85) tiene un mínimo local en el punto $M_0(0, -1, \dots, -1)$.

Cuando $\lambda < 0$, la segunda diferencial (5.87) es positiva para $dx_1 = 0, \dots, dx_{m-1} = 0, dx_m = 1$, y negativa para $dx_1 = 1, dx_2 = 0, \dots, dx_m = 0$. Esto significa que, cuando $\lambda < 0$, la segunda diferencial (5.87) es una forma cuadrática de signo variable. Por eso, cuando $\lambda < 0$, la función (5.85) no tiene un extremo local en el punto $M_0(0, -1, \dots, -1)$.

2) En un plano hay n puntos $M_k(a_k, b_k)$, $k = 1, 2, \dots, n$, en los cuales están concentradas las masas $m_k > 0$. Se pide hallar en dicho plano un punto $M_0(x_0, y_0)$ tal que respecto de él el momento de inercia del sistema indicado de puntos materiales sea mínimo.

Por cuanto el momento de inercia del sistema dado de puntos materiales respecto del punto $M(x, y)$ e

$$I(x, y) = \sum_{k=1}^n m_k [(x - a_k)^2 + (y - b_k)^2], \quad (5.88)$$

el problema consiste en hallar el punto $M_0(x_0, y_0)$, en el cual la función (5.88) alcanza su valor mínimo.

Para hallar los puntos de extremo eventual de la función (5.88) tenemos las ecuaciones siguientes:

$$\frac{\partial I}{\partial x} = 2 \sum_{k=1}^n m_k (x - a_k) = 0, \quad \frac{\partial I}{\partial y} = 2 \sum_{k=1}^n m_k (y - b_k) = 0, \quad (5.89)$$

A partir de las ecuaciones (5.89) concluimos que el único punto de extremo eventual de la función (5.88) es $M_0(x_0, y_0)$, cuyas coordenadas son

$$x_0 = \frac{m_1 a_1 + m_2 a_2 + \dots + m_n a_n}{m_1 + m_2 + \dots + m_n}, \quad y_0 = \frac{m_1 b_1 + m_2 b_2 + \dots + m_n b_n}{m_1 + m_2 + \dots + m_n}. \quad (5.90)$$

Por cuanto $a_{11} = \frac{\partial^2 I}{\partial x^2} = 2 \sum_{k=1}^n m_k > 0$, $a_{12} = \frac{\partial^2 I}{\partial x \partial y} = 0$, $a_{22} = \frac{\partial^2 I}{\partial y^2} = 2 \sum_{k=1}^n m_k > 0$, tenemos $a_{11} a_{22} - a_{12}^2 > 0$, y, de acuerdo con la

afirmación demostrada en el p. 3, la función (5.88) tiene un mínimo local en el punto $M_0(x_0, y_0)$ con las coordenadas (5.90). Es fácil convencerse de que el valor de $I(x, y)$ en este punto es mínimo. Advertimos en conclusión que las fórmulas (5.90) determinan las coordenadas del centro de gravedad del sistema de puntos materiales que se considera.

§ 7. Método gradienta! de búsqueda del extremo de una función fuertemente convexa

En este párrafo se expone la teoría del método gradienta! de búsqueda del extremo de una función fuertemente convexa que es de amplio uso en la práctica.

La idea de este método es extraordinariamente sencilla. Para encontrar aproximadamente los puntos del mínimo de una función de m variables, se emplea el hecho de que la orientación del gradiente de esta función coincide con la dirección de crecimiento máximo de esta función. Por consiguiente, un vector $-\text{grad } f(x_0)$ en todo punto $x_0 = (\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ está orientado según la dirección de decrecimiento máximo de la función $f(x) = f(x_1, x_2, \dots, x_m)$. Esto nos permite esperar que si, partiendo de cierta aproximación nula $x_0 = (\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$, construimos la k -ésima aproximación $x_k = (\overset{k}{x}_1, \overset{k}{x}_2, \dots, \overset{k}{x}_m)$ con arreglo a la fórmula recurrente

$$x_{k+1} = x_k - \alpha \text{grad } f(x_k),$$

entonces, para α positivo y suficientemente pequeño, la sucesión de puntos $\{x_k\}$ convergerá hacia el punto del mínimo de la función $f(x)$.

El presente párrafo está dedicado a la realización estricta de esta idea simple.

1. **Conjuntos convexos y funciones convexas.** Sea $x_1 = (\overset{1}{x}_1, \overset{1}{x}_2, \dots, \overset{1}{x}_m)$ y $x_2 = (\overset{2}{x}_1, \overset{2}{x}_2, \dots, \overset{2}{x}_m)$ dos puntos de un espacio euclídeo m -dimensional E^m , los cuales pueden considerarse como vectores en E^m con las coordenadas correspondientes.

Llamemos *segmento* que une los puntos x_1 y x_2 a un conjunto de puntos del espacio E^m de la forma $x_1 + t(x_2 - x_1)$, donde t es un número cualquiera del segmento $0 \leq t \leq 1$.

Denotemos el segmento que une los puntos x_1 y x_2 con el símbolo $\underline{x_1x_2}$.

Definición 1. Un conjunto Q de puntos del espacio E^m se llama convexo si posee la propiedad siguiente: cualesquiera que sean dos puntos x_1 y x_2 , pertenecientes al conjunto Q , el segmento $\underline{x_1x_2}$ que los une también pertenece a este conjunto.

Como ejemplo de conjunto convexo en el espacio E^m puede servir una esfera m -dimensional (no importa que sea abierta o cerrada), o bien un semiespacio $x_m \geq 0$ (es decir, un conjunto de todos los puntos (x_1, x_2, \dots, x_m) del espacio E^m , cuya m -ésima coordenada satisface la condición $x_m \geq 0$).

Como ejemplo de conjunto Q no convexo puede servir un complemento de la esfera m -dimensional o una esfera m -dimensional, en la cual está suprimido por lo menos un punto.

Sea Q cierto conjunto de puntos del espacio E^m , y sea x cualquier punto fijo de este espacio.

Llamemos *distancia* del punto x al conjunto Q a la cota inferior exacta de las distancias del punto x a toda clase de puntos de dicho conjunto.

Denotemos con el símbolo $\rho(x, Q)$ la distancia del punto x al conjunto Q .

Así, pues, por definición,

$$\rho(x, Q) = \inf_{y \in Q} \rho(x, y).$$

Para cualquier conjunto Q del espacio E^m y para todo punto x de este espacio existe una distancia $\rho(x, Q)$ *). En particular, si el punto x pertenece al conjunto Q , tenemos $\rho(x, Q) = 0$.

No obstante, en el conjunto Q no siempre existe un punto y tal que se verifique $\rho(x, y) = \rho(x, Q)$.

Así, por ejemplo, si el conjunto Q es una esfera *abierta* m -dimensional, y x es un punto de E^m dispuesto fuera de dicha esfera, en tal conjunto Q no existe un punto y tal que $\rho(x, y) = \rho(x, Q)$ (puesto que para todos los puntos y de la esfera abierta Q se cumple la desigualdad $\rho(x, y) > \rho(x, Q)$).

Si, no obstante, Q contiene un punto y tal que $\rho(x, y) = \rho(x, Q)$, el citado punto y se denomina *proyección del punto x sobre el conjunto Q* .

La proyección del punto x sobre el conjunto Q se denota con el símbolo $P_Q(x)$.

Subrayemos que si el punto x pertenece al conjunto Q , tenemos $P_Q(x) = x$.

Así, pues, la proyección $P_Q(x)$ del punto x sobre el conjunto Q se define mediante una relación

$$\rho(x, P_Q(x)) = \rho(x, Q) = \inf_{y \in Q} \rho(x, y).$$

Cabe destacar que pueden existir varias proyecciones del punto x sobre el conjunto Q . Por ejemplo, si Q es una esfera m -dimensional con centro en el punto x , cualquier punto de Q será proyección del punto x sobre el conjunto Q .

* Puesto que el conjunto $\rho(x, y)$ para toda clase de y pertenecientes a Q está siempre acotado inferiormente (mediante el número cero, por ejemplo).

Es válido, sin embargo, el lema siguiente.

Lema 1. *Si un conjunto Q del espacio E^m es convexo y cerrado, y x es cualquier punto de E^m , existe, y, además, la única proyección del punto x sobre el conjunto Q .*

DEMOSTRACION Al principio demostremos la existencia de por lo menos una sola proyección del punto x sobre el conjunto Q . Denotemos con $\rho(x, Q)$ la distancia del punto x al conjunto Q . Por cuanto $\rho(x, Q)$ está definida como cota inferior exacta $\inf_{y \in Q} \rho(x, y)$, se encontrará una sucesión $\{y_n\}$ de puntos del conjunto Q tal que $\rho(x, y_n) \rightarrow \rho(x, Q)$.

Por definición de límite de una sucesión numérica para cualquier $\varepsilon > 0$, todos los elementos y_n satisfacen, a partir de cierto número, la relación

$$\rho(x, Q) - \varepsilon < \rho(x, y_n) < \rho(x, Q) + \varepsilon.$$

De aquí se deduce que la sucesión $\{y_n\}$ de puntos del espacio E^m está en todo caso acotada y, por eso, en virtud del teorema de Bolzano-Weierstrass (véase p. 2, § 2, cap. 5, en esta sucesión puede seleccionarse una subsucesión convergente $\{y_{k_n}\}$, donde $n = 1, 2, \dots$. Denotemos con y el límite de la subsucesión $\{y_{k_n}\}$. Por ser el conjunto Q cerrado, el punto y pertenece a este conjunto. Resta probar que

$$\rho(x, y) = \rho(x, Q) = \lim_{n \rightarrow \infty} \rho(x, y_{k_n}).$$

Advirtamos que en vista de las desigualdades triangulares $\rho(x, y_{k_n}) \leq \rho(x, y) + \rho(y, y_{k_n})$ y $\rho(x, y) \leq \rho(x, y_{k_n}) + \rho(y_{k_n}, y)$, se verifica una relación $|\rho(x, y_{k_n}) - \rho(x, y)| \leq \rho(y, y_{k_n})$. De esta relación y de la convergencia de la subsucesión $\{y_{k_n}\}$ hacia y se desprende que $\lim_{n \rightarrow \infty} \rho(x, y_{k_n}) = \rho(x, y)$, es decir, $\rho(x, Q) = \rho(x, y)$.

De este modo queda demostrada la existencia de por lo menos una proyección del punto x sobre el conjunto Q .

Demostremos ahora que existe sólo una proyección del punto x sobre el conjunto Q . Supongamos que existen dos diferentes proyecciones y_1 y y_2 del punto x sobre el conjunto Q . Por cuanto Q es un conjunto convexo, todo el segmento $\overline{y_1 y_2}$ que une y_1 e y_2 pertenece al conjunto Q . En particular, al conjunto Q le pertenece el punto medio $\frac{y_1 + y_2}{2}$ del citado segmento. Cerciorémonos de que la distancia $\rho\left(x, \frac{y_1 + y_2}{2}\right)$ del punto x al citado punto medio del segmento $\overline{y_1 y_2}$ es estrictamente inferior a la distancia $\rho(x, y_1) = \rho(x, y_2)$.

Excluyamos del análisis un caso trivial en que $\frac{y_1 + y_2}{2} = x$.

En este caso $\rho\left(x, \frac{y_1 + y_2}{2}\right) = 0$, mientras que $\rho(x, y_1) = -\rho(x, y_2) > 0$, pues, de lo contrario (es decir, si $\rho(x, y_1) = -\rho(x, y_2) = 0$), ambos puntos y_1 e y_2 coincidirían con x y no podrían ser diferentes. Así, pues, en el caso trivial $\frac{y_1 + y_2}{2} = x$ la inecuación

$$\rho\left(x, \frac{y_1 + y_2}{2}\right) < \rho(x, y_1) - \rho(x, y_2) \quad (5.91)$$

es evidente.

Demostremos ahora la inecuación (5.91) para el caso en que $\frac{y_1 + y_2}{2} \neq x$.

Haciendo uso de la propiedad del producto escalar de dos vectores en el espacio E^m *, obtendremos la relación

$$\begin{aligned} \rho^2\left(x, \frac{y_1 + y_2}{2}\right) &= \left(\frac{y_1 + y_2}{2} - x, \frac{y_1 - y_2}{2} - x\right) = \\ &= \left(\frac{y_1 - x}{2} + \frac{y_2 - x}{2}, \frac{y_1 - x}{2} - \frac{y_2 - x}{2}\right) = \\ &= \frac{1}{4} [(y_1 - x, y_1 - x) + 2(y_1 - x, y_2 - x) + (y_2 - x, y_2 - x)]. \end{aligned} \quad (5.92)$$

Cerciorémonos ahora de la validez de la inecuación estricta

$$\rho|(y_1 - x, y_2 - x)| < \sqrt{(y_1 - x, y_1 - x)} \sqrt{(y_2 - x, y_2 - x)}. \quad (5.93)$$

Con este fin hagamos uso de que para cualesquiera vectores \mathbf{a} y \mathbf{b} del espacio E^m , no colineales uno respecto del otro (es decir, tales que $\mathbf{a} \neq \lambda \mathbf{b}$, cualquiera que sea λ real) se cumple la inecuación estricta de Cauchy-Buniakovski**)

$$|(\mathbf{a}, \mathbf{b})| < \sqrt{(\mathbf{a}, \mathbf{a}) \cdot (\mathbf{b}, \mathbf{b})}.$$

Esto significa que para demostrar la inecuación (5.93) basta convenirse de que los vectores $y_1 - x$ e $y_2 - x$ no son colineales, es decir, convenirse de que para ningún λ real puede verificarse la igualdad

$$y_1 - x = \lambda (y_2 - x). \quad (5.94)$$

Si la igualdad (5.94) fuera válida para un λ tal que $|\lambda| \neq 1$, sería imposible la igualdad $\rho(y_1, x) = \rho(y_2, x)$.

La validez de la igualdad (5.94) para $\lambda = \pm 1$ contradiría a que los puntos y_1 e y_2 son diferentes.

*) Véase, por ejemplo, § 4, cap. 4 del libro *Algebra lineal* de V. A. Il'in y E. Y. Pozniak (en ruso).

***) En efecto, cuando $\mathbf{a} \neq \lambda \mathbf{b}$, el vector $(\mathbf{a} - \lambda \mathbf{b})$ no es nulo. Por eso, el trinomio cuadrado $(\mathbf{a} - \lambda \mathbf{b}, \mathbf{a} - \lambda \mathbf{b}) = (\mathbf{a}, \mathbf{a}) - 2\lambda (\mathbf{a}, \mathbf{b}) + \lambda^2 (\mathbf{b}, \mathbf{b})$ es estrictamente positivo y su discriminante $4[(\mathbf{a}, \mathbf{b})^2 - (\mathbf{a}, \mathbf{a}) \cdot (\mathbf{b}, \mathbf{b})]$ es estrictamente negativo.

Por fin, la validez de la igualdad (5.94) para $\lambda = -1$ significaría que $= x$, lo que $\frac{y_1 + y_2}{2}$ se ha excluido del análisis.

Así, pues, la igualdad (5.94) no se verifica, cualquiera que sea λ real, por lo cual la demostración de la desigualdad (5.93) se da por terminada.

Al cotejar la igualdad (5.92) con la desigualdad (5.93), llegamos a que

$$\begin{aligned} \rho \left(x, \frac{y_1 + y_2}{2} \right) &< \frac{1}{4} |(y_1 - x, y_1 - x)| \\ &+ 2 \sqrt{(y_1 - x, y_1 - x)(y_2 - x, y_2 - x)} + \\ &+ (y_2 - x, y_2 - x) \cdot \left[\frac{1}{4} \left[\sqrt{(y_1 - x, y_1 - x)} + \sqrt{(y_2 - x, y_2 - x)} \right]^2 \right. \\ &\quad \left. - \frac{1}{4} [\rho(x, y_1) + \rho(x, y_2)]^2 - \rho^2(x, y_1) - \rho^2(x, y_2) \right]. \end{aligned}$$

La demostración de la desigualdad (5.91) queda terminada. Pero, esta desigualdad significa que en el conjunto Q hay un punto $\frac{y_1 + y_2}{2}$ más próximo a x que los puntos y_1 e y_2 , lo que contradice al hecho de que cada uno de los puntos y_1 e y_2 es una proyección del punto x sobre el conjunto Q , es decir, es la cota inferior exacta de la distancia $\rho(x, y)$ para toda clase de y pertenecientes a Q .

La contradicción obtenida señala que la suposición de que existen dos proyecciones diferentes y_1 e y_2 del punto x sobre el conjunto Q es errónea.

La demostración del lema 1 ha concluido.

Pasemos ahora a la definición de función convexa.

Definición 2. Una función $f(x)$, definida en un conjunto convexo Q del espacio E^m , se denomina convexa hacia las y negativas, o simplemente convexa sobre dicho conjunto, si para cualesquiera dos puntos x_1 y x_2 del conjunto Q y para todo número real t del segmento $0 \leq t \leq 1$ se cumple la inecuación

$$f[x_2 + t(x_2 - x_1)] \leq f(x_1) + t[f(x_2) - f(x_1)]. \quad (5.95)$$

Definición 3. Una función $f(x)$, definida en un conjunto convexo Q del espacio E^m , se denomina estrictamente convexa sobre este conjunto si para cualesquiera dos puntos x_1 y x_2 del conjunto Q y para todo número real t del intervalo $0 < t < 1$ se cumple una inecuación estricta

$$f[x_2 + t(x_2 - x_1)] < f(x_1) + t[f(x_2) - f(x_1)]. \quad (5.96)$$

Está claro que toda función $f(x)$, estrictamente convexa en el conjunto Q , es convexa en este conjunto.

Es fácil establecer la condición suficiente de convexidad (de convexidad estricta, respectivamente) de la función $f(x)$ dos veces diferenciable sobre el conjunto convexo Q .

En lo que sigue más adelante siempre supondremos que el conjunto Q tiene por lo menos un punto interior.

Lema 2. Supongamos que una función $f(x)$ está definida y es dos veces diferenciable sobre el conjunto convexo Q . Entonces, para que esta función sea convexa (estrictamente convexa) sobre Q , es suficiente que la segunda diferencial d^2f de esta función en todos los puntos de Q sea una forma cuadrática definida casi positiva (definida estrictamente positiva).

DEMOSTRACION. Sean x_1 y x_2 cualesquiera dos puntos fijos del conjunto Q . Veamos en el segmento $0 \leq t \leq 1$ la siguiente función de una variable independiente t :

$$F(t) = f[x_1 + t(x_2 - x_1)] - f(x_1) - t[f(x_2) - f(x_1)]. \quad (5.97)$$

Recordemos que la segunda diferencial d^2f de la función $f(x) = f(x_1, x_2, \dots, x_m)$ de m variables independientes x_1, x_2, \dots, x_m en un punto dado $x = (x_1, x_2, \dots, x_m)$ es *)

$$d^2f(x) = \sum_{i=1}^m \sum_{h=1}^m \frac{\partial^2 f}{\partial x_i \partial x_h}(x) \cdot \Delta x_i \cdot \Delta x_h. \quad (5.98)$$

Al diferenciar la función (5.97) dos veces respecto de t según la regla de diferenciación de una función compuesta, obtendremos

$$F''(t) = \sum_{i=1}^m \sum_{k=1}^m \frac{\partial^2 f}{\partial x_i \partial x_h}[x_1 + t(x_2 - x_1)] (x_i - x_1)^2 (x_h - x_1)^2, \quad (5.99)$$

donde $(x_1^1, x_2^1, \dots, x_m^1)$ y $(x_1^2, x_2^2, \dots, x_m^2)$ son las coordenadas de los puntos x_1 y x_2 , respectivamente.

Al cotejar las relaciones (5.98) y (5.99), nos convencemos de validez de la igualdad

$$F''(t) = d^2f[x_1 + t(x_2 - x_1)], \quad (5.100)$$

donde en la expresión para d^2f los incrementos Δx_i se han tomado iguales a $x_i^2 - x_i^1$.

Realicemos, los razonamientos posteriores, para concretar, para el caso en que la segunda diferencial d^2f en todos los puntos de Q es una forma cuadrática definida casi positiva. En este caso, para todo t del segmento $0 \leq t \leq 1$, el segundo miembro (y, por lo tanto, el primero) de (5.100) es no negativo, es decir, para todo t del segmento $0 \leq t \leq 1$

$$F''(t) \geq 0. \quad (5.101)$$

*) Véase el p. 2, § 5, cap. 5.

En virtud de la definición 2 y la relación (5.97), basta demostrar que para todo t del segmento $0 \leq t \leq 1$ se cumple la inecuación

$$F(t) \leq 0. \quad (5.102)$$

Para demostrar la inecuación (5.102) hagamos uso de la relación (4.101) y las igualdades fáciles de comprobar

$$F(0) = 0, \quad F(1) = 0. \quad (5.103)$$

Supongamos que en el interior del segmento $0 \leq t \leq 1$ existe por lo menos un punto t_0 , en el cual $F(t) > 0$. Entonces, la función $F(t)$ alcanza su valor máximo en el segmento $0 \leq t \leq 1$ en cierto punto interior t_0 de dicho segmento, con la particularidad de que $F(t_0) > 0$. En este punto t_0 la función $F(t)$ tiene un máximo local, y, por eso, $F'(t_0) = 0$. Pero, de la inecuación (5.101) se deduce que la derivada $F'(t)$ no decrece en todo el segmento $0 \leq t \leq 1$, y, por lo tanto, también en el segmento $t_0 \leq t \leq 1$. De aquí y de la condición $F'(t_0) = 0$ se infiere que la derivada $F'(t)$ es no negativa en todo punto del segmento $t_0 \leq t \leq 1$, y, por eso, la función $F(t)$ no decrece en este segmento. Esto nos lleva a la inecuación

$$F(1) \geq F(t_0) > 0,$$

que contradice a la segunda relación de (5.103).

La contradicción obtenida demuestra que la suposición de que en el segmento $0 \leq t \leq 1$ existe por lo menos un punto t , en el cual $F(t) > 0$, es errónea, esto es, demuestra la validez de la inecuación (5.102) en todo punto del segmento $0 \leq t \leq 1$.

Con esto queda demostrada la primera parte del lema (sobre la convexidad de $f(x)$ a condición de que d^2f sea una forma cuadrática definida casi positiva).

La segunda parte del lema (sobre la convexidad estricta de $f(x)$ a condición de que d^2f sea una forma cuadrática definida estrictamente positiva) se demuestra de un modo análogo. Partiendo de la inecuación (5.101), válida esta vez con el signo $>$, y de las inecuaciones (5.103), suponiendo que en el interior del segmento $0 \leq t \leq 1$ existe por lo menos un punto t , en el cual $F(t) \geq 0$, deduciremos que $F(t)$ tiene en el interior del segmento $0 \leq t \leq 1$ un punto de máximo local t_0 ; además, $F(t_0) \geq 0$. Mas, en tal caso, siendo $F'(t_0) = 0$, de (5.101) obtenemos que $F'(t) > 0$ en todo punto del semiintervalo $t_0 < t \leq 1$, lo que significa que $F(1) > F(t_0) \geq 0$.

De nuevo obtenemos una contradicción con la segunda relación de (5.103), la que demuestra que $F(t) < 0$ en todo punto del intervalo $0 < t < 1$, es decir, demuestra la convexidad estricta de $f(x)$ sobre el conjunto Q .

El lema 2 está completamente demostrado.

El lema demostrado conduce, naturalmente, a la idea de exami-

nar la clase siguiente, más estrecha todavía, de funciones convexas en un conjunto convexo Q y dos veces diferenciables en el mismo.

Definición 4. Una función $f(x)$ dos veces diferenciable en el conjunto convexo Q , se denomina fuertemente convexa en dicho conjunto si existe dos constantes positivas k_1 y k_2 tales que la segunda diferencial d^2f de esta función definida mediante la relación (5.98) satisface en todos los puntos x del conjunto Q las inecuaciones

$$k_1 \cdot (\Delta x)^2 \leq d^2 f \leq k_2 \cdot (\Delta x)^2. \quad (5.104)$$

(En estas inecuaciones Δx designa un vector con las coordenadas $(\Delta x_1, \Delta x_2, \dots, \Delta x_m)$, y el símbolo $(\Delta x)^2$ denota el cuadrado escalar del vector mencionado.)

De la inecuación izquierda (5.104) se deduce inmediatamente que la segunda diferencial de una función fuertemente convexa es una función definida estrictamente positiva en todos los puntos del conjunto Q , y, por eso, (en virtud del lema 2) una función convexa sobre el conjunto Q es a ciencia cierta estrictamente convexa en dicho conjunto.

Además, la clase de funciones fuertemente convexas es bastante amplia e importante en los problemas aplicados y nos limitamos a esta clase al exponer la teoría del método gradiental de búsqueda de un mínimo.

Empecemos aclarando la cuestión de existencia y de unicidad del mínimo.

2. Existencia del mínimo de una función fuertemente convexa y unicidad del mínimo de una función estrictamente convexa. Supongamos que una función $f(x)$ está definida sobre un conjunto convexo Q . Diremos que esta función tiene en un punto x_0 del conjunto Q un mínimo local si existe un δ -entorno de este punto x_0 tal que el valor $f(x_0)$ es el mínimo entre los valores de esta función en todos los puntos de intersección del δ -entorno de x_0 con el conjunto Q .

Admitida esta definición, el concepto de mínimo local incluye también los puntos del mínimo de contorno de la función $f(x)$ en la frontera del conjunto Q .

De este modo, teniendo presente la definición dada, podemos subdividir los puntos de mínimo en puntos de mínimo local interior (cuando dichos puntos son puntos interiores de Q) y puntos de mínimo local de contorno (cuando estos puntos son puntos de frontera de Q).

Para estudiar el problema de existencia y unicidad de un punto de mínimo local nos hará falta el siguiente teorema auxiliar.

Lema 3. Supongamos que en un conjunto convexo Q viene dada una función convexa diferenciable $f(x)$. Para que esta función tenga un mínimo local en el punto x_0 del conjunto Q , es necesario y suficiente que para cualquier vector Δx , para el cual el punto $x_0 + \Delta x$ pertenece al

conjunto Q , se cumpla la inecuación *) que sigue

$$(\text{grad } f(x_0), \Delta x) \geq 0. \quad (5.105)$$

DEMOSTRACIÓN. 1) NECESIDAD. En virtud de la afirmación demostrada en el p. 6, § 4, cap. 5, el primer miembro de (5.105) es igual al producto de la derivada de la función $f(x)$ en el punto x_0 según la dirección del vector Δx por la longitud $|\Delta x|$ de este vector:

$$(\text{grad } f(x_0), \Delta x) = \frac{\partial f}{\partial e}(x_0) |\Delta x|, \quad (14.106)$$

donde $e = \frac{\Delta x}{|\Delta x|}$ es el vector unidad en dirección de Δx .

Por cuanto x_0 es el punto de mínimo local de la función $f(x)$, la derivada $\frac{\partial f}{\partial e}(x_0)$ según cualquier dirección de $e = \frac{\Delta x}{|\Delta x|}$ es no negativa (con mayor precisión, es igual a cero cuando x_0 es un punto de extremo local interior, y no negativa cuando x_0 es un punto de extremo local de contorno).

Así, pues, el segundo miembro de (5.106) (y, por eso, también el primer miembro de (5.105)) es no negativo. La necesidad está demostrada.

2) SUFICIENCIA. Supongamos que para todo vector Δx , para el cual el punto $x_0 + \Delta x$ pertenece a Q , se cumple la inecuación (5.105). Demostremos que el punto x_0 es un punto de mínimo local de la función $f(x)$.

Por cuanto la función $f(x)$ es, de acuerdo con la hipótesis, convexa sobre el conjunto Q , para cualesquiera dos puntos x_1 e x_2 de dicho conjunto y cualquier número t del segmento $0 \leq t \leq 1$ se cumple la desigualdad (5.95). Al suponer en esta desigualdad $x_1 = x_0$, $x_2 = x_0 + \Delta x$, podemos escribir esta desigualdad en la forma

$$f(x_0 + \Delta x) - f(x_0) \geq \frac{f(x_0 + t\Delta x) - f(x_0)}{t}. \quad (5.107)$$

Considerando fijos x_0 y Δx , pasemos en la desigualdad (5.107) al límite para $t \rightarrow 0 + 0$. Por definición de la derivada direccional (véase p. 6, § 4, cap. 5), el límite del segundo miembro de (5.107) para $t \rightarrow 0 + 0$ es exactamente igual al producto del segundo miembro de (5.106). Por eso, en virtud de las relaciones (5.105) y (5.106), este límite es no negativo. Considerando que el primer miembro de (5.107) no depende de t , de la (5.107), en límite para $t \rightarrow 0 + 0$ inecuación se infiere que

$$f(x_0 + \Delta x) - f(x_0) \geq 0.$$

*) En la inecuación (5.105) se toma el producto escalar de los vectores $\text{grad } f(x_0)$ y Δx . Véase en el p. 6, § 4, cap. 5 la definición de $\text{grad } f(x_0)$.

La última desigualdad, lícita para cualquier vector Δx , para el cual el punto $x_0 + \Delta x$ pertenece a Q , demuestra que la función $f(x)$ tiene en el punto x_0 un mínimo local. La suficiencia está demostrada.

El lema 3 está completamente demostrado.

OBSERVACION 1. La demostración aducida evidencia que para el caso en que x_0 es un punto interior del conjunto Q , es decir, cuando se trata de un mínimo local interior, en la formulación del lema 3 el signo \geq de la inequación (5.105) puede ser cambiado por el signo $=$.

OBSERVACION 2. Al demostrar la necesidad del lema 3, no hemos recurrido a la exigencia de convexidad de la función $f(x)$. Por eso la necesidad se demuestra sin la exigencia de que la función $f(x)$ sea convexa. De otras palabras, es válida la afirmación siguiente: si una función $f(x)$ es diferenciable en un conjunto convexo Q y tiene un mínimo local en el punto interior (de frontera) x_0 de dicho conjunto, para cualquier vector Δx , en cuyo caso el punto $x_0 + \Delta x$ pertenece a Q , se cumple la desigualdad

$$(\text{grad } f(x_0), \Delta x) = 0 \quad [(\text{grad } f(x_0), \Delta x) \geq 0].$$

Pasemos a la cuestión de unicidad y existencia de un punto de mínimo local.

Teorema (de la unicidad del mínimo local de una función estrictamente convexa). Si una función $f(x)$ es diferenciable y estrictamente convexa sobre un conjunto convexo Q , puede tener un mínimo local solamente en un único punto de este conjunto.

DEMOSTRACION. Supongamos que la función $f(x)$ tiene un mínimo local en dos puntos diferentes x_1 y x_2 del conjunto Q . Entonces, la condición de convexidad (5.95) para los puntos x_1 y x_2 puede escribirse en la forma

$$f(x_2) - f(x_1) \geq \frac{f[x_1 + t(x_2 - x_1)] - f(x_1)}{t} \quad (5.108)$$

(aquí, t es un número cualquiera del segmento $0 \leq t \leq 1$).

Al cambiar de papel los puntos x_1 y x_2 en la relación (5.108), obtendremos la inequación

$$f(x_1) - f(x_2) \geq \frac{f[x_2 + t(x_1 - x_2)] - f(x_2)}{t}. \quad (5.109)$$

Al pasar en límite para $t \rightarrow 0 + 0$, el segundo miembro en (5.108) (en el segundo miembro de (5.109), respectivamente) proporciona la derivada de la función $f(x)$ según la dirección del vector $x_2 - x_1$ (del vector $x_1 - x_2$, respectivamente), tomada en el punto x_1 (en el punto x_2 , respectivamente), multiplicada por $|x_2 - x_1|$. Por cuanto ambos puntos x_1 y x_2 son puntos de mínimo local, ambas derivadas direccionales citadas son no negativas, es decir, los límites de los segundos miembros de (5.108) y (5.109) para $t \rightarrow 0 + 0$ ambos son no negativos.

De este modo, de las desigualdades (5.108) y (5.109) obtenemos en límite para $t \rightarrow 0 + 0$:

$$f(x_2) - f(x_1) \geq 0, \quad f(x_1) - f(x_2) \geq 0.$$

La comparación de las últimas desigualdades nos conduce a la conclusión de que $f(x_1) = f(x_2)$.

Haciendo uso de la igualdad $f(x_1) = f(x_2)$, obtendremos, partiendo de la condición de convexidad estricta (5.96), que

$$f[x_1 + t(x_2 - x_1)] < f(x_1) \quad (5.110)$$

para todo t del intervalo $0 < t < 1$.

La desigualdad (5.110) contradice al hecho de que la función $f(x)$ tiene un mínimo local en el punto x_1 (en el punto $x_1 + t(x_2 - x_1)$, tan próximo como se quiera al punto x_1 para t pequeño, la función $f(x)$ tiene un valor inferior al de $f(x_1)$).

La contradicción obtenida muestra que nuestra suposición de que la función $f(x)$ tiene un mínimo local en dos puntos diferentes del conjunto Q es errónea. El teorema está demostrado.

Mostremos la existencia de un mínimo local para unas restricciones más rigurosas que las adoptadas al demostrar la unicidad.

Teorema (de la existencia de un mínimo local de una función). Si una función $f(x)$ es fuertemente cóncava sobre un conjunto cerrado convexo Q , dicha función tiene en el conjunto Q un punto x_0 de mínimo local*).

DEMOSTRACION Notemos, al principio, que el teorema es a ciencia cierta válido para el caso en que el conjunto cerrado convexo Q es, además, acotado. Entonces, de acuerdo con el segundo teorema de Weierstrass (véase teorema 5.7), la función $f(x)$, siendo en todo caso continua sobre el conjunto Q , alcanza en cierto punto x_0 de dicho conjunto su valor mínimo en Q . Este punto x_0 será precisamente el punto de mínimo local.

Resta demostrar el teorema para el caso cuando el conjunto cerrado convexo Q no está acotado. Fijemos cierto punto interior x_1 del conjunto Q y desarrollemos la función $f(x)$ según la fórmula de Taylor con centro en el punto x_1 , tomando el término residual $R_2(x)$ en forma de Lagrange **) (véase p. 3, § 5, cap. 5). El desarrollo citado tendrá la forma:

$$f(x) = f(x_1) + df(x_1) + \frac{1}{2} d^2f[x_1 + \theta(x - x_1)] \quad (5.111)$$

*) Por cuanto la función $f(x)$, fuertemente cóncava sobre un conjunto convexo Q , es estrictamente fuerte en dicho conjunto, de acuerdo con el teorema anterior, el punto x_0 será el único punto de mínimo local.

**) Tenemos en cuenta que la función $f(x)$, fuertemente cóncava en el conjunto Q , es dos veces diferenciable en este conjunto.

donde θ es un número del intervalo $0 < \theta < 1$, de suerte que el punto $x_1 + \theta(x - x_1)$ pertenece al segmento que une los puntos x_1 y x^* .

Si designamos con Δx el vector $x - x_1$, para $df(x_1)$ se verificará la igualdad

$$df(x_1) = (\text{grad } f(x_1), \Delta x).$$

De esta igualdad se deduce que

$$|df(x_1)| \leq |\text{grad } f(x_1)| \cdot |\Delta x| \quad (5.112)$$

Ahora, aprovechando la desigualdad izquierda en la definición de convexidad fuerte (5.104), llegaremos a la desigualdad

$$d^2f[x_1 + \theta(x - x_1)] \geq k_1 \cdot (\Delta x)^2. \quad (5.113)$$

De las relaciones (5.111)–(5.113) concluimos que

$$\begin{aligned} f(x) - f(x_1) &\geq -|df(x_1)| + \frac{1}{2} d^2f[x_1 + \theta(x - x_1)] \geq \\ &\geq -|\text{grad } f(x_1)| \cdot |\Delta x| + \frac{k_1}{2} |\Delta x|^2, \end{aligned}$$

de suerte que

$$f(x) - f(x_1) \geq |\Delta x| \cdot \left[\frac{k_1}{2} |\Delta x| - |\text{grad } f(x_1)| \right]. \quad (5.114)$$

Considerando que el punto x_1 es fijo y la magnitud $|\text{grad } f(x_1)|$ es cierto número fijo, podemos a ciencia cierta elegir un número positivo R tan grande que para $|\Delta x| > R$ la expresión encerrada entre corchetes en (5.114) será positiva.

Esto significa que, cuando $|\Delta x| > R$, se cumple la desigualdad $f(x) > f(x_1)$, es decir, fuera de una esfera cerrada C_R de radio R con centro en el punto x_1 los valores de $f(x)$ sobrepasan el valor de $f(x_1)$ (en el centro de la esfera mencionada).

Denotemos con Q_R la intersección del conjunto Q con dicha esfera C_R . Por cuanto ambos conjuntos, Q y C_R , son convexos y cerrados, su intersección Q_R será también convexa y cerrada. Además, puesto que el conjunto Q_R es acotado, entonces, de conformidad con lo demostrado más arriba, la función $f(x)$ tiene en Q_R el único punto x_0 de mínimo local.

Teniendo en cuenta lo demostrado acerca de que en todo punto de Q , dispuesto fuera de los márgenes de Q_R , los valores de $f(x)$ son superiores a los de $f(x_1)$, estos valores con mayor razón son superiores a $f(x_0)$, es decir, el punto x_0 es punto de mínimo local de $f(x)$ tam-

*) Cualquiera que sea el punto x del conjunto Q , el segmento que une los puntos x y x_1 pertenece al conjunto Q , por ser dicho conjunto convexo. En la nota al pie de la página para el teorema de Taylor (5.15) se ha observado que a título de entorno del centro de desarrollo puede tomarse cualquier entorno estelar, es decir, puede tomarse todo el conjunto Q .

bién sobre todo el conjunto Q . El teorema está completamente demostrado.

3. Búsqueda del mínimo de una función fuertemente convexa. Se ha demostrado que una función fuertemente convexa $f(x)$, definida sobre un conjunto cerrado convexo Q , tiene en este conjunto un único punto x_0 de mínimo local.

Volvamos a la construcción y argumentación del algoritmo, con cuya ayuda se busca este punto x_0 .

Fijemos un punto arbitrario x_1 del conjunto Q y un número arbitrario α que satisface las desigualdades

$$0 < \alpha < 2/k_2, \quad (5.115)$$

donde k_2 es la constante de la desigualdad (5.104) que determina la convexidad fuerte de la función $f(x)$.

Considerando x_1 como la primera aproximación, formemos una sucesión iterativa $\{x_h\}$ con ayuda de una relación recurrente

$$x_{h+1} = P_Q(x_h - \alpha \text{grad } f(x_h)), \quad h = 1, 2, \dots \quad (5.116)$$

En el presente punto demos demos la afirmación siguiente.

Teorema fundamental. *Supongamos que la función $f(x)$ es fuertemente convexa sobre el conjunto cerrado convexo Q y sea x_0 un punto arbitrario del conjunto Q . Entonces, la sucesión iterativa $\{x_h\}$, definida mediante la relación recurrente (5.116), converge hacia el punto x_0 de mínimo local de la función $f(x)$, cualquiera que sea α que satisficase las igualdades (5.115).*

Recalquemos que este teorema nos ofrece el algoritmo para hallar cualquier mínimo local (interior o de contorno) de la función $f(x)$ fuertemente convexa sobre todo conjunto arbitrario cerrado y convexo Q (no es preciso que esté acotado).

Antes de demostrar el teorema fundamental aduciremos cuatro lemas.

Lema 4. *Si Q es un conjunto cerrado convexo de puntos en E^m , y x es un punto fijo arbitrario de E^m , siendo y un punto arbitrario de Q , entonces*

$$(x - P_Q(x), y - P_Q(x)) \leq 0. \quad (5.117)$$

Demostración. Supongamos que la desigualdad (5.117) no es cierta. Entonces, existe un punto y del conjunto Q tal que

$$(x - P_Q(x), y - P_Q(x)) > 0. \quad (5.118)$$

De (5.118) se desprende en seguida que el punto y no coincide con $P_Q(x)$.

Por ser el conjunto Q convexo, cualquier punto $z = P_Q(x) + t(y - P_Q(x))$ del segmento que une los puntos $P_Q(x)$ e y , pertenece al conjunto Q . Calculemos la distancia entre cualquier

punto de este género z y el punto x

$$\rho^2(z, x) = (x - P_Q(x) - t(y - P_Q(x)), x - P_Q(x) - t(y - P_Q(x))) = \rho^2(x, P_Q(x)) - 2t(x - P_Q(x), y - P_Q(x)) + t^2\rho^2(y, P_Q(x)). \quad (5.119)$$

Por cuanto x e y son fijos y t es cualquier número del segmento $0 \leq t \leq 1$, entonces, en virtud de las desigualdades (5.118), podemos elegir un t tal que satisfará la desigualdad

$$0 < t < \frac{2(x - P_Q(x), y - P_Q(x))}{\rho^2(y, P_Q(x))}.$$

Eligiendo t de tal manera, obtenemos

$$-2t(x - P_Q(x), y - P_Q(x)) + t^2\rho^2(y, P_Q(x)) < 0,$$

y deducimos de (5.119) que $\rho^2(z, x) < \rho^2(x, P_Q(x))$. La última desigualdad contradice a que el punto $P_Q(x)$ es una proyección del punto x sobre el conjunto Q : en el conjunto Q hay un punto z que está alejado de x a una distancia inferior a la que existe entre $P_Q(x)$ y x . La contradicción obtenida finaliza la demostración del lema.

Lema 5. *Supongamos que $f(x)$ es diferenciable y convexa en un conjunto cerrado convexo Q . Si para cierto α positivo la proyección $P_Q(x_0 - \alpha \cdot \text{grad } f(x_0))$ del punto $x_0 - \alpha \cdot \text{grad } f(x_0)$ sobre el conjunto Q coincide con el punto x_0 de este conjunto, la función $f(x)$ tiene en el punto x_0 un mínimo local.*

DEMOSTRACION. Haciendo uso del lema 4, escribamos la desigualdad (5.117) para los puntos $x = x_0 - \alpha \cdot \text{grad } f(x_0)$ e $y = x_0 + \Delta x$, donde Δx es cualquier vector, para el cual el punto $y = x_0 + \Delta x$ pertenece a Q . Como resultado obtenemos:

$$(x_0 - \alpha \cdot \text{grad } f(x_0) - P_Q(x_0 - \alpha \cdot \text{grad } f(x_0)), x_0 + \Delta x - P_Q(x_0 - \alpha \cdot \text{grad } f(x_0))) \leq 0.$$

Considerando que $P_Q(x_0 - \alpha \cdot \text{grad } f(x_0)) = x_0$, obtenemos de la última desigualdad la relación siguiente:

$$(\text{grad } f(x_0), \Delta x) \geq 0.$$

Esta relación, válida para cualquier vector Δx , para el cual el punto $x_0 + \Delta x$ pertenece a Q , establece, en virtud del lema 3, que la función $f(x)$ tiene en el punto x_0 un mínimo local. El lema 5 está demostrado.

Supongamos que la función $f(x)$ es fuertemente convexa sobre un conjunto convexo cerrado y acotado Q . Denotemos con m el valor mínimo de $f(x)$ sobre el conjunto Q y con μ , un número estrictamente superior a m , de suerte que

$$\mu > m = \min_{x \in Q} f(x).$$

Fijamos un número v , estrictamente superior a μ , y denotemos con \hat{Q} un subconjunto de aquellos puntos x del conjunto Q , para los cuales

$$\mu \leq f(x) \leq v. \quad (5.120)$$

El conjunto \hat{Q} , siendo un subconjunto del conjunto acotado Q , es acotado de por sí.

Cerciorémonos de que el conjunto \hat{Q} es *cerrado*. Sea $\{x_k\}$ una sucesión convergente arbitraria de puntos del conjunto \hat{Q} . Se pide demostrar que el límite x_0^* de esta sucesión también pertenece al conjunto \hat{Q} . Por cuanto cada punto x_k pertenece al conjunto \hat{Q} , tenemos para todo número k

$$\mu \leq f(x_k) \leq v. \quad (5.121)$$

La función estrictamente convexa $f(x)$ es en todo caso continua en Q , y, por eso, dado que la sucesión $\{x_k\}$ converge hacia x_0 se deduce, en virtud de la definición de continuidad de una función, la convergencia de la sucesión $\{f(x_k)\}$ hacia el número $f(x_0)$. Por cuanto todos los elementos de la sucesión numérica convergente $\{f(x_k)\}$ satisfacen las desigualdades (5.121), el límite $f(x_0)$ de esta sucesión también satisface las inecuaciones $\mu \leq f(x_0) \leq v$ (véase cap. 3, corolario 2 del teorema 3.13, t. 1). Precisamente esto último significa que el punto x_0 pertenece al conjunto \hat{Q} . Hemos demostrado, pues, el carácter cerrado del conjunto \hat{Q} .

Así, pues, el conjunto \hat{Q} de todos los puntos x del conjunto Q , para los cuales son válidas las desigualdades (5.120), es cerrado y acotado.

Demostremos el lema siguiente.

Lema 6. *Supongamos que la función $f(x)$ es fuertemente convexa en un conjunto cerrado convexo Q ; x es un punto cualquiera de Q y α , cualquier número positivo; el símbolo Δx denota la diferencia*

$$\Delta x = P_Q(x - \alpha \cdot \text{grad } f(x)) - x. \quad (5.122)$$

Entonces se cumple la desigualdad

$$(\text{grad } f(x), \Delta x) \leq \frac{1}{\alpha} |\Delta x|. \quad (5.123)$$

Si, además, el conjunto Q es acotado y el punto x pertenece al subconjunto \hat{Q} de aquellos puntos de Q , para los cuales se cumplen las desigualdades (5.120) cuando $\mu > \min f(x)$, se encontrará un número estrictamente

*) Por cuanto el conjunto de partida Q es cerrado, el límite x_0 en todo caso pertenece a Q .

tamente positivo y tal que se cumple la desigualdad

$$|\Delta x| \geq \gamma. \quad (5.124)$$

DEMOSTRACION. Demostremos primero la desigualdad (5.123). Fijemos arbitrariamente un punto x del conjunto Q y, recurriendo al lema 4, escribamos la desigualdad (5.117), tomando en la misma en lugar de x un punto $x - \alpha \cdot \text{grad } f(x)$, y en lugar de y , el punto x . Obtendremos la desigualdad

$$(x - \alpha \cdot \text{grad } f(x) - P_Q(x - \alpha \cdot \text{grad } f(x)), x - P_Q(x - \alpha \cdot \text{grad } f(x))) \leq 0,$$

la cual, considerando las designaciones $\Delta x = P_Q(x - \alpha \cdot \text{grad } f(x)) - x$, se escribirá en la forma

$$(-\alpha \cdot \text{grad } f(x) - \Delta x, -\Delta x) \leq 0.$$

De la última desigualdad y de las propiedades del producto escalar se deduce que

$$\alpha (\text{grad } f(x), \Delta x) + |\Delta x|^2 \leq 0,$$

lo que nos conduce a la desigualdad (5.123).

Queda por demostrar que bajo un supuesto adicional de que Q es acotado y que x pertenece al subconjunto \hat{Q} , existe $\gamma > 0$ tal que se cumple la desigualdad (5.124).

Veamos una función no negativa del punto x de la forma

$$|\Delta x| = |P_Q(x - \alpha \cdot \text{grad } f(x)) - x|. \quad (5.125)$$

Cerciorémonos de que esta función es función del punto x , continua sobre el conjunto Q .

Demostremos al principio que la función vectorial $P_Q(x)$ es una función continua del punto x . Con este fin basta probar la desigualdad

$$|P_Q(x + \Delta x) - P_Q(x)| \leq |\Delta x|, \quad (5.126)$$

que se cumple para cualesquiera vectores x y Δx .

En virtud del lema 4, son válidas las desigualdades

$$(x - P_Q(x), P_Q(x + \Delta x) - P_Q(x)) \leq 0,$$

$$(x + \Delta x - P_Q(x + \Delta x), P_Q(x) - P_Q(x + \Delta x)) \leq 0.$$

Haciendo uso de estas desigualdades y de la desigualdad de Cauchy-Buniakovski, obtendremos una cadena de relaciones

$$\begin{aligned} |P_Q(x + \Delta x) - P_Q(x)|^2 &= \\ &= (P_Q(x + \Delta x) - P_Q(x), P_Q(x + \Delta x) - P_Q(x)) = \\ &= (P_Q(x + \Delta x) - x, P_Q(x + \Delta x) - P_Q(x)) + \\ &\quad + (x - P_Q(x), P_Q(x + \Delta x) - P_Q(x)) \leq \\ &\leq (P_Q(x + \Delta x) - x, P_Q(x + \Delta x) - P_Q(x)) = \end{aligned}$$

$$\begin{aligned}
 &= (P_Q(x + \Delta x) - x - \Delta x, P_Q(x + \Delta x) - P_Q(x)) + \\
 &+ (\Delta x, P_Q(x + \Delta x) - P_Q(x)) \leq (\Delta x, P_Q(x + \Delta x) - \\
 &\quad - P_Q(x)) \leq |\Delta x| \cdot |P_Q(x + \Delta x) - P_Q(x)|,
 \end{aligned}$$

de la cual se deduce la desigualdad (5.126).

Se ha demostrado, pues, que $P_Q(x)$ es una función vectorial continua del punto x . De la convexidad fuerte de $f(x)$ en Q proviene que $\alpha \cdot \text{grad } f(x)$ es también una función vectorial continua en Q del punto x . Mas, en tal caso, del teorema de continuidad de una función compuesta y de la continuidad de la diferencia de las funciones continuas se desprende que la función

$$P_Q(x - \alpha \cdot \text{grad } f(x)) - x$$

es en Q una función vectorial continua del punto x .

El módulo de la citada función vectorial, es decir, la función escalar (5.125) es con mayor razón continua sobre el conjunto Q y, por eso, sobre el subconjunto \hat{Q} del mismo.

Así, pues, la función (5.125) es continua y no negativa en todo punto sobre el conjunto acotado cerrado \hat{Q} . En este caso, de acuerdo con el segundo teorema de Weierstrass (véase teorema 5.7), esta función alcanza en el conjunto \hat{Q} su valor mínimo no negativo γ . El valor mínimo mencionado γ es a ciencia cierta estrictamente positivo, pues, si γ fuera igual a cero, en el conjunto \hat{Q} se encontraría un punto x_0 tal que $P_Q(x_0 - \alpha \cdot \text{grad } f(x_0)) - x_0 = 0$, y esto significaría, en virtud del lema 5, que en este punto x_0 del conjunto \hat{Q} la función $f(x)$ tiene sobre el conjunto Q un único mínimo local (este mínimo se dispone, por definición de \hat{Q} , fuera de \hat{Q}). El lema 6 queda completamente demostrado.

Lema 7. *Supongamos que una función $f(x)$ es fuertemente convexa sobre un conjunto cerrado convexo Q : x es un punto cualquiera de Q ; α es cualquier número que satisface las desigualdades (5.115); Δx es la diferencia de tipo (5.122). Entonces, al pasar del punto x al punto $x^* = P_Q(x - \alpha \cdot \text{grad } f(x))$, el valor de la función $f(x)$ no crece, con la particularidad de que **

$$f(x) - f(x^*) \geq \left(\frac{1}{\alpha} - \frac{k_2}{2} \right) |\Delta x|^2. \quad (5.127)$$

Si, además, el conjunto Q está acotado y el punto x pertenece al subconjunto \hat{Q} de aquellos puntos de Q , para los cuales se cumple la desigualdad (5.120) cuando $\mu > \min_{x \in \hat{Q}} f(x)$, la desigualdad (5.127) se convierte

* De (5.115) se deduce que $\left(\frac{1}{\alpha} - \frac{k_2}{2} \right) > 0$.

en una desigualdad

$$f(x) - f(x^*) \geq \left(\frac{1}{\alpha} - \frac{k_2}{2} \right) \gamma^2. \quad (5.128)$$

donde $\gamma > 0$ es la constante del lema 6.

DEMOSTRACIÓN. Basta establecer para cualquier punto x del conjunto Q la desigualdad (5.127), pues de esta desigualdad y de la (5.124) se deduce en seguida la desigualdad (5.128) (para los puntos x pertenecientes a \hat{Q} , a condición de que Q esté acotado).

Demostremos al principio la desigualdad (5.127) para un caso en que el punto x es punto interior del conjunto Q . Considerando que el punto $x^* = P_Q(x - \alpha \cdot \text{grad } f(x))$ pertenece al conjunto Q , sobre el cual la función $f(x)$ es fuertemente convexa, expresemos el valor $f(x^*)$ según la fórmula de Taylor con centro en el punto x , tomando el término residual $R_2(x^*)$ en forma de Lagrange. En este caso obtenemos

$$f(x^*) = f(x) + (\text{grad } f(x), \Delta x) + \frac{1}{2} d^2 f(x + \theta \Delta x), \quad (5.129)$$

donde $\Delta x = x^* - x = P_Q(x - \alpha \cdot \text{grad } f(x)) - x$, $0 < \theta < 1$.

Utilizando la desigualdad (5.123) y la desigualdad derecha de (5.104), de la fórmula de Taylor (5.129) obtenemos

$$f(x^*) - f(x) \leq -\frac{1}{\alpha} |\Delta x|^2 + \frac{k_2}{2} \cdot |\Delta x|^2,$$

de modo que para el caso de un punto interior x la desigualdad (5.127) queda demostrada.

Sea, ahora x un punto de frontera del conjunto Q . Por definición de punto de frontera, existe una sucesión $\{x_n\}$ de puntos interiores del conjunto Q que converge hacia x .

Para todo punto x_n obtenemos, rigiéndonos por la fórmula de Taylor con centro en dicho punto:

$$f(x^*) = f(x_n) + (\text{grad } f(x_n), x^* - x_n) + \frac{1}{2} d^2 f \times \\ \times [x_n + \theta_n(x^* - x_n)], \quad (5.130)$$

donde $0 < \theta_n < 1$.

Teniendo en cuenta que la desigualdad derecha de (5.104) es válida para $d^2 f$ en cualquier punto del conjunto Q y que $\text{grad } f(x)$ es una función vectorial continua del punto x sobre el conjunto Q , llegamos a que en límite, para $n \rightarrow \infty$, de la relación (5.130) se deduce la validez de la desigualdad (5.127) para el punto de frontera x del conjunto Q .

El lema está demostrado.

Pasemos ahora directamente a la demostración del teorema fundamental.

Demostremos primero el teorema fundamental suponiendo adicionalmente que un conjunto cerrado convexo Q es también acotado.

Tomemos un punto arbitrario x_1 del conjunto Q y formemos una sucesión iterativa $\{x_n\}$ de puntos definidos mediante la relación recurrente (5.116), a condición de que α satisfaga las desigualdades (5.115).

Del lema 7, y, con mayor precisión, de la desigualdad (5.127) proviene en seguida que

$$f(x_k) - f(x_{k+1}) \geq \left(\frac{1}{\alpha} - \frac{k_2}{2} \right) \cdot |x_k - x_{k+1}|^2 \geq 0.$$

De este modo, la sucesión $\{f(x_k)\}$ es no creciente. Por cuanto, además, esta sucesión está acotada inferiormente (por el valor mínimo m de la función $f(x)$ sobre el conjunto Q), será convergente (véase teorema 3.15 del cap. 3, t.1. Denotemos con μ el límite de la sucesión $\{f(x_n)\}$. Está claro que $\mu \geq m$, donde m es el valor mínimo de $f(x)$ sobre el conjunto Q . Además, por cuanto todos los términos de la sucesión convergente no creciente no son menores que su límite (véase la observación 3 al teorema 3.15, cap. 3, t.1, para todos los números k se cumple la desigualdad

$$f(x_k) \geq \mu. \quad (5.131)$$

Demostremos que para el límite μ se verifica la igualdad $\mu = m = \min_{x \in Q} f(x)$

Supongamos que esta igualdad no tiene lugar, es decir, admitamos que $\mu > m$. Entonces, si denotamos con v el valor máximo de $f(x)$ en el conjunto Q y con \hat{Q} , un subconjunto de puntos de Q , para los cuales se cumplen las desigualdades (5.120), entonces, en virtud del lema 7, existe una constante estrictamente positiva γ tal que se cumple la desigualdad (5.128) que conduce a la desigualdad siguiente:

$$f(x_k) - f(x_{k+1}) \geq \left(\frac{1}{\alpha} - \frac{k_2}{2} \right) \gamma^2 > 0. \quad (5.132)$$

válida para cualquier número k .

Al sumar las desigualdades (5.132), escritas para los números k iguales a 1, 2, 3, . . . , $(n-1)$, obtenemos que para todo número n

$$f(x_1) - f(x_n) \geq (n-1) \left(\frac{1}{\alpha} - \frac{k_2}{2} \right) \gamma^2$$

o bien, que es lo mismo,

$$f(x_n) \leq f(x_1) - (n-1) \left(\frac{1}{\alpha} - \frac{k_2}{2} \right) \gamma^2. \quad (5.132')$$

Las desigualdades (5.132'), válidas para todo número n , contradicen a la desigualdad (5.131), pues, la magnitud del miembro derecho de

(5.132') se hace inferior al número μ para n suficientemente grande.

La contradicción obtenida quiere decir que nuestra suposición de que $\mu > m$ es errónea, es decir, demuestra que $\mu = m$. Así, pues, se ha demostrado que la sucesión $\{f(x_k)\}$ converge hacia el valor mínimo m de la función $f(x)$ sobre el conjunto Q .

Queda demostrar que la propia sucesión iterativa $\{x_k\}$ converge hacia el punto x_0 , en el cual se alcanza este valor mínimo *).

Fijamos arbitrariamente un número positivo ε y designemos con C_ε una esfera abierta m -dimensional de radio ε con centro en el punto x_0 .

Luego, denotemos con \hat{Q}_ε aquella parte del conjunto Q que no contiene puntos de la esfera C_ε . Está claro que \hat{Q}_ε es un conjunto cerrado acotado, de modo que la función $f(x)$ alcanza en este conjunto (en virtud del segundo teorema de Weierstrass) su valor mínimo que se designará por m_ε .

Podemos afirmar que $m_\varepsilon > m$, puesto que, de lo contrario, se perturbaría la condición de existencia sobre el conjunto Q del único punto de mínimo local de la función $f(x)$.

Ahora, se puede afirmar también que sobre el conjunto \hat{Q}_ε se tienen sólo un número *finito* de puntos de la sucesión $\{x_k\}$ (puesto que la sucesión $\{f(x_k)\}$ que cuenta con un número infinito de elementos que satisfacen la desigualdad $f(x_k) \geq m_\varepsilon > m$, no puede converger hacia el número m).

Por consiguiente, hemos demostrado que para todo $\varepsilon > 0$ existe un número N , a partir del cual todos los elementos de la sucesión $\{x_k\}$ se disponen dentro de la esfera C_ε de radio ε con centro en el punto x_0 . Esto significa precisamente que la sucesión $\{x_k\}$ converge hacia el punto x_0 .

Con ello queda demostrado el teorema fundamental para el caso del conjunto cerrado convexo acotado Q .

Ahora, sea Q un conjunto cerrado convexo *no acotado*. De nuevo fijamos arbitrariamente un punto x_1 de este conjunto y formamos la sucesión iterativa (5.116), a condición de que α satisfaga las desigualdades (5.115).

Al demostrar el teorema de existencia de mínimo local de una función fuertemente convexa (véase p. 2) se ha establecido que el punto x_0 de mínimo local de la función fuertemente convexa $f(x)$ sobre el conjunto cerrado convexo no acotado Q se dispone en aquella parte Q_R del conjunto Q que se contiene en la esfera C_R con centro en el punto x_1 , cuyo radio R está elegido de la condición

$$\frac{k_1}{2} R - |\text{grad } f(x_1)| > 0.$$

*1) Ya hemos demostrado que el valor mínimo de la función $f(x)$ sobre el conjunto Q se alcanza en el único punto de este conjunto.

En aquella circunstancia establecimos también que el subconjunto Q_R del conjunto Q es un conjunto cerrado convexo acotado y que en todo punto fuera de Q_R los valores de $f(x)$ son superiores a $f(x_1)$.

En virtud del lema 7 (y, con mayor precisión, en virtud de la desigualdad (5.127)), la sucesión $\{f(x_k)\}$ es no creciente, mientras que fuera de los márgenes de Q_R todos los valores de $f(x)$ sobrepasan los de $f(x_1)$, razón por la cual todos los puntos de la sucesión iterativa $\{x_k\}$ se disponen en Q_R y, por eso, para todo número k

$$P_Q(x_k - \alpha \cdot \text{grad } f(x_k)) = P_{Q_R}(x_k - \alpha \cdot \text{grad } f(x_k)).$$

Por consiguiente, la sucesión iterativa (5.116) puede ser sustituida por

$$x_{k+1} = P_{Q_R}(x_k - \alpha \cdot \text{grad } f(x_k)),$$

después de lo cual todos los razonamientos posteriores se reducen al conjunto cerrado convexo acotado Q_R , es decir, al caso ya examinado más arriba.

El teorema fundamental está demostrado por completo.

OBSERVACIÓN 1. Con una sencillez particular se expresa la sucesión (5.116) para el caso en que el conjunto Q coincide con todo el espacio E^m . En este caso para todo punto x se verifica la igualdad $P_Q(x) = x$ y, por eso, la fórmula recurrente (5.116) tendrá la forma

$$x_{k+1} = x_k - \alpha \cdot \text{grad } f(x_k)$$

OBSERVACIÓN 2. El método expuesto permite buscar (si se cumplen las condiciones correspondientes) la solución $\overset{\circ}{x}_0 = (\overset{\circ}{x}_1, \overset{\circ}{x}_2, \dots, \overset{\circ}{x}_m)$ de un sistema de m ecuaciones funcionales:

$$\begin{cases} f_1(x) = f_1(x_1, x_2, \dots, x_m) = 0, \\ f_2(x) = f_2(x_1, x_2, \dots, x_m) = 0, \\ \dots \\ f_m(x) = f_m(x_1, x_2, \dots, x_m) = 0. \end{cases}$$

Basta señalar que la solución del sistema citado coincide con el punto de mínimo local de la función

$$f(x) = f_1^2(x) + f_2^2(x) + \dots + f_m^2(x).$$

OBSERVACIÓN 3. La teoría de búsqueda de un mínimo local de una función fuertemente convexa (hacia las y negativas), expuesta aquí, puede aplicarse sin complicaciones algunas al problema de búsqueda del mínimo local de una función fuertemente convexa (hacia las y positivas).

Complemento

Sobre la elección de la partición óptima de un segmento para el cálculo aproximado de una integral

En el capítulo 3 un segmento $[a, b]$ se dividía, con el fin de calcular aproximadamente una integral

$$\int_a^b f(x) dx, \quad (5.133)$$

on un número suficientemente grande n de segmentos parciales iguales y sobre cada uno de estos segmentos la función $f(x)$ se sustituya por un polinomio de órdenes nulo, primero o segundo. El error surgido en este caso despreciaba plenamente las propiedades individuales de la función $f(x)$. Naturalmente, surge, por eso, la cuestión de variación de los puntos de partición del segmento principal $[a, b]$ y de elección, para cada función fija $f(x)$, de tal partición óptima del segmento principal en n segmentos parciales (como regla, no iguales uno a otro) que asegure una magnitud mínima del error de la fórmula aproximada dada.

Detengámonos aquí en la resolución de la cuestión mencionada por un método que se debe a A.N. Tijonov y S.S. Gaisarián*).

Con el fin de calcular aproximadamente la integral (5.133), dividamos el segmento $[a, b]$ en n segmentos parciales mediante unos puntos

$$a = x_0 < x_1 < x_2 < \dots < x_n = b.$$

Denotaremos longitud del k -ésimo segmento parcial $\{x_{k-1}, x_k\}$ ($k = 1, 2, \dots, n$) con el símbolo h_k , (de modo que $h = x_k - x_{k-1}$).

Al introducir la integral (5.133) en forma de una suma de integrales

$$\int_a^b f(x) dx = \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(x) dx, \quad (5.134)$$

aproximemos cada una de las integrales del segundo miembro de (5.134) con ayuda de una de las tres fórmulas aproximadas (de los rectángulos, trapecios o de las parábolas). Cualquiera de las tres fórmulas mencionadas puede ser escrita en la forma

$$\int_{x_{k-1}}^{x_k} f(x) dx = h_k \sum_{i=0}^{m-1} q_i f(x_k + t_i h_k) + R_{s, m, k}(f), \quad (5.135)$$

donde los nudos t_i y los pesos q_i ($i = 0, 1, \dots, m-1$) se eligen de modo que el término residual $R_{s, m, k}(f)$ sea de orden h_k^s para cierto $s > 1$ **).

Al introducir (5.135) en (5.134), obtenemos que

$$\int_a^b f(x) dx = \sum_{k=1}^n h_k \sum_{i=0}^{m-1} q_i f(x_k + t_i h_k) + R_{s, m}(f), \quad (5.136)$$

donde

$$R_{s, m}(f) = \sum_{k=1}^n R_{s, m, k}(f). \quad (5.137)$$

Planteemos la cuestión de una elección de la partición $\{x_k\}$ del segmento $[a, b]$ con la cual el cuadrado del error (5.137) alcance el mínimo, siendo fijos el número n de puntos de partición, la función $f(x)$ y la fórmula aproximada (5.135). Planteada de tal modo la cuestión, el cuadrado del error $R_{s, m}^2(f)$ sólo depende de la elección de los puntos intermedios de partición, es decir, es una

* Véase la obra *Sobre la elección de las redes óptimas en el cálculo aproximado de las cuadraturas* por A.N. Tijonov y S.S. Gaisarián (En: *Revista de matemática de cómputo y física matemática*, v. 9, No. 5, 1969) (en ruso).

** En particular, para la fórmula de los trapecios en (5.135) se debe poner $m = 2$, $s = 2$, $t_0 = 0$, $t_1 = 1$, $q_0 = q_1 = 1/2$.

función de $(n-1)$ variables x_1, x_2, \dots, x_{n-1} , definida en un tetraedro $a < x_1 < x_2 < \dots < x_{n-1} < b$. Por cuanto el tetraedro citado es un dominio abierto, puede suceder que el mínimo de la función $R_{s,m}^2(f)$ en el citado tetraedro, como regla, no se alcance (es decir, no existe, en general, la partición óptima del segmento $[a, b]$).

Podemos, sin embargo, demostrar que si la derivada $f^{(s)}(x)$ conserva invariable su signo en el segmento $[a, b]$, el valor mínimo del cuadrado del error $R_{s,m}^2(f)$ se alcanza con una partición del segmento $[a, b]$ tal que sus nudos \bar{x}_h satisfacen $(n-1)$ ecuaciones

$$\frac{\partial R_{s,m}^2(f)}{\partial x_h} = 0 \quad (k=1, 2, \dots, n-1) \quad (5.138)$$

y las condiciones adicionales $x_0 = a, x_n = b^*$.

Detengámonos más detalladamente en el caso de la fórmula de los trapecios. En este caso, en la fórmula (5.136) hay que adaptar $s=2, m=2, t_0=0, t_1=1, q_0=q_1=1/2$, por lo que la fórmula (5.136) tomará la forma

$$\int_a^b f(x) dx = \sum_{k=1}^n \frac{h_k}{2} [f(x_k) + f(x_{k-1})] - R_{2,2}(f). \quad (5.139)$$

Las ecuaciones (5.138) se reducen, en virtud de (5.139) y (5.134), a la forma

$$f(x_{k+1}) - f(x_{k-1}) = f'(x_k) (x_{k+1} - x_{k-1}) \quad (k=1, 2, \dots, n-1). \quad (5.140)$$

La existencia de la solución del sistema de ecuaciones (5.140) se asegura por la conservación del signo de la segunda derivada $f''(x)$ en el segmento $[a, b]$, es

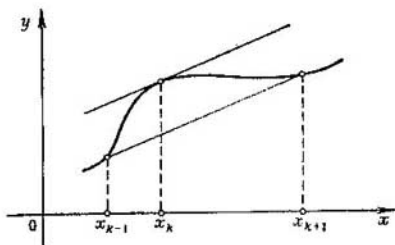


Fig. 5.4

decir, conservando la orientación de la convexidad de la curva $y=f(x)$ para $a \leq x \leq b$. Los nudos $\{\bar{x}_k\}$ de la partición óptima del segmento $[a, b]$ definida mediante las ecuaciones (5.140), poseen la propiedad geométrica siguiente: una secante trazada por los puntos de la gráfica de la función $y=f(x)$ cuyas abscisas son \bar{x}_{k+1} y \bar{x}_{k-1} , es paralela a la tangente a la gráfica mencionada, trazada por su punto de abscisa \bar{x}_k .

* La demostración de esta afirmación puede encontrarse en la obra de A.N. Tijonov y S.S. Gaisarián, citada en la nota al pie de la pág. 241.

Esta propiedad constituye un corolario inmediato de las igualdades (5.140) y se ilustra en la fig. 5.4 (los razonamientos son los mismos que en el § 7, cap. 8, t. 1).

Se puede demostrar que los nudos de «partición aproximada» $\{\tilde{x}_k\}$ ($k = 0, 1, \dots, n$) definidos mediante las igualdades recurrentes

$$\tilde{x}_0 = a, \quad \tilde{x}_{k+1} = \tilde{x}_k + \lambda |f'(\tilde{x}_k)|^{-1/3} \quad (k = 0, 1, \dots, n-1) \quad (5.141)$$

satisfacen para $\lambda = (b-a)/n$ las ecuaciones (5.140) con un error (o, como suele decirse, con un defecto) de orden λ^2 . Esto quiere decir que, siendo n grande, la partición $\{\tilde{x}_k\}$ es próxima a la partición óptima $\{\bar{x}_k\}$. De este modo, cuando n es grande, el cálculo de la integral (5.133) según la fórmula de los trapecios con la partición $\{\tilde{x}_k\}$, cuyos nudos se definen sucesivamente mediante las relaciones recurrentes (5.141), asegura un error próximo al error mínimo.

Capítulo 6

TEORÍA DE LAS FUNCIONES IMPLÍCITAS Y SUS APLICACIONES

§ 1. Concepto de función implícita

En las matemáticas y en sus aplicaciones hemos de enfrentarnos con los problemas en que una variable u , que por el sentido de un problema dado es una función de los argumentos x, y, \dots , se define mediante una ecuación funcional

$$F(u, x, y, \dots) = 0. \quad (6.1)$$

En este caso suele decirse que u , siendo una función de los argumentos x, y, \dots está dada implícitamente. Así, por ejemplo, una función $u = -\sqrt{1-x^2-y^2}$, examinada dentro del círculo $x^2 + y^2 \leq 1$, puede ser dada implícitamente mediante una ecuación funcional

$$F(u, x, y) = u^2 + x^2 + y^2 - 1 = 0. \quad (6.2)$$

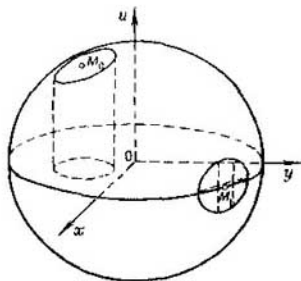


Fig. 6.1

Surge, naturalmente, la cuestión: en qué condiciones la ecuación funcional (6.1) es resoluble unívocamente respecto de u , es decir, define unívocamente la función implícita $u = \varphi(x, y, \dots)$, y una cuestión más específica: en qué condiciones esta función explícita es *continua y diferenciable*. Estas preguntas no son simples. Por ejemplo, en el

caso general, la ecuación funcional (15.2) define, dentro del círculo $x^2 + y^2 \leq 1$, además de la función explícita citada más arriba $u = -\sqrt{1-x^2-y^2}$, una infinidad de otras funciones. Tal es, por ejemplo, una función $u = +\sqrt{1-x^2-y^2}$, así como cualquier otra función u , igual a $+\sqrt{1-x^2-y^2}$ para ciertos puntos (x, y) del círculo $x^2 + y^2 \leq 1$, e igual a $-\sqrt{1-x^2-y^2}$, para los demás puntos de este círculo. Con el fin de aclarar la cuestión de las condiciones que aseguran la resolubilidad unívoca de la ecuación (6.2) respecto de u , recurramos a una ilustración geométrica. La ecuación (6.2) define en un espacio (u, x, y) la esfera S de radio 1 con centro en el origen de las coordenadas (fig. 6.1). Tomemos en la esfera S un

punto $M_0 (\overset{\circ}{u}, \overset{\circ}{x}, \overset{\circ}{y})$ que no esté dispuesto en el plano Oxy , es decir, un punto tal, en cuyo caso $\overset{\circ}{u} \neq 0$. Es obvio que la parte de la esfera S , dispuesta dentro de un entorno suficientemente pequeño del punto M_0 , se proyecta unívocamente sobre el plano Oxy . Desde el punto de vista analítico esto significa que si consideramos la función $F(u, x, y) = u^2 + x^2 + y^2 - 1$ sólo en el entorno citado del punto M_0 , la ecuación (6.2) es unívocamente resoluble respecto de u y define la única función explícita $u = +\sqrt{1 - x^2 - y^2}$ para $\overset{\circ}{u} > 0$ y $u = -\sqrt{1 - x^2 - y^2}$ para $\overset{\circ}{u} < 0$.

En cambio, si tomamos en la esfera S un punto $M_1 (\overset{\circ}{0}, \overset{\circ}{x}, \overset{\circ}{y})$ dispuesto en el plano Oxy (fig. 6.1), entonces, evidentemente, una parte de la esfera S dispuesta en cualquier entorno de M_1 se proyecta sobre el plano Oxy de una manera no unívoca. Analíticamente esto significa que si examinamos la función $F(u, x, y) = u^2 + x^2 + y^2 - 1$ en cualquier entorno del punto M_1 , la ecuación (6.2) no será unívocamente resoluble respecto de u . Prestemos atención a que la derivada parcial $\frac{\partial F}{\partial u} = 2u$ de la función $F(u, x, y) = u^2 + x^2 + y^2 - 1$ no se anula en el punto M_0 , anulándose en M_1 . Más abajo deduciremos que para que la ecuación funcional general (6.1) sea resoluble unívocamente en un entorno del punto M_0 respecto de u un papel de principio lo desempeña el hecho de que la derivada parcial $\frac{\partial F}{\partial u}$ no se reduce a cero en el punto M_0 . Estableceremos, además, las condiciones, bajo las cuales una función explícita que es la única solución de la ecuación (6.1) sea continua y diferenciable.

En adelante denotaremos con el símbolo R el espacio de variables (u, x, y, \dots) y con el R' , el espacio de las variables (x, y, \dots) . Para reducir las notaciones y para hacer más cómoda la ilustración geométrica, vamos a examinar dos variables x e y .

§ 2. Teorema de existencia y de diferenciabilidad de una función implícita y algunas de sus aplicaciones

1. Teorema de existencia y de diferenciabilidad de una función implícita

Teorema 6.1. *Supongamos que una función $F(u, x, y)$ es diferenciable en cierto entorno de un punto $M_0 (\overset{\circ}{u}, \overset{\circ}{x}, \overset{\circ}{y})$ del espacio R , con la particularidad de que la derivada parcial $\frac{\partial F}{\partial u}$ es continua en el punto M_0 . Entonces, si en el punto M_0 la función F se reduce a cero y la derivada parcial $\frac{\partial F}{\partial u}$ no se anula, entonces para un número positivo ε suficientemente pequeño existe tal entorno del punto $M'_0 (\overset{\circ}{x}, \overset{\circ}{y})$ del espacio R' que dentro*

de los márgenes de este entorno existe una única función $u = \varphi(x, y)$ que satisface la condición $|u - \bar{u}| < \varepsilon$ y que es la solución de la ecuación.

$$F(u, x, y) = 0, \quad (6.3)$$

y, además, esta función $u = \varphi(x, y)$ es continua y diferenciable en el entorno mencionado del punto M_0 .

OBSERVACION. 1. En las condiciones del teorema 6.1 la exigencia de continuidad de la derivada parcial $\frac{\partial F}{\partial u}$ en el punto M_0 puede ser omitida, mas habrá que exigir adicionalmente que esta derivada no se anule no sólo en el propio punto M_0 , sino tampoco en cierto entorno de este punto y conserve un signo determinado en dicho entorno.

DEMOSTRACION del teorema 6.1.

1. Demostremos, ante todo, que para $\varepsilon > 0$ suficientemente pequeño en un entorno del punto $M_0(x, y)$ existe una única función $u = \varphi(x, y)$ que satisface la condición $|u - \bar{u}| < \varepsilon$ y que es la solución de la ecuación (6.3). Para

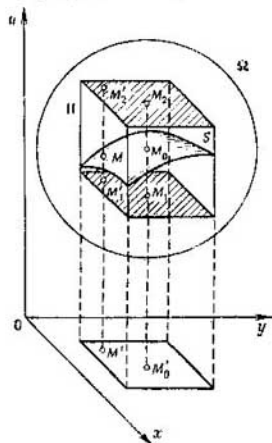


Fig. 6.2

hacer la demostración más evidente, vamos a acompañarla con una ilustración geométrica. Por el curso de geometría analítica sabemos que la ecuación (6.3) define en el espacio R cierta superficie S (fig. 6.2), con la particularidad de que, en virtud de la condición $F(M_0) = 0$, el punto M_0 se dispone en esta superficie. Desde el punto de vista geométrico la resolubilidad unívoca de la ecuación (6.3) respecto de u significa que una parte de la superficie S , dispuesta en proximidad inmediata al punto M_0 , puede ser proyectada unívocamente sobre el plano de coordenadas Oxy .

Convengamos en considerar, para concretar, que la derivada parcial

$\frac{\partial F}{\partial u}$ es positiva en el punto M_0 . Entonces, de la continuidad de la citada derivada en M_0 y del teorema de la estabilidad del signo de una función continua se deduce que existe un entorno del punto M_0 , en cuyos márgenes la derivada $\frac{\partial F}{\partial u}$ es siempre positiva. Podemos imagi-

narnos este entorno en forma de una esfera Ω de radio suficientemente pequeño con centro en el punto M_0 . Ahora, fijamos un número positivo ε tan pequeño que cada uno de los puntos $M_1(\dot{u} - \varepsilon, \dot{x}, \dot{y})$ y $M_2(\dot{u} + \varepsilon, \dot{x}, \dot{y})$ esté dispuesto dentro de la esfera Ω (con el fin de conseguirlo, basta tomar ε inferior al radio de la esfera Ω). Subrayemos que en este caso el número ε está acotado inferiormente sólo por el cero y podemos tomarlo tan pequeño como se quiera. Lo último será útil para nosotros más abajo.

Veamos una función $F(u, x, y)$ de una variable sobre un segmento $\dot{u} - \varepsilon \leq u \leq \dot{u} + \varepsilon$. En el lenguaje geométrico esto quiere decir que estudiamos la función de tres variables $F(u, x, y)$ a lo largo del segmento M_1M_2 (véase fig. 6.2). Por cuanto la derivada $\frac{\partial E}{\partial u}(u, \dot{x}, \dot{y})$ es positiva sobre el segmento $\dot{u} - \varepsilon \leq u \leq \dot{u} + \varepsilon$, entonces la función $F(u, \dot{x}, \dot{y})$ crece en dicho segmento. Mas, en este caso, por cuanto esta función es igual a cero en el centro del segmento citado, (es decir, para $u = \dot{u}$), la función $F(u, \dot{x}, \dot{y})$ tiene valor negativo en el extremo izquierdo y positivo, en el extremo derecho del segmento mencionado, es decir,

$$F(M_1) < 0, F(M_2) > 0.$$

Luego examinaremos las funciones $F(\dot{u} - \varepsilon, x, y)$ y $F(\dot{u} + \varepsilon, x, y)$ de dos variables x y y , es decir, hablando en el lenguaje geométrico, analizaremos la función $F(u, x, y)$ en dos planos paralelos al plano de coordenadas Oxy , el primero de los cuales pasa por el punto M_1 , y el segundo, por el punto M_2 . Por cuanto $F(M_1) < 0$, $F(M_2) > 0$ y la función $F(u, x, y)$ es continua en todo punto de la esfera Ω , entonces, de acuerdo con el teorema de la estabilidad del signo de una función continua, en los planos mencionados se encontrarán entornos de los puntos M_1 y M_2 tales, en cuyos márgenes la función F conserva los mismos signos que en los puntos M_1 y M_2 . Podemos tomar estos entornos en forma de cuadrados abiertos con centros en los puntos M_1 y M_2 y con lado suficientemente pequeño 2δ (en la fig. 6.2 los cuadrados citados están rayados). El hecho de que la función $F(u, x, y)$ conserva el mismo signo en los cuadrados citados se expresa analíticamente mediante las inecuaciones

$$\left. \begin{array}{l} F(\dot{u} - \varepsilon, x, y) < 0, \\ F(\dot{u} + \varepsilon, x, y) > 0 \end{array} \right\} \text{ para } |x - \dot{x}| < \delta, \quad |y - \dot{y}| < \delta. \quad (6.4)$$

Impongamos una condición más a la elección del lado de los cuadrados en consideración: tomemos δ tan pequeño que ambos cuadrados men-

cionados estén dentro de la esfera Ω (esto es factible a ciencia cierta, pues los centros de los cuadrados M_1 y M_2 son puntos interiores de la esfera Ω). Con tal elección de δ cualquier punto del espacio (u, x, y) , cuyas coordenadas satisfacen las inequaciones

$$|x - \overset{\circ}{x}| < \delta, \quad |y - \overset{\circ}{y}| < \delta, \quad |u - \overset{\circ}{u}| < \varepsilon, \quad (6.5)$$

estará dentro de la esfera Ω . Desde el punto de vista geométrico, las inequaciones (6.5) definen un paralelepípedo rectangular abierto con centro en el punto M_0 y con lados paralelos a los ejes coordenados u, x, y , e iguales, respectivamente, a $2\varepsilon, 2\delta$ y 2δ . Denotemos dicho paralelepípedo con el símbolo Π . Por cuanto el paralelepípedo Π está dispuesto en el interior de la esfera Ω , la derivada $\frac{\partial F}{\partial u}$ será positiva en todo punto del paralelepípedo Π^* . Además, en virtud de las inequaciones (6.4), la función $F(u, x, y)$ es negativa en la base inferior y positiva, en la base superior de Π .

Demostremos ahora que la ecuación (6.3) es unívocamente resoluble respecto de u si consideramos la función $F(u, x, y)$ sólo para los valores u, x, y dispuestos en el interior del paralelepípedo Π . Aclaremos qué es lo que necesitamos demostrar. Sea $M'(x, y)$ un punto cualquiera del espacio R' , cuyas coordenadas satisfacen las desigualdades

$$|x - \overset{\circ}{x}| < \delta, \quad |y - \overset{\circ}{y}| < \delta. \quad (6.6)$$

Dicho de otro modo, sea $M'(x, y)$ un punto cualquiera del plano Oxy dispuesto dentro del cuadrado con centro en un punto $M'_0(\overset{\circ}{x}, \overset{\circ}{y})$ con lados iguales a 2δ . Se pide demostrar que para las coordenadas x, y del punto M' se encontrará un número u , el único, del intervalo $\overset{\circ}{u} - \varepsilon < u < \overset{\circ}{u} + \varepsilon$ de tal índole que $F(u, x, y) = 0$. (Desde el punto de vista geométrico, esto quiere decir que cualquier recta paralela al eje u , que corta el paralelepípedo Π , interseca la superficie S en el interior del paralelepípedo Π en uno, y sólo un punto.)

Al fijar los valores x e y que satisfacen las desigualdades (6.6), veamos una función $F(u, x, y)$ del argumento u sobre el segmento $\overset{\circ}{u} - \varepsilon \leq u \leq \overset{\circ}{u} + \varepsilon$, es decir, examinemos la función $F(u, x, y)$ sobre el segmento $M'_1 M'_2$, donde M'_1 y M'_2 son los puntos de intersección de una recta que pasa por el punto $M'(x, y)$ y que es paralela al eje Ou , con las bases del paralelepípedo Π (véase fig. 6.2). Por cuanto la derivada $\frac{\partial F}{\partial u}(u, x, y)$ es positiva sobre el segmento $\overset{\circ}{u} - \varepsilon \leq u \leq \overset{\circ}{u} + \varepsilon$, la función $F(u, x, y)$ crece en este segmento (o bien, lo que es lo mismo, crece en el segmento $M'_1 M'_2$). Mas, en este caso, de las condiciones

* Incluidos los cuadrados abiertos dispuestos en sus bases.

$F(M'_1) < 0$, $F(M'_2) > 0$ se desprende que en el interior del segmento $\dot{u} - \varepsilon \leq u \leq \dot{u} + \varepsilon$ existe un único valor u tal que $F(u, x, y) = 0$ (o bien, en el lenguaje geométrico, en el interior del segmento $M'_1 M'_2$ existe el único punto M dispuesto sobre la superficie S).

Supongamos ahora que la función $u = \varphi(x, y)$ simboliza una regla por cuyo intermedio a todo punto $M'(x, y)$ del entorno (6.6) se le pone en correspondencia el único número u del intervalo $\dot{u} - \varepsilon < u < \dot{u} + \varepsilon$, para el cual $F(u, x, y) = 0$. Se ha demostrado que en el entorno (6.6) existe la única función $u = \varphi(x, y)$ que satisface la condición $|u - \dot{u}| < \varepsilon$ y que es la solución de la ecuación (6.3).

2. Demostremos ahora que la función $u = \varphi(x, y)$ es continua en cualquier punto $M'(x, y)$ del entorno (6.6). Por cuanto para todo punto $M'(x, y)$ del entorno (6.6) quedan cumplidas las mismas condiciones *) que para el punto $M'_0(x, y)$, es suficiente demostrar la continuidad de la función $u = \varphi(x, y)$ sólo en el punto $M'_0(x, y)$. Se pide demostrar que para cualquier ε suficientemente pequeño y positivo existe un número positivo δ tal que con cualesquiera x e y que satisfacen las desigualdades $|x - \dot{x}| < \delta$, $|y - \dot{y}| < \delta$, se cumpla la desigualdad $|u - \dot{u}| < \varepsilon$, donde $u = \varphi(x, y)$, $\dot{u} = \varphi(\dot{x}, \dot{y})$. Al tomar a título de ε el número que se ha elegido más arriba al exponer el p. 1, la existencia de δ se asegurará mediante las desigualdades (6.5). Resta notar que en los razonamientos aducidos en el p. 1 el número positivo ε puede ser elegido tan pequeño como se quiera (esto ya se ha dicho en el p. 1).

Con ello queda establecida la continuidad de la función $u = \varphi(x, y)$. Escribamos la condición de continuidad de la función $u = \varphi(x, y)$ en el punto $M'_0(x, y)$ en forma de diferencias. Al designar con Δu el incremento total de la función $u = \varphi(x, y)$ en el punto $M'_0(x, y)$, correspondiente a los incrementos de los argumentos Δx y Δy , obtenemos que $\Delta u \rightarrow 0$ cuando $\begin{cases} \Delta x \rightarrow 0, \\ \Delta y \rightarrow 0. \end{cases}$

3. Resta demostrar la diferenciabilidad de la función $u = \varphi(x, y)$ en cualquier punto $M'(x, y)$ del entorno (6.6). En virtud de la observación del p. 2, resulta suficiente probar la diferenciabilidad de la función $u = \varphi(x, y)$ en el mismo punto $M'_0(x, y)$. Con el fin de ha-

*) A saber, a todo punto $M'(x, y)$ del entorno (6.6) le corresponde un punto $M(u, x, y)$ del espacio R tal que la función $F(u, x, y)$ se anula en el punto M , es diferenciable en cierto entorno del punto M y tiene en este entorno una derivada parcial $\frac{\partial F}{\partial u}$ distinta de cero.

cerlo, calculemos el incremento total Δu de la función $u = \varphi(x, y)$ en el punto $M'_0(\overset{\circ}{x}, \overset{\circ}{y})$, correspondiente a los incrementos de los argumentos Δx y Δy . Por cuanto $F(\overset{\circ}{u}, \overset{\circ}{x}, \overset{\circ}{y}) = 0$ y $F(\overset{\circ}{u} + \Delta u, \overset{\circ}{x} + \Delta x, \overset{\circ}{y} + \Delta y) = 0$, el incremento total ΔF de la función $F(u, x, y)$ en el punto $M'_0(\overset{\circ}{u}, \overset{\circ}{x}, \overset{\circ}{y})$, correspondiente a los incrementos de los argumentos Δu , Δx y Δy , es igual a cero. Pero, en vista de la condición de diferenciabilidad de la función $F(u, x, y)$ en el punto $M_0(\overset{\circ}{u}, \overset{\circ}{x}, \overset{\circ}{y})$, este incremento total tiene la forma

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

Aquí, todas las derivadas parciales $\frac{\partial F}{\partial u}$, $\frac{\partial F}{\partial x}$ y $\frac{\partial F}{\partial y}$ se toman en el punto $M_0(\overset{\circ}{u}, \overset{\circ}{x}, \overset{\circ}{y})$; $\alpha, \beta, \gamma \rightarrow 0$, cuando $\begin{cases} \Delta x \rightarrow 0, \\ \Delta y \rightarrow 0, \\ \Delta u \rightarrow 0. \end{cases}$

Así, pues, obtenemos

$$0 = \left(\frac{\partial F}{\partial u} + \gamma \right) \Delta u + \left(\frac{\partial F}{\partial x} + \alpha \right) \Delta x + \left(\frac{\partial F}{\partial y} + \beta \right) \Delta y. \quad (6.7)$$

De acuerdo con la forma en diferencias de la condición de continuidad de la función $u = \varphi(x, y)$ en el punto $M'_0(\overset{\circ}{x}, \overset{\circ}{y})$, $\Delta u \rightarrow 0$ cuando $\begin{cases} \Delta x \rightarrow 0, \\ \Delta y \rightarrow 0 \end{cases}$. De este modo, podemos afirmar que de la condición $\begin{cases} \Delta x \rightarrow 0, \\ \Delta y \rightarrow 0 \end{cases}$ se infiere que α, β y $\gamma \rightarrow 0$.

Por la hipótesis del teorema, la derivada parcial $\frac{\partial F}{\partial u}$ es distinta de cero en el punto M_0 . Por cuanto $\gamma \rightarrow 0$ cuando $\begin{cases} \Delta x \rightarrow 0, \\ \Delta y \rightarrow 0, \end{cases}$ la expresión $\frac{\partial F}{\partial u} + \gamma$ no se anula, para Δx y Δy suficientemente pequeños. En tal caso, la fórmula (6.7) puede ser dividida por $\left(\frac{\partial F}{\partial u} + \gamma \right)$ y, como resultado, obtendremos

$$\Delta u = \left(- \frac{\frac{\partial F}{\partial x} + \alpha}{\frac{\partial F}{\partial u} + \gamma} \right) \Delta x + \left(- \frac{\frac{\partial F}{\partial y} + \beta}{\frac{\partial F}{\partial u} + \gamma} \right) \Delta y. \quad (6.8)$$

Según el teorema del valor límite de un cociente de dos funciones, podemos afirmar que

$$-\frac{\frac{\partial F}{\partial x} + \alpha}{\frac{\partial F}{\partial u} + \gamma} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial u}} + \mu, \quad -\frac{\frac{\partial F}{\partial y} + \beta}{\frac{\partial F}{\partial u} + \gamma} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial u}} + \nu, \quad (6.9)$$

donde μ y $\nu \rightarrow 0$ cuando $\begin{cases} \Delta x \rightarrow 0, \\ \Delta y \rightarrow 0. \end{cases}$

Al cotejar las fórmulas (6.8) y (6.9), obtenemos en definitiva

$$\Delta u = \left(-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial u}} \right) \Delta x + \left(-\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial u}} \right) \Delta y + \mu \Delta x + \nu \Delta y. \quad (6.10)$$

La fórmula (6.10) demuestra la diferenciabilidad de la función $u = \varphi(x, y)$ en el punto $M'_0(x, y)$. Con esto queda demostrado completamente el teorema 6.1.

OBSERVACIÓN 2. La demostración aducida se extiende sin complicaciones algunas al caso de una función implícita que depende no de dos, sino de cualquier número finito de argumentos x_1, x_2, \dots, x_m). La ventaja del caso de dos argumentos x e y sólo consiste en que se admite una interpretación geométrica ilustrativa en el espacio (u, x, y) .

2. Cálculo de las derivadas parciales de una función definida implícitamente. Detengámonos en el cálculo de las derivadas parciales de una función definida implícitamente mediante la ecuación (6.3). Supongamos cumplidas las condiciones del teorema 6.1. Entonces, para el incremento total de la función $u = \varphi(x, y)$ es válida la representación (6.10). Esta representación y el teorema 5.9 permiten afirmar que las derivadas parciales de la función $u = \varphi(x, y)$ se definen por las fórmulas

$$\frac{\partial u}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial u}}, \quad \frac{\partial u}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial u}}. \quad (6.11)$$

Fórmulas análogas son válidas también para el caso en que una función definida implícitamente depende no de dos, sino de cualquier número finito de argumentos x_1, x_2, \dots, x_m . En tal caso

$$\frac{\partial u}{\partial x_k} = -\frac{\frac{\partial F}{\partial x_k}}{\frac{\partial F}{\partial u}} \quad (k = 1, 2, \dots, m).$$

*) En particular, de un argumento también.

Si deseamos asegurar la existencia de las *segundas* derivadas de la función $u = \varphi(x, y)$, definida implícitamente, han de reforzarse naturalmente, las exigencias planteadas ante la función $F(u, x, y)$ en el teorema 6.1, a saber: hemos de exigir adicionalmente que la función $F(u, x, y)$ sea dos veces diferenciable en el punto que se considera. Para las suposiciones admitidas detengámonos en el cálculo de las *derivadas parciales de segundo orden*.

Introduzcamos la noción de derivada parcial total de una función, útil para la exposición ulterior. Supongamos definida una función diferenciable de tres argumentos $\Phi(u, x, y)$, con la particularidad de que uno de estos argumentos, u , es de por sí una función diferenciable de otros dos argumentos x o y . Entonces, la función $\Phi(u, x, y)$ puede considerarse como función compuesta de dos argumentos x, y . Las derivadas parciales de esta función compuesta respecto de x e y se llamarán *derivadas parciales totales de la función $\Phi(u, x, y)$ respecto de x e y* , y se denotarán con los símbolos $\frac{D\Phi}{Dx}$ y $\frac{D\Phi}{Dy}$. De acuerdo con la regla de diferenciación de una función compuesta, obtenemos las siguientes fórmulas para las derivadas parciales totales mencionadas:

$$\frac{D\Phi}{Dx} = \frac{\partial\Phi}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial\Phi}{\partial x}, \quad \frac{D\Phi}{Dy} = \frac{\partial\Phi}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial\Phi}{\partial y}.$$

Pasemos al cálculo de las derivadas parciales de segundo orden de una función implícitamente definida. Para concretar, calculemos la derivada $\frac{\partial^2 u}{\partial y \partial x}$. Diferenciando la primera de las fórmulas (6.11) respecto de y y teniendo presente que cada una de las derivadas parciales $\frac{\partial F}{\partial x}$ y $\frac{\partial F}{\partial u}$ depende de tres argumentos u, x, y , el primero de los cuales es de por sí una función de x e y , tendremos

$$\begin{aligned} \frac{\partial^2 u}{\partial y \partial x} &= \frac{D \begin{bmatrix} \frac{\partial F}{\partial x} \\ -\frac{\partial F}{\partial u} \end{bmatrix}}{Dy} = \frac{-\frac{\partial F}{\partial u} \frac{D \left[\frac{\partial F}{\partial x} \right]}{Dy} + \frac{\partial F}{\partial x} \frac{D \left[\frac{\partial F}{\partial u} \right]}{Dy}}{\left(\frac{\partial F}{\partial u} \right)^2} = \\ &= \frac{-\frac{\partial F}{\partial u} \left(\frac{\partial^2 F}{\partial x \partial u} \frac{\partial u}{\partial y} + \frac{\partial^2 F}{\partial x \partial y} \right) + \frac{\partial F}{\partial x} \left(\frac{\partial^2 F}{\partial u^2} \frac{\partial u}{\partial y} + \frac{\partial^2 F}{\partial u \partial y} \right)}{\left(\frac{\partial F}{\partial u} \right)^2}. \end{aligned}$$

Al introducir en la fórmula obtenida la expresión $\frac{\partial u}{\partial y}$, definida por la segunda de las fórmulas (6.11) obtenemos en definitiva

$$\begin{aligned} \frac{\partial^2 u}{\partial y \partial x} &= \\ &= \frac{\frac{\partial^2 F}{\partial x \partial u} \frac{\partial F}{\partial y} \frac{\partial F}{\partial u} - \frac{\partial^2 F}{\partial x \partial y} \left(\frac{\partial F}{\partial u}\right)^2 - \frac{\partial^2 F}{\partial u^2} \frac{\partial F}{\partial x} \frac{\partial F}{\partial y} + \frac{\partial^2 F}{\partial y \partial u} \frac{\partial F}{\partial x} \frac{\partial F}{\partial u}}{\left(\frac{\partial F}{\partial u}\right)^3}. \end{aligned} \quad (6.12)$$

De un modo análogo se calculan las derivadas parciales $\frac{\partial^2 u}{\partial x^2}$ y $\frac{\partial^2 u}{\partial y^2}$. Por un método similar pueden calcularse también las derivadas parciales de tercer orden y de órdenes subsiguientes*).

EjemPLOS. 1) Calcúlese la derivada parcial $\frac{\partial^2 u}{\partial y \partial x}$ de la función $u = \varphi(x, y)$, definida mediante la ecuación

$$x + y + u - e^{-(x+y+u)} = 0.$$

Haciendo uso de las fórmulas (6.11), calculemos, ante todo, las derivadas parciales de primer orden

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = -\frac{1 + e^{-(x+y+u)}}{1 + e^{-(x+y+u)}} = -1.$$

Es evidente, ahora, que $\frac{\partial^2 u}{\partial y \partial x} = 0$.

2) El mismo problema para una función definida mediante la ecuación

$$u^2 + x^2 + y^2 - a^2 = 0.$$

Haciendo uso de la fórmula (6.11), obtendremos

$$\frac{\partial u}{\partial x} = -\frac{x}{u}, \quad \frac{\partial u}{\partial y} = -\frac{y}{u}.$$

Luego tendremos

$$\frac{\partial^2 u}{\partial y \partial x} = \frac{D\left(\frac{\partial u}{\partial x}\right)}{Dy} = \frac{D\left(-\frac{x}{u}\right)}{Dy} = -\frac{x}{u^2} \frac{\partial u}{\partial y} = -\frac{xy}{u^3}.$$

*) A condición de que la función $F(u, x, y)$ sea diferenciable en el punto dado un número correspondiente de veces.

3. Puntos singulares de una superficie y de una curva plana. Examinemos una superficie S (una curva plana L) definida en un sistema dado de coordenadas rectangulares cartesianas mediante la ecuación $F(x, y, z) = 0$ ($F(x, y) = 0$). En lo que se refiere a la función $F(x, y, z)$ ($F(x, y)$), supongamos que ella en todo punto tiene derivadas parciales continuas de primer orden respecto de todos los argumentos en cierto entorno de cualquier punto de la superficie S (de la curva L). Llamemos singular un punto dado de la superficie S (de la curva L) si en este punto se reducen a cero todas las derivadas parciales de primer orden de la función

$$F(x, y, z) \quad (F(x, y)).$$

En el entorno de un punto singular no puede aplicarse a la ecuación $F(x, y, z) = 0$ ($F(x, y) = 0$) el teorema 6.1, es decir, no se puede afirmar que esta ecuación es resoluble por lo menos respecto de una de las variables x, y, z (x, y). De este modo, una parte de la superficie S (un tramo de la curva L), adyacente al punto singular, puede no admitir la proyección unívoca sobre ninguno de los planos coordenados (sobre ningún eje de coordenadas). La estructura de la superficie S (de la curva L) en un entorno de un punto singular puede ser muy compleja y requiere un análisis adicional.

Los puntos de la superficie S (de la curva L) que no son singulares se denominan *ordinarios*. En el entorno de un punto ordinario es válido el teorema 6.1, de suerte que una parte de la superficie S (un tramo de la curva L), adyacente a un punto ordinario, admite la proyección unívoca por lo menos sobre uno de los planos coordenados (por lo menos sobre uno de los ejes coordenados), lo que facilita esencialmente el análisis de dicha parte (tramo).

EjemPLOS. 1) Hállense los puntos singulares de un cono circular $x^2 + y^2 - z^2 = 0$. Por cuanto $F(x, y, z) = x^2 + y^2 - z^2$, resulta que $\frac{\partial F}{\partial x} = 2x$, $\frac{\partial F}{\partial y} = 2y$, $\frac{\partial F}{\partial z} = -2z$. El único punto singular es el origen de las coordenadas. Es bien conocido que en el entorno de este punto la superficie del cono no puede ser proyectado unívocamente sobre ninguno de los planos coordenados (fig. 6.3).

2) El mismo problema con relación a una curva plana $x^2 - y^2 + x^3 = 0$. Las derivadas parciales tienen por expresión $\frac{\partial F}{\partial x} = 2x + 3x^2 = x(2 + 3x)$, $\frac{\partial F}{\partial y} = -2y$. Ambas derivadas parciales se anulan en dos puntos del plano $(0, 0)$ y $(-2/3, 0)$. De los dos puntos mencionados sólo el primero pertenece a la curva en consideración, es decir, es singular. Al construir la curva $x^2 - y^2 + x^3 = 0$ en el entorno del punto $(0, 0)$, nos convencemos de que este punto es punto múltiple de la gráfica (fig. 6.4). Está claro que en el entorno de este punto la curva no puede proyectarse unívocamente ni sobre el eje Ox , ni sobre el eje Oy .

4. Condiciones que aseguran la existencia de la inversa de la función $y = f(x)$. Valgámonos del teorema 6.1. con el fin de aclarar las condiciones cuyo cumplimiento permite que la función $y = f(x)$ tenga en cierto entorno del punto x_0 función inversa $x = f^{-1}(y)$, definida en cierto entorno del punto y_0 , donde $y_0 = f(x_0)$. Analicemos $y = f(x)$ como función definida mediante una ecuación funcional de la forma $F(x, y) = f(x) - y = 0$.

En este caso la cuestión de existencia de la función inversa coincide con la de resolubilidad respecto de x de la citada ecuación funcional. Obtendremos, antes de demostrar este teorema, la siguiente

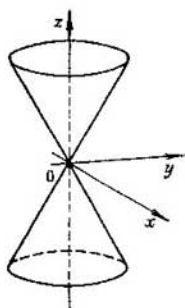


Fig. 6.3

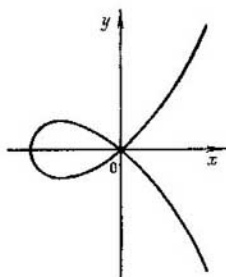


Fig. 6.4

afirmación que viene a ser el corolario del teorema 6.1 y de la observación 1: *si la función $y = f(x)$ tiene en cierto entorno del punto x_0 una derivada distinta de cero, para dicha función existe en el entorno del punto x_0 la inversa $x = f^{-1}(y)$, definida y diferenciable en cierto entorno del punto y_0 , donde $y_0 = f(x_0)$. La derivada de la citada inversa en el punto y_0 es igual, en virtud de la segunda fórmula de (6.11), a $\frac{1}{f'(x_0)}$.*

§ 3. Funciones implícitas definidas por un sistema de ecuaciones funcionales

1. Teorema de resolubilidad de un sistema de ecuaciones funcionales. En el párrafo antecedente hemos estudiado la cuestión de existencia y diferenciability de una función implícita definida mediante una ecuación funcional. En este párrafo examinaremos un problema análogo para una colección de m (m es un número natural cualquiera) funciones implícitas definidas mediante un sistema de ecua-

cierto entorno del punto $N_0(u_1, \dots, u_m)$ la única solución

$$\begin{aligned} x_1 &= \Psi_1(u_1, \dots, u_m), \\ x_m &= \Psi_m(u_1, \dots, u_m). \end{aligned} \tag{6.27'}$$

Es evidente que las funciones (6.17') realizan la aplicación inversa.

Notemos que en las condiciones de la afirmación enunciada, tanto las funciones (6.27) que realizan la aplicación directa, como las (6.27') que realizan la aplicación inversa, son *continuas*. Una aplicación biunívoca que posee esta propiedad se denomina *homeomorfa*.

§ 4. Dependencia de las funciones

1. Concepto de dependencia de las funciones. Condición suficiente de independencia. Supongamos que m funciones de unas mismas n variables

$$\left. \begin{aligned} u_1 &= \varphi_1(x_1, x_2, \dots, x_n), \\ u_2 &= \varphi_2(x_1, x_2, \dots, x_n), \\ &\dots \dots \dots \dots \dots \dots \dots \\ u_m &= \varphi_m(x_1, x_2, \dots, x_n) \end{aligned} \right\} \tag{6.28}$$

están definidas y son diferenciables en cierto dominio *) D n -dimensional abierto.

Diremos que *una de estas funciones, u_h , por ejemplo, depende en el dominio D de las demás funciones si simultáneamente para todos los puntos (x_1, x_2, \dots, x_n) del dominio D*

$$u_h = \Phi(u_1, \dots, u_{h-1}, u_{h+1}, \dots, u_m), \tag{6.29}$$

donde Φ es cierta función definida y diferenciable en un dominio correspondiente de variación de sus argumentos. Las funciones u_1, u_2, \dots, u_m se denominarán *dependientes en el dominio D* si una de estas funciones (no importa, cuál) depende de las demás en el dominio D .

En cambio, si no existe una función diferenciable Φ tal que simultáneamente para todos los puntos del dominio D se verifique una identidad de la forma (6.29), las funciones u_1, u_2, \dots, u_m se denominarán *independientes en el dominio D* .

EJEMPLOS 1. Es fácil convencerse de que tres funciones de cuatro variables

$$\begin{aligned} u_1 &= x_1^2 + x_2^2 + x_3^2 + x_4^2, \\ u_2 &= x_1 + x_2 + x_3 + x_4, \\ u_3 &= 2x_1x_2 + 2x_1x_3 + 2x_1x_4 + 2x_2x_3 + 2x_2x_4 + 2x_3x_4 \end{aligned}$$

*) Es particular, a título de dominio D puede tomarse cierto entorno de un punto fijo M_0 de un espacio n -dimensional.

son dependientes en cualquier dominio D de un espacio cuadrivimensional, pues para todos los puntos (x_1, x_2, x_3, x_4) de este dominio

$$u_1 = u_2^2 - u_3.$$

2. Mostremos ahora que dos funciones de dos variables $u_1 = x + y$ y $u_2 = x - y$ son independientes en cualquier dominio

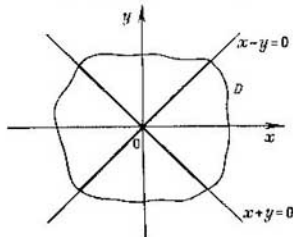


Fig. 6.5

D del plano x, y que contiene el origen de las coordenadas. Está claro que la función u_1 conserva su valor nulo en la recta $x + y = 0$ que pasa por el origen de las coordenadas (fig. 6.5). Pero, sobre esta recta la función u_2 tiene valor variable $u_2 = 2x$. Por eso, en este tramo de la recta, dispuesto en el interior de D , u_2 no depende de u_1 , a ciencia cierta. Es análoga la demostración de que sobre un tramo de la recta $x - y = 0$,

dispuesto en el interior del dominio D , $u_2 = 0$, $u_1 = 2x$, y, por consiguiente, u_1 no depende de u_2 .

OBSERVACIÓN. En el curso de álgebra lineal se introduce la noción de *dependencia lineal* de las funciones: m funciones u_1, u_2, \dots, u_m se llaman linealmente dependientes en un dominio D si para todos los puntos del dominio D una de estas funciones se expresa en forma de una función lineal de las demás funciones. Está claro que la dependencia lineal de las funciones es un caso particular de la dependencia de estas funciones, pues, si las funciones u_1, u_2, \dots, u_m son linealmente dependientes en el dominio D , son dependientes en dicho dominio, pero existen funciones dependientes en el dominio D sin ser en D linealmente dependientes (por ejemplo, las funciones escritas en el ejemplo 1).

Teorema 6.3. (condición suficiente de dependencia de las funciones). Supongamos que m funciones de $n \geq m$ variables

$$u_1 = \varphi_1(x_1, \dots, x_m, x_{m+1}, \dots, x_n),$$

$$u_m = \varphi_m(x_1, \dots, x_m, x_{m+1}, \dots, x_n)$$

están definidas y son diferenciables en un entorno del punto $M_0(x_1, \dots, x_m, x_{m+1}, \dots, x_n)$. Entonces, si el jacobiano formado a partir de estas funciones respecto de algunas m variables es distinto de cero en el punto M_0 , las citadas funciones son independientes en cierto entorno del punto M_0 .

DEMOSTRACIÓN. Sin limitar la generalidad, convengamos en considerar que en el punto M_0 está distinto de cero un jacobiano

$$\frac{D(u_1, u_2, \dots, u_m)}{D(x_1, x_2, \dots, x_m)}. \quad (6.30)$$

Formemos a partir de las derivadas parciales de las funciones (6.28) la siguiente matriz funcional

$$\left\| \begin{array}{ccc} \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \frac{\partial \varphi_2}{\partial x_1} & \frac{\partial \varphi_2}{\partial x_2} & \cdots & \frac{\partial \varphi_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \varphi_m}{\partial x_1} & \frac{\partial \varphi_m}{\partial x_2} & \cdots & \frac{\partial \varphi_m}{\partial x_n} \end{array} \right\|, \quad (6.32)$$

que contiene m filas y n columnas.

Tiene lugar la afirmación siguiente.

Teorema 6.4. *Supongamos que en la matriz funcional (6.32):*

1) *cierto menor de r -ésimo orden *) es distinto de cero en el punto $M_0(x_1, x_2, \dots, x_n)$, 2) todos los menores de $(r+1)$ -ésimo orden son nulos en cierto entorno del punto M_0 **).* Entonces, las r funciones representadas en el citado menor de r -ésimo orden, son independientes en el entorno del punto M_0 ; cada una de las demás funciones depende en dicho entorno de las r funciones mencionadas.

DEMOSTRACION. Sin perder la generalidad de los razonamientos, conpongamos en considerar que en el punto M_0 es distinto de cero el menor dispuesto en la esquina superior izquierda de la matriz (6.32), es decir, es distinto de cero el determinante

$$\left| \begin{array}{ccc} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_r} \\ \dots & \dots & \dots \\ \frac{\partial \varphi_r}{\partial x_1} & \cdots & \frac{\partial \varphi_r}{\partial x_r} \end{array} \right|. \quad (6.33)$$

Entonces, la independencia en el entorno del punto M_0 de las funciones u_1, u_2, \dots, u_r se deduce en seguida del teorema 6.3. Resta por demostrar que cualquiera de las funciones u_{r+1}, \dots, u_m (***) depende en el entorno del punto M_0 respecto de u_1, u_2, \dots, u_r . Demostremos, por ejemplo, que u_{r+1} depende en el entorno del punto M_0 de u_1, u_2, \dots, u_r . Prestaremos nuestra atención a las primeras r funciones (6.28). Si designamos por $\overset{\circ}{u}_1, \dots, \overset{\circ}{u}_r$ los números de tipo $\overset{\circ}{u}_1 = \varphi_1(x_1, x_2, \dots, x_n), \dots, \overset{\circ}{u}_r = \varphi_r(x_1, x_2, \dots, x_n)$, siempre en cierto entorno del punto $N_0(\overset{\circ}{u}_1, \dots, \overset{\circ}{u}_r, x_1, \dots, x_n)$ del es-

*) Recordemos que se llama menor de r -ésimo orden de una matriz dada un determinante formado de los elementos dispuestos en la intersección de algunas r columnas con r filas de la matriz

**) Cuando $r = \min(m, n)$, la exigencia 2) ha de omitirse.

***) Por supuesto, se supone en este caso que $m > r$.

res x_1, \dots, x_r , definidos mediante las ecuaciones (6.35), en la $(r+1)$ -ésima ecuación de (6.28). En este caso u_{r+1} se convierte en una función de los argumentos $u_1, \dots, u_r, x_{r+1}, \dots, x_n$, pues $u_{r+1} = \varphi_{r+1}(x_1, \dots, x_r, x_{r+1}, \dots, x_n) = \varphi_{r+1}[\psi_1(u_1, \dots, u_r, x_{r+1}, \dots, x_n), \dots, \psi_r(u_1, \dots, u_r, x_{r+1}, \dots, x_n), x_{r+1}, \dots, x_n] = \Phi(u_1, \dots, u_r, x_{r+1}, \dots, x_n)$ (esta función fue denotada por el símbolo Φ). Resta por demostrar que para todos los valores de las variables $x_1, \dots, x_r, x_{r+1}, \dots, x_n$, dispuestos en un entorno suficientemente pequeño del punto M_0 , la función Φ no depende de x_{r+1}, \dots, x_n . Con este fin basta probar que para todos los x_1, \dots, x_n de un entorno suficientemente pequeño del punto M_0 se verifican las igualdades

$$\frac{\partial \Phi}{\partial x_l} = 0 \quad (l = r+1, \dots, n). \quad (6.37)$$

Diferenciemos la función Φ respecto de la variable x_l ($l = r+1, \dots, n$) como una función compuesta. Obtendremos

$$\frac{\partial \varphi_{r+1}}{\partial x_1} \frac{\partial \psi_1}{\partial x_l} + \dots + \frac{\partial \varphi_{r+1}}{\partial x_r} \frac{\partial \psi_r}{\partial x_l} + \frac{\partial \varphi_{r+1}}{\partial x_l} = \frac{\partial \Phi}{\partial x_l}. \quad (6.36^{r+1})$$

Veamos ahora el siguiente menor de $(r+1)$ -ésimo orden de la matriz (6.32):

$$\Delta = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1} & \dots & \frac{\partial \varphi_1}{\partial x_r} & \frac{\partial \varphi_1}{\partial x_l} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \varphi_r}{\partial x_1} & \dots & \frac{\partial \varphi_r}{\partial x_r} & \frac{\partial \varphi_r}{\partial x_l} \\ \frac{\partial \varphi_{r+1}}{\partial x_1} & \dots & \frac{\partial \varphi_{r+1}}{\partial x_r} & \frac{\partial \varphi_{r+1}}{\partial x_l} \end{vmatrix}. \quad (6.38)$$

Por la hipótesis del teorema este menor es siempre igual a cero en el entorno del punto M_0 . Multipliquemos las igualdades (6.36¹) — (6.36^{r+1}) por los complementos algebraicos correspondientes $\Delta_1, \dots, \Delta_r, \Delta_{r+1}$ de los elementos de la última columna del menor (6.38) y, a continuación, sumemos todas estas igualdades. En virtud del teorema de que la suma de productos de los elementos de una columna dada por los complementos algebraicos correspondientes de los elementos de esta (otra) columna es igual al determinante (al cero), obtendremos *)

$$\Delta = \frac{\partial \Phi}{\partial x_l} \Delta_{r+1}. \quad (6.39)$$

En la igualdad (6.39) el símbolo Δ denota el menor (6.38), igual a cero siempre en el entorno del punto M_0 , y el complemento algebraico

*) Repetimos los razonamientos descritos detalladamente en la pág. 259.

Δ_{r+1} coincide con el menor (6.33), distinto de cero en el punto M_0 , y, por tanto, en cierto entorno de este punto *). De la igualdad (6.39) concluimos que en cierto entorno del punto M_0 las igualdades (6.37) son siempre válidas. El teorema está demostrado.

EJEMPLO. Volvamos a analizar la dependencia de las funciones

$$u_1 = x_1^2 + x_2^2 + x_3^2 + x_4^2,$$

$$u_2 = x_1 + x_3 + x_3 + x_4,$$

$$u_3 = 2x_1x_2 + 2x_1x_3 + 2x_1x_4 + 2x_2x_3 + 2x_2x_4 + 2x_3x_4.$$

La matriz funcional (6.32) tiene por expresión

$$\begin{vmatrix} 2x_1 & 2x_2 & 2x_3 & 2x_4 \\ 1 & 1 & 1 & 1 \\ 2(x_2 + x_3 + x_4) & 2(x_1 + x_3 + x_4) & 2(x_1 + x_2 + x_3) & 2(x_1 + x_2 + x_3) \end{vmatrix}.$$

Es fácil convencerse de que todos los determinantes de tercer orden son idénticamente iguales a cero, con la particularidad de que en cualquier punto del espacio (x_1, x_2, x_3, x_4) , en el que no coinciden las cuatro coordenadas x_1, x_2, x_3, x_4 , por lo menos uno de los determinantes de segundo orden

$$\begin{vmatrix} 2x_1 & 2x_2 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2x_1 & 2x_2 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2x_1 & 2x_4 \\ 1 & 1 \end{vmatrix}$$

es distinto de cero. Por consiguiente, en el entorno de cualquier punto mencionado las funciones u_1 y u_2 son dependientes, mientras que u_3 depende de u_1 y u_2 .

§ 5. Extremo condicionado

1. **Concepto de extremo condicionado.** En el § 6 del cap. 5 buscamos extremos locales de unas funciones cuyos argumentos no están relacionados con condiciones adicionales algunas. Pero, en las matemáticas y en sus aplicaciones se encuentra a menudo un problema de búsqueda de los extremos de una función cuyos argumentos satisfacen ciertas condiciones adicionales de conexión. Llamaremos los extremos de este género *condicionados* con el fin de distinguirlos de los extremos (incondicionales) estudiados en el § 6, cap. 5.

Demos a conocer un ejemplo de problema de búsqueda de un extremo condicionado. Se pide hallar el extremo de una función $u =$

*) Por cuanto todas las derivadas parciales del menor (6.33) son continuas en el punto M_0 , lo será también en el punto M_0 el propio menor (6.33). Mas, en este caso, de conformidad con el teorema de estabilidad del signo de una función continua, este menor es distinto de cero no sólo en el propio punto M_0 , sino también en cierto entorno del mismo.

$= x^2 + y^2$ a condición de que los argumentos de esta función satisfagan la condición de conexión $x + y - 1 = 0$. De este modo, los extremos de la función $u = x^2 + y^2$ se buscan *no en todo el plano Oxy , sino sólo en la recta $x + y - 1 = 0$* . Para resolver el problema planteado, pongamos en la ecuación de la función $u = x^2 + y^2$ el valor de y determinado a partir de la condición de conexión $x + y - 1 = 0$. De este modo reducimos el problema planteado a un problema de búsqueda del extremo

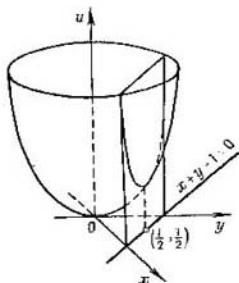


Fig. 6.6

incondicional de la función $u = 2x^2 - 2x + 1$. El último extremo se halla sin dificultades algunas: por cuanto $u' = 4(x - 1/2)$, $u^{(2)} = 4$, la función $u = 2x^2 - 2x + 1$ tiene su mínimo $u = 1/2$ para $x = 1/2$. De este modo, la función $u = x^2 + y^2$ con la condición de conexión $x + y - 1 = 0$ cuenta con el mínimo condicionado $u = 1/2$ en el punto $(1/2, 1/2)$. Advertimos que el mínimo incondicional de la función $u = x^2 + y^2$ se alcanza en el punto $(0, 0)$ y es igual a $u = 0$. Además,

incluso los razonamientos ilustrativos (fig. 6.6) evidencian que el mínimo de la función $u = x^2 + y^2$ (de cuyo gráfico sirve un paraboloide de revolución) en todo el plano Oxy no coincide con su mínimo en la recta $x + y - 1 = 0$.

Pasemos al planteamiento general del problema de búsqueda de extremo condicionado. Supongamos que se pide hallar un extremo de la función de $m + n$ variables

$$u = f(x_1, \dots, x_n, y_1, \dots, y_m), \quad (6.40)$$

al haber m condiciones de conexión

$$\left. \begin{aligned} F_1(x_1, \dots, x_n, y_1, \dots, y_m) &= 0, \\ F_2(x_1, \dots, x_n, y_1, \dots, y_m) &= 0, \\ \dots & \\ F_m(x_1, \dots, x_n, y_1, \dots, y_m) &= 0. \end{aligned} \right\} \quad (6.41)$$

Precisemos, ante todo, el propio concepto de extremo condicionado de la función (6.40), al haber las conexiones (6.41). Diremos que la función (6.40) tiene, al haber las conexiones (6.41), un máximo (mínimo) condicionado en el punto $M_0(x_1, \dots, x_n, y_1, \dots, y_m)$ cuyas coordenadas satisfacen las condiciones de conexión (6.41), si existe un entorno del punto M_0 , en cuyos márgenes el valor de la función (6.40)

cuyo determinante (jacobiano (6.42)) es distinto de cero. En virtud de las igualdades (6.52), la igualdad (6.50) toma la forma

$$\frac{\partial \Psi}{\partial x_1} dx_1 + \dots + \frac{\partial \Psi}{\partial x_n} dx_n = 0. \quad (6.53)$$

Bajo las suposiciones admitidas más arriba las variables x_1, \dots, x_n son *independientes*, razón por la cual de la igualdad (6.53) deducimos que

$$\frac{\partial \Psi}{\partial x_1} = 0, \dots, \frac{\partial \Psi}{\partial x_n} = 0. \quad (6.54)$$

Al adjuntar a las ecuaciones (6.52) y (6.54) las condiciones de conexión (6.41), obtendremos un sistema de $n + 2m$ ecuaciones

$$\left. \begin{aligned} \frac{\partial \Psi}{\partial x_1} = 0, \dots, \frac{\partial \Psi}{\partial x_n} = 0, \quad \frac{\partial \Psi}{\partial y_1} = 0, \dots, \frac{\partial \Psi}{\partial y_m} = 0, \\ F_1 = 0, \dots, F_m = 0 \end{aligned} \right\} \quad (6.55)$$

para determinar $n + m$ coordenadas de los puntos de extremo condicionado eventual y m factores $\lambda_1, \dots, \lambda_m$. En la práctica, al realizar este método, se procede del modo siguiente. Se forma la función de Lagrange (6.51) y se hallan para esta función los puntos de extremo incondicional eventual. Con el fin de excluir los factores $\lambda_1, \dots, \lambda_m$, se atraen las condiciones de conexión (6.41). Este camino de buscar los puntos de extremo condicionado eventual es legítimo, pues conduce precisamente al sistema de $n + 2m$ ecuaciones (6.55). Más abajo, en el p. 4, veremos un ejemplo de aplicación del método de multiplicadores de Lagrange.

3. Condiciones suficientes. En este punto examinaremos uno de los procedimientos de análisis adicional de los puntos de extremo condicionado eventual. Supongamos que en el punto M_0 se cumplen las condiciones necesarias de extremo (6.55). Además, exijamos complementariamente que las funciones (6.40) y (6.41) sean dos veces diferenciables en un entorno del punto M_0 y que sean continuas todas las derivadas parciales de segundo orden en el propio punto M_0 . La construcción de la función de Lagrange (6.51) sugiere, evidentemente, que, al haber las conexiones (6.41), los extremos de la función (6.40) y de la función de Lagrange coinciden *). Pero, en tal caso, de los resultados del § 6, cap. 5 se desprende que para obtener una condición suficiente de extremo de la función (6.40) en el punto M_0 en presencia de las conexiones (6.41), se debe adjuntar a las condiciones (6.55) el requisito de que $d^2\Psi$ sea de signo definido en este punto. Enton-

*) Esto se deduce de que, al haber las conexiones (6.41), la diferencia $f(M) - f(M_0)$ coincide con la diferencia $\Psi(M) - \Psi(M_0)$

ces, de conformidad con los resultados del párrafo 6, cap. 5, podemos constatar un mínimo en el punto M_0 , siempre que, al haber las conexiones (6.41), $d^2\Psi|_{M_0} > 0$, y un máximo, si $d^2\Psi|_{M_0} < 0$. Demos a conocer algunas observaciones más de carácter práctico. Notemos, ante todo, que la segunda diferencial $d^2\Psi$ puede calcularse en el punto dado M_0 de extremo eventual, como si fueran independientes todas las variables $x_1, \dots, x_n, y_1, \dots, y_m$. Efectivamente, la segunda diferencial $d^2\Psi$ de la función Ψ no posee, por regla general, la propiedad de invariación de la forma y tendría que definirse, considerando la dependencia entre y_1, \dots, y_m y x_1, \dots, x_n , mediante la igualdad

$$d^2\Psi = \left(dx_1 \frac{\partial}{\partial x_1} + \dots + dx_n \frac{\partial}{\partial x_n} + dy_1 \frac{\partial}{\partial y_1} + \dots + dy_m \frac{\partial}{\partial y_m} \right)^2 \Psi + \frac{\partial \Psi}{\partial y_1} d^2 y_1 + \dots + \frac{\partial \Psi}{\partial y_m} d^2 y_m.$$

Pero, en el punto de extremo eventual M_0 se verifican las igualdades

$$\frac{\partial \Psi}{\partial y_1} = 0, \dots, \frac{\partial \Psi}{\partial y_m} = 0,$$

de modo que $d^2\Psi$ se define mediante la misma fórmula

$$d^2\Psi = \left(dx_1 \frac{\partial}{\partial x_1} + \dots + dx_n \frac{\partial}{\partial x_n} + \left[dy_1 \frac{\partial}{\partial y_1} + \dots + dy_m \frac{\partial}{\partial y_m} \right]^2 \Psi \right) \quad (6.56)$$

que se emplea cuando todas las variables $x_1, \dots, x_n, y_1, \dots, y_m$ se consideran independientes. Notemos, luego, que cuanto se necesita establecer tiene o no signo definido $d^2\Psi$ sólo en presencia de las conexiones (6.41), al realizar los cálculos, en la fórmula (6.56) para $d^2\Psi$ hay que sustituir dy_1, \dots, dy_m por sus valores definidos a partir del sistema (6.47). A continuación, hay que estudiar la cuestión de es o no de signo definido en el punto dado M_0 la diferencial $d^2\Psi$. Ahora podemos examinar un ejemplo.

4. Ejemplo. Supongamos que hemos de hallar los valores máximo y mínimo de un determinante

$$\Delta = \begin{vmatrix} x_1 & y_1 & \dots & z_1 \\ x_2 & y_2 & \dots & z_2 \\ \dots & \dots & \dots & \dots \\ x_n & y_n & \dots & z_n \end{vmatrix}, \quad (6.57)$$

conociendo la suma de cuadrados de los elementos de cada fila de este determinante. El problema consiste en

El factor constante λ se excluye fácilmente de la condición de conexión (6.60). De esta condición encontramos dos valores:

$$\lambda_1 = \sqrt{\frac{X_k^2 + Y_k^2 + \dots + Z_k^2}{4h_k}} > 0, \quad \lambda_2 = -\sqrt{\frac{X_k^2 + Y_k^2 + \dots + Z_k^2}{4h_k}} < 0.$$

(En este caso omitimos de nuevo el caso trivial cuando todos los X_k, Y_k, \dots, Z_k son nulos, pues en este caso el determinante (6.59) es idénticamente igual a cero.)

De este modo, obtenemos dos puntos de extremo eventual:

$$M_1 \left(-\frac{X_k}{2\lambda_1}, -\frac{Y_k}{2\lambda_1}, \dots, -\frac{Z_k}{2\lambda_1} \right) \text{ y}$$

$$M_2 \left(-\frac{X_k}{2\lambda_2}, -\frac{Y_k}{2\lambda_2}, \dots, -\frac{Z_k}{2\lambda_2} \right).$$

Demostremos que en el punto M_1 se realiza el mínimo condicionado, y en el M_2 , el máximo condicionado. Con este fin, calculemos la segunda diferencial $d^2\Psi$ de la función de Lagrange (6.61). Es fácil ver que

$$d^2\Psi = 2\lambda [(dx_k)^2 + (dy_k)^2 + \dots + (dz_k)^2].$$

De la última fórmula se deduce que $d^2\Psi$ es una forma cuadrática definida positiva para $\lambda = \lambda_1 > 0$ (es decir, en el punto M_1), y definida negativa, para $\lambda = \lambda_2 < 0$ (es decir, en el punto M_2). Así, pues la función (6.59) tiene, al haber la conexión (6.60) mínimo condicionado en el punto M_1 y máximo condicionado en el punto M_2 . Sin calcular los valores máximo y mínimo de la función (6.59), al haber la condición (6.60), notemos que en un punto, donde se alcanzan dichos valores, se verifican las igualdades (6.62), y, por consiguiente, son también válidas las igualdades

$$\frac{x_k}{X_k} = \frac{y_k}{Y_k} = \dots = \frac{z_k}{Z_k}. \quad (6.63)$$

Volvamos ahora al cálculo de los valores extremos del determinante (6.57) al haber n conexiones (6.58). Conservando el sentido de las designaciones admitidas X_k, Y_k, \dots, Z_k , de acuerdo con la propiedad conocida del determinante, podemos escribir para todo i distinto de k

$$x_i X_k + y_i Y_k + \dots + z_i Z_k = 0 \quad (i \neq k). \quad (6.64)$$

De las igualdades (6.63) y (6.64) encontramos que

$$x_i x_k + y_i y_k + \dots + z_i z_k = 0 \quad (\text{para } i \neq k). \quad (6.65)$$

Las igualdades (6.58) y (6.65) y la regla de multiplicación de los determinantes permiten concluir que, al multiplicar el determinante Δ por sí mismo, se obtiene un determinante cuyos elementos son todos

iguales a cero, a excepción de los elementos de la diagonal principal, los cuales son iguales a $h_1, \dots, h_2, \dots, h_n$, respectivamente. De este modo,

$$\Delta \cdot \Delta = \Delta^2 = h_1 h_2 \dots h_n.$$

Por consiguiente, los valores máximo y mínimo de la magnitud del determinante (6.57), al subsistir las condiciones (6.58), son iguales a $\pm \sqrt{h_1 h_2 \dots h_n}$ y $-\sqrt{h_1 h_2 \dots h_n}$, respectivamente. De este modo se obtiene que

$$|\Delta| = \sqrt{h_1 h_2 \dots h_n},$$

o bien

$$|\Delta| \leq \sqrt{(x_1^2 + \dots + z_1^2)(x_2^2 + \dots + z_2^2) \dots (x_n^2 + \dots + z_n^2)}. \quad (6.66)$$

La última desigualdad, válida para un determinante arbitrario (6.57), se llama *desigualdad de Hadamard* *).

OBSERVACIÓN Cuando $n = 3$, la desigualdad de Hadamard (6.66) admite una interpretación geométrica simple. En este caso $|\Delta|$ es el volumen de un paralelepípedo construido sobre los segmentos OA_1, OA_2 y OA_3 que unen el origen de coordenadas O con los puntos $A_1(x_1, y_1, z_1), A_2(x_2, y_2, z_2), A_3(x_3, y_3, z_3)$. La desigualdad (6.66) afirma que de todos los paralelepípedos, cuyas aristas tienen longitud dada, el volumen máximo lo tiene un paralelepípedo rectangular.

Complemento

Cambio de variables

En el análisis y en otros apartados de las matemáticas se encuentra un problema de cambio de variables. Este problema consiste en lo siguiente. Supongamos dada una expresión

$$F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}, \dots\right), \quad (6.67)$$

que contiene las variables independientes x, y , una función $z = z(x, y)$ y las derivadas parciales de la misma. En lugar de las variables independientes x, y y de la función $z = z(x, y)$ se introducen nuevas variables independientes u, v y una nueva función $w = w(u, v)$, con la particularidad de que vienen dadas unas relaciones por cuyo intermedio u, v y w se expresan en términos de x, y y z :

$$\left. \begin{aligned} u &= \varphi(x, y, z), \\ v &= \psi(x, y, z), \\ w &= \chi(x, y, z). \end{aligned} \right\} \quad (6.68)$$

Se requiere transformar la expresión (6.67) respecto de las variables nuevas u, v y w . En este caso supondremos que las funciones (6.68) son diferenciables un

* J. Hadamard (1865—1963), matemático francés.

número suficiente de veces y que el sistema (6.68) puede ser resuelto respecto de x , y y z , mientras que las primeras dos ecuaciones en (6.68) se resuelven respecto de x e y .

Es evidente que para resolver el problema planteado, basta expresar las derivadas $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, ... en términos de u , v , w , $\frac{\partial w}{\partial u}$, $\frac{\partial w}{\partial v}$, Señalemos cómo se puede hacerlo.

Teniendo presente que $z = z(x, y)$, y $w = w(u, v)$, escribamos las primeras diferenciales de las funciones (6.68). Obtendremos

$$du = \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy + \frac{\partial \varphi}{\partial z} dz = \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy + \frac{\partial \varphi}{\partial z} \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right), \quad (6.69)$$

$$dv = \frac{\partial \psi}{\partial x} dx + \frac{\partial \psi}{\partial y} dy + \frac{\partial \psi}{\partial z} dz = \frac{\partial \psi}{\partial x} dx + \frac{\partial \psi}{\partial y} dy + \frac{\partial \psi}{\partial z} \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right), \quad (6.70)$$

$$dw = \frac{\partial w}{\partial u} du + \frac{\partial w}{\partial v} dv = \frac{\partial \chi}{\partial x} dx + \frac{\partial \chi}{\partial y} dy + \frac{\partial \chi}{\partial z} \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right). \quad (6.71)$$

Introduciendo en (6.71) du y dv determinadas por las fórmulas (6.69) y (6.70), e igualando los coeficientes de dx y dy , obtendremos un sistema de dos ecuaciones

$$\left. \begin{aligned} \frac{\partial w}{\partial u} \left[\frac{\partial \varphi}{\partial x} + \frac{\partial \varphi}{\partial z} \frac{\partial z}{\partial x} \right] + \frac{\partial w}{\partial v} \left[\frac{\partial \psi}{\partial z} + \frac{\partial \psi}{\partial z} \frac{\partial z}{\partial x} \right] &= \frac{\partial \chi}{\partial x} + \frac{\partial \chi}{\partial z} \frac{\partial z}{\partial x}, \\ \frac{\partial w}{\partial u} \left[\frac{\partial \varphi}{\partial y} + \frac{\partial \varphi}{\partial z} \frac{\partial z}{\partial y} \right] + \frac{\partial w}{\partial v} \left[\frac{\partial \psi}{\partial y} + \frac{\partial \psi}{\partial z} \frac{\partial z}{\partial y} \right] &= \frac{\partial \chi}{\partial y} + \frac{\partial \chi}{\partial z} \frac{\partial z}{\partial y}, \end{aligned} \right\}$$

a partir del cual se expresan con facilidad $\frac{\partial z}{\partial x}$ y $\frac{\partial z}{\partial y}$:

$$\frac{\partial z}{\partial x} = \frac{\frac{\partial w}{\partial u} \frac{\partial \varphi}{\partial x} + \frac{\partial w}{\partial v} \frac{\partial \psi}{\partial z} - \frac{\partial \chi}{\partial x}}{\frac{\partial w}{\partial u} \frac{\partial \varphi}{\partial z} + \frac{\partial w}{\partial v} \frac{\partial \psi}{\partial z} - \frac{\partial \chi}{\partial z}},$$

$$\frac{\partial z}{\partial y} = \frac{\frac{\partial w}{\partial u} \frac{\partial \varphi}{\partial y} + \frac{\partial w}{\partial v} \frac{\partial \psi}{\partial y} - \frac{\partial \chi}{\partial y}}{\frac{\partial w}{\partial u} \frac{\partial \varphi}{\partial z} + \frac{\partial w}{\partial v} \frac{\partial \psi}{\partial z} - \frac{\partial \chi}{\partial z}}.$$

Si la expresión (6.67) depende también de las derivadas parciales de segundo orden, para determinar estas últimas en términos de las derivadas parciales de w respecto de u y v , se debe escribir las primeras diferenciales de las derivadas de primer orden ya calculadas.

OBSERVACION 1. De un modo análogo se realiza el cambio de las variables cuando las variables antiguas están asociadas con las nuevas no mediante las

relaciones (6.68), sino mediante las relaciones implícitas de la forma*)

$$\left. \begin{aligned} \Phi_1(u, v, w, x, y, z) &= 0, \\ \Phi_2(u, v, w, x, y, z) &= 0, \\ \Phi_3(u, v, w, x, y, z) &= 0. \end{aligned} \right\} \quad (6.72)$$

OBSERVACIÓN 2 Nos hemos limitado al caso de dos variables independientes sólo para abreviar las notaciones. El método aducido es aplicable para el caso de cualquier número de variables independientes (y, en particular, para el de una variable independiente.)

EJEMPLO. Sea z una función de las variables x e y . Transfórmese una expresión**)

$$F = \Delta z = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}$$

respecto de las coordenadas polares u y v :

$$x = u \cos v, \quad y = u \sin v. \quad (6.73)$$

Advertimos que en el ejemplo dado se realiza sólo el cambio de las variables independientes. La función z sigue invariable. De (6.73) se desprende que

$$dx = \cos v \, du - u \sin v \, dv, \quad dy = \sin v \, du + u \cos v \, dv,$$

por lo cual, de las relaciones

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

tenemos

$$\frac{\partial z}{\partial x} (\cos v \, du - u \sin v \, dv) + \frac{\partial z}{\partial y} (\sin v \, du + u \cos v \, dv) = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv.$$

Igualando los coeficientes de du y dv , encontramos

$$\left. \begin{aligned} \frac{\partial z}{\partial x} \cos v + \frac{\partial z}{\partial y} \sin v &= \frac{\partial z}{\partial u}, \\ -\frac{\partial z}{\partial x} u \sin v + \frac{\partial z}{\partial y} u \cos v &= \frac{\partial z}{\partial v}. \end{aligned} \right\}$$

De aquí

$$\left. \begin{aligned} \frac{\partial z}{\partial x} &= \cos v \frac{\partial z}{\partial u} - \frac{1}{u} \sin v \frac{\partial z}{\partial v}, \\ \frac{\partial z}{\partial y} &= \sin v \frac{\partial z}{\partial u} + \frac{1}{u} \cos v \frac{\partial z}{\partial v}. \end{aligned} \right\} \quad (6.74)$$

Hallemos ahora $\frac{\partial^2 z}{\partial x^2}$. Por cuanto $\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \right)$, de la primera fórmula

*) Se admite, por supuesto, la resolubilidad tanto respecto de u , v y w , como también respecto de x , y y z .

**) La expresión citada se llama *operador de Laplace*. Esta desempeña un papel importante en las matemáticas y en sus aplicaciones. P. S. Laplace (1749—1827), matemático, físico y astrónomo francés.

(6.74) encontramos

$$\begin{aligned} \frac{\partial^2 z}{\partial x^2} &= \frac{\partial \left(\frac{\partial z}{\partial x} \right)}{\partial x} = \cos v \frac{\partial \left(\frac{\partial z}{\partial x} \right)}{\partial u} - \frac{1}{u} \operatorname{sen} v \frac{\partial \left(\frac{\partial z}{\partial x} \right)}{\partial v} = \\ &= \cos v \frac{\partial}{\partial u} \left[\cos v \frac{\partial z}{\partial u} - \frac{1}{u} \operatorname{sen} v \frac{\partial z}{\partial v} \right] - \frac{1}{u} \operatorname{sen} v \frac{\partial}{\partial v} \left[\cos v \frac{\partial z}{\partial u} - \frac{1}{u} \operatorname{sen} v \frac{\partial z}{\partial v} \right]. \end{aligned}$$

Después de los cálculos obtenemos

$$\begin{aligned} \frac{\partial^2 z}{\partial x^2} &= \cos^2 v \frac{\partial^2 z}{\partial u^2} - \frac{2}{u} \operatorname{sen} v \cos v \frac{\partial^2 z}{\partial u \partial v} \\ &\quad + \frac{1}{u^2} \operatorname{sen}^2 v \frac{\partial^2 z}{\partial v^2} + \frac{1}{u} \operatorname{sen}^2 v \frac{\partial z}{\partial u} + \frac{2}{u^2} \operatorname{sen} v \cos v \frac{\partial z}{\partial v}. \end{aligned}$$

Análogamente, de la segunda fórmula (6.74), obtenemos

$$\begin{aligned} \frac{\partial^2 z}{\partial y^2} &= \frac{\partial \left(\frac{\partial z}{\partial y} \right)}{\partial y} = \operatorname{sen} v \frac{\partial \left(\frac{\partial z}{\partial u} \right)}{\partial u} + \frac{1}{u} \cos v \frac{\partial \left(\frac{\partial z}{\partial v} \right)}{\partial v} = \\ &= \operatorname{sen}^2 v \frac{\partial^2 z}{\partial u^2} + \frac{2}{u} \operatorname{sen} v \cos v \frac{\partial^2 z}{\partial u \partial v} + \frac{1}{u^2} \cos^2 v \frac{\partial^2 z}{\partial v^2} + \\ &\quad + \frac{1}{u} \cos^2 v \frac{\partial z}{\partial u} - \frac{2}{u^2} \operatorname{sen} v \cos v \frac{\partial z}{\partial v}. \end{aligned}$$

De este modo, en las coordenadas polares u y v el operador de Laplace

$\Delta z = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}$ tiene por expresión:

$$\Delta z = \frac{\partial^2 z}{\partial u^2} + \frac{1}{u} \frac{\partial z}{\partial u} + \frac{1}{u^2} \frac{\partial^2 z}{\partial v^2}.$$

Capítulo 7

ALGUNAS APLICACIONES GEOMÉTRICAS DEL CÁLCULO DIFERENCIAL

§ 1. Envolvente y curva discriminante de una familia monoparamétrica de curvas planas

1. **Observaciones preliminares.** Definiremos las curvas planas con ayuda de las ecuaciones paramétricas

$$x = \varphi(\alpha), \quad y = \psi(\alpha), \quad (7.1)$$

donde α es un parámetro, o bien con ayuda de las ecuaciones de la forma

$$F(x, y) = 0. \quad (7.2)$$

En lo que sigue más abajo nos harán falta los conceptos de puntos *ordinario* y *singular* de una curva.

Supongamos que una curva L se define mediante las ecuaciones paramétricas (7.1), con la particularidad de que las funciones $x = \varphi(\alpha)$ e $y = \psi(\alpha)$ tienen, para $\alpha = \alpha_0$, derivadas continuas. Un punto $M_0(x_0, y_0)$ de la curva L , cuyas coordenadas x_0 e y_0 , son iguales a $\varphi(\alpha_0)$ y $\psi(\alpha_0)$, respectivamente, se llamará *ordinario* si

$$\varphi'^2(\alpha_0) + \psi'^2(\alpha_0) \neq 0. \quad (7.3)$$

En cambio, si para $\alpha = \alpha_0$, se verifica la relación

$$\varphi'^2(\alpha_0) + \psi'^2(\alpha_0) = 0, \quad (7.4)$$

el punto M_0 se llamará punto *singular* de la curva L .

Supongamos que la curva L se define mediante la ecuación (7.2), con la particularidad de que la función $F(x, y)$ es diferenciable en cierto entorno del punto $M_0(x_0, y_0)$ de esta curva y tiene en el mismo derivadas parciales continuas respecto de x e y . El punto $M_0(x_0, y_0)$ se llamará punto *ordinario* de la curva L si en este punto se verifica una relación *)

$$F_x^2 + F_y^2 \neq 0. \quad (7.5)$$

Si, en cambio, en el punto M_0 se verifica la relación

$$F_x^2 + F_y^2 = 0, \quad (7.6)$$

llamaremos M_0 punto *singular* de la curva L . Cerciorémonos de que si el punto M_0 de la curva L es ordinario, en cierto entorno de este punto la curva L es o bien la gráfica de cierta función diferenciable $y = f(x)$, o bien la gráfica de cierta función diferenciable $x = g(y)$. Efectiva-

*) Los conceptos de puntos ordinario y singular para la curva (7.2) ya fueron introducidos en el p. 3, § 2, cap. 6.

mente, supongamos que la curva L se define mediante las ecuaciones paramétricas (7.1), y, además, queda cumplida la condición (7.3). De la continuidad de las derivadas $\varphi'(\alpha)$ y $\psi'(\alpha)$ para $\alpha = \alpha_0$ y de la condición (7.3) se deduce que en cierto entorno de α_0 al menos una de estas derivadas, por ejemplo, $\varphi'(\alpha)$, no es nula. Entonces, la función $x = \varphi(\alpha)$ es en el citado entorno diferenciable y estrictamente monótona. Bajo estas condiciones existe una función inversa diferenciable monótona $\alpha = \varphi^{-1}(x)$. Poniendo esta función en la expresión $y = \psi(\alpha)$, nos convencemos de que la curva L es en cierto entorno del punto M_0 la gráfica de la función diferenciable $y = f(x) = \psi[\varphi^{-1}(x)]$. La validez de la afirmación enunciada para el caso en que la curva se define con ayuda de la ecuación (7.2) se infiere de que en el entorno de un punto ordinario subsiste el teorema 6.4 sobre las funciones implícitas y, por eso, el tramo de la curva adyacente a un punto ordinario es la gráfica de la función diferenciable $y = f(x)$ o de la función $g(y)$.

OBSERVACIÓN 1 En geometría el punto M_0 de la curva L se denomina ordinario si en cierto entorno de este punto la curva L es la gráfica de cierta función diferenciable, y singular, si en cualquier entorno de este punto la curva L no puede ser representada en forma de la gráfica de una función diferenciable. Hemos visto que un punto de la curva L , ordinario según nuestra definición, lo será también desde el punto de vista geométrico. Pueden aducirse ejemplos, cuando un punto de la curva, singular según nuestra definición, será ordinario desde el punto de vista geométrico. De este modo, nuestra definición de punto ordinario es más estrecha que la definición geométrica, pero más cómoda en las aplicaciones.

Introduzcamos ahora el concepto de *tangencia* de las curvas L_1 y L_2 en su punto común M_0 . Diremos que las curvas L_1 y L_2 son *tangentes* en su punto común M_0 , si ambas curvas tienen tangentes en el punto M_0 y estas tangentes coinciden.

En adelante nos hará falta la *condición de tangencia* de dos curvas L_1 y L_2 .

Supongamos que la curva L_1 se define por la ecuación (7.1), y la curva L_2 , por la (7.2), siendo $M_0(x_0, y_0)$ el punto común de dichas curvas (con la particularidad de que las coordenadas x_0 e y_0 corresponden al valor $\alpha = \alpha_0$ del parámetro α). Convergamos en considerar que el punto M_0 es punto *ordinario* de las curvas L_1 y L_2 y que estas curvas son *tangentes* en el punto M_0 . Entonces, las citadas curvas representan en el entorno de M_0 las gráficas de unas funciones diferenciables. Consideraremos, para concretar, que L_1 y L_2 son las gráficas de las funciones $y = f_1(x)$ o $y = f_2(x)$. Puesto que, por hipótesis, las curvas L_1 y L_2 son tangentes en el punto $M_0(x_0, y_0)$, los coeficientes angulares de las tangentes en M_0 a las gráficas de las funciones $f_1(x)$ y $f_2(x)$ son iguales, es decir,

$$f_1'(x_0) = f_2'(x_0). \quad (7.7)$$

Utilizando las fórmulas de diferenciación de las funciones definidas paramétricamente (véase § 11, cap. 5, t. 1) y las fórmulas de diferenciación de las funciones implícitas (véase p. 2, § 2, cap. 6), obtenemos

$$f'_1(x_0) = \frac{\psi'(\alpha_0)}{\varphi'(\alpha_0)}, \quad f'_2(x_0) = -\frac{F'_x(M_0)}{F'_y(M_0)}. \quad (7.8)$$

Las fórmulas (7.8) permiten atribuir a la igualdad (7.7) la forma siguiente

$$\frac{\psi'(\alpha_0)}{\varphi'(\alpha_0)} = -\frac{F'_x(M_0)}{F'_y(M_0)} \quad (7.9)$$

o bien

$$F'_x(M_0)\psi'(\alpha_0) + F'_y(M_0)\varphi'(\alpha_0) = 0. \quad (7.10)$$

En adelante llamaremos la última relación *condición de tangencia en el punto M_0 de las curvas L_1 y L_2* , definidas por las ecuaciones (7.1) y (7.2), respectivamente. Omitiendo los argumentos de las funciones y haciendo uso de las designaciones $\varphi' = \frac{dx}{d\alpha}$ y $\psi' = -\frac{dy}{d\alpha}$, escribiremos la condición de tangencia en la forma

$$F'_x \frac{dx}{d\alpha} + F'_y \frac{dy}{d\alpha} = 0. \quad (7.11)$$

OBSERVACIÓN 2 Si en un punto común M_0 de las curvas L_1 y L_2 , dadas por las ecuaciones (7.1) y (7.2), respectivamente, se cumple la condición de tangencia (7.10) (o. lo que es lo mismo, (7.11)), y si, además, el punto M_0 es punto ordinario de las curvas L_1 y L_2 , entonces, L_1 y L_2 son tangentes en el punto M_0 . En efecto, de las relaciones (7.3), (7.5) y de la condición (7.10) provienen o bien la condición (7.9), o bien la condición $\frac{\psi'(\alpha_0)}{\varphi'(\alpha_0)} = -\frac{F'_y(M_0)}{F'_x(M_0)}$, es decir, la igualdad entre los coeficientes angulares de las tangentes en el punto común M_0 de las curvas L_1 y L_2 . Esto significa precisamente que las curvas L_1 y L_2 son tangentes en el punto M_0 . Notemos que la condición de tangencia se cumple también cuando el punto M_0 es punto singular por lo menos de una de las curvas L_1 y L_2 .

Así, pues, la condición de tangencia (7.10) se cumple tanto cuando las curvas L_1 y L_2 son tangentes en el punto M_0 , como también en el caso en que M_0 es un punto singular por lo menos de una de estas curvas mencionadas.

2. Familias monoparamétricas de curvas planas. Puntos característicos de las curvas de una familia. En diferentes problemas físicos y geométricos se encuentran frecuentemente familias de curvas planas. En óptica geométrica se estudian, por ejemplo, los haces (familias) reflejados y refractados de rayos; en mecánica, las familias de

trayectorias posibles de una partícula material en un campo dado de fuerzas; en geometría, las familias de tangentes a las curvas. Uno de los métodos posibles de definir las familias de líneas de tal género consiste en lo siguiente. Se toma una función $F(x, y, \alpha)$ de tres variables

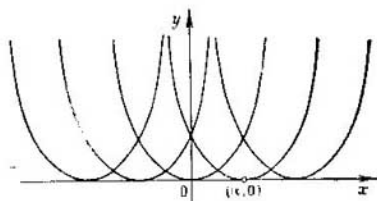


Fig. 7.1

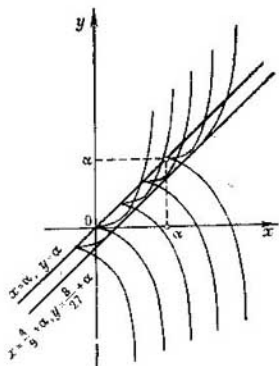


Fig. 7.2

x, y, α y para cada valor del parámetro α del conjunto dado $\{\alpha\}$ *) se define una curva de la familia mediante la ecuación

$$F(x, y, \alpha) = 0. \tag{7.12}$$

Diremos que la relación (7.12) define una familia monoparamétrica de curvas planas. Llamaremos el parámetro α parámetro de la familia

Veamos los ejemplos de familias monoparamétricas de curvas planas

1º La ecuación $y - (x - \alpha)^2 = 0$ define una familia de parábolas que se obtienen desplazando a lo largo del eje Ox de la parábola $y - x^2 = 0$ (fig. 7.1).

2º. La ecuación $(y - \alpha)^2 - (x - \alpha)^3 = 0$ define una familia de parábolas semicúbicas que se obtienen desplazando a lo largo de la bisectriz del primer ángulo coordenado de una parábola semicúbica $y^2 - x^3 = 0$ (fig. 7.2).

Supongamos que la función $F(x, y, \alpha)$ es diferenciable en el dominio de su definición. En este caso podemos introducir la noción de punto característico de la familia de curvas definida mediante la relación (7.12). Un punto $M(x, y)$ se denomina punto característico

*) El conjunto $\{\alpha\}$ representa, corrientemente, un intervalo.

de la familia de curvas (7.12), correspondiente al valor dado α del parámetro de la familia, si las coordenadas x e y de este punto satisfacen un sistema de ecuaciones

$$\left. \begin{aligned} F(x, y, \alpha) &= 0, \\ F'_\alpha(x, y, \alpha) &= 0. \end{aligned} \right\} \quad (7.13)$$

He aquí la interpretación geométrica del punto característico de la familia de curvas (7.12). Para simplificar, limitémonos al caso en que cualesquiera dos curvas de la familia se intersecan *). Sean

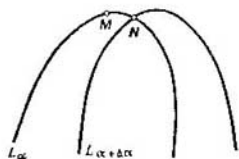


Fig. 7.3

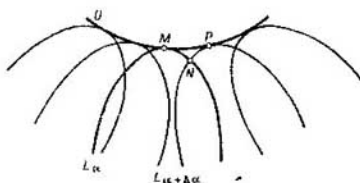


Fig. 7.4

L_α y $L_{\alpha+\Delta\alpha}$ dos curvas de la familia (7.12), correspondientes a los valores del parámetro α y $\alpha + \Delta\alpha$ (fig. 7.3). Las coordenadas del punto N de su intersección satisfacen el siguiente sistema de ecuaciones:

$$F(x, y, \alpha) = 0, \quad F(x, y, \alpha + \Delta\alpha) = 0$$

o bien un sistema equivalente

$$\begin{aligned} F(x, y, \alpha) &= 0, \\ \frac{F(x, y, \alpha + \Delta\alpha) - F(x, y, \alpha)}{\Delta\alpha} &= 0. \end{aligned}$$

Si $\Delta\alpha \rightarrow 0$, el punto N , dispuesto en la curva $F(x, y, \alpha) = 0$, tiende, en el caso general, hacia cierto punto M en esta curva. Por cuanto $\lim_{\Delta\alpha \rightarrow 0} \frac{F(x, y, \alpha + \Delta\alpha) - F(x, y, \alpha)}{\Delta\alpha} = F'_\alpha(x, y, \alpha)$, las coordenadas del punto M satisfacen las ecuaciones (7.13), por lo cual el punto M es punto característico de una curva de la familia. Así, pues, el punto característico de una curva dada es un punto que sirve de límite, para $\Delta\alpha \rightarrow 0$, de los puntos de intersección de la curva dada L_α con la curva $L_{\alpha+\Delta\alpha}$, próxima a ella.

3. Envolvente y curva discriminante de una familia monoparamétrica de curvas planas. Supongamos que una familia monoparamétrica

*) Existen familias de curvas en que dos curvas cualesquiera no se intersecan. Como ejemplo de tal familia puede servir una familia de parábolas definidas mediante la ecuación $y - (x - \alpha)^2 = 0$.

métrica de curvas planas se define mediante la relación (7.12). Conviengamos en considerar que la función $F(x, y, \alpha)$ es diferenciable en el dominio de su definición. Introduzcamos el concepto de *envoltente* de la familia de curvas (7.12).

Llámanse *envoltente de la familia monoparamétrica de curva* (7.12) a una curva O , la cual: 1) en cada uno de sus puntos es tangente a una sola curva de la familia (7.12), 2) en los puntos distintos es tangente a diferentes curvas de la familia citada. Razonamientos geométricos ilustrativos sugieren que la envoltente es tangente a las curvas de la familia en los puntos característicos de estas curvas y, por eso, puede considerarse, bajo condiciones determinadas, como lugar geométrico de puntos característicos de las curvas de la familia. En efecto, sea M un punto de tangencia de la envoltente O y la curva L_α de la familia, correspondiente al valor α del parámetro de la familia (fig. 7.4), y P , un punto de tangencia de la envoltente O y la curva $L_{\alpha+\Delta\alpha}$ de la familia, correspondiente al valor $\alpha + \Delta\alpha$ del parámetro, y sea N el punto de intersección de las curvas L_α y $L_{\alpha+\Delta\alpha}$. La figura 7.4 muestra claramente que cuando $\Delta\alpha \rightarrow 0$, los puntos P y N tienden al punto M , es decir, la envoltente O es tangente a la curva L_α precisamente en el punto característico M de esta curva.

Llamaremos *curva discriminante* de la familia de curvas (7.12) a un lugar geométrico de puntos característicos de las curvas de esta familia. Aclaremos en qué condiciones una curva discriminante es envoltente. Demostremos previamente el lema siguiente.

Lema. Sea $M_0(x_0, y_0)$ un punto característico de la familia $F(x, y, \alpha) = 0$, correspondiente al valor α_0 del parámetro de la familia. Supongamos también que las funciones $F(x, y, \alpha)$ y $F'_\alpha(x, y, \alpha)$ son diferenciables en cierto entorno del punto (x_0, y_0, α_0) y que las derivadas parciales de estas funciones respecto de x e y son continuas en el propio punto (x_0, y_0, α_0) . Entonces, si en el punto (x_0, y_0, α_0) el jacobiano $\frac{D(F, F'_\alpha)}{D(x, y)}$ es distinto de cero, la curva discriminante que pasa por M_0 puede ser definida en cierto entorno de este punto mediante las ecuaciones paramétricas $x = \varphi(\alpha)$ e $y = \psi(\alpha)$, donde $\varphi(\alpha)$ y $\psi(\alpha)$ son unas funciones diferenciables en cierto entorno de α_0 de la función.

DEMOSTRACIÓN. Por cuanto $M_0(x_0, y_0)$ es un punto característico de la curva de la familia, correspondiente al valor α_0 del parámetro de la familia, entonces $F(x_0, y_0, \alpha_0) = 0$ y $F'_\alpha(x_0, y_0, \alpha_0) = 0$, y, por eso, en virtud del lema, el sistema de ecuaciones (7.13) que define los puntos característicos de las curvas de la familia, satisface todas las exigencias del teorema 6.2 de resolubilidad del sistema de ecuaciones respecto de x e y . Por consiguiente, en cierto entorno del punto α_0 están definidas dos funciones

$$x = \varphi(\alpha), \quad y = \psi(\alpha). \quad (7.14)$$

que constituyen una única solución, continua y diferenciable, del sistema (7.13). Una curva, definida mediante las ecuaciones paramétricas (7.14), se compone de los puntos característicos y es, por eso, curva discriminante de la familia que pasa por el punto M_0 . Advertimos que, siendo única la solución del sistema (7.13), los puntos diferentes de la curva discriminante definida mediante las ecuaciones paramétricas (7.54) son puntos característicos de las diferentes curvas de la familia (7.12). Demos a conocer, ahora, las condiciones complementarias cuyo cumplimiento asegura que una curva discriminante que pasa por el punto M_0 , sea *envolvente* en cierto entorno de este punto.

Teorema 7.1. *Supongamos que, además de las condiciones enunciadas en el lema, se cumplen las condiciones siguientes: 1) en cierto entorno del punto (x_0, y_0, α_0) las derivadas $F'_x, F'_y, F''_{\alpha x}, F''_{\alpha y}$ y $F''_{\alpha\alpha}$ son continuas; 2) en el punto (x_0, y_0, α_0) se verifican las relaciones $F''_{x^2} + F''_{y^2} < 0, F''_{y^2} > 0$. Entonces, una curva discriminante que pasa por el punto característico $M_0(x_0, y_0)$ es en cierto entorno de este punto una envolvente de la familia de curvas en consideración.*

DEMOSTRACIÓN. Ya nos hemos convencido de que cierto tramo de la curva discriminante que pasa por el punto M_0 puede ser definido, bajo las condiciones formuladas, mediante las ecuaciones paramétricas (7.14) que constituyen la solución del sistema (7.13). Pongamos esta solución en las ecuaciones (7.13) y diferenciamos respecto de α las identidades obtenidas $F(x, y, \alpha) \equiv 0$ y $F_{\alpha}(x, y, \alpha) \equiv 0$. Obtendremos

$$\left. \begin{aligned} F'_x \frac{dx}{d\alpha} + F'_y \frac{dy}{d\alpha} + F'_{\alpha} &\equiv 0, \\ F''_{\alpha x} \frac{dx}{d\alpha} + F''_{\alpha y} \frac{dy}{d\alpha} + F''_{\alpha\alpha} &\equiv 0 \end{aligned} \right\} \quad (7.15)$$

Por cuanto $F'_{\alpha} \equiv 0$, la primera relación de (7.15) toma una forma

$$F'_x \frac{dx}{d\alpha} + F'_y \frac{dy}{d\alpha} \equiv 0.$$

Vemos que en todo punto de la curva discriminante que se considera se cumple la condición de tangencia (7.11) de esta curva y de la correspondiente curva de la familia. Por eso, en virtud de la observación 2 del p. 1, para dar por terminada la demostración del teorema, basta establecer que cada punto característico de la curva de la familia dispuesto en cierto entorno del punto M_0 y cada punto de la curva discriminante en este entorno son ordinarios. Por hipótesis del teorema, en el punto (x_0, y_0, α_0) se cumple la relación $F''_{x^2} + F''_{y^2} < 0$, la cual se cumplirá también, en virtud de que las derivadas parciales F'_x y F'_y son continuas, en cierto entorno del punto mencionado. Por consiguiente, en este entorno todos los puntos característicos de las curvas de la familia son ordinarios. De la relación $F''_{\alpha\alpha} \neq 0$ (válida

en virtud de que esta derivada es continua en cierto entorno del punto (x_0, y_0, α_0) y de la segunda identidad (7.15) se deduce que en el citado entorno las derivadas $\frac{dx}{d\alpha}$ y $\frac{dy}{d\alpha}$ no se anulan simultáneamente *). De este modo, todos los puntos de la curva discriminante en cierto entorno de M_0 son ordinarios. Del lema que acabamos de demostrar también se deduce que diferentes puntos de la curva discriminante son puntos característicos de las distintas curvas de la familia. El teorema está demostrado.

OBSERVACION 1. Los razonamientos aducidos al demostrar el teorema muestran que en el caso en que quedan cumplidas solamente las condiciones del lema, en cada punto de la curva discriminante se cumple la condición de tangencia de esta curva y de la curva de la familia. Ya se ha notado (véase observación 2, p. 1) que la condición de tangencia se cumple también cuando el punto común de dos curvas es un punto singular por lo menos de una de ellas. De aquí se desprende que la curva discriminante puede ser un lugar geométrico de puntos singulares de las curvas de la familia (si en cada punto característico se cumple la condición $F'_x{}^2 + F'_y{}^2 = 0$). Indiquemos que la propia curva discriminante también puede tener puntos singulares (si $\frac{dx}{d\alpha}$ y $\frac{dy}{d\alpha}$ son simultáneamente iguales a cero).

OBSERVACION 2. El teorema 7.1 puede interpretarse en el lenguaje geométrico del modo siguiente. *Si todas las curvas de la familia y la curva discriminante no tienen puntos singulares, la citada curva discriminante es una envolvente.*

Veamos algunos ejemplos.

1°. Hállese la curva discriminante de una familia $y - (x - \alpha)^2 = 0$.

Tenemos $F(x, y, \alpha) = y - (x - \alpha)^2$, $F'_\alpha(x, y, \alpha) = 2(x - \alpha)$. De este modo, el sistema (7.3) tiene por expresión

$$y - (x - \alpha)^2 = 0, \quad 2(x - \alpha) = 0.$$

De aquí se deduce que los puntos característicos tienen por coordenadas $(\alpha, 0)$ (véase fig. 7.1). Por eso, la curva discriminante se define mediante las ecuaciones paramétricas

$$x = \alpha, \quad y = 0.$$

Luego, se tiene $F'_x = -2(x - \alpha)$, $F'_y = 1$. En los puntos de la curva discriminante $F'_x{}^2 + F'_y{}^2 = 1$. Además, $F''_{\alpha\alpha} = -2 \neq 0$. De

*) La continuidad de estas derivadas se infiere inmediatamente de la continuidad de las derivadas F'_x , F'_y , $F''_{\alpha x}$, $F''_{\alpha y}$, y $F''_{\alpha\alpha}$ y de las relaciones (7.15), a partir de las cuales las derivadas $\frac{dx}{d\alpha}$ y $\frac{dy}{d\alpha}$ pueden hallarse por el método algebraico.

este modo, la curva discriminante está representada por el eje Ox que es la envolvente.

2°. Hállese la curva discriminante de una familia $(y - \alpha)^2 - (x - \alpha)^3 = 0$. Tenemos: $F(x, y, \alpha) = (y - \alpha)^2 - (x - \alpha)^3$, $F'_\alpha(x, y, \alpha) = -2(y - \alpha) + 3(x - \alpha)^2$. El sistema (7.13) tiene por expresión

$$(y - \alpha)^2 - (x - \alpha)^3 = 0, \quad -2(y - \alpha) + 3(x - \alpha)^2 = 0.$$

De aquí proviene que la curva discriminante consiste en dos líneas rectas, definidas mediante las ecuaciones paramétricas

$$x = \alpha, \quad y = \alpha \quad \text{y} \quad x = \frac{4}{9} + \alpha, \quad y = \frac{8}{27} + \alpha$$

(véase fig. 7.2). Es fácil convencerse de que en los puntos de la recta $x = \alpha, y = \alpha$ se cumple la condición $F'_y{}^2 + F'_x{}^2 = 0$, es decir, esta parte de la curva discriminante es un lugar geométrico de puntos singulares de las curvas de la familia. El lector se convencerá fácilmente de que la recta $x = \frac{4}{9} + \alpha, y = \frac{8}{27} + \alpha$ es la envolvente de la familia de curvas en consideración.

4. **Envolvente y superficie discriminante de una familia monoparamétrica de superficies.** Veamos una familia monoparamétrica de superficies que se define mediante una ecuación

$$F(x, y, z, \alpha) = 0. \quad (7.16)$$

Supondremos, además, que la función $F(x, y, z, \alpha)$ es diferenciable en su dominio de definición. Una línea L en la superficie de la familia (7.16), correspondiente al valor α del parámetro de la familia, se llama *característica* si las coordenadas de los puntos de esta línea satisfacen un sistema de ecuaciones

$$\left. \begin{aligned} F(x, y, z, \alpha) &= 0, \\ F'_\alpha(x, y, z, \alpha) &= 0. \end{aligned} \right\} \quad (7.17)$$

Un lugar geométrico de las características recibe el nombre de *superficie discriminante* de la familia (7.17).

Llámanse *envolvente* de una familia

monoparamétrica de superficies (7.16) a una superficie O , tangente a todas las superficies de la familia.

Podemos demostrar la afirmación siguiente.

Si ninguna de las superficies de una familia ni la superficie discriminante tienen puntos singulares, la citada superficie discriminante es una envolvente.

Advirtamos que una superficie discriminante puede representar el lugar geométrico de puntos singulares de las superficies de la familia y puede de por sí contar con puntos singulares.

Veamos el ejemplo siguiente. Hállese la envolvente de una familia de esferas de radio constante R , cuyos centros se ubican en los puntos de una curva dada L^* (la fig. 7.5 muestra una de semejantes esferas con centro en el punto M de la curva L), definida mediante las ecuaciones

$$x = \varphi(\alpha), \quad y = \psi(\alpha), \quad z = \chi(\alpha).$$

* Las envolventes de semejantes familias de esferas se llaman *superficies en canal*.

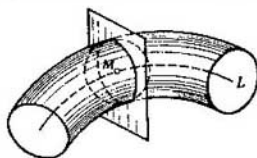


Fig. 7.5

La familia monoparamétrica de esferas que se considera se define mediante una ecuación:

$$[x - \varphi(\alpha)]^2 + [y - \psi(\alpha)]^2 + [z - \chi(\alpha)]^2 - R^2 = 0. \quad (7.18)$$

Las características de la familia mencionada de esferas se definen a partir de la ecuación (7.18) y de la ecuación

$$[x - \varphi(\alpha)] \varphi'(\alpha) + [y - \psi(\alpha)] \psi'(\alpha) + [z - \chi(\alpha)] \chi'(\alpha) = 0. \quad (7.19)$$

La ecuación (7.19) expresa un plano que pasa por el centro M de la esfera, perpendicularmente a la tangente a la curva L . Por eso, las características son unas circunferencias que constituyen las líneas de intersección de las esferas en consideración con los planos (7.19) (véase fig. 7.5). Notemos que si L es una circunferencia, la envolvente será un toro.

§ 2. Osculación de las curvas planas

1. Concepto de orden de osculación de las curvas planas. Supongamos que dos curvas L_1 y L_2 se tocan en cierto punto M_0 *) (fig. 7.6). Supongamos también que M es un punto arbitrario en la tangente común a las curvas L_1 y L_2 , y M_1, M_2 son los puntos de intersección con las curvas L_1 y L_2 de la perpendicular a la tangente mencionada,alzada en el punto M **) (véase fig. 7.6). Diremos que *dos curvas L_1 y L_2 tienen en el punto M_0 un orden de osculación n si existe un límite distinto de cero*

$$\lim_{M \rightarrow M_0} \frac{|M_1 M_2|}{|M M_0|^{n+1}}. \quad (7.20)$$

(En este caso $|M_1 M_2|$ denota la longitud del segmento $M_1 M_2$, y $|M M_0|$, la longitud del segmento $M M_0$).

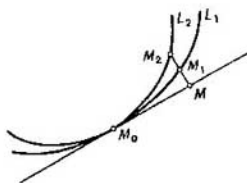


Fig. 7.6

OBSERVACIÓN 1. Si el límite (7.20) es igual a cero, suele decirse que las curvas L_1 y L_2 tienen un orden de osculación superior a n .

OBSERVACIÓN 2. Si dos curvas L_1 y L_2 tienen en el punto M_0 un orden de osculación superior a *cualquier* n , suele decirse que tienen en dicho punto un *orden infinito de osculación*.

EJEMPLOS 1°. Las curvas que representan las gráficas de las funciones $y = x^2$ e $y = 3x^2$ son tangentes en el origen de coordenadas O ; además, el eje Ox es su tangente común. Al tomar en el eje Ox un punto M de abscisa x , obtendremos que $|OM| = |x|$, y $|M_1 M_2| = |3x^2 - x^2| = 2x^2$.

*) Es decir, pasan por el punto M_0 y tienen en él tangentes coincidentes.

***) Se supone en este caso que si el punto M es suficientemente próximo a M_0 , la perpendicular alzada en el punto M a la tangente corta cada una de las curvas L_1 y L_2 solamente en un punto.

Por cuanto

$$\lim_{M \rightarrow 0} \frac{|M_1 M_2|}{|OM|^2} = \lim_{x \rightarrow 0} \frac{2x^2}{|x|^2} = 2 \neq 0,$$

las curvas en consideración tienen en el punto O un orden de osculación igual a la unidad.

2°. Examinemos ahora dos curvas L_1 y L_2 , la primera de las cuales coincide con el eje Ox , y la segunda sirve de gráfica para la función

$$y = \begin{cases} e^{-1/|x|} & \text{para } x \neq 0, \\ 0 & \text{para } x = 0. \end{cases}$$

Cerciorémonos de que las citadas curvas tienen un orden infinito de osculación en el origen de coordenadas O . Por cuanto Ox sirve de tangente común en el punto O , entonces, al tomar en este eje un punto M de abscisa x , llegamos a que $|OM| = |x|$, y $|M_1 M_2| = e^{-1/|x|}$. Resulta suficiente demostrar que para cualquier n el

límite $\lim_{M \rightarrow 0} \frac{|M_1 M_2|}{|OM|^{n+1}} = \lim_{x \rightarrow 0} \frac{e^{-1/|x|}}{|x|^{n+1}}$ es igual a cero. Suponiendo

$t = 1/|x|$, reduzcamos este límite al límite $\lim_{t \rightarrow +\infty} \frac{t^{n+1}}{e^t}$. Al final del p. 2 § 12, cap. 8 (t. 1) se ha demostrado que el último límite es igual a cero.

2. Orden de osculación de las curvas que sirven de gráfica para las funciones. Supongamos que dos curvas L_1 y L_2 representan las gráficas de las funciones $y = f_1(x)$ e $y = f_2(x)$, respectivamente. Admitamos también que dichas curvas son tangentes en cierto punto $M_0(x_0, f_1(x_0))$; además, el punto M_0 es ordinario para cada una de estas curvas*). Demostraremos que en estas condiciones la definición de orden de osculación de las curvas L_1 y L_2 , ofrecida en el punto antecedente, puede ser sustituida por otra definición equivalente, más cómoda para las aplicaciones.

Sea Δx un incremento arbitrario del argumento en el punto x_0 , y $x = x_0 + \Delta x$. Diremos que las curvas L_1 y L_2 tienen en el punto $M_0(x_0, f_1(x_0))$ un orden de osculación n si existe un límite distinto de cero

$$\lim_{\Delta x \rightarrow 0} \frac{|f_2(x_0 + \Delta x) - f_1(x_0 + \Delta x)|}{|\Delta x|^{n+1}} = \lim_{x \rightarrow x_0} \frac{|f_2(x) - f_1(x)|}{|x - x_0|^{n+1}}. \quad (7.21)$$

Para poder hablar del orden de osculación de las curvas en consideración L_1 y L_2 en el sentido de la definición dada en el punto

*) La definición de punto ordinario de una curva, definida mediante la ecuación $F(x, y) = 0$, fue propuesta en el p. 1, § 1. En particular, para una curva, definida mediante la ecuación $y = f(x)$, el punto $M_0(x_0, f(x_0))$ será ordinario si la derivada $f'(x)$ es solamente continua en un punto x_0 .

anterior, es necesario, ante todo, demostrar que en cierto entorno de M_0 dichas curvas se proyectan *unívocamente* sobre su tangente común. De esto se trata en el lema 1 que viene más abajo. La parte restante del presente punto se dedica a la demostración de dos lemas más, de los cuales se deduce inmediatamente la equivalencia de dos definiciones de orden de osculación de las curvas L_1 y L_2 .

Lema 1. Si el punto $M_0(x_0, f(x_0))$ es un punto ordinario de una curva L que sirve de gráfica para la función $y = f(x)$, la curva L en cierto entorno de M_0 se proyecta unívocamente sobre su tangente.

DEMOSTRACIÓN Notemos, ante todo, que la existencia de la tangente a la curva L en el punto M_0 se desprende de la existencia de la derivada $f'(x_0)$. Pasemos ahora a un nuevo sistema de coordenadas rectangulares cartesianas XY , colocando el nuevo origen en el punto de tangencia M_0 y dirigiendo el eje

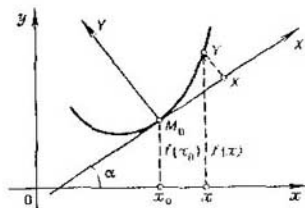


Fig. 7.7

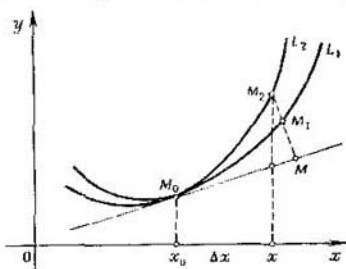


Fig. 7.8

X' por la tangente en M_0 (fig. 7.7). Si denotamos con α el ángulo entre el eje x y el eje X' , el nuevo sistema de coordenadas se obtiene, evidentemente, del antiguo por la traslación del origen de coordenadas al punto $M_0(x_0, f(x_0))$ y el giro de los ejes a un ángulo α . Haciendo uso de las fórmulas conocidas de transformación de las coordenadas al trasladar y al girar los ejes (véase el fascículo 8), obtendremos la expresión siguiente para las nuevas coordenadas X, Y de los puntos de la curva L en términos de las coordenadas antiguas x e $y = f(x)$:

$$\begin{aligned} X &= (x - x_0) \cos \alpha + [f(x) - f(x_0)] \sin \alpha, \\ Y &= -(x - x_0) \sin \alpha + [f(x) - f(x_0)] \cos \alpha. \end{aligned} \quad (7.22)$$

Alcanzaremos nuestro objetivo si mostramos que a base de las ecuaciones (7.22) podemos expresar Y como una función de X (esto significará precisamente que el tramo de la curva L , adyacente a M_0 , se proyecta unívocamente sobre el eje X , es decir, sobre la tangente). Para ello basta probar que en un entorno del punto M_0 la primera de las ecuaciones (7.22) es unívocamente resoluble respecto de x . (Efectivamente, expresando x a través de X de la primera ecuación (7.22) y colocando la expresión obtenida en la segunda ecuación (7.22), definiremos Y como una función de X). De conformidad con el teorema 6.1, es suficiente demostrar que la derivada respecto de x del miembro derecho de la primera de las ecuaciones (7.22) es distinta de cero en el punto x_0 y es continua en este punto. Pero esto es evidente, pues la derivada mencionada se expresa como $\cos \alpha + f'(x) \sin \alpha$, y en el punto x_0 , donde $f'(x_0) = \operatorname{tg} \alpha$, es igual a $1/\cos \alpha$. Por cuanto $\alpha \neq \pi/2$, resulta que $\cos \alpha \neq 0$. El lema está demostrado.

Supongamos que las curvas L_1 y L_2 que sirven de gráficas de las funciones $y = f_1(x)$ e $y = f_2(x)$ son tangentes en el punto $M_0(x_0, f_1(x_0))$, ordinario para cada una de estas curvas, y $x = x_0 + \Delta x$. Supongamos también que desde el punto $M_2(x, f_2(x))$ está trazada una perpendicular sobre la tangente en M_0 (fig. 7.8); M_1 es el punto de intersección de esta perpendicular con la curva L_1^* , y M , con la tangente. En estas condiciones resultan válidas las afirmaciones siguientes.

Lema 2. Existe un límite distinto de cero

$$\lim_{\Delta x \rightarrow 0} \frac{|f_2(x) - f_1(x)|}{|M_1M_2|}. \quad (7.23)$$

Lema 3. Existe un límite distinto de cero

$$\lim_{\Delta x \rightarrow 0} \frac{|M_0M|}{|\Delta x|}. \quad (7.24)$$

Notemos que de los lemas 2 y 3 se infiere que el límite (7.21), distinto de cero, existe cuando, y sólo cuando, existe el límite (7.20), distinto de cero, lo que significa precisamente que las dos definiciones de orden de osculación son equivalentes. Pasamos a la demostración de los lemas 2 y 3.

DEMOSTRACIÓN DEL LEMA 2. Denotemos, para simplificar, la distancia M_1M_2 con ρ , y, para concretar, convengamos en considerar que las curvas L_1 y L_2 están según muestra la fig. 7.8. Si X e Y son coordenadas del punto M_1 , y α es el ángulo de inclinación de la tangente en M_0 al eje Ox , entonces evidentemente,

$$X = x + \rho \operatorname{sen} \alpha,$$

$$Y = f_2(x) - \rho \operatorname{cos} \alpha.$$

Por cuanto el punto M_1 se dispone en la curva L_1 , resulta que $Y = f_1(x)$, o bien $f_2(x) - \rho \operatorname{cos} \alpha = f_1(x + \rho \operatorname{sen} \alpha)$. Aplicando al segundo miembro de la última igualdad la fórmula de Lagrange con relación al segmento $[x, x + \rho \operatorname{sen} \alpha]$ tendremos

$$f_2(x) - \rho \operatorname{cos} \alpha = f_1(x) + \rho \operatorname{sen} \alpha f'(\xi),$$

donde ξ es cierto punto dispuesto en el interior del segmento indicado. Por cuanto $\rho \rightarrow 0$ para $\Delta x \rightarrow 0$ y la derivada $f'(x)$ es continua en el punto x_0 , la última igualdad puede ser escrita en la forma

$$f_2(x) - \rho \operatorname{cos} \alpha = f_1(x) + \rho \operatorname{sen} \alpha [f'(x_0) + \varepsilon], \quad (7.25)$$

donde $\varepsilon \rightarrow 0$ para $\Delta x \rightarrow 0$. Considerando que $f'(x_0) = \operatorname{tg} \alpha$, de (7.25) obtendremos

$$\frac{f_2(x) - f_1(x)}{\rho} = \frac{1}{\operatorname{cos} \alpha} + \varepsilon \operatorname{sen} \alpha.$$

El lema 2 está demostrado.

DEMOSTRACIÓN DEL LEMA 3. Denotemos con β el ángulo comprendido entre la cuerda M_0M_2 y la tangente M_0M (fig. 7.9). Evidentemente, $\beta \rightarrow 0$ para $\Delta x \rightarrow 0$. De los triángulos rectángulos M_0M_2P y M_0M_2M (aquí, M_0P es paralelo al eje Ox) encontramos

$$|M_0M_2| = \frac{|\Delta x|}{\operatorname{cos}(\alpha + \beta)} = \frac{|M_0M|}{\operatorname{cos} \beta}.$$

*) En virtud del lema 1, para Δx suficientemente pequeño habrá un solo punto de este género.

La última fórmula evidencia que

$$\lim_{\Delta x \rightarrow 0} \frac{|M_0 M|}{|\Delta x|} = \lim_{\Delta x \rightarrow 0} \frac{\cos \beta}{\cos(\alpha + \beta)} = \frac{1}{\cos \alpha}.$$

El lema 3 está demostrado.

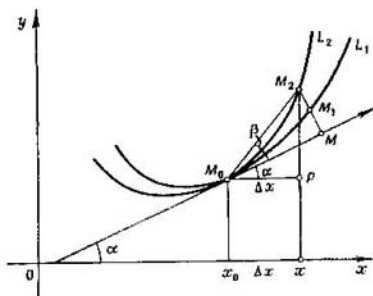


Fig. 7.9

3. Condiciones suficientes de osculación de orden n . Supongamos, al igual que antes, que las curvas L_1 y L_2 , que sirven de gráficas para las funciones $f_1(x)$ y $f_2(x)$, son tangentes en el punto $M_0(x_0, f_1(x_0))$. Es válida la afirmación siguiente.

Teorema 7.2. Sean $y = f_1(x)$ e $y = f_2(x)$ funciones $(n + 1)$ veces diferenciables en cierto entorno del punto x_0 ; sean, además, las derivadas de $(n + 1)$ -ésimo orden continuas en el propio punto x_0 . Entonces, si en el punto x_0 se verifican las relaciones

$$f_1(x_0) = f_2(x_0), f_1'(x_0) = f_2'(x_0), \dots \\ \dots, f_1^{(n)}(x_0) = f_2^{(n)}(x_0), f_1^{(n+1)}(x_0) \neq f_2^{(n+1)}(x_0), \quad (7.26)$$

las curvas L_1 y L_2 tienen en el punto $M_0(x_0, f_1(x_0))$ un orden de osculación n .

DEMOSTRACION Sea $F(x) = f_2(x) - f_1(x)$. Es suficiente demostrar que existe un límite distinto de cero

$$\lim_{x \rightarrow x_0} \frac{|F(x)|}{|x - x_0|^{n+1}}.$$

Por cuanto, en virtud de (7.26), $F(x_0) = F'(x_0) = \dots = F^{(n)}(x_0) = 0$, $F^{(n+1)}(x_0) \neq 0$, entonces, al escribir para la función $F(x)$ la fórmula de Taylor con el término residual en forma

de Lagrange, tendremos

$$F(x) = \frac{1}{(n+1)!} F^{(n+1)}(x_0 + \theta(x-x_0))(x-x_0)^{n+1}, \quad 0 < \theta < 1. \quad (7.27)$$

Por ser continua la derivada de $(n+1)$ -ésimo orden en el punto x_0 ,

$$F^{(n+1)}(x_0 + \theta(x-x_0)) = F^{(n+1)}(x_0) + \varepsilon, \quad (7.28)$$

donde $\varepsilon \rightarrow 0$ cuando $x \rightarrow x_0$. De las relaciones (7.27) y (7.28) obtenemos

$$\lim_{x \rightarrow x_0} \frac{|F(x)|}{|x-x_0|^{n+1}} = \frac{1}{(n+1)!} |F^{(n+1)}(x_0)| \neq 0.$$

El teorema está demostrado.

OBSERVACION 1. Si se cumplen todas las condiciones del teorema 7.2, a excepción, quizás, de la condición $f_1^{(n+1)}(x_0) \neq f_2^{(n+1)}(x_0)$, podemos afirmar que las curvas L_1 y L_2 tienen en el punto M_0 un orden de osculación no inferior a n .

OBSERVACION 2. Aclaremos el orden de osculación de la curva L , que sirve de gráfica de la función $y = f(x)$, con su tangente T en el punto $M_0(x_0, f(x_0))$. Consideraremos, además, que la función $f(x)$ es tres veces diferenciable en un entorno del punto x_0 y que su tercera derivada es continua en el punto x_0 . Recordemos que la tangente T sirve de gráfica de la función lineal $y = \varphi(x) = f'(x_0)(x-x_0) + f(x_0)$. Por cuanto $\varphi(x_0) = f(x_0)$, $\varphi'(x_0) = f'(x_0)$ y $\varphi''(x_0) = 0$, entonces, si $f''(x_0) \neq 0$, la curva L tiene con su tangente T un orden de osculación $n = 1$, y cuando $f''(x_0) = 0$, el orden mencionado de osculación no es inferior a dos.

4. Circunferencia osculatriz. Supongamos que $M_0(x_0, y_0)$ es un punto de la curva L que sirve de gráfica de la función $y = f(x)$, la cual tiene tercera derivada continua en el punto x_0 . Por el punto M_0 se pueden trazar un número infinito de circunferencias tangentes a la curva L en dicho punto. Es fácil convencerse de que la parte de cada una de semejantes circunferencias, dispuesta en cierto entorno del punto M_0 , es una gráfica de la función de la forma $y = y(x)$. Por eso, podemos hablar del orden de osculación de la curva L y cualquiera de las circunferencias citadas en su punto común M_0 . Aquella de las citadas circunferencias que tiene con la curva L un orden de osculación no inferior a dos se denomina *circunferencia osculatriz a la curva L en el punto M_0* . La afirmación siguiente establece las condiciones suficientes para que la curva L tenga en el punto M_0 una circunferencia osculatriz.

Teorema 7.7. *Supongamos que una curva L sirve de gráfica de la función $y = f(x)$, con la particularidad de que $f(x)$ tiene en el punto x_0 una derivada de segundo orden, distinta de cero, y tercera derivada, continua. Entonces, en el punto $M_0(x_0, f(x_0))$ existe una circunferencia osculatriz a la curva L .*

DEMOSTRACION. Buscaremos la ecuación de la circunferencia osculatriz en la forma

$$(x - a)^2 + (y - b)^2 = \rho^2, \quad (7.29)$$

donde a , b y ρ son unas constantes que han de ser definidas. Resolvamos la ecuación (7.29) respecto de y y pongamos la solución hallada $y = y(x)$ en el primer miembro de (7.29). Entonces, podemos considerar (7.29) como una identidad respecto de x (considerando, por supuesto, que $y = y(x)$). Diferenciamos la identidad (7.29) dos veces respecto de x y exijamos que en las relaciones que se obtienen como resultado de la diferenciación y en la propia relación (7.29) los valores y_0 , y'_0 , y''_0 de la función $y = y(x)$ y de dos primeras derivadas suyas en el punto x_0 sean iguales a $f(x_0)$, $f'(x_0)$, $f''(x_0)$, respectivamente:

$$y_0 = f(x_0), \quad y'_0 = f'(x_0), \quad y''_0 = f''(x_0).$$

De este modo, obtendremos el siguiente sistema de ecuaciones respecto de a , b y ρ :

$$\begin{aligned} (x_0 - a)^2 + (y_0 - b)^2 &= \rho^2, \\ (x_0 - a) + (y_0 - b) y'_0 &= 0, \\ 1 + (y'_0)^2 + (y_0 - b) y''_0 &= 0. \end{aligned}$$

A condición de que $y'' = f''(x_0) = 0$, $\rho > 0$, este sistema tiene una solución única:

$$a = x_0 - \frac{1 + y_0'^2}{y_0''} y_0', \quad b = y_0 + \frac{1 + y_0'^2}{y_0''}, \quad \rho = \frac{(1 + y_0'^2)^{3/2}}{|y_0''|}. \quad (7.30)$$

De los razonamientos aducidos se desprende que para una circunferencia, cuyo centro tiene por coordenadas (a, b) y cuyo radio ρ se define mediante las fórmulas (7.30), y para la curva L se cumplen todas las condiciones de la observación 1 al teorema 7.2. Por eso, la citada circunferencia será osculatriz.

OBSERVACION. Si, en las condiciones del teorema 7.3, el requisito $f''(x_0) \neq 0$ se sustituye por el opuesto, $f''(x_0) = 0$, la curva L en el punto M_0 no posee circunferencia osculatriz.

No obstante, en virtud de la observación 2 al teorema 7.2. en este caso la curva L y la tangente a la misma en el punto M_0 tienen en el punto mencionado un orden de osculación no inferior a dos. De este modo, podemos considerar que en el caso que se examina la circunferencia osculatriz degenera en el punto M_0 a una recta.

§ 3. Curvatura de una curva plana

1. **Concepto de curvatura de una curva plana.** Supongamos que una curva L está dada mediante las ecuaciones paramétricas $x = x(t)$, $y = y(t)$. Convengamos en considerar que para la curva L se cumplen las condiciones especificadas en el § 1, cap. 2*). En tal caso, podemos elegir una dirección positiva en cualquier punto M de la curva L . Pongámonos de acuerdo llamar dirección positiva en el punto dado M de la curva L aquella, en la cual se desplazará el punto M al aumentar el parámetro t .

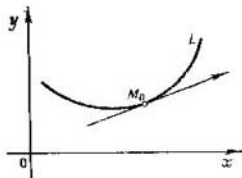


Fig. 7.10

Fijemos ahora en la curva L un punto M_0 , correspondiente al valor del parámetro t_0 , y supongamos que este punto es un punto ordinario de la curva L .

Entonces, el tramo de la curva L , adyacente al punto M_0 , será la gráfica de una función de la forma o bien $y = f(x)$, o bien

$x = f_1(y)$, y, además, existe una tangente a la curva L en el punto M_0 . Introduzcamos en esta tangente la dirección positiva correspondiente a la dirección positiva en el punto M_0 de la curva L **).

En adelante vamos a analizar sólo la tangente dirigida. En la fig. 7.10 la dirección de la tangente se indica con una flecha.

Ahora, supongamos que todos los puntos de la curva L , dispuestos en cierto entorno del punto fijo M_0 , son ordinarios. Sea M uno de los puntos indicados. Introduzcamos la noción de ángulo de contingencia del tramo de la curva M_0M . Para concretar, consideraremos que el punto M corresponde a un valor mayor del parámetro que el punto M_0 .

Llamaremos ángulo de contingencia del tramo de la curva M_0M a un ángulo formado por las tangentes dirigidas a la curva L en los puntos M y M_0 , el cual se toma con el signo más, si la tangente en el punto M_0 ha de ser girada en el sentido contrahorario para hacerla coincidir con la tangente en el punto M por la vía más corta, y con el signo menos, en el caso contrario.

En la fig. 7.11 está expuesto el tramo M_0M con un ángulo de contingencia positivo, y en la fig. 7.12, el tramo M_0M con un ángulo de contingencia negativo.

*) Es decir, consideraremos que la curva L no tiene puntos múltiples ni tramos de autocolocación.

***) Es decir, convengamos en llamar dirección positiva en la tangente aquella, en la que se desplazará, al aumentar el parámetro, un punto que representa la proyección de un punto de la curva L sobre la tangente.

Introduzcamos también el concepto de *curvatura media del tramo* M_0M de una curva. Llámase *curvatura media del tramo* M_0M de una curva a la razón entre el ángulo de contingencia de este tramo y la longitud del mismo.

Por cuanto el punto M_0 de la curva L se considera fijo, la curvatura media del tramo M_0M será una función del punto M , o una fun-

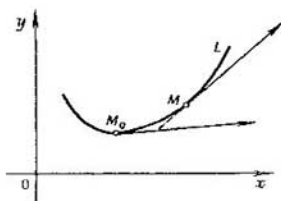


Fig. 7.11

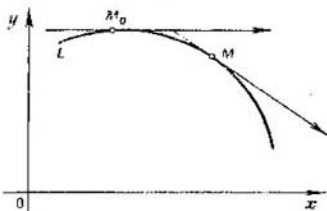


Fig. 7.12

ción del parámetro t . Denotaremos esta función con el símbolo $\kappa_{M_0}(M)$, o bien $\kappa_{M_0}(t)$.

Surge, naturalmente, una cuestión sobre el análisis del límite de esta función cuando el punto M tiende a lo largo de la curva L al punto M_0 (o bien, lo que es lo mismo, al tender el parámetro t hacia t_0).

Definición. El valor límite de la curvatura media del tramo M_0M de una curva, cuando el punto M tiende a lo largo de la curva al punto M_0 , lleva el nombre de *curvatura en el punto dado* M_0 de la curva L y se designa por el símbolo $k(M_0)$.

De este modo, por definición,

$$k(M_0) = \lim_{M \rightarrow M_0} \kappa_{M_0}(M) = \lim_{t \rightarrow t_0} \kappa_{M_0}(t).$$

Queda a cargo del lector cerciorarse de que: 1) tanto la curvatura media de cualquier tramo de una línea recta, como también la curvatura en cualquier punto de dicha línea son iguales a cero, 2) tanto la curvatura media de cualquier tramo de una circunferencia de radio R , como también la curvatura en cualquier punto de dicha circunferencia son iguales a $1/R$.

Establezcamos una fórmula para calcular la curvatura en cualquier punto arbitrario de la curva L .

2. Fórmula para calcular la curvatura. Sea L una curva dada por las ecuaciones paramétricas $x = x(t)$, $y = y(t)$, y sea M_0 un punto fijo de esta curva, correspondiente al valor del parámetro t_0 . Supongamos que todos los puntos de la curva L de cierto entorno de M_0 son ordinarios y que las funciones $x(t)$ e $y(t)$ tienen en el

punto t_0 segundas derivadas. En estas condiciones establezcamos la fórmula general para el cálculo de la curvatura de la curva L en el punto M_0 .

Sean \dot{x} y \ddot{x} los valores de las derivadas primera y segunda de la función $x = x(t)$ en el punto t_0 , y sea $\dot{x} + \Delta\dot{x}$, el valor de la primera derivada de esta función en el punto $t_0 + \Delta t$ (Δt es un incremento arbitrario del parámetro t). De este modo, Δx es el incremento de la primera derivada de la función $x = x(t)$. Supongamos que \dot{y} , \ddot{y} e

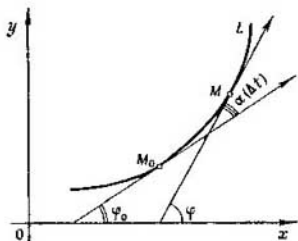


Fig. 7.13

$\dot{y} + \Delta\dot{y}$ son los valores correspondientes de las derivadas de la función $y = y(t)$.

Al considerar que el punto M_0 , correspondiente al valor del parámetro t_0 , es fijo, mientras que M corresponde al valor del parámetro $t = t_0 + \Delta t$, el ángulo de contingencia del tramo M_0M y la longitud de este tramo pueden considerarse como funciones del argumento Δt . Denotemos estas funciones con $\alpha(\Delta t)$ y $l(\Delta t)$, respectivamente.

Por definición, la curvatura $k(M_0)$ en el punto M_0 de la curva es igual al valor límite

$$k(M_0) = \lim_{\Delta t \rightarrow 0} \frac{\alpha(\Delta t)}{l(\Delta t)}. \quad (7.31)$$

Demostremos que el valor límite (7.31) existe y calculemoslo.

Denotemos con φ_0 y φ los ángulos de inclinación al eje Ox de las tangentes a la curva L trazadas por los puntos M_0 y M respectivamente (fig. 7.13). Entonces, evidentemente, cualquiera que sea la disposición de los puntos M_0 y M , para el ángulo de contingencia $\alpha(\Delta t)$ será válida la relación

$$\alpha(\Delta t) = \varphi - \varphi_0.$$

De la última relación se deduce que

$$\operatorname{tg} \alpha(\Delta t) = \frac{\operatorname{tg} \varphi - \operatorname{tg} \varphi_0}{1 + \operatorname{tg} \varphi \operatorname{tg} \varphi_0}. \quad (7.32)$$

Partiendo del sentido geométrico de la derivada (véase el p. 4, § 1, cap. 5, t. 1) y de la expresión para la derivada de una función dada paramétricamente (véase § 11, cap. 5, t. 1) podemos escribir

$$\operatorname{tg} \varphi_0 = \frac{\dot{y}}{\dot{x}}, \quad \operatorname{tg} \varphi = \frac{\dot{y} + \Delta\dot{y}}{\dot{x} + \Delta\dot{x}}.$$

De este modo, la fórmula (7.32) puede ser escrita en la forma

$$\operatorname{tg} \alpha(\Delta t) = \frac{\frac{\dot{y} + \Delta \dot{y}}{\dot{x} + \Delta \dot{x}} - \frac{\dot{y}}{\dot{x}}}{1 + \frac{\dot{y}(\dot{y} + \Delta \dot{y})}{\dot{x}(\dot{x} + \Delta \dot{x})}} = \frac{\dot{x} \cdot \Delta \dot{y} - \dot{y} \cdot \Delta \dot{x}}{\dot{x}(\dot{x} + \Delta \dot{x}) + \dot{y}(\dot{y} + \Delta \dot{y})}. \quad (7.33)$$

Por cuanto las funciones $x = x(t)$ e $y = y(t)$ tienen en el punto t_0 segundas derivadas, las primeras derivadas de estas funciones en el punto t_0 son continuas y, por lo tanto,

$$\lim_{\Delta t \rightarrow 0} \Delta \dot{y} = 0, \quad \lim_{\Delta t \rightarrow 0} \Delta \dot{x} = 0.$$

PERO, en este caso, en virtud de la igualdad (7.33),

$$\lim_{\Delta t \rightarrow 0} \operatorname{tg} \alpha(\Delta t) = 0.$$

De la última igualdad y de que el arco tangente es continuo se desprende que

$$\lim_{\Delta t \rightarrow 0} \alpha(\Delta t) = 0.$$

Por eso se verifica la igualdad

$$\lim_{\Delta t \rightarrow 0} \frac{\alpha(\Delta t)}{\operatorname{tg} \alpha(\Delta t)} = \lim_{\Delta t \rightarrow 0} \left[\cos \alpha(\Delta t) \frac{\alpha(\Delta t)}{\operatorname{sen} \alpha(\Delta t)} \right] = 1. \quad (7.34)$$

La igualdad (7.34) y el teorema del valor límite del producto reduce el problema de cálculo del valor límite (7.31) al cálculo del valor límite siguiente:

$$k(M_0) = \lim_{\Delta t \rightarrow 0} \frac{\operatorname{tg} \alpha(\Delta t)}{l(\Delta t)}. \quad (7.35)$$

Para el cálculo de este valor límite notemos que la longitud $l(\Delta t)$ del tramo de la curva M_0M se define por la fórmula

$$l(\Delta t) = \int_{t_0}^{t_0 + \Delta t} \sqrt{\dot{x}^2(\tau) + \dot{y}^2(\tau)} \, d\tau.$$

Al aplicar la fórmula del valor medio a la integral del segundo miembro de la última igualdad, obtendremos

$$l(\Delta t) = \Delta t \sqrt{\dot{x}^2(\tau^*) + \dot{y}^2(\tau^*)}, \quad (7.36)$$

donde $t_0 \leq \tau^* \leq t_0 + \Delta t$.

Las relaciones (7.35), (7.33) y (7.36) permiten concluir que el cálculo de la curvatura se reduce al cálculo del valor límite

$$k(M_0) = \lim_{\Delta t \rightarrow 0} \frac{\dot{x} \frac{\Delta y}{\Delta t} - \dot{y} \frac{\Delta x}{\Delta t}}{[\dot{x}(x + \Delta x) + \dot{y}(y + \Delta y)] \sqrt{\dot{x}^2(t^*) + \dot{y}^2(t^*)}}. \quad (7.37)$$

Por cuanto las funciones $x = x(t)$ e $y = y(t)$ tienen en el punto t_0 segundas derivadas, existen los valores límites

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta \dot{x}}{\Delta t} = \ddot{x} \quad \text{y} \quad \lim_{\Delta t \rightarrow 0} \frac{\Delta \dot{y}}{\Delta t} = \ddot{y}. \quad (7.38)$$

Luego, como las primeras derivadas de las funciones $x = x(t)$ e $y = y(t)$ son continuas, se deduce que

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \Delta \dot{x} &= 0, & \lim_{\Delta t \rightarrow 0} \Delta \dot{y} &= 0, \\ \lim_{\Delta t \rightarrow 0} \sqrt{\dot{x}^2(t^*) + \dot{y}^2(t^*)} &= \sqrt{\dot{x}^2 + \dot{y}^2}. \end{aligned} \quad (7.39)$$

La existencia de los valores límites (7.38) y (7.39) condiciona el valor límite del segundo miembro de (7.37), igual a

$$\frac{\ddot{x}\dot{y} - \dot{x}\ddot{y}}{[\dot{x}^2 + \dot{y}^2]^{3/2}}.$$

De este modo, queda demostrado que en el punto M_0 la curvatura de la curva L existe y se determina por la fórmula

$$k(M_0) = \frac{\ddot{x}\dot{y} - \dot{x}\ddot{y}}{[\dot{x}^2 + \dot{y}^2]^{3/2}}. \quad (7.40)$$

OBSERVACIÓN. Supongamos que se pide calcular la curvatura $k(M_0)$ en un punto dado M_0 de la curva L que sirve de gráfica para una función dos veces diferenciable $y = f(x)$.

Adoptando en la fórmula (7.40) $x = t$, $y = f(t)$, obtendremos para la curvatura buscada la fórmula siguiente

$$k(M_0) = \frac{f''(x)}{[1 + (f'(x))^2]^{3/2}}.$$

Por fin, si la curva está dada mediante la ecuación polar $r = r(\theta)$, donde $r(\theta)$ es una función dos veces diferenciable del ángulo polar θ , entonces, tomando por el parámetro t el ángulo polar θ y teniendo presente que $x = r(\theta) \cos \theta$, $y = r(\theta) \sin \theta$, obtendremos

mos la expresión siguiente para la curvatura:

$$k = \frac{r^2(\theta) + 2[r'(\theta)]^2 - r(\theta)r''(\theta)}{\{r^2(\theta) + [r'(\theta)]^2\}^{3/2}}.$$

Calculemos, a título de ejemplo, la curvatura en un punto arbitrario de la catenaria $y = a \operatorname{ch} \frac{x}{a}$.

Por cuanto

$$1 + [f'(x)]^2 = \operatorname{ch}^2 \frac{x}{a} = \frac{y^2}{a^2}, \quad f''(x) = \frac{1}{a} \operatorname{ch} \frac{x}{a} = \frac{y}{a^2},$$

la curvatura será igual a

$$k = \frac{a}{y^2} = \frac{1}{a \operatorname{ch}^2 \frac{x}{a}}.$$

§ 4. Evoluta y evolvente

1. Normal a una curva plana. Sea L una curva plana definida mediante las ecuaciones paramétricas

$$x = \varphi(t), \quad y = \psi(t). \quad (7.41)$$

Convengamos en considerar que la curva L no tiene puntos múltiples ni tramos de autocolocación. Además, consideraremos que todo punto de la curva L es ordinario*).

Introduzcamos la noción de normal a una curva en un punto dado M .

Una recta que está dispuesta en el plano de la curva L y que pasa por el punto M de la curva perpendicularmente a la tangente a L en el punto M , se llama normal a L en el punto M (fig. 7.14).

Hallemos la ecuación de la normal a una curva. Sean x e y las coordenadas del punto M de la curva L , y sean X e Y las coordenadas de cualquier punto N de la normal a la curva. Por definición, la normal es perpendicular a la tangente, y, por eso, el coeficiente angular k_n de la normal está asociado con el coeficiente angular k_t de la tangente

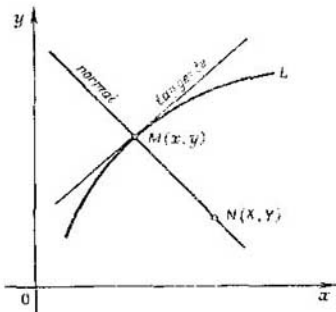


Fig. 7.14

*) Advertimos que en estas condiciones existe una tangente en cada punto de la curva L .

mediante una relación *)

$$k_n k_t = -1. \quad (7.42)$$

Por cuanto el coeficiente angular k_t de la tangente es igual a $\frac{\psi'}{\varphi'}$ (si la curva (7.41) está definida paramétricamente), de (7.42) obtenemos

$$k_n = -\frac{\varphi'}{\psi'}. \quad (7.43)$$

Haciendo uso de la ecuación de una recta con el coeficiente angular dado k_n , que se conoce por el curso de geometría analítica, obtendremos la siguiente ecuación para la normal a la curva L :

$$Y - y = -\frac{\varphi'}{\psi'} (X - x). \quad (7.44)$$

Teniendo presente que $x = \varphi(t)$ e $y = \psi(t)$, escribamos la ecuación (7.44) en la forma

$$\varphi' (X - \varphi) + \psi' (Y - \psi) = 0. \quad (7.45)$$

OBSERVACION. Cuando la tangente en un punto M es paralela al eje de abscisas, su coeficiente angular k_t es nulo y, por eso, las relaciones (7.42), (7.43) y (7.44) carecen de sentido. No obstante, en este caso la ecuación (7.45) es la ecuación de la normal. En efecto, si $k_t = 0$ (la tangente es paralela al eje de abscisas), $\psi' = 0$, y $\varphi' \neq 0$ **, por lo cual la relación (7.44) toma la forma: $X - \varphi = 0$. La última forma es la ecuación de una recta perpendicular al eje Ox , la cual, al cortar el eje Ox , deja en el mismo un segmento igual a φ . Está claro que esta recta coincide en el caso que se considera con la normal en el punto M .

2. Evoluta y evolvente de una curva plana. Supongamos que una curva L satisface las mismas condiciones que en el punto antecedente. Volvamos a la ecuación (7.45). Si consideramos t en esta ecuación como un parámetro, ésta es una ecuación *monoparamétrica de una familia de todas las normales* a la curva plana L . Una idea de una familia de normales a una curva plana la ofrece la fig. 7.15.

Bajo determinadas condiciones una familia monoparamétrica de normales cuenta con una envolvente que se denomina *evoluta* de la curva L .

Así, pues, *llámase evoluta de la curva plana L a la envolvente de una familia monoparamétrica de normales de dicha curva. La curva L , considerada con relación a su evoluta, se denomina evolvente.*

*) Esta relación se conoce por el curso de geometría analítica (véase, por ejemplo, el fascículo *Geometría analítica* del presente curso).

**) Recordemos que $k_t = \frac{\psi'}{\varphi'}$ y para un punto ordinario $\varphi'^2 + \psi'^2 \neq 0$.

Aclaremos las condiciones de existencia de la evoluta de la curva plana L y hallemos sus ecuaciones paramétricas.

Convengamos en considerar que la curva L sin puntos singulares viene dada mediante las ecuaciones paramétricas (7.41). Supongamos, para simplificar, que el parámetro t varía en el intervalo $(0, 1)$. Admitamos también que las funciones $\varphi(t)$ y $\psi(t)$ de las relaciones

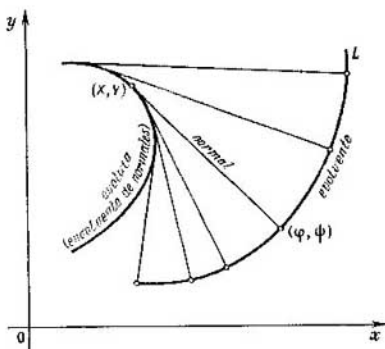


Fig. 7.15

(7.41), tienen terceras derivadas continuas sobre el intervalo $(0, 1)$. Entonces, resulta válido el teorema siguiente.

Teorema 7.4. *Supongamos que en todos los puntos de la curva L su curvatura k y la derivada de la curvatura *) no son nulas. Entonces, existe una evoluta de la curva L , con la particularidad de que las ecuaciones paramétricas de la evoluta tienen por expresión*

$$\left. \begin{aligned} X &= \varphi - \frac{\varphi'^2 + \psi'^2}{\varphi'\psi'' - \varphi''\psi'} \psi', \\ Y &= \psi + \frac{\varphi'^2 + \psi'^2}{\varphi'\psi'' - \varphi''\psi'} \varphi'. \end{aligned} \right\} \quad (7.46)$$

DEMOSTRACIÓN Hemos definido la evoluta como envolvente de una familia monoparamétrica de normales a la curva L . Esta familia se da mediante la ecuación (7.45). De este modo, la función

*) Para las condiciones impuestas en las funciones φ y ψ la curvatura k es una función diferenciable del parámetro t .

$F(X, Y, t)$ que define la familia se expresa con ayuda de la relación

$$F(X, Y, t) = \varphi'(X - \varphi) + \psi'(Y - \psi), \quad (7.47)$$

donde t desempeña el papel de parámetro.

Aprovechemos ahora las deducciones del § 1 de este capítulo sobre la existencia de una envolvente para aclarar la cuestión de existencia de una evoluta (es decir, de una envolvente de la familia (7.45)). Hemos de comprobar el cumplimiento de las condiciones del lema del p. 3, § 1 de este capítulo y de las condiciones del teorema 7.1. Pasemos a la comprobación de las condiciones mencionadas. Detengámonos primero en la comprobación de las condiciones del lema. Es evidente que con las exigencias formuladas para las funciones φ y ψ las funciones $F(X, Y, t)$ y $F'_t(X, Y, t)$ son diferenciables. Cerciorémonos ahora de que el jacobiano $\frac{D(F, F'_t)}{D(X, Y)}$ es distinto de cero. Haciendo uso de la expresión (7.40) para la curvatura k de la curva L , podemos representar este jacobiano en la forma siguiente:

$$\frac{D(F, F'_t)}{D(X, Y)} = k[\varphi'^2 + \psi'^2]^{3/2}.$$

Por hipótesis del teorema, la curvatura k es distinta de cero. Además, por cuanto todos los puntos de la curva L son ordinarios, tenemos $\varphi'^2 + \psi'^2 \neq 0$. De este modo, el jacobiano mencionado es distinto de cero y, por consiguiente, todas las condiciones del lema están cumplidas.

Comprobemos ahora las condiciones del teorema 7.1. Es evidente que con las exigencias formuladas para las funciones φ y ψ las derivadas $F'_X, F'_Y, F''_{tX}, F''_{tY}, F''_{tt}$ son continuas. Resta sólo convencerse de que la relación $F''_{tt} \neq 0$ es válida para todos los valores de t del intervalo $(0, 1)$ y para los puntos característicos en las normales a la curva L .

Por cuanto

$$F''_{tt} = \varphi'''(X - \varphi) + \psi'''(Y - \psi) - 3(\varphi'\varphi'' + \psi'\psi'') \quad (7.48)$$

y, según el p. 2, § 1 de este capítulo y las fórmulas (7.13), los puntos característicos para el caso en cuestión se determinan por las relaciones

$$\begin{aligned} \varphi'(X - \varphi) + \psi'(Y - \psi) &= 0, & (7.49) \\ \varphi''(X - \varphi) + \psi''(Y - \psi) - (\varphi'^2 + \psi'^2) &= 0, \end{aligned}$$

el valor de la derivada F''_{tt} en los puntos característicos será igual, en virtud de (7.48) y (7.49), a

$$F''_{tt} = \frac{\varphi'\psi''' - \varphi'''\psi'}{\varphi'\psi'' - \varphi''\psi'}(\varphi'^2 + \psi'^2) - 3(\varphi'\psi'' + \psi'\psi'''). \quad (7.50)$$

Volvamos a la expresión (7.40) para la curvatura k de la curva L . Habida cuenta de las designaciones (7.41), obtenemos de la fórmula (7.40) por diferenciación:

$$k' = -\frac{\varphi'\psi'' - \varphi''\psi'}{[\varphi'^2 + \psi'^2]^{5/2}} \left\{ \frac{\varphi'\psi'' - \varphi''\psi'}{\varphi'\psi'' - \varphi''\psi'} (\varphi'^2 + \psi'^2) - 3(\varphi'\varphi'' + \psi'\psi'') \right\}.$$

Con ayuda de esta relación y de la expresión (7.40) para la curvatura demos a la expresión (7.50) la forma siguiente:

$$F''_{tt} = \frac{k'}{k} (\varphi'^2 + \psi'^2).$$

Dado que, por hipótesis del teorema, k y k' no son nulos y, además, $\varphi'^2 + \psi'^2 \neq 0$, de la última expresión para F''_{tt} proviene que para todos los valores de t del intervalo $(0, 1)$ y para los puntos característicos de las normales a $LF''_{tt} \neq 0$. De este modo, las condiciones del teorema 7.1 también están cumplidas.

Hallemos ahora las ecuaciones paramétricas de la evoluta. Por cuanto, para las condiciones formuladas, la evoluta es un lugar geométrico de puntos característicos de la familia de normales, las coordenadas X e Y de los puntos de la evoluta se determinan a partir de las relaciones (7.49). Encontrando a partir de estas relaciones X e Y , obtendremos precisamente las ecuaciones paramétricas (7.46). El teorema queda demostrado.

OBSERVACIÓN. En geometría se usan frecuentemente los conceptos de radio de curvatura y de centro de curvatura. Llámase *radio de curvatura* R de una curva L a una magnitud $1/k$, donde k es la curvatura de L , y llámase *centro de curvatura* a un punto de la normal a la curva que dista de un punto dado de la curva L en dirección de la concavidad de L a una magnitud igual a R .

Notemos que el radio de curvatura de una curva es igual al radio de la circunferencia osculadora y el centro de curvatura coincide con el centro de la circunferencia osculatriz.

Cerciorémonos de que la evoluta es un lugar geométrico de centros de curvatura de la curva. En efecto, de las relaciones (7.46) obtenemos

$$(X - \varphi)^2 + (Y - \psi)^2 = \frac{1}{k^2} = R^2.$$

es decir, un punto de la evoluta con las coordenadas (X, Y) dista del punto de la curva L con las coordenadas (φ, ψ) a una magnitud R . El sentido geométrico de la evoluta (véase fig. 7.15) muestra que el punto (X, Y) está dispuesto en la normal orientada hacia la concavidad de la curva L .

EJEMPLO. Hallemos las ecuaciones de la evoluta de una elipse $ABCD$ (fig. 7.16) definida mediante las ecuaciones paramétricas

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

Por cuanto $\varphi' = -a \operatorname{sen} t$, $\psi' = b \operatorname{cos} t$, $\varphi'' = -a \operatorname{cos} t$, $\psi'' = -b \operatorname{sen} t$, tenemos, pues, $\varphi'^2 + \psi'^2 = a^2 \operatorname{sen}^2 t + b^2 \operatorname{cos}^2 t$, y $\varphi'\psi'' - \psi'\varphi'' = ab$. Por eso, de acuerdo con (7.46), las ecuaciones

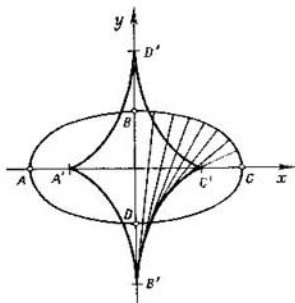


Fig. 7.16

paramétricas de la evolvente de la elipse tienen por expresión

$$X = \frac{a^2 - b^2}{a} \operatorname{cos}^3 t, \quad Y = \frac{b^2 - a^2}{b} \operatorname{sen}^3 t.$$

De este modo, la evolvente de una elipse es la llamada *astroide alargada* (véase fig. 7.16).

Advirtamos que en los puntos A , B , C , D de la elipse (es decir, en sus vértices) la derivada de la curvatura es igual a cero. Por eso, en dichos puntos no se cumplen las condiciones del teorema 7.4. En los puntos correspondientes A' , B' , C' , D' la evolvente de la elipse tiene singularidades, es decir, los llamados puntos de retroceso.

INDICE ALFABÉTICO DE MATERIAS

- Abel 46
- Aplicación biunívoca 262
 - — homeomorfa 263
- Area inferior 65
 - del sector curvilíneo 67
 - de la superficie de revolución 73
 - superior 65
 - del trapecio curvilíneo 67
- Astroide 73

- Bonnet 33
- Borel 21
- Buniakovski 42

- Cambio de variable bajo el signo de la integral definida 37
- Característica 290
- Centro de curvatura 307
- Cesaro 151
- Cicloide 63
- Circunferencia osculatriz 296
- Condición de Cauchy 167
- Condición necesaria y suficiente de existencia del valor límite de una función (criterio de Cauchy) 167
 - — — — de integrabilidad 17
- Condiciones necesarias de un extremo local de la función 209
- Condiciones suficientes de un extremo local de la función 210
- Conjunto cerrado 160
 - conexo 161
 - inconexo 161
- Continuidad de una función de varias variables 168
 - uniforme de una función 19
- Convergencia absoluta del producto infinito 145
- Criterio de d'Alembert 116
 - de Cauchy 118
 - de Dirichlet-Abel 137
 - integral de Cauchy-Maclaurin 120
 - de Leibniz 136
 - de Raabe 123
- Criterios de comparación 113
- Cuerpo cubicable 71

- Curva discriminante 287
 - rectificable 54
 - simple 49
- Curvatura de una curva plana 298

- D'Alembert 116
- Derivada según la dirección 188
 - parcial de la función 175
- Derivadas parciales de orden superior 191
- Desigualdad de Buniakovski para las sumas 42
 - de Cauchy—Buniakovski para las integrales 45
 - de Hölder para las integrales 44
 - — — — sumas 41
 - de Minkowski para las integrales 45
 - — — — sumas 43
- Diámetro del conjunto 175
- Diferenciación de una función compuesta 183
- Diferencial 183
 - de un arco 62
 - de orden superior 191
- Distancia del punto al conjunto 221
- Dominio 160
 - cerrado 161
 - de definición de la función 162
 - limitado 161

- Elipse 63
- Entorno del punto 160, 162
 - rectangular del punto 160
- Envolvente de la familia monoparamétrica de la curva 287
 - — — — de superficies 290
- Espacio coordenado 156, 158
 - euclídeo 156, 158
 - métrico 159
- Estrofoide 50
- Evoluta 304
- Evolvente 304
- Existencia del mínimo de una función fuertemente convexa 227

- Extremo local de la función 209
 — — de una función de m variables 209
- Familia monoparamétrica de curvas planas 285
- Figura plana cuadrable 64
- Forma cuadrática de signo definido 211
 — en diferencias de la condición de continuidad de la función 169
 — invariante de la diferencial 187
- Fórmula de Newton-Leibniz 36
- de Simpson (método de las parábolas) 102, 103
 — de Taylor para una función de m variables 202
 — de los trapecios 101
 — de Wallis 448
- Función compuesta 172
 — continua sobre el conjunto 169
 — — en el punto 169
 — — sobre una recta 171
 — diferenciable 177
 — fuertemente convexa 227
 — homogénea 186
 — implícita 244
 — infinitésima en un punto 166
 — integrable 10
 — inversa 255
 — de Lagrange 273
 — uniformemente continua 19
 — de m variables 162
 — n veces diferenciable 193
- Gradiente 188
- Gráfica de la función 157
- Heine 21
- Hölder 41
- Incremento parcial de la función 170
 — total de la función 169
- Integración por partes 38
- Integral definida 40
- Integrales de Darboux 14
- Jacobi 250
- Jacobiano 256
- Jordan 64
- Laplace 280
- Lebesgue 22
- Lema de Darboux 16
 — de Heine—Borel 21
- Límite reiterado 168
 — de la sucesión 163
- Línea catenaria 63
- Longitud del arco de la curva 54
 — máxima de los segmentos parciales 10
- Máximo local de la función 209
- Método combinado 94
 — de las cuerdas 85, 92
 — gradiente de búsqueda del extremo 220
 — de la «horquilla» 83
 — de multiplicadores indeterminados de Lagrange 273
 — de las parábolas (fórmula de Simpson) 102, 103
 — de los rectángulos 97
 — de la sumación de Cesaro 151
 — — — de Poisson—Abel 152
- Métodos generalizados de la sumación 151
 — de las iteraciones (aproximaciones sucesivas) 86
- Mínimo local de la función 209
- Minkowski 43
- Normal a la superficie 181
- Operador de Laplace 280
- Orden de osculación de las curvas planas 291
- Oscilación de la función 18, 175
- Parámetro natural 55
- Plano coordenado 156
 — euclídeo 156
 — tangente 179
- Poisson 152
- Primer teorema de Weierstrass 174
- Primera fórmula del valor medio 31, 32
- Producto infinito 140
 — — absolutamente convergente 145
 — — condicionalmente convergente 145
 — — convergente 141

- Progresión geométrica 108
 Proyección del punto sobre el conjunto 221
 Punto característico 285
 — de frontera del conjunto 160
 — interior del conjunto 160
 — ordinario de la curva 254, 283
 — — de la superficie 254
 — singular de la curva 254, 283
 — — de la superficie 254

 Raabe 123
 Radio de curvatura 307
 Resto de la serie 110
 Riemann 10

 Segmento del espacio E^m 220
 Segunda fórmula del valor medio 33
 Segundo teorema de Weierstrass 174
 Serie absolutamente convergente 126
 — alternada 136
 — armónica 112
 — — generalizada 116
 — condicionalmente convergente 127
 — convergente 107
 — divergente 107
 — de Leibniz 136
 — numérica 107
 Simpson 103

 Sucesión acotada 164
 — de Cauchy 163
 — convergente 162
 — fundamental 163
 — iterativa 86
 Suma inferior 11
 — integral 9
 — parcial de la serie 107
 — de la serie 107
 — superior 12
 Superficie de nivel 190
 Sylvester 211

 Tangencia de las curvas 283
 Teorema de Bolzano-Weierstrass 164
 — de Euler sobre las funciones homogéneas 186
 — de Riemann 129
 Término residual de la fórmula de Taylor en forma integral 39
 Transformación de Abel 47, 125
 Trapecio curvilíneo 67

 Valor límite de la función 165
 — — reiterado 167
 — particular de la función 162
 Volumen inferior del cuerpo 71
 — superior del cuerpo 71

 Wallis 146

A NUESTROS LECTORES:

Mir edita libros soviéticos traducidos al español, inglés, francés, árabe y otros idiomas extranjeros. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica, manuales para los centros de enseñanza superior y escuelas tecnológicas, literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia-ficción.

Dirijan sus opiniones a la Editorial Mir, 1 Rizhskí per., 2, 129820, Moscú, I-110, GSP, URSS.