CURSO DE ANÁLISIS MATEMÁTICO

L.D. KUDRIÁVTSEV

2

3 s. antes de n.e.
ARQUÍMEDES

1596 R.DESCARTES 1650

1601 P. FERMAT 1665

1643 I.NEWTON 1727

1646 G. LEIBNIZ 1716

1707 L. EULER 1783

1685 B.TAYLOR 1731

1717 J. D'ALEMBERT 1783

1736 J.LAGRANGE 1813

1749 P. LAPLACE 1827

1777 C. GAUSS 1855

1772 J. FOURIER 1837

1789 A.CAUCHY 1857

1781 B.BOLZANO 1848

1792 N.I. LOBACHEVSKI 1856

1902 M.V. OSTROGRADSKI 1862

Л. Д. КУДРЯВЦЕВ

КУРС МАТЕМАТИЧЕСКОГО АНАЛИЗА

Том 2

МОСКВА
«ВЫСШАЯ ШКОЛА»

CURSO DE ANÁLISIS MATEMÁTICO

L. D. KUDRIÁVTSEV

2

EDITORIAL MIR MOSCÚ

Traducido del ruso por el ingeniero K. P. Medkov

Impreso en la URSS

На испанском языке

Este libro es la segunda parte del Curso de análisis matemático en dos tomos. En el libro vienen expuestas las cuestiones, estudiadas corrientemente por los estudiantes de segundo año. La numeración de los capítulos, párrafos y figuras en este tomo continúa la numeración correspondiente del primer tomo.

El quinto capítulo con el que empieza este tomo está dedicado al cálculo diferencial de las funciones de varias variables y representa, en esencia, la continuación inmediata del segundo capítulo del primer tomo. Los capítulos ulteriores contienen la exposición del cálculo integral de las funciones de varias variables, la teoría de las series y de la integral de Fourier. La transformación de Fourier se expone al principio en la forma clásica, y luego se dan sus generalizaciones para el espacio L_2 y para las funciones generalizadas. El tomo se termina por un "Complemento", cuya parte principal concierne a los métodos numéricos destinados para calcular los valores aproximados de las funciones, las soluciones aproximadas de ecuaciones y a los cálculos aproximados de las integrales.

Indice

CAPÍTULO QUINTO

Cálculo diferencial de la función de varias variables (continuación)

CAPÍTULO SEXTO

Cálculo integral de las funciones de varias variables

CAPÍTULO VII

Series de Fourier. Integral de Fourier

Complemento

CAPÍTULO QUINTO

CÁLCULO DIFERENCIAL DE LA FUNCIÓN DE VARIAS VARIABLES (CONTINUACIÓN)

§ 39. FÓRMULA DE TAYLOR Y SERIE DE TAYLOR PARA LAS FUNCIONES DE VARIAS VARIABLES

39.1. FÓRMULA DE TAYLOR PARA LAS FUNCIONES DE VARIAS VARIABLES

Si una función de varias variables tiene un número suficiente de derivadas continuas en un entorno de cierto punto, dicha función puede ser representada dentro del entorno citado (al igual que se hizo en el caso de la función de una sola variable) en forma de una suma de cierto polinomio y un resto que es, en determinado sentido, "pequeño".

Teorema 1. *Supongamos que la función $z = f(x, y)$ está definida y es continua, lo mismo que todas sus derivadas parciales hasta el orden m inclusive ($m \geqslant 1$), en un δ-entorno**) *del punto (x_0, y_0). Entonces, para cualesquiera Δx y Δy, que satisfacen la condición $\rho = \sqrt{\Delta x^2 + \Delta y^2} < \delta$, existe tal $\theta = \theta(\Delta x, \Delta y)$, $0 < \theta < 1$, que resulta ser válida la fórmula*

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

$$+ \frac{1}{2!}\left[\frac{\partial^2 f(x_0, y_0)}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f(x_0, y_0)}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f(x_0, y_0)}{\partial y^2}\Delta y^2\right] +$$

$$+ \frac{1}{3!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[3]} f(x_0, y_0) + \ldots$$

$$\ldots + \frac{1}{(m-1)!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[m-1]} f(x_0, y_0) + r_{m-1}(\Delta x, \Delta y)$$

*) En algunas obras se emplea el término "δ-vecindad". (*N. del Tr.*)

o, en la forma más breve,

$$\Delta z = \sum_{k=1}^{m-1} \frac{1}{k!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[k]} f(x_0, y_0) + r_{m-1}(\Delta x, \Delta y), \tag{39.1}$$

donde

$$r_{m-1}(\Delta x, \Delta y) = \frac{1}{m!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[m]} f(x_0 + \theta\Delta x, y_0 + \theta\Delta y). \tag{39.2}$$

La fórmula (39.1) se denomina *fórmula de Taylor* (de orden $m - 1$) para la función f y la función $r_{m-1}(\Delta x, \Delta y)$, *su término residual*, mientras que la inscripción (39.2) de éste lleva el nombre de término residual de la fórmula de Taylor en la *forma de Lagrange*.

Cuando $m = 1$, en (39.1) requiere explicaciones el sentido del primer término del segundo miembro, puesto que en este caso el supraíndice de sumación es igual a cero. En el caso dado se supone, según la definición, que el término citado es igual a cero, es decir, la fórmula (39.1) adopta la forma

$$\Delta z = r_0(\Delta x, \Delta y).$$

En adelante, siempre cuando se encuentre una expresión escrita con ayuda del símbolo $\sum$, en la que el valor del supraíndice de sumación sea inferior al valor del subíndice, convendremos en considerar también que esta expresión es igual a cero.

DEMOSTRACIÓN. Supongamos que Δx y Δy están fijados de una manera tal que $\rho = \sqrt{\Delta x^2 + \Delta y^2} < \delta$, entonces todos los puntos del tipo $(x_0 + t\Delta x, y_0 + t\Delta y)$, donde $0 \leqslant t \leqslant 1$, se disponen en un segmento que une los puntos (x_0, y_0) y $(x_0 + \Delta x, y_0 + \Delta y)$, por lo cual todos ellos pertenecen al δ-entorno del punto (x_0, y_0). Por esta razón tiene sentido la composición de funciones

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

y

$$x = x_0 + t\Delta x, \quad y = y_0 + t\Delta y, \quad 0 \leqslant t \leqslant 1,$$

es decir, la función compuesta

$$F(t) = f(x_0 + t\Delta x, y_0 + t\Delta y), \quad 0 \leqslant t \leqslant 1. \tag{39.3}$$

Es evidente que

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = F(1) - F(0). \tag{39.4}$$

Como la función f tiene en el δ-entorno del punto (x_0, y_0) m derivadas parciales continuas, de acuerdo con el teorema sobre las derivadas de una función compuesta (véase el p. 20.3), la función F tendrá también en el segmento $[0, 1]$ m derivadas continuas, por lo cual para F será válida la fórmula de Taylor de orden $m - 1$ con el término residual en la forma de Lagrange:

$$F(t) - F(0) = F'(0)t + \frac{F''(0)}{2!}t^2 + \ldots + \frac{F^{(m-1)}(0)}{(m-1)!}t^{m-1} + \frac{F^{(m)}(\theta t)}{m!}t^m,$$

$$0 < \theta < 1, \quad (39.5)$$

y en el entorno considerado del punto (x_0, y_0) la función (39.3) puede derivarse m veces conforme a la regla para derivar funciones compuestas (véase la observación 2 en p. 20.4), con la particularidad de que los valores de las derivadas parciales mixtas que se obtienen *no dependen* del orden en que se lleva a cabo la derivación (véase el p. 21.1).

Al expresar las derivadas $F^{(k)}(t)$ en términos de las derivadas de la función $f(x, y)$ y al poner en la fórmula (39.5) $t = 1$ (véase (39.4)), obtendremos la función requerida de Taylor correspondiente a la función $f(x, y)$. En efecto, de (39.3) se deduce que

$$F'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dx}{dt} = \\ = \frac{\partial f(x_0 + t\Delta x, y_0 + t\Delta y)}{\partial x}\Delta x + \frac{\partial f(x_0 + t\Delta x, y_0 + t\Delta y)}{\partial y}\Delta y.$$

De aquí, al omitir para abreviar las designaciones de los argumentos, obtendremos para $F''(t)$

$$F''(t) = \frac{d}{dt}\left(\frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y\right) = \frac{\partial^2 f}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f}{\partial y^2}\Delta y^2.$$

En general, por inducción es fácil establecer que

$$F^{(k)}(t) = \left(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y}\right)^{[k]} f(x_0 + t\Delta x, y_0 + t\Delta y), \\ k = 1, 2, \ldots, m. \quad (39.6)$$

Al poner en las fórmulas (39.6) $t = 0$ para $k = 1, 2, \ldots, m - 1$, tendremos:

$$F'(0) = \frac{\partial f(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f(x_0, y_0)}{\partial y}\Delta y;$$

$$F''(0) = \frac{\partial^2 f(x_0, y_0)}{\partial x^2}\Delta x^2 + 2\frac{\partial^2 f(x_0, y_0)}{\partial x \partial y}\Delta x \Delta y + \frac{\partial^2 f(x_0, y_0)}{\partial y^2}\Delta y^2,$$

y, en general,

$$F^{(k)}(0) = \left(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y}\right)^{[k]} f(x_0, y_0), \quad k = 1, 2, \ldots, m - 1. \quad (39.7)$$

Cuando $k = m$, al sustituir t por θt,

$$F^{(m)}(\theta t) = \left(\Delta x\frac{\partial}{\partial x} + \Delta y\frac{\partial}{\partial y}\right)^{[m]} f(x_0 + \theta t\Delta x, y_0 + \theta t\Delta y). \quad (39.8)$$

Sustituyamos ahora (39.7) y (39.8) en (39.5) y pongamos $t = 1$; en este caso, en virtud de la relación (39.4)

$$\Delta z = F(1) - F(0) = \sum_{k=1}^{m-1} \frac{F^{(k)}(0)}{k!} + \frac{F^{(m)}(0)}{m!} =$$

$$= \sum_{k=1}^{m-1} \frac{1}{k!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[k]} f(x_0, y_0) +$$

$$+ \frac{1}{m!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[m]} f(x_0 + \theta\Delta x), y_0 + \theta\Delta y), \quad 0 < \theta < 1. \qquad \square$$

Corolario. *En vista de las suposiciones del teorema* 1 *es lícita la fórmula*

$$\Delta z = \sum_{k=1}^{m} \frac{1}{k!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^{[k]} f(x_0, y_0) + r_m(\Delta x, \Delta y), \tag{39.9}$$

con la particularidad de que el término residual $r_m(\Delta x, \Delta y)$ *puede ser escrito en cualquiera de las siguientes formas*:

$$r_m(\Delta x, \Delta y) = \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y)\, \Delta x^k\, \Delta y^{m-k}, \tag{39.10}$$

donde

$$\lim_{\rho\to 0} \varepsilon_k(\Delta x, \Delta y) = 0, \quad k = 0, 1, \ldots, m, \quad \rho = \sqrt{\Delta x^2 + \Delta y^2},$$

o bien

$$r_m(\Delta x, \Delta y) = \varepsilon(\Delta x, \Delta y)\rho^m, \tag{39.11}$$

donde $\lim\limits_{\rho\to 0} \varepsilon(\Delta x, \Delta y) = 0$, *es decir*,

$$r_m(\Delta x, \Delta y) = o(\rho^m). \tag{39.12}$$

La representación del término residual de la fórmula de Taylor tal como se indica en (39.12) se denomina su inscripción en la *forma de Peano*.

DEMOSTRACIÓN. Pongamos

$$\varepsilon_k^{\cdot}(\Delta x, \Delta y) = \frac{\partial^m f(x_0 + \theta\Delta x, y_0 + \theta\Delta y)}{\partial x^k \partial y^{m-k}} - \frac{\partial^m f(x_0, y_0)}{\partial x^k \partial y^{m-k}}. \tag{39.13}$$

Debido a la continuidad de todas las derivadas parciales de orden m

$$\lim_{\rho\to 0} \varepsilon_k^{\cdot}(\Delta x, \Delta y) = 0$$

Al hacer uso de la expresión (39.13), transformemos el resto $r_{m-1}(\Delta x, \Delta y)$ (véase (39.2)), de la manera siguiente:

$$r_{m-1}(\Delta x, \Delta y) = \frac{1}{m!} \sum_{k=0}^{m} C_m^k \frac{\partial^m f(x_0 + \theta\Delta x, y_0 + \theta\Delta y)}{\partial x^k \partial y^{m-k}} \Delta x^k \Delta y^{m-k} =$$

$$= \frac{1}{m!} \sum_{k=0}^{m} C_m^k \frac{\partial^m f(x_0, y_0)}{\partial x^k \partial y^{m-k}} \Delta x^k \Delta y^{m-k} +$$

$$+ \frac{1}{m!} \sum_{k=0}^{m} C_m^k \varepsilon_k' (\Delta x, \Delta y) \Delta x^k \Delta y^{m-k} =$$

$$= \frac{1}{m!} \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{[m]} f(x_0, y_0) + \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \Delta x^k \Delta y^{m-k}, \tag{39.14}$$

donde $\varepsilon_k(\Delta x, \Delta y) = \frac{C_m^k}{m!} \varepsilon_k' (\Delta x, \Delta y)$, y por esta razón

$$\lim_{\rho \to 0} \varepsilon_k(\Delta x, \Delta y) = 0. \tag{39.15}$$

Sustituyendo (39.14) en (39.1), obtenemos la fórmula de Taylor (39.9) con el término residual en la forma (39.10).

Mostremos que el término residual (39.10) puede ser escrito en la forma (39.11). Con este fin pongamos

$$\varepsilon(\Delta x, \Delta y) = \sum_{k=0}^{m} \varepsilon_k (\Delta x, \Delta y) \left(\frac{\Delta x}{\rho}\right)^k \left(\frac{\Delta y}{\rho}\right)^{m-k} \tag{39.16}$$

En este caso

$$r_m(\Delta x, \Delta y) = \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \Delta x^k \Delta y^{m-k} =$$

$$= \rho^m \sum_{k=0}^{m} \varepsilon_k(\Delta x, \Delta y) \left(\frac{\Delta x}{\rho}\right)^k \left(\frac{\Delta y}{\rho}\right)^{m-k} = \varepsilon(\Delta x, \Delta y)\rho^m,$$

y, como $\left|\frac{\Delta x}{\rho}\right| \leqslant 1$ y $\left|\frac{\Delta y}{\rho}\right| \leqslant 1$, de (39.15) se desprende que

$$\lim_{\rho \to 0} \varepsilon (\Delta x, \Delta y) = 0. \qquad \square$$

Aprovechando la noción de diferenciales de órdenes superiores, podemos atribuir a la fórmula de Taylor una forma más compacta que es idéntica en apariencia a la fórmula de Taylor para las funciones de una sola variable, escrita asimismo con el empleo de diferenciales. Efectivamente, puesto que (véase el p. 21.2)

$$d^k f(x, y) = \left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y} \right)^{[k]} f(x, y), \quad k = 0, 1, 2, \ldots, m,$$

entonces, suponiendo, para abreviar, $M_0 = (x_0, y_0)$ y $M = (x_0 + \Delta x, y_0 + \Delta y)$, podemos escribir la fórmula (39.9) en la forma

$$\Delta z = \sum_{k=1}^{m} \frac{1}{k!} d^k f(M_0) + r_m(M). \tag{39.17}$$

La forma indicada de la fórmula de Taylor es más simple a consecuencia de lo cual resulta cómoda para recordarla.

He aquí algunas observaciones referentes a las demostraciones del teorema 1 y de su corolario. Ante todo, en las condiciones del teorema se ha exigido que la función f tenga derivadas continuas de orden hasta m inclusive en cierto δ-entorno del punto (x_0, y_0). Se podría exigir que en el entorno citado sean continuas *sólo* las derivadas de orden m, puesto que de la continuidad de éstas se infiere que son continuas también en el entorno dado todas las derivadas inferiores de la función en consideración, es decir, las derivadas de órdenes $k = 0, 1, \ldots, m - 1$ (véase el p. 20.2).

Subrayemos que la continuidad de las derivadas parciales en el δ-entorno del punto (x_0, y_0) se ha usado, en primer lugar, para que las derivadas parciales con las que nos encontramos no dependan del orden en que se realiza la derivación (esto se ha usado tanto en la demostración de la fórmula de Taylor (39.1), como en la propia forma de notación de dicha fórmula) y, en segundo lugar, para que la función (39.3) pueda ser derivada m veces según la regla de derivación de la función compuesta. Fijemos la atención en que para $m = 1$ las derivadas mixtas están ausentes; entre tanto, para que haya posibilidad de derivar la función (39.3) una sola vez, según la regla de la función compuesta y, por consiguiente, para que sea válido el teorema 1, resulta suficiente una suposición más débil sobre la función en consideración f. A saber, en lugar de la suposición sobre la derivabilidad continua de la función f en el δ-entorno arriba mencionado del punto (x_0, y_0), es suficiente su derivabilidad en este entorno (véanse las definiciones 2 y 4 en el p. 20.2).

La continuidad de las derivadas parciales de orden m (en el *punto* (x_0, y_0)) se ha empleado también en la demostración del corolario del teorema 1: es necesaria para que las funciones $\varepsilon_k' (\Delta x, \Delta y)$, definidas mediante las fórmulas (39.13), tiendan a cero cuando $\rho \to 0$.

Recalquemos, además, que con las suposiciones admitidas en la fórmula (39.9) se ha demostrado que $r_m(\Delta x, \Delta y) = o(\rho^m)$ para $\rho \to 0$ no en el sentido del límite según cualquier dirección fijada, como podría parecer, a primera vista, proveniente de la demostración aducida, sino en el sentido más fuerte, esto es, en el sentido del límite en el punto (x_0, y_0) (¿por qué?).

La fórmula (39.1) puede ser un tanto generalizada, si no se tienen aspiraciones de que sea válida para todos los puntos $(x_0 + \Delta x_0, y_0 + \Delta y)$ del δ-entorno del punto (x_0, y_0), sino que se considera la fórmula citada sólo para Δx y Δy fijados. A saber, si la función f está definida y tiene derivadas parciales continuas de orden m en un conjunto abierto que contiene un segmento con los extremos (x_0, y_0) y $(x_0 + \Delta x, y_0 + \Delta y)$, entonces la fórmula (39.1) queda también justa, igual que su demostración. De esto se deduce que si la función f está definida en la región convexa G (véase el p. 18.2) y tiene en G derivadas parciales continuas de orden m, para cualesquiera dos puntos $(x_0, y_0) \in G$ y $(x_0 + \Delta x, y_0 + \Delta y) \in G$ es válida la fórmula de Taylor (39.1).

Ejercicio 1. Sea la función $f(x, y)$ continua, lo mismo que sus derivadas parciales hasta el orden m inclusive, en cierto entorno del punto (x_0, y_0). Demuéstrese que su polinomio de Taylor de orden m, es decir,

$$P(x,y) = \sum_{k=0}^{m} \frac{1}{k!}\left[(x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}\right]^{(k)} f(x_0, y_0)$$

es un polinomio de mejor aproximación de la función $f(x, y)$ en un "entorño infinitamente pequeño del punto (x_0, y_0)". Esto significa lo siguiente: cualquiera que sea el polinomio $Q(x, y)$ de grado no superior a m (es decir, en cada término suyo la suma de exponentes de las potencias de las variables x e y no debe ser mayor que el número m) tal que

$$f(x,y) = Q(x,y) + o(\rho^n), \quad n \geqslant m, \quad \text{cuando} \quad \rho \to 0,$$

donde

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

coincide con el polinomio citado de Taylor $P(x, y)$ de la función $f(x, y)$.

Todo lo dicho se extiende también al caso de una función de cualquier número de variables.

Teorema 1´. *Si una función de n variables $y = f(x_1, \ldots, x_n)$ está definida y es continua junto con todas sus derivadas parciales hasta el orden m, $m \geqslant 1$, inclusive en cierto δ-entorno del punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(n)})$, entonces es válida la fórmula*

$$\Delta y = f(x_1^{(0)} + \Delta x_1, \ldots, x_n^{(0)} + \Delta x_n) - f(x_1^{(0)}, \ldots, x_n^{(0)}) =$$

$$= \sum_{k=1}^{m-1} \frac{1}{k!}\left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n}\right)^{(k)} f(x^{(0)}) + r_{m-1}(\Delta x), \qquad (39.18)$$

donde

$$r_{m-1}(\Delta x) =$$

$$= \frac{1}{m!}\left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n}\right)^{(m)} f(x_1^{(0)} + \theta\Delta x_1, \ldots, x_n^{(0)} + \theta\Delta x_n),$$

$$0 < \theta < 1, \quad \Delta x = (\Delta x_1, \ldots, \Delta x_n), \qquad (39.19)$$

y también la fórmula

$$\Delta y = \sum_{k=1}^{m} \frac{1}{k!}\left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n}\right)^{(k)} f(x^{(0)}) + r_m(\Delta x), \qquad (39.20)$$

donde $r_m(\Delta x)$ puede escribirse en cada una de las siguientes formas: o bien

$$r_m(\Delta x) = \sum_{m_1+\ldots+m_n=m} \varepsilon_{m_1\ldots m_n}(\Delta x)\,\Delta x_1^{m_1}\ldots\Delta x_n^{m_n}, \qquad (39.21)$$

donde $\lim\limits_{\rho\to 0} \varepsilon_{m_1,\ldots,m_n}(\Delta x) = 0$, $\rho = \sqrt{\sum\limits_{i=1}^{n} \Delta x_i^2}$, *o bien*

$$r_m(\Delta x) = \varepsilon(\Delta x)\rho^m, \quad \lim_{\rho\to 0} \varepsilon(\Delta x) = 0, \qquad (39.22)$$

es decir,

$$r_m(\Delta x) = o(\rho^m), \quad \rho \to 0.$$

Por fin, en términos de las diferenciales la fórmula (39.20) *puede escribirse en la forma*

$$\Delta y = \sum_{k=1}^{m} \frac{1}{k!} d^k f(x^{(0)}) + r_m(\Delta x). \tag{39.23}$$

Ahora abrimos los paréntesis en las fórmulas (39.18) y (39.19), haciendo uso de la fórmula algebraica

$$\left(\sum_{i=1}^{n} a_i\right)^k = \sum_{k_1+\ldots+k_n=k!} \frac{k!}{k_1!k_2!\ldots k_n!} a_1^{k_1} a_2^{k_2} \ldots a_n^{k_n}.$$

Para anotar el resultado en la forma más breve introduzcamos unas designaciones nuevas. Pongamos $k = (k_1, \ldots, k_n)$, $|k| = k_1 + \ldots + k_n$, $k! = k_1! \ldots k_n!$,

$$f^{(k)} = \frac{\partial^{|k|} f}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}}, \quad (x - x^{(0)})^k = (x_1 - x_1^{(0)})^{k_1} \ldots (x_n - x_n^{(0)})^{k_n};$$

$k = (k_1, \ldots, k_n)$ se llama *multiíndice.*

Introducidas las designaciones indicadas, la fórmula de Taylor (39.18) con el término residual en la forma (39.19) se escribirá así:

$$f(x) = \sum_{|k|<m} \frac{f^{(k)}(x^{(0)})}{k!}(x - x^{(0)})^k + \sum_{|k|=m} \frac{f^{(k)}(x^{(0)} + \theta(x - x^{(0)}))}{k!}(x - x^{(0)})^k.$$

Aquí, como siempre, $x = (x_1, \ldots, x_n)$, $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ y

$$x^{(0)} + \theta(x - x^{(0)}) = (x_1^{(0)} + \theta(x_1 - x_1^{(0)}), \ldots, x_n^{(0)} + \theta(x_n - x_n^{(0)})).$$

La fórmula de Taylor que acabamos de exponer para las funciones de cualquier número de variables tiene la misma forma que para las funciones de una sola variable.

A veces, particularmente en el caso de las funciones de varias variables, para las derivadas se utiliza la designación

$$D^k \overset{\text{def}}{=} \frac{\partial^{k_1+\ldots+k_n}}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}},$$

donde $k = (k_1, \ldots, k_n)$ es un multiíndice. Si se usa dicha notación, la fórmula de Taylor adoptará la forma

$$\begin{aligned} f(x) = &\sum_{|k|<m} \frac{1}{k!} D^k f(x^{(0)})(x - x^{(0)})^k + \\ &+ \sum_{|k|=m} \frac{1}{k!} D^k f(x^{(0)} + \theta(x - x^{(0)}))(x - x^{(0)})^k, \quad 0 < \theta < 1. \end{aligned}$$

39.2. FÓRMULA DE INCREMENTOS FINITOS PARA LAS FUNCIONES DE VARIAS VARIABLES

El caso particular de la fórmula de Taylor (39.18), cuando $m = 1$, se donomina corrientemente fórmula de incrementos finitos de Lagrange para las funciones de varias variables. Debido a las observaciones que en el punto anterior siguen el teorema 1 sobre las suposiciones bajo las cuales resultan válidas las fórmulas (39.1) y (39.18), del teorema 1′ obtenemos la siguiente afirmación.

Teorema 2. *Si la función $f(x_1, \ldots, x_n)$ es derivable en todo punto de cierta región convexa $G \subset R^n$, para cualquier par de puntos $(x_1, \ldots, x_n)$ y $(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n)$ de G existe tal θ, $0 < \theta < 1$, que*

$$f(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n) - f(x_1, \ldots, x_n) = \sum_{i=1}^{n} \frac{\partial f(x_1 + \theta\Delta x_1, \ldots, x_n + \theta\Delta x_n)}{\partial x_i} \Delta x_i$$

o, en la forma más breve,

$$f(x + \Delta x) - f(x) = \sum_{i=1}^{n} \frac{\partial f(x + \theta\Delta x)}{\partial x_i} \Delta x_i, \tag{39.24}$$

donde $x = (x_1, \ldots, x_n)$, $x + \Delta x = (x_1 + \Delta x_1, \ldots, x_n + \Delta x_n)$ y

$$x + \theta\Delta x = (x_1 + \theta\Delta x_1, \ldots, x_n + \theta\Delta x_n).$$

Como ya se ha indicado, la fórmula (39.24) expresa precisamente la *fórmula de incrementos finitos de Lagrange.*

Dicha fórmula, al igual que la fórmula de Taylor en general, tiene muchas y varias aplicaciones en diferentes problemas del análisis matemático.

Prestemos nuestra atención a que el teorema 2 no es el caso particular del teorema 1, pues no es la derivabilidad continua de la función en consideración en todo punto del conjunto G lo que se exige de él, sino sólo su derivabilidad. No obstante, la demostración del teorema 2 se contiene de hecho en la demostración del teorema 1. En efecto, como ya se ha indicado en las observaciones referentes a la demostración del teorema 1 y al corolario de éste (véase el p. 39.1), para $m = 1$ la demostración del teorema 1, aducida arriba, conserva rigor también bajo las suposiciones del teorema 2, es decir, cuando sólo se supone la derivabilidad (y no derivabilidad continua) de la función f.

Demostremos la siguiente afirmación como ejemplo de la aplicación de la fórmula (39.24).

Teorema 3. *Si una función es derivable en todo punto de la región convexa G y tiene en G derivadas parciales acotadas, será uniformemente continua en la región mencionada.*

DEMOSTRACIÓN. Si

$$\left|\frac{\partial f(x)}{\partial x_i}\right| \leqslant c, \quad i = 1, 2, \ldots, n, \quad x \in G$$

(c es una constante), entonces para cualesquiera dos puntos $x' = (x_1', \ldots, x_n')$ y $x'' = (x_1'', \ldots, x_n'')$ de (39.24) se infiere que

$$|f(x'') - f(x')| \leqslant \sum_{i=1}^{n} \left| \frac{\partial f(\xi)}{\partial x_i} \right| |x_i'' - x_i'| \leqslant cn\rho(x', x'')$$

(aquí ξ es algún punto del segmento que tiene por extremos los puntos x' y x''). Por eso, si está prefijado $\varepsilon > 0$, basta tomar $\delta = \frac{\varepsilon}{cn}$ para que se cumpla la desigualdad

$$|f(x'') - f(x')| < \varepsilon, \tag{39.25}$$

cualesquiera que sean los puntos $x' \in G$ y $x'' \in G$ tales que $\rho(x', x'') < \delta$, lo que atestigua la continuidad uniforme de la función f en la región G. □

39.3. SOBRE LA ESTIMACIÓN DEL TÉRMINO RESIDUAL DE LA FÓRMULA DE TAYLOR EN TODO EL DOMINIO DE DEFINICIÓN DE LA FUNCIÓN

El término residual en la fórmula de Taylor depende, evidentemente, no sólo de los incrementos de los argumentos, sino también del mismo punto en cuyo entorno se examina el desarrollo de la función y que en el p. 39.1 se consideraba fijado. Ahora serán de interés para nosotros el comportamiento y la estimación del término residual en dependencia del cambio del punto mencionado. Con el fin de subrayar dicha dependencia, el termino residual de orden m se designará en este párrafo por $r_m(x, \Delta x)$, donde $x = (x_1, \ldots, x_n)$ es un punto en cuyo entorno se desarrolla la función dada según la fórmula de Taylor. Como hasta ahora, $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$.

En las fórmulas (39.21) y (39.22), en lugar de $\varepsilon_{m_1 \ldots m_n}(\Delta x)$ y $\varepsilon(\Delta x)$ se escribirán $\varepsilon_{m_1 \ldots m_n}(x, \Delta x)$ y $\varepsilon(x, \Delta x)$, respectivamente. En lo sucesivo nos hará falta la estimación del término residual de la fórmula de Taylor en la forma de Peano para todo el dominio en el que existe el desarrollo según la fórmula indicada.

Introduzcamos primero el concepto de continuidad de las derivadas parciales en la clausura de un conjunto abierto. Esto requiere una definición especial, puesto que en un punto límite del conjunto abierto G el concepto de derivada parcial no está definido, en el caso general, aún cuando la función quede definida en la clausura $\overline{G}$ del conjunto G (véase, por ejemplo, el punto M de la frontera de la región G en la fig. 156).

Definición 1. *Una función f, definida en el conjunto abierto $G \subset R^n$, se llama continuamente prolongable a la clausura $\overline{G}$ de éste, si existe tal función F, continua en $\overline{G}$, que $F = f$ en G.*

La función F se denomina prolongación continua de la función f (a $\overline{G}$) y se designará, para simplificar, también con el símbolo f.

En virtud de la unicidad del límite de una función es evidente que si una función definida en G tiene prolongación continua a $\overline{G}$, ésta es única.

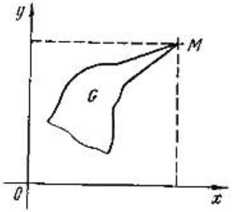

Fig. 156

Definición 2. *Una función f se denomina continuamente derivable (m veces continuamente derivable) en $\overline{G}$, si f está definida en G y todas sus derivadas parciales de primer orden (derivadas parciales hasta el orden m inclusive) son continuamente prolongables de G a $\overline{G}$.*

Ejerciocios. 2. Demuéstrese que si la función f está definida en el conjunto abierto $G \subset R^n$ y tiene en éste una derivada $\frac{\partial f}{\partial x_1}$, continuamente prolongable a la clausura $\overline{G}$ del conjunto, y si, además, en cierto punto de la frontera del conjunto G existe la derivada parcial (unilateral) $\frac{\partial f}{\partial x_1}$, entonces la función coincide con la prolongación contínua de la derivada parcial a dicho punto.

3. Demuéstrese que para que una función continua, definida en el conjunto abierto acotado $G \subset R^n$, sea continuamente prolongable a la clausura de éste, es necesario y suficiente que sea uniformemente continua en G. Pruébese que en el caso de un conjunto abierto no acotado la condición de continuidad uniforme de una función prolongable, siendo suficiente para la prolongación continua, no es necesaria.

4. Constrúyase un ejemplo de la función, continua y acotada en una región, que no pueda ser continuamente prolongada a la clausura de dicha región.

Volveremos ahora a la fórmula de Taylor. Supongamos que la función f es m veces continuamente derivable en la clausura $\overline{G}$ de un conjunto acotado abierto G. En este caso, de acuerdo con los resultados obtenidos en el p. 39.1, en todo punto $x \in G$ tiene lugar el desarrollo (39.20) de la función f según la fórmula de Taylor, con la particularidad de que $\varepsilon_{m_1 \ldots m_n}(x, \Delta x)$ en la fórmula (39.21) y $\varepsilon(x, \Delta x)$ en la fórmula (39.22) tienden a cero, para $\rho \to 0$, uniformemente en el conjunto G (véase la definición en el p. 20.2), es decir, para cualquier $\varepsilon > 0$ existe tal $\delta = \delta(\varepsilon) > 0$ que si

$$\rho = \sqrt{\sum_{i=1}^{n} \Delta x_i^2}, < \delta \tag{39.26}$$

se tiene

$$|\varepsilon_{m_1 \ldots m_n}(x, \Delta x)| < \varepsilon \quad \text{y} \quad |\varepsilon(x, \Delta x)| < \varepsilon$$

para todos los puntos $x \in G$.

En el caso dado esto se deduce inmediatamente del método por cuyo intermedio se obtienen las funciones $\varepsilon_{m_1\ldots m_n}$ y $\varepsilon(\Delta x)$. Efectivamente, por ser la clausura $\bar{G}$ del conjunto abierto G acotada y cerrada, las prolongaciones continuas de las derivadas parciales de orden m de la función dada a $\bar{G}$ son en ella uniformemente continuas, razón por la cual (véase la fórmula (39.13) para el caso de $n = 2$; en el caso general es válida la fórmula análoga), cumplida la condición (39.26), tenemos

$$|\varepsilon_{m_1\ldots m_n}(x, \Delta x)| \leqslant \omega\left(\delta, \frac{\partial f}{\partial x_1^{m_1}\ldots\partial x_n^{m_n}}, \bar{G}\right). \tag{39.27}$$

Aquí el segundo miembro (módulo de continuidad de la derivada correspondiente) no depende del punto del conjunto G y tiende a cero cuando $\delta \to 0$. Por ello, de (39.27) se desprende que $\varepsilon_{m_1\ldots m_n}$ tiende a cero en G de modo uniforme.

Ahora podemos estimar el infinitésimo $\varepsilon(\Delta x, \Delta y)$ en la fórmula (39.22). Para n natural arbitrario podemos representarla, por analogía con el caso en que $n = 2$ (véase 39.16), en la forma

$$\varepsilon(x, \Delta x) = \sum_{m_1+\ldots+m_n=m} \varepsilon_{m_1\ldots m_n}(x, \Delta x)\left(\frac{\Delta x_1}{\rho}\right)^{m_1}\ldots\left(\frac{\Delta x_n}{\rho}\right)^{m_n}.$$

De aquí tenemos:

$$|\varepsilon(x, \Delta x)| \leqslant \sum_{m_1+\ldots+m_n=m} |\varepsilon_{m_1\ldots m_n}(x, \Delta x)|. \tag{39.28}$$

En el segundo miembro de la desigualdad (39.28) figura cierto número fijo de sumandos; designémoslo con N. En virtud de que la función $\varepsilon_{m_1\ldots m_n}(x, \Delta x)$ tiende a cero en G uniformemente (lo que ya se ha demostrado), para cualquier $\varepsilon > 0$ prefijado existe tal $\delta = \delta(\varepsilon) > 0$, que, cumplida la condición $\rho(x, x + \Delta x) < \delta$, tenemos

$$|\varepsilon_{m_1\ldots m_n}(x, \Delta x)| < \frac{\varepsilon}{N}, \quad m_1 + \ldots + m_n = m.$$

De aquí y también de la desigualdad (39.28) se deduce que

$$|\varepsilon(x, \Delta x)| < \varepsilon. \ \square$$

Demos a conocer una estimación más en total del término residual de la fórmula de Taylor que se obtiene de la inscripción de éste en la forma de Lagrange (39.19).

Si la función f está definida en un conjunto abierto G y tiene en dicho conjunto derivadas parciales acotadas de orden m, es decir, si existe una constante $M > 0$ tal que

$$\left|\frac{\partial^m f(x)}{\partial x_1^{m_1}\ldots\partial x_n^{m_n}}\right| \leqslant M, \quad m_1 + \ldots + m_n = m, \quad x \in G, \tag{39.29}$$

entonces, cumplida la condición $\rho(x, x + \Delta x) < \delta$, para todo $x \in G$ se verifica la desigualdad

$$|r_{m-1}(x, \Delta x)| \leqslant \frac{Mn^m\delta^m}{m!}.$$

Esto proviene directamente de la fórmula (39.19), si los valores absolutos de cada sumando de su segundo miembro se estiman por medio de la desigualdad (39.29) y otra desigualdad evidente $|\Delta x_i| \leqslant \delta$.

39.4. CONVERGENCIA UNIFORME SEGÚN EL PARÁMETRO DE UNA FAMILIA DE FUNCIONES

En el punto anterior hemos tropezado con la noción de convergencia uniforme en el conjunto dado de una familia de funciones dependientes de cierto parámetro cuando este último tiende hacia los valores determinados. En calidad de tales funciones en nuestro caso intervenían $\varepsilon_{m_1 \ldots m_n}(x, \Delta x)$ y $\varepsilon(x, \Delta x)$, donde el papel del parámetro lo desempeñaba Δx. Con este caso, en su forma más sencilla, ya chocamos antes, en el p. 20.2.

Enunciemos la definición de la convergencia uniforme de una familia de funciones en el caso general.

Definición 3. *Supongamos que X es un conjunto arbitrario, $Y \subset R^m$, $y^{(0)}$ es un punto del espacio R^m o uno de los infinitos*) ∞, $+\infty$, $-\infty$ (los dos últimos infinitos merecen ser considerados sólo en el caso cuando $m = 1$), con la particularidad de que la intersección de cualquier entorno reducido y_0 con el conjunto Y es no vacía. Supongamos luego que la función $\varphi(x)$ está definida para todo $x \in X$ y la función $f(x, y)$, para cualesquiera $x \in X$ e $y \in Y$.*

Suele decirse que $f(x, y)$ tiende uniformemente en el conjunto X hacia la función $\varphi(x)$ para $y \to y^{(0)}$ y se escribe en este caso

$$f(x, y) \underset{X}{\rightrightarrows} \varphi(x), \quad y \to y^{(0)}$$

si para cualquier $\varepsilon > 0$ existe tal entorno reducido $\mathring{U}(y^{(0)})$ del punto $y^{(0)}$ que se verifica la desigualdad

$$|f(x, y) - \varphi(x)| < \varepsilon, \tag{39.30}$$

cualesquiera que sean $x \in X$ e $y \in Y \cap \mathring{U}(y^{(0)})$.

En este caso la variable y se llama con frecuencia parámetro y la función $f(x, y)$, $y \in Y$, "familia de las funciones de x" (en el sentido de que esta función prefija las funciones de la variable x para diferentes valores fijos de $y \in Y$).

Por analogía con el caso de la convergencia uniforme de una sucesión de funciones (véase el p. 36.1), la condición de convergencia uniforme de las funciones según un parámetro puede enunciarse, empleando la noción habitual de límite de una función, de la manera siguiente.

La función $f(x, y)$ tiende uniformemente en el conjunto X hacia la función $\varphi(x)$ para $y \to y^{(0)}$ cuando, y sólo cuando,

*) En lo sucesivo los infinitos ∞, $+\infty$, $-\infty$ se llamarán también, para simplificar, puntos ("infinitamente alejados").

$$\lim_{y \to y^0} \sup_{x \in X} |f(x, y) - \varphi(x)| = 0. \tag{39.31}$$

Así pues, la condicion $f(x, y) \underset{X}{\rightrightarrows} \varphi(x)$, $y \to y^{(0)}$, es equivalente a que la función $F(y) \overset{\text{def}}{=} \sup_{x \in X} |f(x, y) - \varphi(x)|$ tiende a cero cuando $y \to y^{(0)}$. La demostración de esta afirmación no es del todo dificil y es análoga al caso de la convergencia uniforme de una sucesión de funciones. Esta demostración queda al cargo del lector.

En el caso que se considera es justo también el análogo del criterio de Cauchy para la convergencia uniforme de las sucesiones.

Teorema 4 (criterio de Cauchy). *Para que la función $f(x, y)$, para $y \to y^{(0)}$, tienda uniformemente en el conjunto X hacia cierta función, es necesario y suficiente que para cualquier $\varepsilon > 0$ exista tal entorno reducido $\mathring{U}(y^{(0)})$ del punto $y^{(0)}$ que se verifique la desigualdad*

$$|f(x, y'') - f(x, y')| < \varepsilon, \tag{39.32}$$

cualesquiera que sean

$$y' \in \mathring{U}(y^{(0)}) \cap Y, \quad e \quad y'' \in \mathring{U}(y^{(0)}) \cap Y \quad y \quad x \in X.$$

En efecto, la necesidad en la condición (39.32) proviene fácilmente, como siempre en las situaciones semejantes, de la condición (39.30). Con el fin de demostrar la suficiencia, se debe probar que de la condición (39.32) se infiere que para cualquier $x \in X$ fijo existe $\lim\limits_{y \to y^{(0)}} f(x, y)$ y que la función $f(x, y)$ tiende hacia este límite, para $y \to y^{(0)}$, uniformemente.

Se recomienda que el lector mismo compruebe todas estas afirmaciones.

Ejercicio 5. Demuéstrese: para que la función $f(x, y)$, $x \in X$, $y \in X$, tienda uniformemente en el conjunto X, cuando $y \to y^{(0)}$, hacia la función $\varphi(x)$, $x \in X$, es necesario y suficiente que para toda sucesión $y^{(n)} \in Y$, $y^n \neq y^{(0)}$, $n = 1, 2, \ldots$, que tiende a $y^{(0)}$, la sucesión $f(x, y^{(n)})$, $n = 1, 2, \ldots$, converja uniformemente sóbre el conjunto X hacia la función $\varphi(x)$.

Ejemplos. 1. Examinemos una familia de funciones $f(x, y) = e^{-xy}$, donde $0 \leqslant \leqslant x \leqslant 1$, $0 \leqslant y < +\infty$. Es evidente que

$$\lim_{y \to +\infty} f(x, y) = \begin{cases} 0, & \text{si} \quad x > 0, \\ 1, & \text{si} \quad x = 0 \end{cases}$$

(de este modo, la variable y es un parámetro, si se usa la terminología indicada anteriormente). Designaremos la función límite mediante $\varphi(x)$,

$$\varphi(x) = \begin{cases} 0, & \text{si} \quad x > 0, \\ 1, & \text{si} \quad x = 0. \end{cases} \tag{39.33}$$

Demostremos que la función $f(x, y)$ tiende hacia $\varphi(x)$, cuando $y \to +\infty$, de manera no uniforme. Para esto es suficiente probar que existe tal $\varepsilon_0 > 0$ que, cualquiera que sea el entorno $\mathring{U}(+\infty)$, se encontrarán $x \in [0, 1]$ e $y \in \mathring{U}(+\infty)$ tales que se verifique la desigualdad $|e^{-xy} - \varphi(x)| \geqslant \varepsilon_0$. Elijamos ε_0 tal que sea $0 < \varepsilon_0 < 1$, y un entorno arbitrario $\mathring{U}(+\infty)$. Entonces, cualquiera que sea $y \in \mathring{U}(+\infty)$, para él se verifica

$\lim\limits_{x\to 0} e^{-xy} = 1$, y, por ende, existe tal $x \in (0, 1]$ que

$$|e^{-xy} - \varphi(x)| = |e^{-xy} - 0| > \varepsilon_0.$$

De este modo, en el caso dado las condiciones del criterio de Cauchy no se cumplen (véase el teorema 4).

Sin embargo, para todo a, $0 < a < 1$, la familia de funciones $f(x, y) = e^{-xy}$ tiende uniformemente, cuando $y \to +\infty$, hacia cero en el segmento $[a, 1]$. Comprobemos el cumplimiento de las condiciones del criterio de Cauchy en este caso (véase el teorema 4). Para cualquier $\varepsilon > 0$ existe un número $\eta_\varepsilon > 0$ tal que $e^{-a\eta_\varepsilon} < \varepsilon$ (basta tomar cualquier $\eta > \dfrac{|\ln \varepsilon|}{a}$); por esto para cualesquiera $y > \eta_\varepsilon$ y todos los $x \in [a, 1]$ tendremos

$$|e^{-xy} - 0| = e^{-xy} < e^{-a\eta_\varepsilon} < \varepsilon.$$

Por supuesto, la investigación de la convergencia uniforme de la familia de funciones en consideración podría ser ejecutada también aplicando el criterio (39.31). Efectivamente, al emplear la fórmula (39.33), obtendremos

$$\sup_{0 \leqslant x \leqslant 1} |e^{-xy} - \varphi(x)| \geqslant \sup_{0 < x \leqslant 1} e^{-xy} = 1,$$

por lo cual la condición (39.31) no se cumple a ciencia cierta. En cambio, si $0 < a < 1$, entonces

$$\lim_{y\to+\infty} \sup_{a \leqslant x \leqslant 1} |e^{-xy} - \varphi(x)| = \lim_{y\to+\infty} \sup_{a \leqslant x \leqslant 1} e^{-xy} = \lim_{y\to+\infty} e^{-ay} = 0.$$

De este modo,

$$e^{-xy} \underset{[0,1]}{\not\rightrightarrows} \varphi(x), \quad e^{-xy} \underset{[0,a]}{\rightrightarrows} \varphi(x), \quad y \to +\infty, \quad 0 < a < 1.$$

2. En el caso cuando Y es un conjunto de los números naturales, $Y = \{1, 2, 3, \ldots\}$, e $y^{(0)} = +\infty$, la definición aducida de la convergencia uniforme según un parámetro se convierte en la definición de convergencia uniforme de una sucesión de funciones $f_n(x) = f(x, n)$, $n = 1, 2, \ldots$, en el conjunto X.

3. Sea la función $f(x, y)$ continua en el rectángulo $Q = \{(x, y): -\infty < a \leqslant x \leqslant b < +\infty,\ -\infty < c \leqslant y \leqslant d < +\infty\}$ y sea $y_0 \in [c, d]$.

Designemos mediante $\omega(\delta, f)$ el módulo de continuidad de la función f en el rectángulo Q; entonces

$$|f(x, y) - f(x, y_0)| \leqslant \omega(|y - y_0|; f), \quad (x, y) \in Q. \tag{39.34}$$

El segundo miembro de esta desigualdad no depende de x, y, siendo uniforme la continuidad de la función f en el rectángulo Q, $\lim\limits_{\delta\to 0} \omega(\delta; f) = 0$. Por esta razón, de la desigualdad (39.34) se desprende que en el segmento $[a, b]$ la función $f(x, y)$ tiende uniformemente hacia la función $f(x, y_0)$ cuando $y \to y_0$.

Ejercicio 6. Demuéstrese que si la familia de funciones $f(x, y)$, $x \in X \subset R^n$, $y \subset Y \subset R^m$, es tal que para todo $y \in Y$ fijo las funciones $f(x, y)$ son continuas respecto de x sobre el conjunto X y tienden uniformemente sobre dicho conjunto hacia $\varphi(x)$ para $y \to y^{(0)}$, entonces $\varphi(x)$ es también continua sobre el conjunto X.

39.5. OBSERVACIONES ACERCA DE LAS SERIES DE TAYLOR PARA LAS FUNCIONES DE VARIAS VARIABLES

Si la función $f(x)$ está definida y es derivable un número infinito de veces en cierto δ-entorno del punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in R^n$, entonces, para dicha función la fórmula de Taylor (39.20) será, evidentemente, válida para cualquier n natural, $n = 1, 2, \ldots$, y, además, $\sum\limits_{i=1}^{n} \Delta x_i^2 < \delta^2$. Si, en este caso, la serie

$$\sum_{k=1}^{\infty} \frac{1}{k!}\left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n}\right)^{[k]} f(x^{(0)})$$

convergerá hacia $\Delta y = f(x) - f(x^{(0)})$ (véase el p. 38.2), se obtendrá la siguiente fórmula

$$\Delta y = f(x) - f(x^{(0)}) = \sum_{k=1} \frac{1}{k!}\left(\Delta x_1 \frac{\partial}{\partial x_1} + \ldots + \Delta x_n \frac{\partial}{\partial x_n}\right)^{[k]} f(x^{(0)}),$$

donde $x = (x_1, \ldots, x_n)$ y $x_i - x_i^{(0)} = \Delta x_i$, $i = 1, 2, \ldots, n$. De aquí, trasladando $f(x^{(0)})$ al segundo miembro, obtenemos el desarrollo de la función en serie de potencias llamada *serie de Taylor* de la función f:

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!}\left[\left(x_1 - x_1^{(0)}\right)\frac{\partial}{\partial x_1} + \ldots + \left(x_n - x_n^{(0)}\right)\frac{\partial}{\partial x_n}\right]^{[k]} f(x^{(0)}),$$

o, abriendo los corchetes,

$$f(x) = \sum_{|k|=0}^{\infty} \frac{1}{k!} D^k f(x^{(0)})(x - x^{(0)})^k,$$

donde $k = (k_1, \ldots, k_n)$ es un multiíndice.

Ejercicio 7. Desarróllese en serie de Taylor la función $f(x, y) = e^{x+y}$.

§ 40. EXTREMOS DE LAS FUNCIONES DE VARIAS VARIABLES

40.1. CONDICIONES NECESARIAS DE UN EXTREMO

Los problemas que se estudian en este párrafo y en algunos otros que siguen llevan un carácter analítico y sus demostraciones no se hacen más difíciles al aumentar el número de las variables. Es por eso que nuestra intención es considerar dichos problemas en el caso general n-dimensional, subrayando, si es necesario, sus pecualiaridades específicas para los casos de $n = 2$ y $n = 3$.

Definición 1. *Supongamos que la función $f(x)$ está definida en el conjunto $X \subset R^n$. El punto $x^{(0)} \in X$ se llama punto de máximo estricto (de mínimo estricto), si*

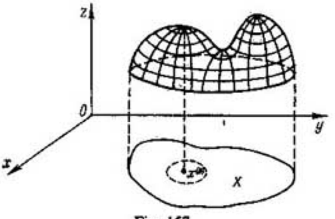

Fig. 157

existe tal entorno $U(x^{(0)})$ *del punto* $x^{(0)}$ *que para todo* $x \in U(x^{(0)}) \cap X, x \neq x^{(0)}$, *se verifique la desigualdad* $f(x) < f(x^{(0)})$*(la desigualdad* $f(x) > f(x^{(0)})$ respectivamente).

De este modo, el punto de máximo estricto (de mínimo estricto) se caracteriza por que $\Delta f = f(x) - f(x^{(0)}) < 0$ $(\Delta f > 0)$, cualquiera que sea $x \in U(x^{(0)}) \cap X, x \neq x^{(0)}$) (fig. 157).

En cambio, si para el punto $x^{(0)}$ *existe un entorno* $U(x^{(0)})$ *tal que para todo* $x \in U(x^{(0)}) \cap X$ *se verifica la condición* $f(x) \leqslant f(x^{(0)})$ $(f(x) \geqslant f(x^{(0)}))$, *entonces* $x^{(0)}$ *se denomina simplemente punto de máximo (punto de mínimo).*

Definición 2. *Los puntos de máximo y mínimo (estrictos) de una función llevan el nombre de puntos de extremo (estricto).*

Teorema 1. *Supongamos que la función* $f(x)$, $x = (x_1, x_2, \ldots, x_n)$ *está definida en cierto entorno del punto* $x^{(0)}$; *si dicho punto es un punto de extremo de la función* $f(x)$ *y si en él existe cualquiera de las derivadas* $\frac{\partial f}{\partial x_j}$ *(j puede asumir uno de los valores* $1, 2, \ldots, n$*), esta última es nula,* $\frac{\partial f(x^{(0)})}{\partial x_j} = 0$.

Corolario. *Si la función* $f(x)$ *es derivable en el punto de extremo* $x^{(0)}$, *su diferencial en este punto es igual a cero,* $df(x^{(0)}) = 0$.

DEMOSTRACIÓN (del teorema y del corolario). Sea, para concretar, $j = 1$. Si $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ es el punto de extremo para la función $f(x) = f(x_1, \ldots, x_n)$, entonces $x_1^{(0)}$ es el punto de extremo para la función $f(x_1, x_2^{(0)}, \ldots, x_n^{(0)})$ de una sola variable x_1 (fig. 158), por lo cual si en este punto existe derivada $\frac{\partial f}{\partial x_1}$, entonces, de acuerdo con el teorema de Fermat (véase el p. 11.1), ésta es igual a cero, es decir,

$$\frac{\partial f(x^{(0)})}{\partial x_1} = \left.\frac{df(x_1, x_2^{(0)}, \ldots, x_n^{(0)})}{dx_1}\right|_{x_1 = x_1^{(0)}} = 0.$$

Lo mismo ocurre en el caso de cualquier variable x_j $(j = 2, \ldots, n)$.

Si la función $f(x)$ es derivable en el punto de extremo $x^{(0)}$, en este punto existen todas las derivadas $\frac{\partial f}{\partial x_i}$, $i = 1, 2, \ldots, n$, y, de conformidad con lo demostrado, to-

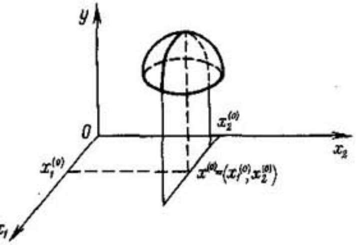

Fig. 158

das ellas son nulas, por lo cual también

$$df(x^{(0)}) = \sum_{i=1}^{n} \frac{df(x^{(0)})}{\partial x_i} dx_i = 0. \square$$

Ejemplos. 1. Hallemos los puntos de extremo de la función $z = x^2 + y^2$. En virtud de lo demostrado los puntos de extremo se disponen entre aquellos, para los cuales $dz = 0$. Puesto que $dz = 2xdx + 2ydy$, la condición $dz = 0$ se cumple en un único punto $(0, 0)$. En dicho punto $z = 0$, en todos los demás $z = x^2 + y^2 > 0$. Por esta razón $(0, 0)$ es el punto de mínimo estricto para la función $z = x^2 + y^2$ (fig. 159).

2. Investiguemos los puntos de extremo de la función $z = x^2 - y^2$. Procediendo igual que en el caso anterior, encontramos que esta vez también la condición $dz = 0$ se cumple en el punto $(0, 0)$ y en dicho punto $z = 0$. No obstante, aquí tenemos $z > > 0$ para $y = 0$ y todo $x \neq 0$, mientras que para $x = 0$ y cualquier $y \neq 0$ se tiene $z < 0$. Por esto, el punto $(0, 0)$ no es un punto de extremo y, consecuentemente, la función $z = x^2 - y^2$ no tiene en general puntos extremos (fig. 160).

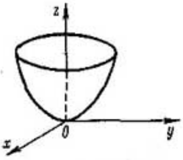

Fig. 159

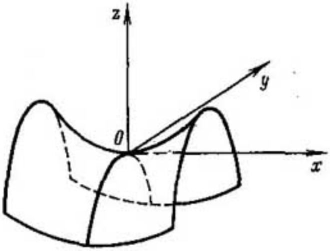

Fig. 160

40.2. CONDICIONES SUFICIENTES DE UN EXTREMO ESTRICTO

Recordemos algunas definiciones referentes al Curso del álgebra.

Definición 3. *Una forma cuadrática* $A(x) = A(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$, $a_{ji} = a_{ij}$, $i, j = 1, 2, \ldots, n$, *se denomina definida positiva (definida negativa), si* $A(x) > 0$ ($A(x) < 0$, *respectivamente) para cualquier punto* $x \in R^n$, $x \neq 0$.

Una forma cuadrática, que es definida positiva o definida negativa, lleva, además, el nombre de forma cuadrática de signo definido.

Definición 4. *Una forma cuadrática que asume tanto valores positivos como negativos se llama de signo indefinido.*

Lema 1. *Sea S una esfera unidad en* R^n:

$$S = \{x: x_1^2 + \ldots + x_n^2 = 1\},$$

y supongamos que $A(x)$ *es una forma cuadrática de signo definido; entonces*

$$\inf_{x \in S} |A(x)| = \mu > 0.$$

DEMOSTRACIÓN. La función $A(x)$ es un polinomio de segundo grado respecto de las variables $x_1, \ldots, x_n$, por lo cual $A(x)$ y, consecuentemente, también $|A(x)|$ son continuas en todo el espacio R^n. De aquí se desprende que la función $|A(x)|$ es continua en el compacto S. Conforme al teorema de Weierstrass, la función $|A(x)|$ alcanza en S su cota inferior, es decir, existe tal punto $x^{(0)} \in S$ que

$$\mu \stackrel{\text{def}}{=} \inf_{x \in S} |A(x)| = |A(x^{(0)})|.$$

Por definición de la forma cuadrática de signo definido, $|A(x)| > 0$ para todo punto $x \in S$, quiere decir, en particular, $\mu = |A(x^{(0)})| > 0$. □

Definición 5. *Sea f una función derivable en el punto* $x^{(0)} \in R^n$. *Si* $df(x^{(0)}) = 0$, *entonces* $x^{(0)}$ *se llama punto estacionario de la función f.*

Es evidente que el punto $x^{(0)}$ en el que la función f es derivable será estacionario, si, y sólo si,

$$\frac{df(x^{(0)})}{dx_i} = 0, \quad i = 1, 2, \ldots, n. \tag{40.1}$$

Según el corolario del teorema 1, el punto de extremo, donde la función f es derivable, es estacionario; lo recíproco, por supuesto, no es cierto en el caso general: no todo punto estacionario, en el que la función es derivable, es un punto de extremo (véase el ejemplo 2 al final del p. 40.1).

Teorema 2 (condiciones suficientes de un extremo estricto). *Supongamos que la función f está definida y tiene derivadas continuas de segundo orden en cierto entorno del punto* $x^{(0)}$. *Sea* $x^{(0)}$ *un punto estacionario de la función f; en este caso, si la forma cuadrática*

$$A(dx_1, \ldots, dx_n) = \sum_{i,j=1}^{n} \frac{\partial^2 f(x^{(0)})}{\partial x_i \partial x_j} dx_i, dx_j, \tag{40.2}$$

(es decir, la segunda diferencial de la función f en el punto $x^{(0)}$) es definida positiva (definida negativa), entonces $x^{(0)}$ es el punto de mínimo estricto (de máximo estricto, respectivamente); en cambio, si la forma cuadrática (40.2) es indefinida, en el punto $x^{(0)}$ no hay extremo.

DEMOSTRACIÓN. Sea $U(x^{(0)}, \delta_0)$ un δ_0-entorno del punto $x^{(0)}$, estacionario para la función f, en el cual la función f tiene segundas derivadas continuas. Supongamos que el punto

$$x^{(0)} + dx = (x_1^{(0)} + dx_1, \ldots, x_n^{(0)} + dx_n)$$

pertenece a dicho entorno.

Rigiéndose por la fórmula de Taylor (véase (39.23)) y tomando en consideración las condiciones (40.1) referentes al carácter estacionario de un punto, obtenemos

$$\Delta f = f(x^{(0)} + dx) - f(x^{(0)}) = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f(x^{(0)})}{\partial x_i \partial x_j} dx_i\, dx_j + \varepsilon(dx)\rho^2,$$

donde $dx = (dx_1, \ldots, dx_n)$, $\rho^2 = dx_1^2 + \ldots + dx_n^2$, y

$$\lim_{\rho \to 0} \varepsilon(dx) = 0, \tag{40.3}$$

o bien

$$\Delta f = \frac{\rho^2}{2}\left[\sum_{i,j=1}^{n} \frac{\partial^2 f(x^{(0)})}{\partial x_i \partial x_j} \frac{dx_i}{\rho} \frac{dx_j}{\rho} + 2\varepsilon(dx)\right] =$$

$$= \frac{\rho^2}{2}\left[A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) + 2\varepsilon(dx)\right]. \tag{40.4}$$

El punto $\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right)$ se dispone en la esfera unidad S (es desir, en la esfera con centro en el origen de coordenadas y radio igual a 1), pues $\left(\frac{dx_1}{\rho}\right)^2 + \ldots + \left(\frac{dx_n}{\rho}\right)^2 = 1$.

Supongamos que la forma cuadrática (40.2) es de signo definido. En este caso, de acuerdo con el lema, $\inf_S |A| = \mu > 0$. Elijamos δ, $0 < \delta < \delta_0$, de un modo tal que sea $2|\varepsilon(dx)| < \mu$ cuando $\rho < \delta$. Entonces, para $\rho < \delta$, es decir, para $x^{(0)} + dx \in U(x^{(0)}, \delta)$ y $dx \neq 0$, toda la expresión entre corchetes en el segundo miembro de la fórmula (40.4) será del mismo signo que tiene el primer sumando $A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right)$:

$$\operatorname{sign} \Delta f = \operatorname{sign} A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right).$$

Por eso, si la forma cuadrática (40.2) es definida positiva, tenemos $\Delta f > 0$, y si definida negativa, entonces $\Delta f < 0$, cuando $x^{(0)} + dx \in U(x^{(0)}, \delta)$. Quiere decir, en el

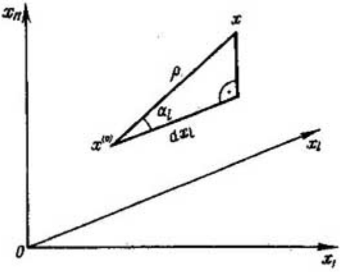

Fig. 161

primer caso $x^{(0)}$ es un punto de mínimo estricto y en el caso segundo, un punto de máximo estricto.

Supongamos ahora que la forma cuadrática (40.2) es indefinida. Esto significa que existen dos puntos $dx' = (dx_1', \ldots, dx_n')$ y $dx'' = (dx_1'', \ldots, dx_n'')$ tales que $A(dx_1', \ldots, dx_n') > 0$ y $A(dx_1'', \ldots, dx_n'') < 0$. Basándonos sobre este hecho no podemos decir de inmediato que el incremento de la función Δf cambia de signo en cualquier entorno del punto $x^{(0)}$, puesto que los puntos $x^{(0)} + dx' = (x_1^{(0)} + dx_1', \ldots, x_n^{(0)} + dx_n')$ y $x^{(0)} + dx'' = (x_1^{(0)} + dx_1'', \ldots, x_n^{(0)} + dx_n'')$ pueden, en general, incluso no pertenecer al dominio de definición de la función f. No obstante, el resultado deseado se deducirá de que la forma cuadrática $A(dx)$ conserva invariable un mismo signo o igualdad a cero en toda recta que pasa por el punto $x^{(0)}$ de la que está extraído este mismo punto, mientras que el valor $A\left(\dfrac{dx}{\rho}\right)$, $dx \neq 0$, no depende, en general, de la elección del punto en dicha recta.

Examinemos el punto $dx' = (dx_1', \ldots, dx_n')$. Tracemos una semirrecta que tiene su origen en el punto $x^{(0)}$ y pasa por el punto $x^{(0)} + dx'$. Para cualquier punto $x = (x_1, \ldots, x_n)$ de esta semirrecta pongamos $dx_i = x_i - x_i^{(0)}$, $i = 1, 2, \ldots, n$, y $\rho = \sqrt{\sum_{i=1}^{n} dx_i^2}$. En este caso (fig. 161)

$$\frac{dx_i}{\rho} = \cos\alpha_i, \quad i = 1, 2, \ldots, n, \tag{40.5}$$

donde $\cos\alpha_i$ son cosenos directores de la semirrecta en consideración. Por esta razón el punto

$$\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) = (\cos\alpha_1, \ldots, \cos\alpha_n), \tag{40.6}$$

dispuesto, evidentemente, en la esfera unidad*) S con centro $x^{(0)}$, será el mismo para todos los puntos x de esta semirrecta, es decir, el punto (40.6) no depende de la distancia ρ entre x y $x^{(0)}$.

Por consiguiente, el valor de la forma cuadrática (40.2) en el punto (40.6), es decir, $A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right)$ tampoco depende de ρ. De aquí, para todo punto (40.6) tenemos

$$A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) = A\left(\frac{dx'_1}{\rho'}, \ldots, \frac{dx'_n}{\rho}\right) = \frac{1}{\rho'^2} A(dx'_1, \ldots, dx'_n) > 0.$$

Sea $A\left(\frac{dx_1}{\rho}, \ldots, \frac{dx_n}{\rho}\right) = \mu' > 0$. Elijamos $\rho_0 > 0$ de modo tal que para $\rho < \rho_0$ tenga lugar la desigualdad $2|\varepsilon(dx)| < \mu'$, lo que es posible en vista de (40.3). Entonces, para cualquier punto $x^{(0)} + dx$, que se dispone en la semirrecta (40.5) y es tal que $0 < \rho = \sqrt{\sum_{i=1}^{n} dx_i^2} < \rho_0$, la expresión entre corchetes en la fórmula (40.4) tendrá el signo del primer término y, por ende, $\Delta f > 0$. Así pues, en cualquier entorno del punto $x^{(0)}$ hay puntos para los cuales $\Delta f > 0$.

Análogamente, partiendo del valor negativo de la forma cuadrática (40.2) en el punto (dx_i''), se demuestra que en todo entorno del punto $x^{(0)}$ existen puntos, para los cuales $\Delta f < 0$. Esto es precisamente el indicio de que en el caso que se considera $x^{(0)}$ no es un punto de extremo. $\square$

Cuando el teorema citado se aplica en la práctica, surge una pregunta: ¿cómo se establece, si es definida positiva o definida negativa la forma cuadrática (40.2)? Con este fin puede aprovecharse, por ejemplo, el así llamado *criterio de Sylvester* de la definición positiva de una forma cuadrática el cual se demuestra en el curso de álgebra. Este criterio consiste en lo siguiente.

Para que una forma cuadrática

$$A(x) = A(x_1, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \tag{40.7}$$

en la que $a_{ij} = a_{ji}$, $i, j = 1, 2, \ldots, n$, sea definida positiva, es necesario y suficiente que se verifiquen las desigualdades

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0 \ldots, \quad \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0.$$

*) Recordemos que para los cosenos directores se verifica la igualdad $\cos^2\alpha_1 + \ldots + \cos^2\alpha_n = 1$.

Al observar que la forma cuadrática $A(x)$ es definida negativa cuando, y sólo cuando, la forma cuadrática $-A(x) = \sum_{i,j=1}^{n} (-a_{ij}) x_i x_j$ es definida positiva, obtenemos, haciendo uso de las propiedades conocidas del determinante, el siguiente criterio para distinguir la definición negativa.

Para que la forma cuadrática (40.7) *sea definida negativa, es necesario y suficiente que se verifiquen las desigualdades*

$$a_{11} < 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} < 0 \ldots, (-1)^n \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} > 0.$$

Enunciemos ahora el teorema 2 para el caso de dos variables, expresando las condiciones que se imponen en la forma cuadrática (40.2) de una manera explícita, en términos de las segundas derivadas parciales.

Teorema 3. *Supongamos que la función $f(x, y)$ está definida y tiene derivadas parciales continuas de segundo orden en cierto entorno del punto (x_0, y_0), que es un punto estacionario para $f(x, y)$, es decir, en este punto*

$$f_x = f_y = 0. \tag{40.8}$$

Entonces, si en (x_0, y_0)

$$f_{xx}f_{yy} - f_{xy}^2 > 0, \tag{40.9}$$

es un punto de extremo estricto, a saber, de máximo estricto, si en este punto

$$f_{xx} < 0^{*)},$$

y de mínimo estricto, si

$$f_{xx} > 0.$$

Si en el punto (x_0, y_0)

$$f_{xx}f_{yy} - f_{xy}^2 < 0, \tag{40.10}$$

el extremo en él está ausente.

En fin, cuando

$$f_{xx}f_{yy} - f_{xy}^2 = 0 \tag{40.11}$$

en el punto (x_0, y_0), puede ocurrir que haya un extremo en él y puede ocurrir que no lo haya.

Efectivamente, si $f_{xx} \neq 0$ en el punto (x_0, y_0), la forma cuadrática (40.2) en nuestro caso puede escribirse así:

$$A(dx, dy) = f_{xx}dx^2 + 2f_{xy}dx\,dy + f_{yy}dy^2 =$$

---

*) De la condición (40.9) proviene, evidentemente, que $f_{xx} \neq 0$ en el punto (x_0, y_0).

$$= \frac{1}{f_{xx}} \left[(f_{xx}dx + f_{xy}dy)^2 + (f_{xx}f_{yy} - f_{xy}^2)dy^2 \right]. \tag{40.12}$$

Todas las derivadas parciales aquí y en adelante se refieren al punto (x_0, y_0).

Vemos inmediatamente que si se cumplen las condiciones (40.9), la expresión entre corchetes en la fórmula (40.12) es positiva para $dx^2 + dy^2 > 0$, es decir, $A(dx, dy)$ es una forma cuadrática definida, a saber, definida positiva cuando $f_{xx} > 0$ y definida negativa, cuando $f_{xx} < 0$. Por supuesto, dicha deducción proviene también del criterio de Sylvester. En el primer caso, según el teorema 2, (x_0, y_0) es el punto de mínimo estricto y en el segundo, de máximo estricto. Si, además, queda cumplida la condición (40.10), entonces para $dy = 0$, $dx \neq 0$ tenemos de (40.12): $\operatorname{sign} A(dx, 0) = \operatorname{sign} f_{xx}$, y para $dx = f_{xy}$, $dy = -f_{xx}$ se obtiene $\operatorname{sign} A(f_{xy}, -f_{xx}) = -\operatorname{sign} f_{xx}$, de donde se deduce que la forma cuadrática $A(dx, dy)$, es, cumplida la condición (40.10), indefinida.

Así pues, hemos investigado por completo el caso

$$f_{xx} \neq 0 \quad \text{y} \quad f_{xx}f_{yy} - f_{xy}^2 \neq 0.$$

El caso en que

$$f_{xx} = 0, \quad f_{yy} \neq 0 \quad \text{y} \quad f_{xx}f_{yy} - f_{xy}^2 \neq 0$$

se investiga de modo análogo.

Si, en cambio, $f_{xx} = f_{yy} = 0$, pero, como hasta ahora, $f_{xx}f_{yy} - f_{xy}^2 \neq 0$, entonces, evidentemente, $f_{xy} \neq 0$, por lo tanto en este caso se cumple la condición (40.10) y $A(dx, dy) = 2f_{xy}\,dx\,dy$. Vemos de inmediato que la forma cuadrática $A(dx, dy)$ es indefinida bajo las suposiciones adoptadas, pues $\operatorname{sign} A(dx, dy) = -\operatorname{sign} A(dx, -dy)$. Por eso, con el objeto de obtener los valores de la forma cuadrática de signos contrarios, basta tomar, al principio, dx y dy de un mismo signo y después, de signos opuestos. Según el teorema 2, (x_0, y_0) no es en el caso dado un punto de extremo.

Finalmente, el caso $f_{xx} = f_{yy} = f_{xy} = 0$ no es compatible con la suposición de que $f_{xx}f_{yy} - f_{xy} \neq 0$. De este modo, hemos examinado todos los casos posibles, cuando se cumple la desigualdad $f_{xx}f_{yy} - f_{xy}^2 \neq 0$.

Para acabar con la demostración del teorema nos basta mostrar con unos ejemplos que, cuando tiene lugar la correlación (40.11), puede existir un extremo y puede no existir.

El punto (0, 0) para la función $z = x^2 + 2xy + y^2$ es estacionario y en él $z_{xx} = z_{xy} = z_{yy} = 2$, y, por tanto, se cumple la condición (40.11). Al advertir que $z = (x + y)^2$, vemos que siempre $z \geqslant 0$, con la particularidad de que $z = 0$ en la recta $x + y = 0$; por ello el punto (0, 0) es un punto de extremo, aunque sea no estricto.

Para la función $z = xy^3$ el punto (0, 0) es también estacionario y en este punto $z_{xx} = z_{yy} = z_{xy} = 0$, por lo cual la condición (40.11) queda asimismo cumplida. Sin embargo, debido a que la fórmula, que representa dicha función, contiene potencias impares de las variables x e y, la función cambia de signo en cualquier entorno de cero, a consecuencia de lo cual, (0, 0) no es un punto de extremo.

40.3. OBSERVACIONES SOBRE LOS EXTREMOS EN LOS CONJUNTOS

Sea f una función derivable en un conjunto acotado abierto G y continua en la clausura $\bar{G}$ de dicho conjunto. Se pide hallar los valores máximo y mínimo de la función f en el conjunto $\bar{G}$ (de acuerdo con el teorema 3, p. 19.5, estos valores existen). Con este objeto podemos, por ejemplo, hallar todos los puntos estacionarios de la función f en G, calcular en éstos los valores de la función y escoger, siempre que sea posible (desde el punto de vista teórico, esto es posible, por ejemplo, cuando el número de los puntos estacionarios sea finito), aquellos en los que la función asume los valores máximo y mínimo entre todos los valores en los puntos estacionarios. A continuación, conviene comparar estos valores con los que toma la función en la frontera del conjunto abierto G, hallando, por ejemplo (en el caso de poder hacerlo), los valores máximo y mínimo de la función f en la frontera de la región G. Al comparar los valores máximo y mínimo en los puntos estacionarios con los valores correspondientes máximo y mínimo en la frontera del conjunto G, podemos, evidentemente, hallar el máximo y el mínimo buscados de f en $\bar{G}$.

Cuando G es una región plana y su frontera está constituida por una curva definida por cierta representación $x = x(t)$, $y = y(t)$, $\alpha \leqslant t \leqslant \beta$, la cuestión sobre la búsqueda de los valores extremos de la función $f(x, y)$ en la frontera de G se reduce a la investigación del extremo de la función de una variable $f(x(t), y(t))$, que se realiza con ayuda de los métodos ya conocidos.

Los métodos que pueden emplearse en el caso multidimensional para buscar puntos extremos en la frontera de una región serán considerados en el § 43.

Ejercicios. 1. Hállense los extremos de la función $z = x^3 + 12xy^2 - 15x - 24y$.

2. ¿Tendrá un extremo la función $z = x^4y^2 - 3x^2y + 2x + y$ en el punto $(1, 1)$?

3. Hállense los valores máximo y mínimo de la función $z = x^2 + y^2 - 4x - 2y + 4$ en una región cerrada limitada por las líneas $x = 4$, $y = -1$, $x - y = 3$.

4. Sea $a = \text{const} > 0$, $X = \{(x, y) : |x| < a, y \in R\}$. Hállense todos los extremos de la función $z = \frac{3}{2a}x^2 + \sqrt{6(a^2 - x^2)}\cos y$ en X y todos los valores máximos y mínimos de ella en X.

5. La superficie total de un paralelepípedo rectangular es igual a $6a^2$. ¿Para qué valores de la longitud de sus aristas el volumen del paralelepípedo será máximo?

§ 41. FUNCIONES IMPLÍCITAS

41.1. FUNCIONES IMPLÍCITAS DEFINIDAS POR UNA ECUACIÓN

Aclaremos las condiciones bajo las cuales una ecuación de varias variables define una función unívoca, es decir, define una de dichas variables como función de las demás. Empezaremos nuestras consideraciones por el estudio de una ecuación que contiene dos incógnitas

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

Si la función de dos variables $F(x, y)$ está definida en cierto subconjunto A del plano R^2_{xy}, $A \subset R^2_{xy}$, y si existe tal función de una variable $y = f(x)$, definida en el conjunto $B \subset R_x$ contenido en la proyección del conjunto A sobre el eje Ox, que para todo $x \in B$ tenga lugar $(x, f(x)) \in A$ y sea válida la identidad $F(x, f(x)) = 0$, entonces f se denomina *función implícita* definida por la ecuación $F(x, y) = 0$.

Lema. *Sea $F(x; y)$ una función continua en cierto entorno rectangular*

$$U(x_0, y_0) = \{(x, y) : |x - x_0| < \xi, \quad |y - y_0| < \eta\}^{*)}$$

del punto (x_0, y_0) y supongamos que dicha función es, para todo $x \in (x_0 - \xi, x_0 + \xi)$ fijo, estrictamente monótona respecto de y en el intervalo $(y - \eta, y_0 + \eta)$. En este caso, si

$$F(x_0, y_0) = 0,$$

existen tales entornos $U(x_0) = (x_0 - \delta, x_0 + \delta)$, del punto x_0 y $U(y_0) = (y - \varepsilon\, y_0 + \varepsilon)$ del punto y_0, que para todo $x \in U(x_0)$ se tiene una solución, y sólo una, $y \in U(y_0)$ de la ecuación $F(x, y) = 0$. Esta solución, que es una función de x y se designa por $y = f(x)$, es continua en el punto x_0 y

$$f(x_0) = y_0.$$

De este modo, el lema afirma, en particular, que para las suposiciones asumidas, la función implícita $y = f(x)$, definida por la ecuación $F(x, y) = 0$, existe y posee la propiedad de que las igualdades

$$F(x, y) = 0 \quad \text{e} \quad y = f(x)$$

son equivalentes a condición de que $x \in U(x_0)$, $y \in U(y_0)$.

DEMOSTRACIÓN. Por hipótesis del lema, la función $F(x, y)$ es, para todo $x \in (x_0 - \xi, x_0 + \xi)$ fijo, estrictamente monótona respecto de la variable y en el intervalo $(y_0 - \eta, y_0 + \eta)$, en particular, en éste es estrictamente monótona la función $F(x_0, y)$. Supongamos, para concretar, que es estrictamente creciente. Elijamos arbitrariamente $\varepsilon > 0$, subordinado sólo a la condición $0 < \varepsilon < \eta$. Como la función $F(x_0, y)$ de la variable y es, en el segmento $[y_0 - \varepsilon, y_0 + \varepsilon]$, estrictamente creciente y, por hipótesis, $F(x_0, y_0) = 0$, entonces

$$F(x_0, y_0 - \varepsilon) < 0, \quad F(x_0, y_0 + \varepsilon) > 0.$$

Mas la función de dos variables $F(x, y)$ es, conforme a la suposición, continua en el conjunto abierto $U(x_0, y_0)$ y $(x_0, y_0 - \varepsilon) \in U(x_0, y_0)$, $(x_0, y_0 + \varepsilon) \in U(x_0, y_0)$, por lo cual existe tal $\delta > 0$, $0 < \delta < \xi$, que en el δ-entorno del punto $(x_0, y_0 - \varepsilon)$ se verifica la desigualdad $F(x, y) < 0$, y en δ-entorno del punto $(x_0, y_0 + \varepsilon)$, la desigualdad $F(x, y) > 0$ (véase el lema 1 en el p. 19.3). En particular, para todo $x \in (x_0 - \delta, x_0 + \delta)$ (fig. 162) quedan válidas las desigualdades

$$F(x, y_0 - \varepsilon) < 0, \quad F(x, y_0 + \varepsilon) > 0. \tag{41.1}$$

---

*) En concordancia con las designaciones aceptadas en este libro, sería más correcto denotar el entorno del punto (x_0, y_0) mediante $U((x_0, y_0))$ en lugar de $U(x_0, y_0)$. Para simplificar las designaciones, convendremos en omitir el segundo paréntesis.

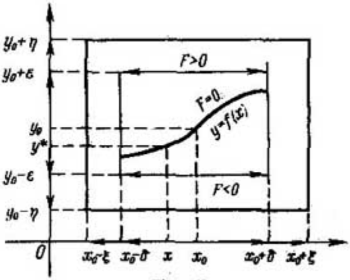

Fig. 162

Pongamos

$$U(x_0) \overset{\text{def}}{=} (x_0 - \delta, x_0 + \delta), \quad U(y_0) \overset{\text{def}}{=} (y_0 - \varepsilon, y_0 + \varepsilon).$$

Dado que, siendo fijado $x \in U(x_0)$, la función $F(x, y)$ de la variable y es continua en el segmento $[y_0 - \varepsilon, y_0 + \varepsilon]$, entonces, de la condición (41.1) se deduce (de acuerdo con el teorema de Cauchy sobre los valores intermedios de una función continua, véase el teorema 2 en el p. 6.2) que existe tal $y^* \in U(y_0)$ (véase la fig. 162) que $F(x, y^*) = 0$. Por ser $F(x, y)$ estrictamente monótona en el segmento $[y_0 - \varepsilon, y_0 + \varepsilon]$ respecto de la variable y, el valor indicado de y^* es único.

De este modo, se ha obtenido una correspondencia unívoca (función unívoca) $x \mapsto y^*$, $x \in U(x_0)$, $y^* \in U(y_0)$ que se designará mediante f: $y^* = f(x)$.

Por definición de esta correspondencia, para cualquier $x \in U(x_0)$ e $y^* = f(x)$ tenemos

$$F(x, y^*) = 0, \quad y^* \in U(y_0)$$

con la particularidad de que el punto y^*, que posee dicha propiedad, es único. Hemos demostrado, pues, la existencia y unicidad de la función buscada f.

Luego, por hipótesis del lema, $F(x_0, y_0) = 0$, y, como $x_0 \in U(x_0)$, $y_0 \in U(y_0)$, entonces, debido a la unicidad de la función f, tenemos $y_0 = f(x_0)$.

Al final notemos que $\varepsilon > 0$ se ha fijado arbitrariamente a condición de que $\varepsilon < \eta$, y que se ha encontrado para él tal $\delta > 0$, que de $|x - x_0| < \delta$ (es decir, de la condición $x \in U(x_0)$) provenía la inclusión $f(x) \in U(y_0)$, es decir, la desigualdad $|f(x) - f(x_0)| < \varepsilon$. Esto precisamente significa la continuidad de la función f en el punto x_0. $\square$

Las condiciones suficientes, cómodas para aplicarlas, de la resolubilidad unívoca de la ecuación $F(x, y) = 0$ en cierto entorno del punto (x_0, y_0), para el cual $F(x_0, y_0) = 0$, las proporciona el siguiente teorema.

Teorema 1. *Supongamos que la función $F(x, y)$ es continua en cierto entorno del punto (x_0, y_0) y tiene en dicho entorno una derivada parcial $F_y(x, y)$ que es continua en el punto (x_0, y_0). En este caso, si*

$$F(x_0, y_0) = 0, \quad F_y(x_0, y_0) \neq 0,$$

existen tales entornos $U(x_0)$ y $U(y_0)$ de los puntos respectivos x_0 e y_0, que para todo $x \in U(x_0)$ se tiene una y sólo una, solución $y = f(x) \in U(y_0)$ de la ecuación $F(x, y) = 0$[*]. *Esta solución es continua en todo punto de $U(x_0)$ e $y_0 = f(x_0)$.*

Si suponemos complementariamente que la función F tiene en cierto entorno del punto (x_0, y_0) una derivada parcial $F_x(x, y)$, continua en el punto (x_0, y_0), entonces la función $f(x)$ tendrá también en el punto x_0 una derivada y para ésta queda válida la fórmula

$$f'(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}.$$

DEMOSTRACIÓN. Debido a que la función $F(x, y)$ es continua en cierto entorno del punto (x_0, y_0) y es también continua en el mismo punto la derivada parcial $F_y(x, y)$, existe un entorno rectangular

$$U(x_0, y_0) = \{(x, y) : |x - x_0| < \xi, \quad |y - y_0| < \eta\}$$

del punto (x_0, y_0), donde la propia función $F(x, y)$ es continua y los valores de la derivada parcial $F_y(x, y)$ son del mismo signo que tiene su valor en el punto (x_0, y_0). Por esto, para todo $x \in (x_0 - \xi, x_0 + \xi)$ fijo la función $\varphi(y) \overset{\text{def}}{=} F(x, y)$ es derivable en el intervalo $(y_0 - \eta, y_0 + \eta)$, mientras que su derivada $\varphi'(y) = F_y(x, y)$ conserva constante el signo. Por consiguiente, la función $\varphi(y)$ es estrictamente monótona en el intervalo indicado.

Así pues, todas las condiciones del lema para la función $F(x, y)$ vienen cumplidas en el entorno rectangular construido $U(x_0, y_0)$. Por lo tanto, existen los entornos $U(x_0) = (x_0 - \delta, x_0 + \delta)$, $U(y_0) = (y_0 - \varepsilon, y_0 + \varepsilon)$ y la única función $y = f(x)$, definida en $U(x_0)$, tales que para todo $x \in U(x_0)$ tienen lugar la inclusión $f(x) \in U(y_0)$ y la igualdad $F(x, f(x)) = 0$, con la particularidad de que la función f es continua en el punto x_0.

Dado que para todo punto (x, y), para el cual $x \in U(x_0)$, $y \in U(y_0)$, existe su entorno rectangular $U(x, y)$ contenido en otro entorno rectangular

$$U_0(x_0, y_0) = \{(x, y) : |x - x_0| < \delta, \quad |y - y_0| < \varepsilon\}$$

(fig. 163), entonces para $U(x, y)$ también se cumplen todas las condiciones del lema. Por consiguiente, siendo única la solución $f(x)$ de la ecuación $F(x, y) = 0$ en el entorno $U_0(x_0, y_0)$, la función $y = f(x)$ es continua en todo punto $x \in U(x_0)$, de conformidad con el mismo lema.

Demostremos ahora la última afirmación del teorema. En vista de la continuidad de las derivadas parciales F_x y F_y en el punto (x_0, y_0), la función F es derivable en este punto:

$$F(x_0 + \Delta x, y_0 + \Delta y) - F(x_0, y_0) =$$
$$= F_x(x_0, y_0)\,\Delta x + F_y(x_0, y_0)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y, \quad (41.2)$$

[*] En este caso suele decirse también que la ecuación $F(x, y) = 0$ es resoluble unívocamente en el entorno $U(x_0, y_0) = \{(x, y) : x \in U(x_0),\ y \in U(y_0)\}$ del punto (x_0, y_0).

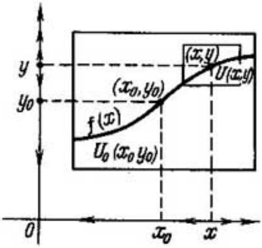

Fig. 163

donde

$$\lim_{\rho\to 0}\varepsilon_1 = \lim_{\rho\to 0}\varepsilon_2 = 0, \quad \rho = \sqrt{\Delta x^2 + \Delta y^2}.$$

Tomemos en la fórmula (41.2)

$$x_0 + \Delta x \in U(x_0), \quad \Delta y = f(x_0 + \Delta x) - f(x_0).$$

Entonces, en vista de la condición $F(x, f(x)) = 0$, obtenemos

$$F(x_0 + \Delta x, y_0 + \Delta y) = F(x_0 + \Delta x, f(x_0 + \Delta x)) = 0,$$

y, como $F(x_0, y_0) = 0$, de (41.2) tenemos

$$F_x(x_0, y_0)\,\Delta x + F_y(x_0, y_0)\,\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y = 0.$$

De aquí

$$\frac{\Delta y}{\Delta x} = -\frac{F_x(x_0, y_0) + \varepsilon_1}{F_y(x_0, y_0) + \varepsilon_2}. \tag{41.3}$$

Sea ahora $\Delta x \to 0$; entonces, siendo la función f continua, $\Delta y \to 0$, y esto significa que cuando $\Delta x \to 0$, se tiene $\rho = \sqrt{\Delta x^2 + \Delta y^2} \to 0$, de donde se desprende que en la fórmula (41.3) $\lim\limits_{\Delta x\to 0}\varepsilon_1 = \lim\limits_{\Delta x\to 0}\varepsilon_2 = 0$. Es por esto que cuando $\Delta x \to 0$, el límite del segundo miembro de la igualdad (41.3) existe y es igual a $-\dfrac{F_x(x_0, y_0)}{F_y(x_0, y_0)}$ (recordemos que $F_y(x_0, y_0) \neq 0$), por consiguiente, cuando $\Delta x \to 0$, existe también el límite del primer miembro, es decir, existe la derivada

$$f'(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}.\ \square \tag{41.4}$$

OBSERVACIÓN. Si las funciones F_x y F_y son continuas en el entorno $U_0(x_0, y_0)$ del punto (x_0, y_0), la derivada f' es continua en el intervalo $U(x_0)$. Efectivamente, al

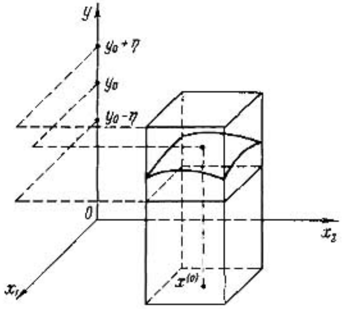

Fig. 164

aplicar la fórmula (41.4) a un punto arbitrario $x \in U(x_0)$, obtenemos

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))},$$

de donde proviene, según el teorema sobre la composición de funciones continuas, la continuidad de la función $f'(x)$ en $U(x_0)$.

Análogamente se introduce la noción de función implícita definida por la ecuación

$$F(x_1, \ldots, x_n, y) = 0, \tag{41.5}$$

como también se enuncia y se demuestra el teorema análogo al teorema 1. Para poder enunciarlo, es sólo suficiente que en el enunciado del teorema 1 por x se entienda un punto de un espacio n-dimensional, $x = (x_1, \ldots, x_n) \in R^n$, en particular, $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$.

Teorema 1′. *Supongamos que la función $F(x, y) = F(x_1, \ldots, x_n, y)$ es continua en cierto entorno del punto $(x^{(0)}, y^{(0)})$ y tiene en este entorno la derivada parcial F_y, continua en el punto $(x^{(0)}, y^{(0)})$.*

Si $F(x^{(0)}, y^{(0)}) = 0$ y $F_y(x^{(0)}, y^{(0)}) \neq 0$, existen tales entornos U_x y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)}$, que para todo $x \in U(x)$ existe una solución, y sólo una,

$$y = f(x) = f(x_1, \ldots, x_n) \in U_y$$

de la ecuación $F(x, y) = 0$), con la particularidad de que esta solución es continua en U_x y, además, $y^{(0)} = f(x^{(0)})$.*

Si, en adición, en cierto entorno del punto $(x^{(0)}, y^0)$, existen todas las derivadas parciales F_{x_i}, continuas en el punto $(x^{(0)}, y^{(0)})$, entonces en el punto $x^{(0)}$ existen tam-

*) En la fig. 164 se expone el caso en que $n = 2$ y el entorno U_x es rectangular.

bién las derivadas parciales f_{x_i}, $i = 1, 2, \ldots, n$, con la particularidad de que si las derivadas parciales F_{x_i}, $i = 1, 2, \ldots, n$, y F_y son continuas en el entorno del punto $(x^{(0)}, y^{(0)})$, las derivadas parciales f_{x_i} existen y son continuas en cierto entorno del punto $x^{(0)}$.

En este caso las fórmulas para las derivadas parciales de la función implícita, definida por la ecuación (41.5), tienen por expresión

$$\frac{\partial y}{\partial x_i} = -\frac{\dfrac{\partial F}{\partial x_i}}{\dfrac{\partial F}{\partial y}}, \quad i = 1, 2, \ldots, n.$$

Ejercicios. 1. Enúnciense las condiciones, bajo las cuales la función $f(x)$, definida por la ecuación $F(x, y) = 0$ (teorema 1), tiene en el punto (x_0, y_0) derivadas continuas de orden hasta n inclusive. Hállense las fórmulas para $f''(x_0)$ y $f'''(x_0)$.

2. Sirviéndose del teorema 1 y las respuestas de los ejercicios anteriores, hállense las condiciones suficientes para que exista la función $x = \varphi(y)$ que sea inversa respecto de $y = f(x)$ y que tenga en el punto y_0 derivadas continuas de orden hasta n inclusive. Demuéstrese que

$$\frac{d^2x}{dy^2} = -\frac{f''(x)}{[f'(x)]^3}, \quad \frac{d^3x}{dy^3} = \frac{3[f''(x)]^2 - f'(x)f'''(x)}{[f'(x)]^5}.$$

3. Hállense $\dfrac{dy}{dx}$ y $\dfrac{d^2y}{dx^2}$, si y es una función definida por la ecuación $\cos x^2y^2 + xy = 2$.

41.2. PRODUCTOS DE LOS CONJUNTOS

Antes de considerar la cuestión referente a la resolubilidad de los sistemas de ecuaciones, introduzcamos unas nociones adicionales.

Sea R_x^n un espacio euclídeo n-dimensional cuyos puntos se designarán mediante $x = (x_1, \ldots, x_n)$, sea R_y^m un espacio euclídeo m-dimensional cuyos puntos se designarán mediante $y = (y_1, \ldots, y_m)$ y sea R_{xy}^{n+m} un espacio euclídeo $(n + m)$- dimensional de los puntos

$$(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_m).$$

Definición 1. *Supongamos que $A \subset R_x^n$ y $B \subset R_y^m$ El conjunto de tales puntos (x, y) del espacio R_{xy}^{n+m} que $x \in A$ e $y \in B$ se llama producto*[*)] *de los conjuntos A y B y se denota por $A \times B$* (véase el p.1.2*). *De este modo,*

$$A \times B = \{(x, y) : x \in A, y \in B\}.$$

Ejemplos. 1. Si $A = R_x^n$, $B = R_y^m$ entonces

$$A \times B = R_x^n \times R_y^m = R_{xy}^{n+m}.$$

2. Supongamos que $n = 2$ y A es un círculo; $m = 1$ y B es un segmento. En este caso $A \times B$ es un cilindro circular recto (fig. 165).

[*)] Se emplea también el término "*producto cartesiano*".

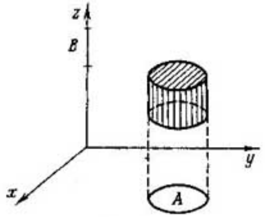

Fig. 165

3. Supongamos que $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in R_x^n$ y $A = P(x^{(0)}; \delta_1, \ldots, \delta_n) = \{x: |x_i - x_i^{(0)}| < \delta_i; i = 1, 2, \ldots, n\}$ es un entorno rectangular del punto $x^{(0)}$; supongamos también que $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)}) \in R_y^m$ y $B = P(y^{(0)}; \eta_1, \ldots, \eta_m) = \{y: |y_j - y_j^{(0)}| < \eta_j, j = 1, 2, \ldots, m\}$ es un entorno rectangular del punto $y^{(0)}$. En este caso

$$\begin{aligned} A \times B = \{(x, y): |x_i - x_i^{(0)}| < \delta_i, \quad i = 1, 2, \ldots, n; \\ |y_j - y_j^{(0)}| < \eta_j, \quad j = 1, 2, \ldots, m\} = \\ = P((x^{(0)}, y^{(0)}); \delta_1, \ldots, \delta_n, \quad \eta_1, \ldots, \eta_m) \end{aligned} \quad (41.6)$$

es un entorno rectangular del punto $(x^{(0)}, y^{(0)})$.

Lo recíproco es también obvio: puesto que todo entorno rectangular del punto $(x^{(0)}, y^{(0)})$ se escribe mediante la fórmula que está en medio de la igualdad (41.6), siempre puede ser representado como un producto de los entornos rectangulares de los puntos $x^{(0)}$ e $y^{(0)}$.

Ejercicio 4. Demuéstrese que si los conjuntos $A \subset R_x^n$ y $B \subset R_y^m$ son abiertos en los espacios respectivos R_x^n y R_y^m, su producto $A \times B$ será también un conjunto abierto en el espacio R_{xy}^{n+m}.

41.3. FUNCIONES IMPLÍCITAS DEFINIDAS POR UN SISTEMA DE ECUACIONES

Consideraremos las condiciones bajo las cuales el sistema de ecuaciones

$$F_i(x, y) = 0, \quad i = 1, 2, \ldots, m, \quad x \in R^n, \quad y \in R^m, \qquad (41.7)$$

o bien, en la forma desarrollada

$$\begin{aligned} F_1(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0, \\ F_2(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ F_m(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0, \end{aligned} \qquad (41.8)$$

es unívocamente resoluble respecto de $y_1, \ldots, y_m$ en cierto entorno del punto $(x^{(0)}, y^{(0)})$ en el que $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$.

Definición 2. *Sea dado un sistema de funciones $u_i = u_i(t_1, \ldots, t_n)$, $i = 1, 2, \ldots, m$, que en cierto punto $t^{(0)}$ tienen todas las derivadas parciales de primer orden. En este caso una matriz, compuesta de las derivadas parciales de dichas funciones en el punto $t^{(0)}$*

$$\left\| \begin{matrix} \frac{\partial u_1}{\partial t_1} & \frac{\partial u_1}{\partial t_2} & \ldots & \frac{\partial u_1}{\partial t_n} \\ \frac{\partial u_2}{\partial t_1} & \frac{\partial u_2}{\partial t_2} & \ldots & \frac{\partial u_2}{\partial t_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial u_m}{\partial t_1} & \frac{\partial u_m}{\partial t_2} & \ldots & \frac{\partial u_m}{\partial t_n} \end{matrix} \right\|$$

o, en la forma más breve,

$$\left\| \frac{\partial u_i}{\partial t_j} \right\| \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n.$$

se denomina matriz de Jacobi) del sistema de funciones dado.*
Si $m = n$, el determinante de la matriz de Jacobi se llama *determinante de Jacobi* o *jacobiano* del sistema de funciones $u_1, \ldots, u_n$ según las variables $t_1, \ldots, t_n$ y se designa así**):

$$\frac{\partial(u_1, \ldots, u_n)}{\partial(t_1, \ldots, t_n)}.$$

Veremos en adelante que el jacobiano de un sistema de funciones surge de un modo natural en las más diversas cuestiones de la teoría de la función de varias variables.

Antes de pasar a la exposición del teorema fundamental, mostremos con un ejemplo sencillo (sin profundizarnos en las detalles) la idea de su demostración y señalemos de qué modo surge en sus condiciones el jacobiano del sistema que se considera. Supongamos que en un entorno del punto (x_0, y_0, z_0) se han dado las funciones continuamente derivables F y Φ, con la particularidad de que

$$F(x_0, y_0, z_0) = 0,$$
$$\Phi(x_0, y_0, z_0) = 0.$$

Supongamos también que es preciso resolver el sistema de ecuaciones

$$F(x, y, z) = 0,$$
$$\Phi(x, y, z) = 0$$

*) K. Jacobi (1804 — 1851), matemático alemán.

**) Se emplea también la designación $\dfrac{D(u_1, \ldots, u_n)}{D(t_1, \ldots, t_n)}$.

en cierto entorno del punto indicado, hallando, a partir del sistema, las variables $y = \varphi(x)$ y $z = \psi(x)$ como funciones continuas φ y ψ de la variable x tales que sea $\varphi(x_0) = y_0$, $\psi(x_0) = z_0$. Al resolver con este fin, por ejemplo, la primera ecuación respecto de z, obtendremos $z = f(x, y)$. Al sustituir esta expresión en la segunda y al resolverla respecto de y, tendremos $y = \varphi(x)$. Poniendo $\psi(x) = f[x, \varphi(x)]$, obtendremos la solución buscada

$$y = \varphi(x)$$
$$z = \psi(x).$$

Surge, naturalmente, una pregunta sobre las condiciones que han de cumplirse para que se puedan ejecutar las operaciones indicadas, o, con más precisión, cuándo existen y están unívocamente definidas todas las funciones mencionadas. (Se debe aclarar, por supuesto, ¿dónde, es decir, para qué valores de las variables x e y, están definidas estas funciones? Esta cuestión no será el objeto del análisis detallado inmediato, pues no queremos apartarnos de la idea principal. Volveremos a considerarla en la demostración del teorema 2 de este punto.)

Para que una de las ecuaciones dadas, la primera, por ejemplo, sea resoluble en cierto entorno del punto (x_0, y_0, z_0) respecto de la variable z, es suficiente que (véase el teorema 1′ en el p. 41.1) $\dfrac{\partial F(x_0, y_0, z_0)}{\partial z} \neq 0$. Si $z = f(x, y)$ es la solución correspondiente, entonces, para que la ecuación $\Phi[x, y, f(x, y)] = 0$, que se obtiene como resultado de sustituir dicha solución en la segunda ecuación, sea resoluble respecto de la variable y, resulta suficiente que la derivada parcial total respecto de y en el primer miembro de la igualdad obtenida no se anule en el punto (x_0, y_0), es decir, que en dicho punto se verifique

$$\frac{\partial \Phi}{\partial y} + \frac{\partial \Phi}{\partial z}\frac{\partial f}{\partial y} \neq 0.$$

Pero, de acuerdo con el p. 41.1,

$$\frac{\partial f}{\partial y} = -\frac{\dfrac{\partial F}{\partial y}}{\dfrac{\partial F}{\partial z}},$$

por consiguiente, al sustituir esta expresión en la desigualdad antecedente, llegamos a que la condición de resolubilidad puede escribirse en la forma

$$\frac{\partial(F, \Phi)}{\partial(y, z)} = \frac{\partial F}{\partial y}\frac{\partial \Phi}{\partial z} - \frac{\partial F}{\partial z}\frac{\partial \Phi}{\partial y} \neq 0 \text{ en el punto } (x_0, y_0, z_0).$$

De esta condición se deduce, obviamente, que en el punto (x_0, y_0, z_0) o bien $\dfrac{\partial F}{\partial z} \neq 0$, o bien $\dfrac{\partial \Phi}{\partial z} \neq 0$, es decir, una de las ecuaciones dadas es resoluble respecto de z.

De este modo, el hecho de que el jacobiano $\dfrac{\partial(F, \Phi)}{\partial(y, z)}$ es distinto de cero en el punto (x_0, y_0, z_0) asegura, para el sistema dado de las ecuaciones, la existencia en cierto entorno del punto (x_0, y_0, z_0) de una solución en la forma

$$y = \varphi(x)$$
$$z = \psi(x).$$

Enunciemos ahora el teorema fundamental de este párrafo.

Teorema 2. *Supongamos que las funciones* $F_i(x, y) = F_i(x_1, \ldots, x_n, y_1, \ldots, y_m)$, $i = 1, 2, \ldots, m$, *son continuamente derivables en cierto entorno del punto* $(x^{(0)}, y^{(0)})$, *donde* $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$, $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)})$. *En este caso, si* $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$, *y en el punto* $(x^{(0)}, y^{(0)})$ *el jacobiano* $\dfrac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)}$ *no es igual a cero, entonces existen tales entornos* U_x *y* U_y *de los puntos* $x^{(0)}$ *e* $y^{(0)}$ *de los espacios respectivos* R_x^n *y* R_y^m, *que para todo* $x \in U_x$ *existe una única solución*

$$y = f(x) \in U_y$$

del sistema de ecuaciones (41.7):

$$y = f(x) = [y_k = f_k(x_1, \ldots, x_n), \quad k = 1, 2, \ldots, m]^{*)},$$

con la particularidad de que las funciones $f_k(x)$, $k = 1, 2, \ldots, m$, *que forman dicha solución son continuamente derivables en* U_x *y, además,* $f(x^{(0)}) = y^{(0)}$.

Así pues, si se cumplen las suposiciones del teorema, la condición

$$F_i(x, y) = 0, \quad i = 1, 2, \ldots, m, \quad (x, y) \in U_x \times U_y$$

es equivalente a la siguiente

$$y = f(x), \quad x \in U_x, \quad y \in U_y.$$

DEMOSTRACIÓN. Indiquemos ante todo que la afirmación: la solución $y = f(x)$ del sistema de ecuaciones (41.7) satisface la condición $f(x^{(0)}) = y^{(0)}$ proviene, evidentemente y de inmediato, de la afirmación sobre la unicidad de la solución $y = f(x) \in U_y$ para $x \in U_x$ y de las condiciones $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$, $x^{(0)} \in U_x$, $y^{(0)} \in U_y$.

Para demostrar el teorema apliquemos el método de la inducción matemática. Para el caso de una ecuación, es decir, cuando $m = 1$, el teorema se ha enunciado en el p. 41.1. Supongamos ahora que es también lícito para $m - 1$ ecuaciones $(m > 1)$. Demostremos que en tal caso tendrá lugar también para m ecuaciones.

Mostremos primero que cada una de las ecuaciones (41.8), la última, por ejemplo,

$$F_m(x_1, \ldots, x_n, \quad y_1, \ldots, y_m) = 0,$$

---

*) El sistema de funciones $f_k(x_1, \ldots, x_n)$, $k = 1, 2, \ldots, m$, viene designado con un símbolo $f(x)$, puesto que define una correspondencia determinada: a los puntos de cierto conjunto del espacio R_x^n el sistema indicado les pone en correspondencia los puntos determinados del espacio R_y^m, o, cómo suele decirse, aplica el conjunto mencionado del espacio R_x^n en el espacio R_y^m.

puede ser resuelta en el entorno del punto $(x^{(0)}, y^{(0)})$ por lo menos respecto de una variable. En efecto, por hipótesis del teorema, en el punto $(x^{(0)}, y^{(0)})$

$$\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)} = \begin{vmatrix} \dfrac{\partial F_1}{\partial y_1} & \ldots & \dfrac{\partial F_1}{\partial y_m} \\ \ldots & \ldots & \ldots \\ \dfrac{\partial F_m}{\partial y_1} & \ldots & \dfrac{\partial F_m}{\partial y_m} \end{vmatrix} \neq 0,$$

y, por ende, en dicho punto por lo menos un elemento en la última fila del jacobiano es distinto de cero. Supongamos, para concretar, que este elemento será el último:

$$\frac{\partial F_m(x^{(0)}, y^{(0)})}{\partial y_m} \neq 0.$$

De aquí, en virtud del teorema 1′ del p. 41.1, se desprende que la ecuación $F_m(x, y) = 0$ puede ser resuelta respecto de y_m en cierto entorno del punto (x_0, y_0). Daremos una definición más exacta de esta idea. Designemos mediante U el entorno del punto $(x^{(0)}, y^{(0)})$, en el que las funciones F_i, $i = 1, 2, \ldots, m$, son continuamente derivables y pongamos $\tilde{y} \overset{\text{def}}{=} (y_1, \ldots, y_{m-1})$. Entonces, existen un entorno rectangular U^{m+n-1} del punto

$$(x^{(0)}, \tilde{y}^{(0)}) = (x_1^0, \ldots, x_n^{(0)}, y_1^{(0)}, \ldots, y_{m-1}^{(0)}) \tag{41.9}$$

y un entorno U^1 del punto $y_m^{(0)}$ tales que $U^{m+n-1} \times U^1 \subset U$, y se tiene la única función

$$y_m = \varphi(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}), \tag{41.10}$$

que está definida en U^{m+n-1} y que satisface las siguientes condiciones: si

$$(x, \tilde{y}) = (x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) \in U^{m+n-1},$$

entonces

$$\varphi(x, \tilde{y}) = \varphi(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) \in U^1, \tag{41.11}$$

$$F_m(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}, \quad \varphi(x, \tilde{y})) = 0. \tag{41.12}$$

Además, de acuerdo con el mismo teorema 1′, la función $\varphi(x, \tilde{y})$ es continuamente derivable en U^{m+n-1} y

$$\varphi(x^{(0)}, \tilde{y}^{(0)}) = y_m^{(0)}. \tag{41.13}$$

En este caso, si $(x, \tilde{y}) \in U^{m+n-1}$ e $y_m \in U^1$, entonces el sistema (41.8) es equivalente al sistema

$$F_j(x, y) = 0, \quad j = 1, 2, \ldots, m-1, \quad y_m = \varphi(x, y). \tag{41.14}$$

Sustituyamos en las primeras $m - 1$ ecuaciones del sistema (41.14) la expresión (41.10). Al introducir las designaciones

$$\Phi_i(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) = F_i(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}, \varphi(x_1, \ldots, x_n, y_1, \ldots, y_{m-1})) \quad i = 1, 2, \ldots, m-1. \tag{41.15}$$

obtendremos, en este caso, el siguiente sistema de $m - 1$ ecuaciones con $m + n - 1$ incógnitas:

$$\begin{gathered} \Phi_1(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \Phi_{m-1}(x_1, \ldots, x_n, y_1, \ldots, y_{m-1}) = 0. \end{gathered} \tag{41.16}$$

Además, para $(x, \bar{y}) \in U^{m+n-1}$, $y_m \in U^1$ el sistema de ecuaciones

$$\begin{gathered} \Phi_j(x, \bar{y}) = 0, \quad j = 1, 2, \ldots, m-1, \\ y_m = \varphi(x, \bar{y}) \end{gathered} \tag{41.17}$$

es equivalente al sistema (41.14).

Mostremos que el sistema (41.16) satisface las condiciones que se diferencian de las que son satisfechas por el sistema (41.8), sólo en que $m - 1$ se ha sustituido por m. En efecto, las funciones Φ_k, $k = 1, 2, \ldots, m-1$, son continuamente derivables en el entorno U^{m+n-1} como composición de las funciones continuamente derivables. De las condiciones $F_i(x^{(0)}, y^{(0)}) = 0$, $i = 1, 2, \ldots, m$, y (41.15), (41.13) proviene que $\Phi_k(x^{(0)}, y^{(0)}) = 0$, $k = 1, 2, \ldots, m-1$.
Demostremos que en el punto $(x^{(0)}, \bar{y}^{(0)})$ (véase (41.9))

$$\frac{\partial(\Phi_1, \ldots, \Phi_{m-1})}{\partial(y_1, \ldots, y_{m-1})} \neq 0.$$

Indiquemos previamente que de (41.10) y (41.15) se deduce que

$$\frac{\partial \Phi_i}{\partial y_k} = \frac{\partial F_i}{\partial y_k} + \frac{\partial F_i}{\partial y_m}\frac{\partial \varphi}{\partial y_k}, \quad i, k = 1, 2, \ldots, m-1, \tag{41.18}$$

y de (41.12), que

$$\frac{\partial F_m}{\partial y_k} + \frac{\partial F_m}{\partial y_m}\frac{\partial \varphi}{\partial y_k} = 0, \quad k = 1, 2, \ldots, m-1. \tag{41.19}$$

Ahora, en el determinante $\dfrac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)}$ agreguemos a la k-ésima columna la última columna multiplicada por $\dfrac{\partial \varphi}{\partial y_k}$, $k = 1, \ldots, m-1$, de lo cual, como se sabe, el valor del determinante no cambiará. Por eso, utilizando (41.18) y (41.19) y desarrollando el determinante obtenido por los elementos de la última fila, obtenemos

$$\left.\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m)}\right|_{(x^{(0)}, y^{(0)})} =$$

$$=\left|\begin{matrix} \dfrac{\partial F_1}{\partial y_1}+\dfrac{\partial F_1}{\partial y_m}\dfrac{\partial \varphi}{\partial y_1} \ldots \dfrac{\partial F_1}{\partial y_{m-1}}+\dfrac{\partial F_1}{\partial y_m}\dfrac{\partial \varphi}{\partial y_{m-1}} & \dfrac{\partial F_1}{\partial y_m} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ \dfrac{\partial F_m}{\partial y_1}+\dfrac{\partial F_m}{\partial y_m}\dfrac{\partial \varphi}{\partial y_1} \ldots \dfrac{\partial F_m}{\partial y_{m-1}}+\dfrac{\partial F_m}{\partial y_m}\dfrac{\partial \varphi}{\partial y_{m-1}} & \dfrac{\partial F_m}{\partial y_m} \end{matrix}\right|_{(x^{(0)},\, y^{(0)})} =$$

$$=\left|\begin{matrix} \dfrac{\partial \Phi_1}{\partial y_1} \ldots \dfrac{\partial \Phi_1}{\partial y_{m-1}} & \dfrac{\partial F_1}{\partial y_m} \\ \cdots\cdots\cdots\cdots & \\ \dfrac{\partial \Phi_{m-1}}{\partial y_1} \ldots \dfrac{\partial \Phi_{m-1}}{\partial y_{m-1}} & \dfrac{\partial F_{m-1}}{\partial y_m} \\ 0 \quad \ldots \quad 0 & \dfrac{\partial F_m}{\partial y_m} \end{matrix}\right|_{(x^{(0)},\, \bar{y}^{(0)})} =$$

$$= \frac{\partial F_m(x^{(0)}, y^{(0)})}{\partial y_m} \left.\frac{\partial(\Phi_1, \ldots, \Phi_{m-1})}{\partial(y_1, \ldots, y_{m-1})}\right|_{(x^{(0)},\, \bar{y}^{(0)})}$$

y como el primer miembro de la igualdad es distinto de cero, será no nulo también el segundo miembro, de donde

$$\left.\frac{\partial(\Phi_1, \ldots, \Phi_{m-1})}{\partial(y_1, \ldots, y_{m-1})}\right|_{(x^0,\, \bar{y}^0)} \neq 0.$$

En vista de que las condiciones, análogas a las condiciones para las funciones F_i, $i = 1, 2, \ldots, m$, se cumplen para las funciones Φ_i, $i = 1, 2, \ldots, m - 1$, y, de acuerdo con la suposición sobre la inducción, el sistema de ecuaciones (41.16) es unívocamente resoluble respecto de las variables $y_1, \ldots, y_{m-1}$ en cierto entorno del punto $(x^{(0)}, \bar{y}^{(0)})$. Con más precisión, sea U^{m+n-1} un entorno rectangular del punto $(x^{(0)}, \bar{y}^{(0)})$ obtenido como resultado de la resolución de la ecuación $F_m = 0$ respecto de la variable y_m. Desarrollémoslo en un producto de los entornos rectangulares U'_x y $U'_{\bar{y}}$ de los puntos $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ e $\bar{y}^{(0)} = (y_1^{(0)}, \ldots, y_{m-1}^{(0)})$ en los espacios respectivos R_x^n y $R_{\bar{y}}^{m-1}$ (aquí $\bar{y} = (y_1, \ldots, y_{m-1})$): $U^{m+n-1} = U'_x \times U'_{\bar{y}}$. Entonces, existen un entorno $U_x \subset U'_x$ del punto $x^{(0)}$, un entorno $U_{\bar{y}} \subset U'_{\bar{y}}$ del punto $\bar{y}^{(0)}$ y el único sistema de funciones

$$\begin{gathered} y_1 = f_1(x) = f_1(x_1, \ldots, x_n), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_{m-1} = f_{m-1}(x) = f_{m-1}(x_1, \ldots, x_n), \end{gathered} \tag{41.20}$$

que están definidas en el conjunto U_x y que satisfacen las siguientes condiciones: si $x \in U_x$, se tiene

$$(f_1(x), \ldots, f_{m-1}(x)) \in U_{\tilde{y}} \tag{41.21}$$

y en U_x las funciones (41.20) son continuamente derivables y satisfacen el sistema de ecuaciones (41.16):

$$\Phi_i(x_1, \ldots, x_n, f_1(x), \ldots, f_{m-1}(x)) = 0, \qquad i = 1, 2, \ldots, m-1. \tag{41.22}$$

Es importante notar que en virtud de la unicidad de la solución (41.20) del sistema (41.16) para $x \in U_x$, $\tilde{y} \in U\tilde{y}$ e $y_m \in U^1$, el sistema de ecuaciones

$$\begin{aligned} y_k &= f_k(x), \quad k = 1, 2, \ldots, m-1, \\ y_m &= \varphi(x, y) \end{aligned} \tag{41.23}$$

es equivalente al sistema (41.17).

Sustituyendo las expresiones (41.20) en (41.10), obtendremos una función de x, definida en U_x; designémosla con f_m:

$$y_m = \varphi(x_1, \ldots, x_n, f_1(x), \ldots, f_{m-1}(x)) = f_m(x_1, \ldots, x_n) = f_m(x). \tag{41.24}$$

Señalemos que el sistema de funciones

$$y_k = f_k(x_1, \ldots, x_n), \quad k = 1, 2, \ldots, m, \tag{41.25}$$

(véanse (41.20) y (41.24)) es precisamente el sistema buscado de funciones que satisface las exigencias enunciadas en el teorema. En efecto, sea $U_y = U_{\tilde{y}} \times U^1$; entonces si $x \in U_x$, de conformidad con (41.21) y (41.11), $f(x) = (f_1(x), \ldots, f_m(x)) \in U_y$. De (41.15), (41.22), (41.24) y (41.12) se deduce que $F_i(x, f(x)) = 0$, $i = 1, 2, \ldots, m$, para cualquier $x \in U_x$. Debido al teorema 1′ y la suposición de inducción, las funciones (41.10) y (41.20) y, por lo tanto, la función (41.24), son continuamente derivables.

Hemos demostrado de este modo que la aplicación $f(x)$, definida por las funciones (41.25), es una solución continuamente derivable del sistema de ecuaciones (41.8) en el conjunto U_x, y si, en este caso, $x \in U_x$, entonces $y = f(x) \in U_y$. Observemos, además, que si $x \in U_x$, el sistema (41.25) es equivalente al sistema (41.23).

Resta demostrar la unicidad de la solución del sistema de ecuaciones (41.8). Con este fin representemos en forma del esquema que sigue los pasos (realizados en el transcurso de la demostración) de unos sistemas de ecuaciones a los otros, equivalentes a los primeros, es decir, a aquellos que tienen exactamente las mismas soluciones:

$$F_i(x, y) = 0, \quad i = 1, 2, \ldots, m.$$
$$\Updownarrow$$
$$F_j(x, y) = 0, \quad j = 1, 2, \ldots, m-1,$$
$$y_m = \varphi(x, y).$$
$$\Updownarrow \qquad \Phi_j(x, \tilde{y}) = F_j(x, \tilde{y}, \varphi(x, \tilde{y})), \quad j = 1, 2, \ldots, m-1,$$
$$\Phi_j(x, y) = 0, \quad j = 1, 2, \ldots, m-1,$$
$$y_m = \varphi(x, y).$$
$$\Updownarrow$$
$$y_j = f_j(x), \quad j = 1, 2, \ldots, m-1,$$

$$y_m = \varphi(x, \bar{y}). \qquad f_m(x) = \varphi(x, f_1(x), \ldots, f_{m-1}(x))$$
$$\Updownarrow$$
$$y_i = f_i(x), \quad i = 1, 2, \ldots, m.$$

Las flechas dobles significan la equivalencia de los sistemas de ecuaciones en consideración, la cual en todo caso tiene lugar para $x \in U_x$, $y \in U_y$. De esta equivalencia deriva precisamente la unicidad de la solución (41.25) del sistema (41.8) en los entornos que se consideran, de lo cual, según se ha observado anteriormente, en virtud de la condición $F_i(x^{(0)}, y^{(0)}) = 0, i = 1, 2, \ldots, m$, se deduce que $f(x^{(0)}) = y^{(0)}$. □

El teorema demostrado sobre las funciones implícitas es uno de los más importantes teoremas del análisis matemático y tiene toda una serie de aplicaciones en diferentes apartados de éste. Con algunas de ellas nos encontraremos en los capítulos ulteriores de nuestro curso. El teorema citado es un "teorema puro de existencia": tanto su enunciado, como la demostración aducida no dan origen, en el caso general, a un método concreto de resolución del sistema (41.8). Por ejemplo, si todas las F_k, $k = 1, 2, \ldots, m$, en el sistema citado de ecuaciones son funciones elementales, entonces, siguiendo el esquema de la demostración del teorema, no tendremos éxito en "hallar (se trata de un caso general) en la forma explícita" todas aquellas funciones cuya existencia se ha utilizado en la demostración mencionada, y obtener una solución del sistema que tenga también la forma de las funciones elementales. Realmente, en este caso la solución del sistema de ecuaciones (41.8), la que existe en virtud del teorema citado, no es, en el caso general, una lista de las funciones elementales (si incluso dicho sistema consta de una sola ecuación).

Naturalmente, si las funciones F_k son elementales y, por consiguiente, vienen definidas mediante ciertas fórmulas, la solución del sistema (41.8) puede hallarse con cualquier grado de exactitud, es decir, en principio, las tablas de los valores de estas soluciones pueden ser compuestas con cualquier grado de exactitud. En realidad, sin embargo, la exactitud, con la que se calculan las soluciones, se determina, desde luego, por el objetivo concreto, en aras del cual se resuelve el sistema en consideración. El propio teorema 2 en este caso nos da una certeza objetiva de que, al realizar correctamente los cálculos correspondientes, calculamos, de hecho, la solución buscada del sistema. No nos detendremos en los métodos numéricos que se emplean para solucionar los sistemas de ecuaciones; sólo algunas de las cuestiones de la resolucion numerica de las ecuaciones se consideran en "Complemento" al final de este tomo.

Indiquemos como una circunstancia esencial que el teorema 2, al igual que todos los teoremas de este tipo, proporciona métodos cualitativos, en el caso dado, para estudiar las propiedades de las soluciones del sistema de ecuaciones.

Resulta interesante notar que las derivadas parciales de la solución del sistema (41.8) se expresan con facilidad, si se cumplen las condiciones del teorema 2, en forma explícita a través de las derivadas parciales de las funciones F_k, $k = 1, 2, \ldots, m$. Efectivamente, con el fin de hallar la derivada parcial $\dfrac{\partial y_j}{\partial x_i}$, se deben derivar las igualdades (41.8) respecto de x_i, considerándolas identidades según $x_1, \ldots, x_n$, es decir, sustituyendo en ellas sus soluciones $y_j = y_j(x_1, \ldots, x_n)$, $j = 1, \ldots, m$. En este caso obtendremos

$$\frac{\partial F_k}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial F_k}{\partial y_j}\frac{\partial y_j}{\partial x_i} = 0, \quad k = 1, 2, \ldots, m.$$

Este sistema de ecuaciones, lineales respecto de $\frac{\partial y_j}{\partial x_i}$, en virtud de que su determinante en el punto considerado es distinto de cero:

$$\frac{\partial(F_1, \ldots, F_m)}{\partial(y, \ldots, y_m)} \neq 0,$$

tiene una solución, y sólo una, la cual puede hallarse, por ejemplo, según la regla de Cramer*).

Si es preciso hallar todas las derivadas

$$\frac{\partial y_j}{\partial x_i}, \quad i = 1, 2, \ldots, n \quad j = 1, 2, \ldots, m,$$

conviene calcular las diferenciales de ambos miembros de las identidades mencionadas arriba (41.8). Haciendo uso de la invariación de la forma de la primera diferencial respecto de la elección de las variables, obtendremos

$$\sum_{j=1}^{n} \frac{\partial F_k}{\partial x_i} dx_i + \sum_{j=1}^{m} \frac{\partial F_k}{\partial y_j} dy_j = 0, \quad k = 1, 2, \ldots, m.$$

En vista de la misma condición $\frac{\partial(F_1, \ldots, F_m)}{\partial(y_1, \ldots, y_m} \neq 0$, este sistema de las ecuaciones, lineales respecto de $dy_1, \ldots, dy_m$, tiene una solución y ésta es única. Si la hallamos, el coeficiente de dx_i en la expresión para dy_j será precisamente la derivada parcial $\frac{\partial y_j}{\partial x_i}$.

Los dos métodos son aplicables también para el cálculo de las derivadas de órdenes superiores de las funciones $y_j(x_1, \ldots, x_n)$ que son soluciones del sistema de las ecuaciones (41.8) (por ejemplo, bajo el supuesto de que todas las funciones F_k, $k = 1, 2, \ldots, m$, tienen derivadas continuas de órdenes correspondientes). Aplicando el método de diferenciales, se debe recordar, por supuesto, que las diferenciales de orden superior al primero, si se expresan en términos de las diferenciales de las funciones, son más complejas, en lo que se refiere a su expresión, en comparación con el caso cuando ellas se expresan sólo en términos de las diferenciales de las variables independientes (véase el p. 21.2).

Las derivadas de órdenes superiores de las funciones $y_j(x_1, \ldots, x_n)$ pueden obtenerse también, por derivación sucesiva, a partir de las expresiones para las prime-

*) G. Cramer (1704 — 1752), matemático suizo.

ras derivadas $\frac{\partial y_j}{\partial x_i}$, determinadas según las fórmulas de Cramer a base del sistema de ecuaciones citado anteriormente

$$\frac{\partial F_k}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial F_k}{\partial y_j}\frac{\partial y_j}{\partial x_i} = 0, \quad k = 1, 2, \ldots, m,$$

en forma de una razón entre dos determinantes. Dicha razón puede derivarse tantas veces cuantas veces son derivables las funciones F_k, $k = 1, \ldots, m$. En este caso, si todas las derivadas de las funciones F_k, $k = 1, \ldots, m$, de orden hasta r inclusive son continuas, lo serán también todas las derivadas parciales de las funciones $y_j(x_1, \ldots, x_n)$, $j = 1, \ldots, m$, de orden hasta el mismo r.

Un conjunto (llamado también a menudo una clase) de todas las funciones que son r veces continuamente derivables en la región G se denota con $C^r(G)$. De este modo: *si, adicionalmente a las condiciones del teorema* 2, $F_k \in C^r(U)$, $k = 1, \ldots, m$, *donde U es un cierto entorno del punto* $(x^{(0)}, y^{(0)})$, *las soluciones* $y_j = y_j(x_1, \ldots, x_n)$ *del sistema de ecuaciones* (41.7) *pertenece también a la clase* $C^r(U_x)$ *en cierto entorno* U_x *del punto* $x^{(0)}$.

Ejercicios. 5. ¿Bajo qué condiciones impuestas sobre f y g, la ecuación $y = xf(z) + g(z)$ define, en cierto entorno U del punto (x_0, y_0), la función $z(x, y) \in C^2(U)$? Demuéstrese que, cumplidas estas condiciones, para cualesquiera $(x, y) \in U$

$$z_y^2 z_{xx} - 2z_x z_y z_{xy} + z_x^2 z_{yy} = 0.$$

6. Sea dado un sistema de ecuaciones

$$uf'(v) = [y - f(v)]^2$$

$$(x + v)f'(v) = y - f(v).$$

Hállense las condiciones, impuestas sobre la función f, para las cuales este sistema define en cierto entorno U del punto (x_0, y_0) las funciones $u = u(x, y)$, $v = v(x, y)$ de la clase $C^1(U)$. Demuéstrese que en este caso $u_x u_y = u$ en todo punto de U.

41.4. APLICACIONES

En este punto se estudiarán las aplicaciones $f: X \to R^m$, $X \subset R^n$, es decir, aplicaciones de tal índole que a todo punto $x = (x_1, \ldots, x_n)$ del conjunto X, dispuesto en el espacio puntual aritmético n-dimensional R^n (véase el p. 18.1), le ponen en correspondencia el punto $y = (y_1, \ldots, y_m)$ del espacio puntual aritmético m-dimensional R^m. De este modo, $f: (x_1, \ldots, x_n) \to (y_1, \ldots, y_m)$, $(x_1, \ldots, x_n) \in X$. Obviamente, la definición de tal aplicación f es equivalente a la definición de m funciones $f_j: X \to R$ tales que sea $f_j: x \mapsto y_j$, $j = 1, \ldots, m$, $x \in X$, $y_j \in R$. Estas funciones

$$f_j(x) = f_j(x_1, \ldots, x_n), \quad i = 1, 2, \ldots, m, \quad x \in X, \tag{41.26}$$

se denominan *funciones coordenadas de la aplicación* f y se escribe

$$f = (f_1, \ldots, f_m).$$

A las aplicaciones que se consideran se generaliza el concepto de continuidad.

Definición 3. *Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$, si para cualquier entorno $V(y^{(0)})$ del punto $y^{(0)} = f(x^{0})$ existe tal entorno $U(x^{(0)})$ del punto $x^{(0)}$ que*

$$f(U(x^{(0)}) \cap X \subset V(y^{(0)}).$$

Por cuanto en todo entorno de un punto*) se contiene su entorno esférico, dicha definición es equivalente a la siguiente.

Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$, si para cualquier ε-entorno del punto $y^{(0)} = f(x^{(0)})$ existe tal δ-entorno del punto $x^{(0)}$ que

$$f(U(x^{(0)}, \delta) \cap X) \subset U(y^{(0)}, \varepsilon).$$

Esto, a su vez, puede ser parafraseado, con ayuda de las desigualdades, de la manera siguiente.

Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$, si para cualquier $\varepsilon > 0$ existe tal $\delta > 0$ que para todos los puntos $x \in X$, que satisfacen la condición $\rho(x, x^{(0)}) < \delta$, se cumple la desigualdad

$$\rho(f(x), f(x^{(0)})) < \varepsilon.$$

La definición de continuidad puede enunciarse también en términos de las sucesiones.

Definición 3'. *Una aplicación $f: X \to R^m$, $X \subset R^n$, se denomina continua en el punto $x^{(0)} \in X$ si para cualquier sucesión $x^{(k)} \in X$, $k = 1, 2, \ldots$, tal que $\lim_{k\to\infty} x^{(k)} = x^{(0)}$, tiene lugar*

$$\lim_{k\to\infty} f(x^{(k)}) = f(x^{(0)}).$$

La equivalencia de estas dos definiciones se demuestra por analogía con la demostración de la equivalencia de las definiciones del límite de funciones según Cauchy y según Heine. Demos a conocer esta demostración.

Supongamos que la aplicación f *es continua en el punto* $x^{(0)}$ en el sentido de la definición 3, $x^{(k)} \in X$, $k = 1, 2, \ldots$, y

$$\lim_{k\to\infty} x^{(k)} = x^{(0)}. \tag{41.27}$$

Prefijemos $\varepsilon > 0$. Para ε existe un $\delta > 0$ tal que si $x \in X$, $\rho(x, x^{(0)}) < \delta$, se verifica la desigualdad $\rho(f(x), f(x^{(0)})) < \varepsilon$.

En virtud de la condición (41.27), existe tal número k_0 que para cualesquiera $k \geqslant k_0$ tenemos $x^{(k)} \in U(x^{(0)}, \delta)$ y, por consiguiente, $\rho(f(x^{(k)}), f(x^{(0)})) < \varepsilon$. Esto precisamente significa que $\lim_{k\to\infty} f(x^{(k)}) = f(x^{(0)})$.

Supongamos ahora que la aplicación f es continua en el punto $x^{(0)}$ en el sentido de la definición 3' y que las condiciones de la definición 3 no están cumplidas, es decir, existe tal $\varepsilon_0 > 0$, que para cualquier $\delta > 0$ existe un $x_\delta \in U(x^{(0)}, \delta) \cap X$, para el

---

*) Recordemos que se llama entorno de un punto cualquier conjunto abierto que contiene dicho punto (véase la definición 14 en el p. 18.2).

cual $\rho(f(x_\delta), f(x^{(0)})) \geqslant \varepsilon_0$. Tomando sucesivamente $\delta = \dfrac{1}{k}$, $k = 1, 2, \ldots$, y poniendo, para abreviar, $x^{(k)} = x_{1/k}$, obtenemos $x^{(k)} \in U\left(x^{(0)}, \dfrac{1}{k}\right) \cap X$, es decir, $\rho\ (x^{(k)}, x^{(0)}) < \dfrac{1}{k}$. Por consiguiente, $\lim\limits_{k\to\infty} x^{(k)} = x^{(0)}$ y $x^{(k)} \in X$; sin embargo, $\rho(f(x^{(k)}), f(x^{(0)}) \geqslant \geqslant \varepsilon_0$ y, de este modo, la sucesión $\{f(x^{(k)})\}$ no tiene el punto $f(x^{(0)})$ en calidad de su límite. La contradicción obtenida demuestra la afirmación enunciada. $\square$

Lema 1. *La aplicación* $f = (f_1, \ldots, f_m)$: $X \to R^m$, $X \subset R^n$, *es continua en el punto* $x^{(0)}$ *cuando, y sólo cuando, en dicho punto son continuas todas las funciones coordenadas* $f_1, \ldots, f_m$.

DEMOSTRACIÓN DE LA NECESIDAD. Supongamos que la aplicación f es continua en el punto $x^{(0)} \in X$, $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)}) \stackrel{\text{def}}{=} f(x^{(0)})$. De acuerdo con la definición 3, para todo entorno $V(y^{(0)})$ del punto $y^{(0)}$, en particular, para cada uno de sus entornos cúbicos (véase el p. 18.1)

$$P(y^{(0)}, \varepsilon) = \{y: |y_i - y_i^{(0)}| < \varepsilon\}$$

existe tal entorno $U(x^{(0)})$ del punto $x^{(0)}$ que

$$f(U(x^{(0)}) \cap X) \subset P(y^{(0)}, \varepsilon).$$

Por consiguiente, para todo $x \in U(x^{(0)}) \cap X$ se verifican las desigualdades

$$|f_j(x) - y_j^{(0)}| < \varepsilon, \quad j = 1, 2, \ldots, m.$$

Esto es testimonio de que todas las funciones coordenadas $f_1, \ldots, f_m$ son continuas en el punto $x^{(0)}$.

DEMOSTRACIÓN DE LA SUFICIENCIA. Supongamos que todas las funciones coordenadas $f_1, \ldots, f_m$ son continuas en el punto $x^{(0)} \in X$, $y^{(0)} = (y_1^{(0)}, \ldots, y_m^{(0)}) = f(x^{(0)})$ y está dado un entorno $V(y^{(0)})$ del punto $y^{(0)}$. En este caso existe tal $\varepsilon > 0$ que el entorno ε-cúbico $P(y^{(0)}, \varepsilon)$ del punto $y^{(0)}$ está contenido en $V(y^{(0)})$,

$$P(y^{(0)}, \varepsilon) \subset V(y^{(0)}).$$

Por ser continua cada una de las funciones f_j, $j = 1, 2, \ldots, m$, en el punto $x^{(0)}$ existen tales entornos $U_j = U(x^{(0)})$ que para $x \in U_j \cap X$ se verifica la desigualdad

$$|f_j(x) - y_j^{(0)}| < \varepsilon. \tag{41.28}$$

Pongamos $U = \bigcap\limits_{j=1}^{m} U_j$. Entonces, U, siendo una intersección de un número finito de los conjuntos abiertos U_j, será conjunto abierto, con la particularidad de que como todos los U_j contenían el punto $x^{(0)}$, lo contiene también U. De este modo, el conjunto U es un entorno del punto $x^{(0)}$. Además, si $x \in U \cap X$, entonces para todo $j = 1, 2, \ldots, m$, se verifican las desigualdades (41.28). Esto quiere decir que

$$f(x) \in P(y^{(0)}, \varepsilon),$$

y, por lo tanto, $f(x) \in V(y^{(0)})$. Así pues, para un entorno arbitrario $V(y^{(0)})$ se ha encontrado tal entorno U del punto $x^{(0)}$ que

$$f(U \cap X) \subset V(y^{(0)}). \square$$

El lema 1 muestra, en particular, que las definiciones de las aplicaciones continuas de un segmento introducidas al considerar el concepto de curva en el p. 16.1 (cuando el segmento se aplica en un espacio tridimensional) y en el p. 18.2 (cuando el segmento se aplica en un espacio euclídeo arbitrario n-dimensional) como aplicaciones cuyas funciones coordenadas son continuas, son equivalentes a la definición de las aplicaciones continuas de un segmento como aplicaciones de tal género que en todo punto del segmento satisfacen las condiciones de la definición 3 de este párrafo.

La aplicación: $f: X \to R^m_y$, $X \subset R^n_x$ se llama *continua en el conjunto* X, si es continua en todo punto del conjunto X.

Lema 2. *La aplicación f de un conjunto abierto del espacio R^n_x en el espacio R^m_y es continua en dicho conjunto cuando, y sólo cuando, la preimagen de cada conjunto abierto del espacio R^m_y, realizándose la aplicación f, será un conjunto abierto del espacio R^n_x.*

DEMOSTRACIÓN DE LA NECESIDAD. Supongamos que f aplica continuamente el conjunto abierto $G \subset R^n_x$ en el espacio R^m_y y sea U un conjunto abierto del espacio R^m_y: $U \subset R^m_y$. Mostremos que la preimagen $f^{-1}(U)$ de este conjunto es un conjunto abierto en el espacio R^n_x. Si el conjunto $f^{-1}(U)$ es vacío, la afirmación queda obvia, puesto que todo conjunto vacío es abierto.

Supongamos que el conjunto $f^{-1}(U)$ no es vacío, es decir, existe un punto $x^{(0)} \in$ $\in f^{-1}(U)$ y, por lo tanto, $f(x^{(0)}) \in U$. Como U es un conjunto abierto, será un entorno del punto $y^{(0)} = f(x^{(0)})$. Por ello, por ser la aplicación f continua en el punto $x^{(0)}$ (véase la definición 3'), existe tal entorno U_x de este punto que $f(U_x \cap G) \subset U$, por consiguiente, $U_x \cap G \subset f^{-1}(U)$. Como el conjunto $U_x \cap G$, representando una intersección de dos conjuntos abiertos U_x y G, es abierto y puesto que $x^{(0)} \in U_x \cap G$, entonces $x^{(0)}$ es un punto interior del conjunto $f^{-1}(U)$.

De este modo, todo punto de la preimagen del conjunto abierto U es un punto interior de esta preimagen, quiere decir, la preimagen es un conjunto abierto.

DEMOSTRACIÓN DE LA SUFICIENCIA. Sea f una aplicación del conjunto abierto G del espacio R^n_x en R^m_y y supongamos que, al realizarse esta aplicación, la preimagen de cada conjunto abierto en el espacio R^m_y es un conjunto abierto en R^n_x. Sea $x^{(0)} \in G$. Mostremos que la aplicación f es continua en el punto $x^{(0)}$.

Sea U_y cierto entorno del punto $y^{(0)} = f(x^{(0)})$. Como la preimagen $f^{-1}(U_y)$ del conjunto abierto U_y es, por hipótesis, un conjunto abierto, y, evidentemente, $x^{(0)} \in f^{-1}(U_y) \subset G$, entonces el conjunto $U_x = f^{-1}(U_y)$ es un entorno del punto $x^{(0)}$ y, además, $f(U_x) = U_y$. De aquí se deduce de inmediato la continuidad de la aplicación f en el punto $x^{(0)}$ (véase la definición 3). $\square$

Ejemplo. Examinemos la aplicación $f: R^2 \to R$, definida por la fórmula $f(x, y) =$ $= \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$. De acuerdo con el lema 2, la preimagen del conjunto abierto $(-\infty, 0)$, es decir, un conjunto de puntos (x, y) que satisfacen la desigualdad $\frac{x^2}{a^2} + \frac{y^2}{b^2} <$ < 1 (y, por consiguiente, forman el interior de una elipse), como también la preimagen del conjunto abierto $(0, +\infty)$, o sea, un conjunto de tales puntos (x, y) que $\frac{x^2}{a^2} + \frac{y^2}{b^2} >$ > 1 (estos puntos forman el exterior de una elipse), son conjuntos abiertos.

En general, si $f: R^n \to R$ es una función continua en R^n, para cualquier número $a \in \in R$ los conjuntos $\{x: f(x) < a, x \in R^n\}$ y $\{x: f(x) > a, x \in R^n\}$ son conjuntos abiertos, siendo preimágenes de los conjuntos abiertos $(-\infty, a)$ y $(a, +\infty)$.

El teorema de Weierstrass en el cual se demuestra que las funciones continuas en un compacto son acotadas y pueden alcanzar sus cotas inferior y exterior se extiende también al caso de las aplicaciones continuas. Con más precisión, resulta válida la siguiente afirmación.

Lema 3. *Sea $f: A \to R^m$, $A \subset R^n$ una aplicación continua de un compacto A en el espacio R^m. El conjunto $f(A)$ es también un compacto en este caso.*

En la forma más breve: *la preimagen continua de un compacto es un compacto.*

DEMOSTRACIÓN. Sea $y^{(k)} \in f(A)$ una sucesión arbitraria de los puntos pertenecientes a $f(A)$. Por definición de la preimagen de un conjunto para la aplicación dada, existe tal punto $x^{(k)} \in A$, que $f(x^{(k)}) = y^{(k)}$, cualquiera que sea $k = 1, 2, \ldots$. Como A es un compacto, de la sucesión $\{x^{(k)}\}$ se puede separar una subsucesión convergente $\{x^{(k_s)}\}$ cuyo límite $x^{(0)}$ pertenece al compacto A: $\lim\limits_{s\to\infty} x^{(k_s)} = x^{(0)} \in A$.

Por ser la función f continua en el punto $x^{(0)}$, tenemos

$$\lim_{s\to\infty} f(x^{(k_s)}) = f(x^{(0)}), \text{ es decir, } \lim_{s\to\infty} y^{(k_s)} = f(x^{(0)}) \in f(A).$$

De este modo, de toda sucesión de puntos, pertenecientes al conjunto $f(A)$, se puede separar una sucesión convergente cuyo límite pertenece a dicho conjunto. Esto precisamente significa que $f(A)$ es un compacto. $\square$

OBSERVACIÓN. Del lema 3 se infiere el teorema demostrado anteriormente de que una función real continua en un compacto alcanza sus cotas inferior y exterior (véase el p. 19.5). En efecto, de conformidad con el lema 3, el conjunto de valores de tal función es un compacto en una recta numérica, mientras que todo compacto tiene en una recta númerica los puntos finitos maximal y minimal. Esto proviene de que un compacto es un conjunto acotado y, por ende, tiene la cota superior (inferior) finita la cual, en virtud de su definición, es un punto adherente del conjunto. Ya que el compacto está cerrado, el punto pertenece a él y es, evidentemente, su punto maximal (minimal).

La noción de continuidad uniforme se generaliza también para el caso de las aplicaciones.

Definición 4. *La aplicación f del conjunto $X \subset R^n_x$ en el espacio R^m_y se llama uniformemente continua, si para cualquier $\varepsilon > 0$ existe tal $\delta = \delta(\varepsilon) > 0$, que para cualesquiera puntos $x' \in X$ y $x'' \in X$, que satisfacen la condición $\rho(x', x'') < \delta$, se verifica la desigualdad $\rho(f(x'), f(x'')) < \varepsilon$.*

Para las aplicaciones es lícita también una afirmación análoga al teorema de Cantor (véase el p. 19.6) para las funciones continuas.

Lema 4. *Una aplicación continua de un compacto es uniformemente continua.*

DEMOSTRACIÓN. Hagamos uso del mismo método que se aplicó al demostrar el teorema de Cantor sobre la continuidad uniforme de las funciones reales continuas en un compacto (véase el teorema 5 en el p. 19.6).

Admitamos que existe una aplicación $f: A \to R^m$, $A \subset R^n$, continua en el compacto A, pero de manera no uniforme. En este caso existe tal $\varepsilon_0 > 0$ que para cualquier $\delta > 0$ existen unos puntos $x'_\delta \in A$ y $x''_\delta \in A$, para los cuales tienen lugar las desigualdades

$$\rho(x_\delta'', x_\delta') < \delta \quad \text{y} \quad \rho(f(x_\delta''),\ f(x_\delta')) \geqslant \varepsilon_0.$$

Sea $\delta = \dfrac{1}{k}$, $x'^{(k)} \overset{\text{def}}{=} x'_{1/k}, x''^{(k)} \overset{\text{def}}{=} x''_{1/k}$, $k = 1, 2, \ldots$. Como A es un compacto, de la sucesión $\{x'^{(k)}\}$ se puede separar una subsucesión convergente $\{x'^{(k_s)}\}$ cuyo límite $x^{(0)}$ está contenido en el conjunto A: $\lim\limits_{s\to\infty} x'^{(k_s)} = x^{(0)} \in A$. En este caso de

$$\rho(x''^{(k_s)}, x^{(0)}) \leqslant \rho(x'^{(k_s)}, x''^{(k_s)}) + \rho(x'^{(k_s)}, x^{(0)}) < \frac{1}{k_s} + \rho(x'^{(k_s)}, x^{(0)}) \to 0,$$

cuando $s \to \infty$, se deduce que la subsucesión $\{x''^{(k_s)}\}$ de la segunda sucesión $\{x''^{(k)}\}$ también converge hacia el punto $x^{(0)}$.

Cabe señalar ahora que de la continuidad de la aplicación f en el punto $x^{(0)}$ se desprende que

$$\lim_{s\to\infty} f(x'^{(k_s)}) = \lim_{s\to\infty} f(x''^{(k_s)}) = f(x^{(0)}),$$

y puesto que

$$\rho(f(x''^{(k_s)}), f(x'^{(k_s)})) \leqslant \rho(f(x''^{(k_s)}),\ f(x^{(0)})) + \rho(f(x^{(0)}), f(x'^{(k_s)})) \to 0, \text{ para } s \to \infty,$$

entonces, $\lim\limits_{s\to\infty} \rho(f(x''^{(k_s)}),\ f(x'^{(k_s)})) = 0$. Esto contradice la condición $\rho(f(x''^{(k_s)}), f(x'^{(k_s)})) \geqslant \varepsilon_0$. □

Con ayuda de las propiedades demostradas de las aplicaciones continuas se puede obtener una propiedad de las regiones (es decir, de los conjuntos abiertos linealmente conexos, véase el p. 18.2) que será útil en lo que sigue. Enunciemos dicha propiedad también en forma de un lema.

Lema 5. *Un conjunto abierto constituye una región cuando, y sólo cuando, sus dos puntos cualesquiera pueden ser unidos mediante una línea quebrada íntegramente dispuesta en el conjunto.*

DEMOSTRACIÓN. La suficiencia de la condición enunciada no requiere demostración. En efecto, si en cierto conjunto abierto $G \subset R^n$ cualesquiera dos puntos pueden unirse mediante cierta quebrada, íntegramente dispuesta en el conjunto, entonces, puesto que toda quebrada es una curva (véase el p. 16.5), cualesquiera dos puntos del conjunto G resultan unidos en él por una curva lo que testimonia, según la definición (véase la definición 25 en el p. 18.2), que el conjunto abierto G es linealmente conexo, es decir, es una región (véase la definición 26 en el mismo punto).

Demostremos la necesidad de las condiciones del lema. Sea G una región del espacio R^n. Examinemos los puntos $x \in G$ e $y \in G$. Por definición de la región, existe una curva $\gamma = \{r(t), a \leqslant t \leqslant b\}$, que une en G los puntos x e y, es decir, $r(a) = x$, $r(b) = y$ y $r(t) \in G$, $a \leqslant t \leqslant b$. La curva γ representa en sí la imagen continua del segmento $[a, b]$ que es un compacto y, por lo tanto (véase el lema 3), la curva también es un compacto. Por cuanto el compacto γ y el conjunto cerrado $R^n \setminus G$ no se intersecan, la distancia entre ellos es superior a cero (véase el lema 7 en el p. 18.2). Por consiguiente, existe un número $\eta > 0$ tal que $\rho(\gamma, R^n \setminus G) > \eta$.

La aplicación $r(t)$, $a \leqslant t \leqslant b$, del segmento $[a, b]$, siendo continua, es también uniformemente continua (véase el lema 4). Por esto existe tal $\delta > 0$, que para cualesquiera dos puntos $t' \in [a, b]$ y $t'' \in [a, b]$, que satisfacen la condición $|t'' - t'| < \delta$, se verifica la desigualdad

$$\rho(r(t''), \quad r(t')) < \eta.$$

De aquí se deduce que para cualquier partición $\tau = \{t_i\}_{i=0}^{i=k}$ del segmento $[a, b]$ de finura $\delta_\tau < \delta$ todos los puntos de la quebrada λ_τ cos vértices $r(t_i)$, $i = 0, 1, \ldots, k$ estarán contenidos en G (¿por qué?). Por consiguiente, $\lambda_\tau \subset G$.

Puesto que el origen y el extremo de la quebrada λ_τ los consituyen el origen y el extremo, respectivamente, de la curva γ, es decir, los puntos prefijados arbitrariamente x e y de G, hemos demostrado, pues, que cualesquiera dos puntos de una región pueden unirse mediante una quebrada. $\square$

Supongamos ahora que $X \subset R^n_x$, $D \subset R^m_y$, $y = f(x)$ es la aplicación del conjunto X en R^m_y, con la particularidad de que $f(X) \subset D$ y $z = g(y)$ es la aplicación de D en R^p_z, es decir, $f\colon X \to D$, $g\colon D \to R^p_z$. En este caso tiene sentido la composición $g \circ f\colon X \to R^p_z$, que aplica el conjunto $X \subset R^n_x$ en el espacio p-dimensional R^p_z: $(g \circ f) \stackrel{\text{def}}{=} g(f(x))$, $x \in X$.

Hemos de notar que si la aplicación $f(x)$ del conjunto X es continua en el punto $x^{(0)} \in X$, y $g(y)$ está definida en cierto entorno del punto $y^{(0)} = f(x^{(0)})$, siempre existe tal entorno U_x del punto $x^{(0)}$ que en el conjunto $X \cap U_x$ tiene sentido la composición $g \circ f$. En efecto, sea U_y un entorno del punto $y^{(0)}$ en el que viene definida la aplicación $g(y)$; de acuerdo con la definición 3, para dicho entorno existe tal entorno U_x que $f(U_x \cap X) \subset U_y$. Es evidente que para todos los puntos $x \in U_x \cap X$ tiene sentido precisamente la composición $g \circ f$.

Recordemos, además, que, de conformidad con la terminología introducida para las funciones (véase el p. 1.2*), la aplicación $f\colon X \to R^m_y$, $X \subset R^n_x$ se llama *biunívoca* o bien *inyección* si a los puntos diferentes del conjunto X les corresponden en esta aplicación distintos puntos. En este caso suele decirse también que el conjunto X se aplica biunívocamente mediante dicha aplicación sobre el conjunto $f(X)$, es decir, $f\colon X \to f(X)$ es una biyección. Cumplida esta condición, en el conjunto $f(X)$ existe una aplicación inversa unívoca (una función inversa) $f^{-1}(y) = x$, donde x es tal que $f(x) = y$. Por ello, $f^{-1}[f(x)] = x$, es decir, ésta es una *aplicación idéntica* (se denomina aplicación idéntica del conjunto X aquella que a todo punto $x \in X$ le pone en correspondencia el mismo punto).

Definición 5. *Si la aplicación f del conjunto $X \subset R^n_x$ en el espacio R^m_y es biunívoca y continua sobre X y una aplicación inversa f^{-1} es continua sobre $f(X)$, entonces f se llama aplicación homeomorfa o homeomorfismo, y el conjunto $f(X)$, imagen homeomorfa del conjunto X, o bien, que es lo mismo, un conjunto homeomorfo respecto del conjunto X.*

Obviamente, si f es un homeomorfismo del conjunto X, f^{-1} es el homeomorfismo del conjunto $f(X)$.

Cuando se realiza la aplicación homeomorfa de un conjunto abierto sobre otro conjunto abierto, las imágenes de los subconjuntos abiertos son también abiertas. En efecto, si f es una aplicación homeomorfa del conjunto abierto G sobre otro conjunto abierto Γ, V es un subconjunto abierto del conjunto G, $W = f(V)$, entonces $V = f^{-1}(W)$, es decir, V es una imagen del conjunto W, al realizarse la aplicación continua f^{-1} del conjunto abierto Γ, y, por consiguiente, W es una preimagen del conjunto abierto V en esta aplicación. Por ello, de acuerdo con el lema 2, el conjunto W es abierto.

Consideraremos ahora la composición de las aplicaciones continuas.

Lema 6. *Sea $f: X \to R^m$, $X \subset R^n$; $g: D \to R^s$, $D \supset f(X)$. Si la aplicación f es continua en el punto $x^{(0)} \in X$, y g es continua en el punto $f(x^{(0)})$, entonces la composición $g \circ f$ es también continua en el punto $x^{(0)}$.*

La demostración de dicha afirmación puede realizarse por un método análogo al empleado en la demostración del teorema 6, p. 5.16 y del teorema 2, p. 19.5. Este último método se basa en la definición de la continuidad en términos de los entornos. Con el fin de evitar la repitición, esta vez demostraremos el lema partiendo de la definición de continuidad en términos de las sucesiones.

Sea $x^{(k)} \in X$, $k = 1, 2, \ldots$, y $\lim\limits_{k\to\infty} x^{(k)} = x^{(0)}$. En este caso $f(x^{(k)}) \in D$ y, por ser continua la aplicación f en el punto $x^{(0)}$, tenemos

$$\lim_{k\to\infty} f(x^{(k)}) = f(x^{(0)}). \tag{41.29}$$

En vista de que la aplicación g es continua en el punto $f(x^0)$, para cualquier sucesión $y^{(k)} \in D$, $k = 1, 2, \ldots$, $\lim\limits_{k\to\infty} y^{(k)} = f(x^{(0)})$ tiene lugar $\lim\limits_{k\to\infty} g(y^{(k)}) = g(f(x^{(0)}))$. En particular, debido a (41.29), para $y^{(k)} = f(x^{(k)})$, se tiene

$$\lim_{k\to\infty} g(f(x^{(k)})) = g(f(x^{(0)})).$$

Esto precisamente implica la continuidad de la composición $g \circ f$ en el punto $x^{(0)}$. $\square$

Ejercicio 7. Demuéstrese que la aplicación biunívoca continua de un compacto del espacio R^n en cierto espacio R^m es un homeomorfismo.

Como conclusión determinemos qué se entenderá por imagen de una curva, para una aplicación continua dada, y demostremos el lema sobre las imágenes continuas de los conjuntos linealmente conexos.

Sea f una aplicación continua del conjunto $X \subset R^n_x$ en el espacio R^m_y y Γ, una curva *dispuesta íntegramente en el conjunto X*, es decir, se ha dado una clase de las aplicaciones equivalentes de segmentos en el conjunto X (véase el § 16).

Sea

$$x(t), \quad a \leqslant t \leqslant b,$$

una de las representaciones de la curva Γ. Una curva en el espacio R^m_y cuya representación es la aplicación

$$f[x(t)], \quad a \leqslant t \leqslant b,$$

se denomina *imagen de la curva* Γ en la aplicación f y se designa por $f(\Gamma)$.

Esta definición es correcta, puesto que, con las suposiciones adoptadas, $f(x(t))$, $a \leqslant t \leqslant b$, es una aplicación continua de un segmento en el espacio y, por consiguiente, define cierta curva.

Lema 7. *Sea $f: X \to R^m$ una aplicación continua del conjunto linealmente conexo $X \subset R^n$ en el espacio R^m. En este caso el conjunto $f(X)$ es también linealmente conexo.*

En la forma más breve: *la imagen continua de un conjunto linealmente conexo es linealmente conexa.*

DEMOSTRACIÓN. Supongamos que X es un conjunto linealmente conexo y f es su aplicación continua en R^m. Para demostrar que el conjunto $f(X)$ es linealmente conexo, hay que demostrar que sus dos puntos cualesquiera pueden unirse en $f(X)$ mediante una curva continua (véase la definición 25 en el p. 18.2). Sea $y^{(1)} \in f(X)$ e $y^{(2)} \in f(X)$; elijamos dos puntos cualesquiera $x^{(1)} \in f^{-1}(y^{(1)})$ y $x^{(2)} \in f^{-1}(y^{(2)})$. Como $x^{(1)} \in X$, $x^{(2)} \in X$ y X es linealmente conexo, existe tal curva Γ, que su origen coincide con el punto $x^{(1)}$, el extremo coincide con el punto $x^{(2)}$ y todos los puntos de ella pertenecen al conjunto X.

La curva $f(\Gamma)$ es la curva buscada. En efecto, su origen es el punto $y^{(1)} = f(x^{(1)})$ y el extremo, el punto $y^{(2)} = f(x^{(2)})$. Todos los demás puntos de la curva pertenecen al conjunto $f(X)$. De este modo, $f(X)$ es un conjunto linealmente conexo. $\square$

Ejercicios. 8. La aplicación $f\colon R^2 \to R^2$ viene dada de modo siguiente: $(x, y) \mapsto (2x, 3y)$. ¿En qué ésta transforma la circunferencia $x^2 + y^2 = 1$?

9. Hállese la imagen de la recta $x = 2$ del plano Oxy, realizándose la aplicación $f\colon R^2 \to R^2$, definida del modo siguiente: $(x, y) \mapsto (xy, y)$.

10. En el plano Oxy se ha dado una recta $x = c$ ($c = \text{const} \neq 0$). Hállese su imagen en la aplicación $f\colon R^2 \to R^2$ con las funciones coordenadas ($e^x \cos y$, $e^x \operatorname{sen} y$).

41.5. APLICACIONES VECTORIALES

Al estudiar las aplicaciones derivables (su definición se dará en el p. 41.7) resulta más conveniente que el espacio R^n, donde se dispone el conjunto que se aplica, y el espacio R^m, en el cual se realiza la aplicación, sean considerados como espacios euclídeos vectoriales (véase el p. 18.4). Para simplificar, un vector n-dimensional con coordenadas $(x_1, \ldots, x_n)$ se designará mediante el mismo símbolo x que se ha usado para designar un punto del espacio puntual n-dimensional con las mismas coordenadas. Esto no nos conducirá a una equivocación, pues tanto el punto del espacio n-dimensional como el vector n-dimensional representan un surtido ordenado de n números naturales.

Sea $X \subset R^n$ y $f\colon X \to R^m$, donde, en nuestro caso, la aplicación f pone en correspondencia a todo vector $x \in X$ cierto vector $y = f(x) \in R^m$. Las aplicaciones de este tipo se llamarán *vectoriales*.

Si $e_1, \ldots, e_n$ son los vectores coordenados en el espacio R^n (véase el p. 18.4), $\varepsilon_1, \ldots, \varepsilon_m$ son los vectores coordenados en el espacio R^m y $x = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i e_i$, $y = (y_1, \ldots, y_m) = \sum_{j=1}^{m} y_j \varepsilon_j$ e $y = f(x)$, entonces cualquier coordenada y_j, $j = 1, 2, \ldots, m$, del vector y es también una función del vector $x \in X$ y, por lo tanto, una función de sus coordenadas $x_1, \ldots, x_n$:

$$y_j = f_j(x) = f_j(x_1, \ldots, x_n), \quad j = 1, 2, \ldots, m. \tag{41.30}$$

Igual que en el caso de un espacio puntual (véase (41.26)), las funciones (41.30) se denominan *funciones coordenadas de la aplicación* f y se escribe $f = (f_1, \ldots, f_m)$

La interpretación de los puntos n-dimensionales $(x_1, \dots, x_n)$ en calidad de vectores no impide, por supuesto, que se consideran tales propiedades de las aplicaciones como son su continuidad y continuidad uniforme. Por ello, todo lo dicho en el punto anterior sobre las aplicaciones queda en vigor para las aplicaciones vectoriales. Recordemos, además, que para la distancia $\rho(x, y)$ entre los vectores x e y es válida la fórmula (véase la vórmula (18.37) en el punto 18.4) $\rho(x, y) = |x - y|$.

A título de ejemplo señalemos que la longitud $|x|$ del vector $x \in R^n$ es una función continua en R^n. Esto proviene de la desigualdad (18.36): como para todo $x_0 \in R^n$ y todo $x \in R^n$ se verifica la desigualdad $||x| - |x_0|| \leqslant |x - x_0|$, entonces

$$\lim_{x \to x_0} |x| = \lim_{|x - x_0| \to 0} |x| = |x_0|.$$

41.6. APLICACIONES LINEALES

Consideraremos una clase especial de las aplicaciones del espacio R^n en R^m, denominadas *lineales*.

Definición 6. *Una aplicación f: $R^n \to R^m$ se llama lineal (o, en forma más completa, lineal homogénea), si para cualesquiera dos vectores $x' \in R^n$, $x'' \in R^n$ y dos números $\lambda' \in \mathbf{R}$, $\lambda'' \in \mathbf{R}$ se verifica la igualdad*

$$f(\lambda' x' + \lambda'' x'') = \lambda' f(x') + \lambda'' f(x'').$$

De esta definición se deduce por inducción que, realizándose la aplicación lineal f, cualquier combinación lineal finita de vectores $x^{(j)} \in R^n$ se aplica en la misma combinación lineal de imágenes $f(x^{(j)})$, $j = 1, 2, \dots, k$, de dichos vectores

$$f\left(\sum_{j=1}^{k} \lambda_j x^{(j)}\right) = \sum_{j=1}^{k} \lambda_j f(x^{(j)}).$$

Las aplicaciones homogéneas lineales se llaman corrientemente *operadores lineales*. Del operador lineal $f: R^n \to R^m$ se dice que él actúa de R^n en R^m.

De la definición del operador lineal se desprende inmediatamente que la composición $g \circ f$ de operadores lineales $f: R^n \to R^m$ y $g: R^m \to R^s$ es también un operador lineal $g \circ f: R^n \to R^s$.

Sea f: $R^n \to R^m$ un operador lineal. La imagen de todo vector coordenado $e_j \in R^n$, $j = 1, 2, \dots, n$, es, en la aplicación f, un vector del espacio R^m y, por ende, se descompone según los vectores coordenados $\varepsilon_i \in R^m$, $i = 1, 2, \dots, m$. Designemos los coeficientes de esta descomposición con a_{ij}:

$$f(e_j) = \sum_{i=1}^{m} a_{ij}\, \varepsilon_i.$$

Sea $y = f(x)$, $x = \sum_{j=1}^{n} x_j e_j$ e

$$y = \sum_{i=1}^{m} y_i \varepsilon_i. \tag{41.31}$$

Por ser lineal la aplicación f, obtenemos

$$y = f(x) = f\left(\sum_{j=1}^{n} x_j e_j\right) = \sum_{j=1}^{n} x_j f(e_j) =$$

$$= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\varepsilon_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j\right)\varepsilon_i. \qquad (41.32)$$

Al comparar los coeficientes de la descomposición del vector y según los vectores coordenados $\varepsilon_1, \dots, \varepsilon_m$ en (41.31) y (41.32), obtendremos

$$\begin{aligned} y_1 &= a_{11}x_1 + \dots + a_{1n}x_n \\ &\dots\dots\dots\dots\dots \\ y_m &= a_{m1}x_1 + \dots + a_{mn}x_n. \end{aligned} \qquad (41.33)$$

Viceversa, es fácil comprobar que toda aplicación $f: R^n \to R^m$, cuyas funciones coordenadas tienen la forma (41.33), es un operador lineal.

La matriz

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \qquad (41.34)$$

se llama *matriz del operador lineal f*.

Evidentemente, si (41.34) es una matriz del operator lineal f, para cualquier $x = \sum_{j=1}^{n} x_j e_j$ tiene lugar (véase (41.32)) la descomposición

$$f(x) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j\right)\varepsilon_i. \qquad (41.35)$$

Ejemplo. Sea π_i un operador de proyección sobre el i-ésimo eje coordenado, es decir,

$$\pi_i(y) = \pi_i(y_1, \dots, y_m) = y_i \qquad (41.36)$$

(i es un número fijo entre los números $1, 2, \dots, m$). En este caso π_i es un operador lineal con la matriz cuadrada de orden m compuesta sólo por ceros salvo el i-ésimo elemento de la diagonal principal el cual es igual a uno:

$$\begin{pmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & \dots & \dots & \dots & & \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & \dots & \dots & \dots & & \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Con ayuda de los operadores de proyección π_i, $i = 1, 2, \dots, n$, se establece fácilmente la relación entre una aplicación vectorial arbitraria $f: X \to R^m$, $X \subset R^n$ y sus funciones coordenadas f_i (véase (41.30)):

$$f_i = \pi_i \circ f, \tag{41.37}$$

es decir, toda función coordenada f_i, $i = 1, 2, \dots, m$, es una composición de la aplicación f con el operador de proyección π_i.

Si $m = 1$, es decir, el operador lineal $f: R^n \to R$ aplica el espacio R^n en el conjunto de todos los números reales, se denomina corrientemente *funcional lineal*.

En virtud de (41.33), toda funcional lineal tiene por expresión

$$y = a_1x_1 + \dots + a_nx_n, \tag{41.38}$$

donde $a_1, \dots, a_n$ son ciertos números reales.

Al designar con a un vector de coordenadas $(a_1, \dots, a_n)$, llegamos a que toda funcional lineal $f: R^n \to R$ tiene por expresión

$$f(x) = (a, x),$$

donde con (a, x) está designado el producto escalar de los vectores a y x. Es evidente también lo recíproco: toda aplicación de la forma $x \mapsto (a, x)$ es una funcional lineal $f: R^n \to R$.

Ejercicio 11. Establézcase cuáles de las aplicaciones a seguir son lineales:
a) $f: R^3 \to R^2$, siendo $f(x, y, z) = (x, z)$;
b) $f: R^4 \to R^4$, siendo $f(x) = -x$, donde x es un vector arbitrario en R^4;
c) $f: R^3 \to R^3$, siendo $f(x) = x + (0, -1, 0)$ donde x es un vector arbitrario en R^3;
d) $f: R^2 \to R^2$, seindo $f(x, y) = (2x + y, y)$;
e) $f: R^2 \to R^2$, siendo $f(x, y) = 2x, y - x)$;
f) $f: R^2 \to R^2$, siendo $f(x, y) = (y, x)$;
g) $f: R^2 \to R$, siendo $f(x, y) = xy$.

Recordemos las definiciones de algunas operaciones sobre matrices (conocidas del álgebra). Si $A = (a_{ij})$ y $B = (b_{ij})$ son matrices rectangulares con el mismo número de columnas y filas, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, la *suma* de ellas se determina como una matriz cuyo elemento c_{ij} es la suma de los elementos correspondientes de las matrices A y B, es decir,

$$c_{ij} \overset{\text{def}}{=} a_{ij} + b_{ij}, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n.$$

Se llama *producto de la matriz A* por el número λ una matriz cuyos elementos c_{ij} se obtienen todos de los elementos correspondientes de la matriz A multiplicándolos por λ

$$c_{ij} \overset{\text{def}}{=} \lambda a_{ij}, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n.$$

Si el número de las columnas en la matriz $A = (a_{ij})$ es igual al número de las filas de la matriz $B = (b_{jk})$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, $k = 1, 2, \dots, s$, entonces el *producto AB* de las matrices A y B se determina como una matriz compuesta de los elementos c_{ik}, que se determinan según las fórmulas:

$$c_{ik} \overset{\text{def}}{=} \sum_{j=1}^{n} a_{ij}b_{jk}, \quad i = 1, 2, \dots, m, \quad k = 1, 2, \dots, s.$$

He aquí dos propiedades de los operadores lineales las cuales nos harán falta en lo sucesivo.

1°. *Si f y g son operadores lineales, $f: R^n \to R^m$, $g: R^n \to R^m$ y λ y μ, números arbitrarios, entonces $\lambda f + \mu g$ es también un operador lineal que actúa de R^n en R^m, con la particularidad de que si A y B son las matrices de los operadores lineales f y g, la suma $\lambda A + \mu B$ será la matriz del operador $\lambda f + \mu g$.*

La demostración de esta afirmación se efectúa por comprobación inmediata: si

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, \dots, m,$$

son funciones coordenadas de la aplicación f, mientras que

$$z_i = \sum_{j=1}^{n} b_{ij}x_j, \quad i = 1, 2, \dots, m,$$

son funciones coordenadas de la aplicación g, entonces para las funciones coordenadas de la aplicación $\lambda f + \mu g$ tendremos (al sumar y multiplicar por números los vectores, sus coordenadas se suman y se multiplican por los mismos números)

$$\lambda y_i + \mu z_i = \lambda \sum_{j=1}^{n} a_{ij}x_j + \mu \sum_{j=1}^{n} b_{ij}x_j = \sum_{j=1}^{n} (\lambda a_{ij} + \mu b_{ij})x_j,$$

es decir, primero, las funciones coordenadas de la aplicación $\lambda f + \mu g$ son funciones lineales y, segundo, los elementos c_{ij} de la matriz de la aplicación $\lambda f + \mu g$ los constituyen los números $c_{ij} = \lambda a_{ij} + \mu b_{ij}$, es decir, los elementos de la matriz $\lambda A + \mu B$, donde $A = (a_{ij})$, $B = (b_{ij})$. □

2°. *Si f y g son operadores lineales, $f: R^n \to R^m$, $g: R^m \to R^s$, su composición $g \circ f$ es también un operador lineal $R^n \to R^s$ y la matriz de la composición es igual al producto de matrices de las aplicaciones g y f.*

Realicemos nuevamente la comprobación inmediata de la afirmación. Si

$$y_i = \sum_{j=1}^{n} a_{ij}x_j, \quad i = 1, 2, \dots, m,$$

son funciones coordenadas de la aplicación f y

$$z_k = \sum_{i=1}^{m} b_{ki}y_i, \quad k = 1, 2, \dots, s,$$

son funciones coordenadas de la aplicación g, entonces

$$z_k = \sum_{i=1}^{m} b_{ki}y_i = \sum_{i=1}^{m} b_{ki} \sum_{j=1}^{n} a_{ij}x_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} b_{ki}a_{ij} \right) x_j,$$

es decir, primero, las funciones coordenadas de la composición $g \circ f$ son funciones lineales, y, segundo, los elementos c_{kj} de la matriz de la composición se obtienen de los elementos de las matrices a_{ij} y b_{ki} de los operadores f y g según la regla

$$c_{kj} = \sum_{i=1}^{m} b_{ki} a_{ij}. \tag{41.39}$$

Según se ha dicho, tal matriz (c_{kj}) se llama precisamente producto de las matrices (b_{ki}) y (a_{ij}). □

Observemos que todo operador lineal $f: R^n \to R^m$ es una aplicación continua del espacio R^n, dado que todas sus funciones coordenadas (41.33) son continuas, pues son lineales.

La longitud del vector $x \in R^n$, como se ha notado en el p. 41.5, es una función continua en el espacio R^n. Por esta razón, si $f: R^n \to R^m$ es un operador lineal, entonces la función $|f(x)|$, representando una composición de dos funciones continuas, será también continua en R^n.

Por cuanto la bola unidad $Q^n = \{x \in R^n : |x| \leqslant 1\}$ es un compacto, para todo operador lineal $f: R^n \to R^m$ la restricción de la función continua $|f|: R^n \to R$ a la bola Q^n, es decir, la función $|f|: Q^n \to R$, está acotada:

$$\sup_{|x| \leqslant 1} |f(x)| < +\infty. \tag{41.40}$$

Definición 7. *Para el operador lineal (en particular, para una funcional lineal, cuando $m = 1$) $f: R^n \to R^m$ el número $\sup\limits_{|x| \leqslant 1} |f(x)|$ lleva el nombre de norma*[*)] *del operador y se denota con $\|f\|$:*

$$\|f\| \overset{\text{def}}{=} \sup_{|x| \leqslant 1} |f(x)|. \tag{41.41}$$

En virtud de la desigualdad (41.40), la norma de cualquier operador lineal es finita.

Estimemos la longitud de la imagen del vector $x \in R^n$ en términos de la norma del operador f y la longitud x del propio vector. Para todo $x \neq 0, x \in R^n$, el vector $\xi = \dfrac{x}{|x|}$ tiene la longitud $1: |\xi| = \left|\dfrac{x}{|x|}\right| = \dfrac{1}{|x|}|x| = 1$. Por ello, al utilizar la linealidad del operador f, la propiedad (18.34) de la longitud del vector y la definición (41.41), obtendremos

$$|f(x)| = \left|f\left(|x|\frac{x}{|x|}\right)\right| = \left||x| f\left(\frac{x}{|x|}\right)\right| = |x|\left|f\left(\frac{x}{|x|}\right)\right| \leqslant$$
$$\leqslant |x| \sup_{|\xi| \leqslant 1} |f(\xi)| = |x| \, \|f\|,$$

es decir,

$$|f(x)| \leqslant \|f\| \, |x|. \tag{41.42}$$

---

*) La definición general de la norma se dará en el p. 57.3.

De esta desigualdad se infiere que para $|x| < 1$ es válida la desigualdad $|f(x)| < \|f\|$. Recordemos que una función, continua en el compacto, alcanza en éste su valor máximo (véase el teorema 3 en el p. 19.6). Por eso una función $|f| : Q^n \to \mathbf{R}$, siendo continua en el compacto Q^n, alcanza en éste su valor máximo

$$\|f\| \sup_{|x| \leqslant 1} |f(x)| = \max_{|x| \leqslant 1} |f(x)|,$$

y, como para $|x| < 1$ tiene lugar la desigualdad $|f(x)| < \|f\|$, el máximo mencionado se alcanza cuando $|x| = 1$, es decir, en la esfera unidad $S^{n-1} = \{x : |x| = 1\}$. De este modo

$$\|f\| = \max_{|x| = 1} |f(x)|. \tag{41.43}$$

Demos a conocer una expresión más para la norma de un operador lineal

$$\|f\| = \sup_{x \in R^n,\, x \neq 0} \frac{|f(x)|}{|x|}. \tag{41.44}$$

Demostrémosla. Empleando otra vez la propiedad de la longitud (18.34) de un vector, la linealidad de la aplicación f y la fórmula (41.41), obtenemos

$$\sup_{x \in R^n,\, x \neq 0} \frac{|f(x)|}{|x|} = \sup_{x \in R^n,\, x \neq 0} \left| \frac{1}{|x|} f(x) \right| =$$

$$= \sup_{x \in R^n,\, x \neq 0} \left| f\left(\frac{x}{|x|} \right) \right| = \sup_{|\xi| = 1} |f(\xi)| = \|f\|.$$

No es difícil estimar la norma $\|f\|$ del operador lineal f en términos de los elementos de su matriz (41.34). Fijándonos en que el cuadrado de longitud del vector es igual a la suma de cuadrados de sus coordenadas y aplicando la fórmula (41.35) y la desigualdad de Cauchy-Schwarz (18.2), tendremos

$$|f(x)|^2 = \left| \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right) e_i \right|^2 =$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right)^2 \leqslant \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}^2 \right) \left(\sum_{j=1}^{n} x_j^2 \right) =$$

$$= \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right) |x|^2.$$

De aquí, para todo $x \neq 0$, $x \in R^n$:

$$\frac{|f(x)|}{|x|} \leqslant \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}.$$

Por ello, en virtud de (41.44),

$$\|f\| = \sup_{x \in R^n,\, x \neq 0} \frac{|f(x)|}{|x|} \leqslant \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} \tag{41.45}$$

De este modo queda demostrada la desigualdad (41.40), esta vez, de "modo algebraico".

41.7. APLICACIONES DERIVABLES

Pasemos ahora a la definición de aplicaciones vectoriales derivables. Recordemos previamente que una función de n variables $f: X \to R, X \subset R^n$*), definida en el entorno del punto $x = (x_1, \ldots, x_n) \in X$, se llama *derivable* en este punto, si existen tales constantes $a_1, \ldots, a_n$ (son derivadas parciales de la función f en este punto: $a_i = \dfrac{\partial f}{\partial x_i}(x)$), que

$$f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, \ldots, x_n) =$$
$$= a_1 h_1 + \ldots + a_n h_n + o(h),\ h \to 0, \tag{41.46}$$

donde $h = (h_1, \ldots, h_n)$.

La aplicación lineal (funcional lineal)

$$(h_1, \ldots, h_n) \mapsto a_1 h_1 + \ldots + a_n h_n$$

en la fórmula (41.46) se denomina *diferencial de la función* f en el punto x. Al designarla mediante $D(x)$, obtenemos

$$D(x)(h) = a_1 h_1 + \ldots + a_n h_n.$$

De este modo, la definición de la derivabilidad (41.46) puede representarse en la forma

$$f(x + h) = f(x) + D(x)(h) + o(h), \quad h \to 0.$$

De modo análogo se determina, en el caso general, la derivabilidad de la aplicación.

Para las aplicaciones que se realizan de un espacio n-dimensional a un espacio m-dimensional "o pequeño" se determina de la manera siguiente: sea U un entorno del punto $x_0 \in X$, $\alpha: U \to R^m$; diremos que $\alpha = o(x)$ para $x \to x_0$, siempre que $|\alpha| = o(|x|)$, $x \to x_0$, es decir, si existe tal función $\varepsilon: U \to R$, que

$$|\alpha(x)| = \varepsilon(x)|x|, \tag{41.47}$$
$$x \in U \text{ y } \lim_{x \to x_0} \varepsilon(x) = 0.$$

Sin ninguna duda $|\alpha(x)|$ es la longitud de un vector en el espacio R^m y $|x|$ es la longitud del vector en el espacio R^n. Para las expresiones del tipo $o(x)$, que son vecto-

*) Con R se designa, como siempre, el conjunto de todos los números reales.

res, quedan vigentes las reglas habituales de operaciones con el símbolo "o pequeño", por ejemplo, $o(x) + o(x) = o(x)$ para $x \to x_0$, etc.

Definición 8. *Sea U un entorno del punto $x \in R^n$ en el espacio R^n. Un aplicación $f: U \to R^m$ se llama derivable en el punto x, si existe tal aplicación lineal (operador lineal) $l: R^n \to R^m$ que*

$$f(x + h) = f(x) + l(h) + o(h), \quad h \to 0, h \in R^n. \tag{41.48}$$

El operador lineal l se llama diferencial de la aplicación f en el punto x y se denota con $D(x)$ o, más detalladamente, con $D_f(x)$.

Haciendo uso de esta designación, la definición de la derivabilidad (41.48) puede escribirse en la forma

$$f(x + h) = f(x) + D_f(x)(h) + o(h), \quad h \to 0. \tag{41.49}$$

La matriz de la diferencial $D_f(x)$ (véase (41.34)) *se llama derivada de la aplicación f en el punto x y se designa mediante $f'(x)$.*

Observemos que de la fórmula (41.48) proviene inmediatamente que una aplicación, derivable en el punto x, es continua en él:

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

Teorema 3. *Si la aplicación $f: X \to R^m$, $X \subset R^n$, es derivable en el punto $x \in X$, entonces su diferencial en este punto se define unívocamente.*

Corolario. *La diferencial de una aplicación lineal coincide con la misma aplicación.*

DEMOSTRACIÓN DEL TEOREMA. Supongamos que a la par con la igualdad (41.48) se cumple también la igualdad

$$f(x + h) = f(x) + l_1(h) + o(h), \quad h \to 0, \tag{41.50}$$

donde $l_1: R^n \to R^m$, l_1 es un operador lineal. Restando una de estas igualdades de la otra, obtenemos

$$l(h) - l_1(h) = o(h) \quad \text{para} \quad h \to 0,$$

es decir, existe tal función $\varepsilon(h)$, definida en cierto entorno V del origen de coordenadas del espacio R^n, es decir, $\varepsilon: V \to R$, que $\lim\limits_{h \to 0} \varepsilon(h) = 0$ y que para todo $h \in V$ tiene lugar la igualdad

$$|l(h) - l_1(h)| = \varepsilon(h)\,|h|. \tag{41.51}$$

Tomemos ahora arbitrariamente $k \in R^n$; para todo t suficientemente pequeño tendremos $tk \in V$. Por eso, en (41.51) para tales t podemos poner $h = tk$:

$$|l(tk) - l_1(tk)| = \varepsilon(tk)\,|tk|.$$

Como $|tk| = |t|\,|k|$ y las aplicaciones l y l_1 son lineales, tendremos

$$l(tk) - l_1(tk) = t[l(k) - l_1(k)],$$

y, por ende,

$$|l(k) - l_1(k)| = \varepsilon(tk)\,|k|. \tag{41.52}$$

Pero $\lim_{t \to 0} tk = 0$, por lo cuàl, en virtud de la propiedad de la función ε, tenemos también $\lim_{t \to 0} \varepsilon(tk) = 0$. Pasando al límite para $t \to 0$ en (41.52), obtendremos $|l(k) - l_1(k)| = 0$, es decir, para cualquier $k \in R^n$

$$l(k) = l_1(k).$$

Esto precisamente indica que $l = l_1$. □

DEMOSTRACIÓN DEL COROLARIO. Sea $f: R^n \to R^m$ un operador lineal. Por ser lineal para cualesquiera $x \in R^n$ y $h \in R^n$, tenemos

$$f(x + h) = f(x) + f(h),$$

es decir, la igualdad (41.48) se cumple para $l = f$ y $o(h) \equiv 0$. Debido a la unicidad de la diferencial, $D_f(x) = f$. □

Teorema 4 ((**linealidad de la diferencial**). *Si las aplicaciones* $f: \mathbf{X} \to R^m$ *y* $g: X \to R^m$, $X \subset R^n$, *son derivables en el punto* $x \in X$, *entonces para cualesquiera números* λ *y* μ *la combinación lineal* $\lambda f + \mu g$ *será también derivable en el punto* x *y*

$$D_{\lambda f + \mu g}(x) = \lambda D_f(x) + \mu D_g(x).$$

DEMOSTRACIÓN. Debido a la derivabilidad de las aplicaciones f y g en el punto x, tenemos (véase (41.49)):

$$f(x + h) = f(x) + D_f(x)(h) + o(h), \quad h \to 0$$

$$g(x + h) = g(x) + D_g(x)(h) + o(h), \quad h \to 0;$$

de aquí

$$(\lambda f(x + h) + \mu g(x + h)) =$$
$$= [\lambda f(x) + \mu g(x)] + [\lambda D_f(x) + \mu D_g(x)]\,(h) + o(h), \quad h \to 0.$$

Como $\lambda D_f(x) + \mu D_g(x)$ es una aplicación lineal (véase el p. 41.6), entonces, de conformidad con la definición 8, la aplicación lineal $\lambda D_f(x) + \mu D_g(x)$ es la diferencial de la aplicación $\lambda f + \mu g$. □

Teorema 5. *Supongamos que* $X \subset R^n$, $D \subset R^m$, $f: X \to D$, $g: D \to R^r$, *con la particularidad de que la aplicación* f *es derivable en el punto* $x \in X$ *y* g, *en el punto* $f(x)$. *En este caso la composición* $g \circ f$ *es derivable en el punto* x *y su diferencial en el mismo es igual a la composición de diferenciales de las aplicaciones* f *y* g:

$$D_{g \circ f}(x) = D_g(f(x)) \circ D_f(x). \tag{41.53}$$

Corolario. *Cumplidas las condiciones del teorema, la derivada de la composición de aplicaciones es igual al producto de las derivadas:*

$$(g \circ f)'(x) = g'(f(x))\, f'(x). \tag{41.54}$$

Como se ve de las fórmulas aducidas, gracias a la elección adecuada de las definiciones y anotaciones, en las formulaciones de los teoremas tiene lugar la analogía completa con el caso unidimensional.

DEMOSTRACIÓN. Por ser derivable la aplicación f, tenemos

$$(g \circ f)(x + h) = g(f(x + h)) = g(f(x) + D_f(x)(h) + o(h)), \quad h \to 0. \tag{41.55}$$

De este modo, el argumento de la función g en el punto $y = f(x)$ ha recibido un incremento

$$k = D_f(x)(h) + o(h). \tag{41.56}$$

Por ello, de (41.55) tenemos en virtud de que la función g es derivable:

$$(g \circ f)(x + h) = g(y + k) = g(y) + D_g(y)(k) + o(k), \quad k \to 0. \tag{41.57}$$

Dado que (véase la desigualdad (41.42)).

$$|D_f(x)(h)| \leqslant \|D_f(x)\| \, |h|, \tag{41.58}$$

donde la norma $\|D_f(x)\|$ del operador lineal $D_f(x)$ es un número no negativo, entonces para la función $k = k(h)$, definida por la igualdad (41.56), obtendremos

$$\lim_{h \to 0} k = 0. \tag{41.59}$$

Más aún, resulta justa la estimación

$$|k| \leqslant \|D_f(x)\| \, |h| + |o(h)|, \quad h \to 0,$$

y como, para h suficientemente pequeños, tiene lugar la desigualdad $|o(h)| < |h|$, para h de este género es lícita también la estimación

$$|k| \leqslant (\|D_f(x)\| + 1) \, |h|. \tag{41.60}$$

Luego, de la definición de $o(k)$ (véase (41.47)) se desprende que existe una función $\varepsilon(k)$ tal que

$$\lim_{k \to 0} \varepsilon(k) = 0 \tag{41.61}$$

y $|o(k)| = \varepsilon(k) \, |k|$. Por esta razón, en virtud de (41.60), para h mencionados suficientemente pequeños, se verifica la desigualdad

$$|o(k)| = \varepsilon(k) \, |k| \leqslant \varepsilon(k)(\|D_f(x)\| + 1) \, |h|. \tag{41.62}$$

Por cuanto de (41.59) y (41.61) se infiere que $\lim\limits_{h \to 0} \varepsilon(k) = 0$, se tiene, por consiguiente,

$$\varepsilon(k)(\|D_f(x)\| + 1) \, |h| = o(h) \quad \text{para} \quad h \to 0,$$

entonces de (41.62) tenemos $|o(k)| \leqslant o(h)$, $h \to 0$, de donde

$$o(k) = o(h) \quad \text{para} \quad h \to 0.$$

Esto significa que la fórmula (41.57) puede escribirse en la forma

$$(g \circ f)(x + h) = g(y) + D_g(y)(k) + o(h), \quad h \to 0, \tag{41.63}$$

donde k se determina por la fórmula (41.56).

Consideremos ahora el sumando medio en el segundo miembro de la igualdad (41.63). Puesto que la aplicación $D_g(y)$ es lineal, tenemos

$$D_g(y)(k) = D_g(y)(D_f(x)(h) + o(h)) =$$
$$= D_g(y)(D_f(x)(h)) + D_g(y)o(h), \quad h \to 0. \tag{41.64}$$

Debido a la desigualdad (41.42), tendremos $|D_g(y)o(h)| \leqslant \|D_g(y)\| \, |o(h)|$, por lo cual

$$D_g(y)o(h) = o(h), \quad h \to 0;$$

por consiguiente, de (41.64) obtendremos:

$$D_g(y)(k) = D_g(y)(D_f(x)(h)) + o(h) = (D_g(y) \circ D_f(x))(h) + o(h), \quad h \to 0.$$

Sustituyendo la expresión obtenida para $D_g(y)$ en (41.63) y tomando en consideración que $y = f(x)$, tendremos en definitiva

$$(g \circ f)(x + h) = g(f(x)) + (D_g(f(x)) \circ D_f(x))(h) + o(h), \quad h \to 0.$$

Como la composición de operadores lineales es un operador lineal, entonces, por la unicidad de la diferencial, el operador $D_g(f(x)) \circ D_f(x)$ es la diferencial de la composición $f \circ g$, es decir, la fórmula (41.53) queda demostrada.

La fórmula (41.54) se deduce de ésta en seguida, puesto que en la composición de los operadores lineales sus matrices se multiplican. □

Teorema 6. *La aplicación $f = (f_1, \dots, f_m)$: $X \to R^m$, $X \subset R^n$, es derivable en el punto $x \in X$, si, y sólo si, todas las funciones coordenadas suyas f_i: $X \to R$, $i = 1, 2, \dots, m$, son derivables en dicho punto. En este caso los elementos a_{ij} de la matriz de la diferencial $D_f(x)$ son derivadas parciales correspondientes de las funciones coordenadas:*

$$a_{ij} = \frac{\partial f_i(x)}{\partial x_j}, \, i = 1, 2, \dots, m, \, j = 1, 2, \dots, n.$$

En otras palabras, la derivada $f'(x)$ es la matriz de Jacobi del sistema de funciones f_i (véase la definición 2 del p. 41.3),

$$f'(x) = \begin{pmatrix} \dfrac{\partial f_1(x)}{\partial x_1} & \dots & \dfrac{\partial f_1(x)}{\partial x_n} \\ \dots & \dots & \dots \\ \dfrac{\partial f_m(x)}{\partial x_1} & \dots & \dfrac{\partial f_m(x)}{\partial x_n} \end{pmatrix} \tag{41.65}$$

y se llama también *matriz de Jacobi de la aplicación* f en el punto x.

DEMOSTRACIÓN 1. Las funciones coordenadas $f_i = \pi_i \circ f$ (véase (41.37)) son una composición de dos aplicaciones derivables: la aplicación f, que es derivable en el punto f por hipótesis, y la proyección π_i (véase (41.36)) que es derivable en todo el espacio R^m como cualquier operador lineal $R^m \to R$. Por consiguiente, de acuerdo con el teorema 5, las funciones f_i, $i = 1, 2, \dots, m$, son derivables en el punto x.

2. Supongamos que todas las funciones coordenadas $f_i = \pi_i \circ f$ de la aplicación son derivables en el punto x. Teniendo presente (41.46), esto implica que existen tales constantes a_{ij}, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$ que

$$f_i(x + h) = f_i(x) + a_{i1}h_1 + \dots + a_{in}h_n + o(h), \, h \to 0, \, i = 1, 2, \dots, m. \tag{41.66}$$

De aquí, como se sabe (véase (20.2)), proviene que los coeficientes a_{ij} de los incrementos h_j que adquieren los argumentos x_j son las derivadas parciales correspondientes de las funciones f_i:

$$a_{ij} = \frac{\partial f_i(x)}{\partial x_j}, \, i = 1, 2, \dots, m, \, j = 1, 2, \dots, n. \tag{41.67}$$

Designemos mediante l: $R^n \to R^m$ un operador lineal con la matriz (a_{ij}). Por cuanto $(o(h), \ldots, o(h)) = o(h)$ *), las igualdades (41.66) pueden escribirse en la forma

$$f(x + h) = f(x) + l(h) + o(h), \quad h \to 0.$$

Esto precisamente significa la derivabilidad de la aplicación f, además de (41.67) se infiere la validez de la fórmula (41.65). □

OBSERVACIÓN 1. En vista de las fórmulas (41.54) y (41.65), el corolario del teorema 5 significa que la matriz de Jacobi de la composición de aplicaciones f y g es igual al producto de las matrices de Jacobi de las aplicaciones mencionadas.

Esto, por otra parte, se deduce también directamente de la fórmula de derivación de una función compuesta: si $z_k = g_k(y_1, \ldots, y_m)$, $k = 1, 2, \ldots, s$, e $y_i = f_i(x_1, \ldots, x_n)$, $i = 1, 2, \ldots, m$, entonces (véase (20.26))

$$\frac{\partial z_k}{\partial x_j} = \sum_{i=1}^{m} \frac{\partial z_k}{\partial y_i} \frac{\partial y_i}{\partial x_j}, \; k = 1, 2, \ldots, s, \; j = 1, 2, \ldots, n,$$

que, de acuerdo con la regla de multiplicación de las matrices (véase el p. 41.6) significa que la matriz $\left(\frac{\partial z_k}{\partial x_j}\right)$ es un producto de las matrices $\left(\frac{\partial z_k}{\partial y_i}\right)$ y $\left(\frac{\partial y_i}{\partial x_j}\right)$:

$$\left(\frac{\partial z_k}{\partial x_j}\right) = \left(\frac{\partial z_k}{\partial y_i}\right) \left(\frac{\partial y_i}{\partial x_j}\right).$$

Definición 9. *En el caso cuando $m = n$, el determinante*

$$\det \left(\frac{\partial f_i(x)}{\partial x_j}\right)$$

de la matriz de Jacobi (41.65) *se llama determinante de Jacobi o jacobiano de la aplicación $f: X \to R^n$, $X \subset R^n$, en el punto $x \in X$ y se denota* (véase el p. 41.3)

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \quad o \quad \frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)}.$$

OBSERVACIÓN 2. Por el curso de álgebra conocemos que al multiplicar las matrices cuadradas. sus determinantes se multiplican, razón por la cual, cumplidas las condiciones del teorema 5 en el caso de $m = n = s$, el jacobiano de la composición de las aplicaciones f y g es igual al producto de jacobianos de las aplicaciones f y g

*) La notación $(o(h), \ldots o(h)) = o(h)$ significa que el vector cuyas coordenadas son infinitésimos de orden superior a h, es de por sí un infinitésimo de orden más elevado que h cuando $h \to 0$. La coincidencia que en el caso dado tiene lugar entre las designaciones del vector y de sus coordenadas se debe a que para el vector n-dimensional se ha elegido, para no complicar la notación, la designación x en la que no está reflejada la dimensión del vector. Ésta se pone de manifiesto sólo cuando el vector se escribe con ayuda de las coordenadas $x = (x_1, \ldots, x_n)$.

$$\frac{\partial(z_1, \dots, z_n)}{\partial(x_1, \dots, x_n)} = \frac{\partial(z_1, \dots, z_n)}{\partial(y_1, \dots, y_n)} \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}. \tag{41.68}$$

En efecto,

$$\frac{\partial(z_1, \dots, z_n)}{\partial(x_1, \dots, x_n)} = \det\left(\frac{\partial z_k}{\partial x_j}\right) = \det\left(\frac{\partial z_k}{\partial y_i}\right)\left(\frac{\partial y_i}{\partial x_j}\right) =$$

$$= \det\left(\frac{\partial z_k}{\partial y_i}\right)\det\left(\frac{\partial y_i}{\partial x_j}\right) = \frac{\partial(z_1, \dots, z_n)}{\partial(y_1, \dots, y_n)} \quad \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}.$$

OBSERVACIÓN 3. Supongamos que $X \subset R^n$ e Id: $X \to X$ es una aplicación idéntica del conjunto X sobre sí mismo. En la forma coordenada se escribe como la condición de igualdad de las cordenadas de los puntos correspondientes en esta aplicación a la imagen v la preimagen, es decir, las funciones coordenadas tienen por expresión

$$f_i(x) = x_i,\ i = 1, 2, \dots, n,\ (x_1, \dots, x_n) \in X.$$

Si $x^{(0)}$ es un punto interior del conjunto X, estas funciones pueden derivarse en dicho punto y, puesto que $\dfrac{\partial x_i}{\partial x_j} = 0$ para $i \neq j$ y $\dfrac{\partial x_i}{\partial x_i} = 1$, entonces la matriz de Jacobi de la aplicación idéntica es una matriz unidad

$$X = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}.$$

Supongamos ahora que $U \subset R^n$, $V \subset R^n$ y f: $U \to V$ es una aplicación biunívoca (inyectiva), mientras que f^{-1}: $f(U) \to U$ es la aplicación inversa a la primera. En este caso, para todo punto $x \in U$ tenemos $f^{-1}(f(x)) = x$, es decir, la composición $f^{-1} \circ f$ es una aplicación idéntica.

Supongamos que la aplicación f es derivable en el punto $x_0 \in U$ (por tanto, x_0 es un punto interior del conjunto U, pues solo para tales puntos viene determinada la noción de derivabilidad), mientras que la aplicación inversa f^{-1} es derivable en el punto $f(x_0)$. Como $f^{-1} \circ f$ es un aplicación idéntica, en vista de la fórmula (41.54) tenemos

$$(f^{-1})'f' = (f^{-1} \circ f)' = (\mathrm{Id})' = X. \tag{41.69}$$

Pasando de esta igualdad de las matrices a los jacobianos de éstas, obtendremos

$$\det(f^{-1})' \det f' = 1, \tag{41.70}$$

ya que $\det X = 1$.

Si la aplicación f viene dada por las funciones coordenadas (41.30), la fórmula (41.70) puede escribirse en la forma

$$\frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)} = 1. \tag{41.71}$$

De esta fórmula se deduce que con las suposiciones admitidas tanto el jacobiano de la aplicación f en el punto x, como el de la aplicación inversa f^{-1} en el punto $f(x)$ no se anulan.

Escribamos la fórmula (41.71) en otra forma

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(y_1, \ldots, y_n)} = \frac{1}{\dfrac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}}. \tag{41.72}$$

Dicha fórmula representa una generalización evidente de la fórmula para la derivada de la función inversa de una sola variable: $\dfrac{dx}{dy} = \dfrac{1}{\dfrac{dy}{dx}}$.

Como conclusión enunciaremos dos definiciones útiles.

Definición 10. *Una aplicación $f: X \to R^m$, $X \subset R^n$, derivable en todo punto $x \in X$, se denomina aplicación derivable del conjunto X.*

Obviamente, si una aplicación es derivable en el conjunto X, entonces, cualquiera que sea el punto $x \in X$, de acuerdo con la definición 8, la aplicación f queda definida en cierto entorno del punto, es decir, X es un conjunto abierto.

De conformidad con el teorema 6, la aplicación $f = (f_1, \ldots, f_n)$ es derivable en el conjunto X cuando, y sólo cuando, son derivables en dicho conjunto todas las funciones coordenadas suyas $f_1, \ldots, f_n$. Si todas las funciones coordenadas son continuamente derivables en X, es decir, si todas sus primeras derivadas parciales son continuas en X, entonces la aplicación f se llama *aplicación continuamente derivable del conjunto X*.

Definición 11. *Una aplicación homeomorfa $f: G \to D$, donde G y D son conjuntos abiertos del espacio R^n, se llama difeomorfa o difeomorfismo, si son derivables tanto la propia aplicación citada como la aplicación inversa $f^{-1}: D \to G$.*

41.8. APLICACIONES CON UN JACOBIANO DISTINTO DE CERO. PRINCIPIO DE CONSERVACIÓN DE LA REGIÓN

Consideraremos ante todo la cuestión acerca de la existencia de una aplicación inversa a la dada. Como se sabe, en el caso de $n = 1$, para una función continuamente derivable en cierto segmento, la condición de que la derivada de ésta no se reduzca a cero (la que conlleva a su monotonía estricta) resulta suficiente para que exista una función unívoca continuamente derivable que sea inversa respecto de la dada. Cuando n es arbitrario, el caso se complica considerablemente: las correspondientes condiciones puntuales impuestas sobre las propiedades de derivabilidad de la aplicación sólo permiten afirmar que la aplicación inversa existe localmente, es decir, en un entorno del punto. Con más precisión, queda válido el siguiente teorema.

Teorema 7. *Sea*

$$y = f(x) = \begin{cases} y_1 = f_1(x_1, \ldots, x_n), \\ \dots\dots\dots\dots\dots\dots \\ y_n = f_n(x_1, \ldots, x_n) \end{cases} \tag{41.73}$$

una aplicación continuamente derivable del conjunto abierto $G \subset R^n$ en el espacio R^n. Si el jacobiano de esta aplicación no se anula en el punto $x^{(0)} \in G$, existen, pues, tales entornos U_x y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)} = f(x^{(0)})$, que $f(x)$, $x \in U_x$, es la aplicación biunívoca del entorno U_x sobre el entorno U_y, y la aplicación inversa es continuamente derivable en el conjunto U_y.

Corolario. *Sea f una aplicación continuamente derivable del conjunto abierto $G \subset R^n$ en el espacio R^n. Si el jacobiano de la aplicación f no es nulo en G, la imagen del conjunto G en esta aplicación es también un conjunto abierto.*

DEMOSTRACIÓN. Examinemos las funciones

$$F_i(x, y) = f_i(x_1, \dots, x_n) - y_i, \quad i = 1, 2, \dots, n.$$

Están definidas para todo $y = (y_1, \dots, y_n) \in R^n_y$ y para todo $x = (x_1, \dots, x_n) \in G \subset R^n_x$. Con su ayuda el sistema de ecuaciones (41.73) que prefijan la aplicación f, se escribirá en la forma

$$F_i(x, y) = 0, \quad i = 1, 2, \dots, n. \tag{41.74}$$

Las funciones $F_i(x, y)$ están definidas y son continuamente derivables en cierto entorno del punto $(x^{(0)}, y^{(0)})$ (como tal entorno puede ser tomado, por ejemplo, $G \times R^n_y$),

$$F_i(x^{(0)}, y^{(0)}) = 0 \text{ y } \left.\frac{\partial(F_1, \dots, F_n)}{\partial(x_1, \dots, x_n)}\right|_{(x^{(0)}, y^{(0)})} = \left.\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}\right|_{x^{(0)}} \neq 0.$$

De este modo, resultan cumplidas todas las condiciones del teorema 2 de este párrafo sobre la resolubilidad del sistema de ecuaciones.

En virtud de este teorema, las ecuaciones (41.74), o, que es lo mismo, el sistema (41.73) pueden ser resueltos y, además, de un modo único, respecto de las variables $x_1, \dots, x_n$ en cierto entorno del punto $(x^{(0)}, y^{(0)})$. Más detalladamente esto significa que existen tales entornos U^*_x y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)}$, $x^{(0)} \in U^*_x$, $y^{(0)} \in U_y$, y tal aplicación única

$$x = g(y) = \begin{cases} x_1 = g_1(y_1, \dots, y_n), \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ x_n = g_n(y_1, \dots, y_n) \end{cases} \tag{41.75}$$

que aplica el entorno U_y en el entorno U^*_x, que para todo $y \in U_y$ tiene lugar la identidad

$$f[g(y)] = y.$$

En otras palabras, para todo punto $y \in U_y$ existe, y además, el único punto $x = g(y) \in U^*_x$ el que, al realizarse la aplicación f, pasa al punto y. De este modo, $x \in f^{-1}(y) \cap U^*_x$ y $g(y)$ es una aplicación unívoca, continuamente derivable e inversa respecto de f en U_y: $g = f^{-1}$.

Pongamos $U_x = U^*_x \cap f^{-1}(U_y)$. Entonces, U_x es un conjunto abierto, pues representa una intersección de dos conjuntos abiertos U^*_x y $f^{-1}(U_y)$ (el carácter abierto del conjunto $f^{-1}(U_y)$ se desprende de que es una preimagen del conjunto abierto U_y en la aplicación continua f, véase el lema 2 en el p. 41.4). Es evidente que

U_x se aplica biunívocamente sobre U_y y, dado que $x^{(0)} \in U_x^*$ y $f(x^{(0)}) = y^{(0)} \in U_y$, entonces $x^{(0)} \in U_x$, es decir, U_x es el entorno buscado del punto $x^{(0)}$. □

OBSERVACIÓN 1. Los entornos U_x y U_y que figuran en las condiciones del teorema 4 poseen, además, la propiedad adicional de que el jacobiano de la aplicación f del entorno U_x sobre el entorno U_y no se anula en el entorno U_x, mientras que el jacobiano de la aplicación inversa f^{-1} no se anula en el entorno U_y. Esto proviene inmediatamente de la fórmula (41.71). En efecto, en vista de que la aplicación f transforma biunívocamente el entorno U_x en el U_y y que f y f^{-1} son continuamente derivables, la fórmula mencionada puede emplearse para el caso de la aplicación f, considerada sobre el conjunto U_x. Conforme a esta fórmula, un producto de los jacobianos de las aplicaciones f y f^{-1} es igual a uno y, por lo tanto, cada uno de ellos no es igual a cero.

DEMOSTRACIÓN DEL COROLARIO. Supongamos que $y = f(x)$ es una aplicación continuamente derivable del conjunto abierto G en el espacio R^n y que $y^{(0)}$ es un punto arbitrario del conjunto $f(G)$. Elijamos un punto $x^{(0)}$ en la preimagen del punto $y^{(0)}$: $x^{(0)} \in f^{-1}(y^{(0)})$. Por consiguiente, $f(x^{(0)}) = y^{(0)}$. En virtud del teorema 4, existen tales entornos $U_x \subset G$ y U_y de los puntos respectivos $x^{(0)}$ e $y^{(0)}$ que $f(U_x) = U_y$. Por consiguiente, $U_y \subset f(G)$. En otras palabras, para todo punto $y_{(0)} \in f(G)$ existe su entorno contenido en el conjunto $f(G)$. De este modo, cualquier punto del conjunto $f(G)$ es interior para este conjunto lo que es un indicio de que $f(g)$ es un conjunto abierto. □

OBSERVACIÓN 2. Si en cierta aplicación f para los puntos $x^{(0)}$ e $y^{(0)} = f(x^{(0)})$ existen unos entornos respectivos U_x y U_y que se aplican biunívocamente por f uno sobre el otro, suele decirse que *la aplicación f es localmente biunívoca en el punto $x^{(0)}$*.

Si en este caso la aplicación f es continua en U_x y f^{-1} es continua en U_y, entonces f se llama *aplicación localmente homeomorfa* en el punto $x^{(0)}$ u *homeomorfismo local*. Por fin, si el homeomorfismo local citado es un difeomorfismo, la aplicación en consideración recibe el nombre de *difeomorfismo local* en el punto dado (véanse las definiciones de homeomorfismo y difeomorfismo en el p. 41.4 y el p. 41.7).

Empleando la terminología mencionada, podemos decir que la aplicación f, considerada en el teorema 4, es una aplicación localmente difeomorfa en todo punto en el que el jacobiano es destinto de cero.

Teorema 8 (principio de conservación de la región). *La imagen de una región n-dimensional de un espacio n-dimensional en la aplicación continuamente derivada con un jacobiano que no se reduce a cero, es una región.*

DEMOSTRACIÓN. Sea G una región, $G \subset R^n$, e $y = f(x)$ una aplicación de G en R^n que satisface las condiciones del teorema. De acuerdo con el corolario del teorema 4, el conjunto $f(G)$ es abierto y, según el lema 7 del p. 41.4, es linealmente conexo. Por ello, si G es una región, el conjunto $f(G)$ será también una región, siempre que se cumplen las condiciones del teorema. □

Ejercicio 12. Constrúyase un ejemplo de una aplicación continuamente derivable de cierta región plana cuyo jacobiano nunca se reduce a cero y no es biunívoc

41.9. FUNCIONES IMPLÍCITAS DEFINIDAS POR UNA ECUACIÓN EN LA QUE SE TRASTORNAN LAS CONDICIONES DE UNICIDAD. PUNTOS SINGULARES DE LAS CURVAS PLANAS

Ya sabemos que si las coordenadas de cierto punto $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ satisfacen la ecuación

$$F(x_1, \dots, x_n) = 0 \tag{41.76}$$

y si en este punto la derivada $\frac{\partial F}{\partial x_i}$ no es nula, entonces bajo ciertas condiciones correspondientes impuestas sobre la continuidad de la misma función F y de la derivada citada, la ecuación (41.76) resulta resoluble en cierto entorno del punto $x^{(0)}$ respecto de x_i y la solución es una función continuamente derivable de las demás coordenadas.

Surge, naturalmente, la cuestión: ¿qué sucederá, cuando en el punto $x^{(0)}$ las derivadas parciales respecto de todos los argumentos se reduzcan a cero: definirá o no, en este caso, la ecuación (41.76) algunas funciones? Detendrémonos en esta cuestión, limitándonos a la consideración de un caso bidimensional, teniendo en cuenta la enorme dificultad con la que se resuelve.

Así pues, consideraremos una ecuación

$$F(x, y) = 0, \tag{41.77}$$

donde la función F está definida y es continuamente derivable en cierto entorno del punto (x_0, y_0) de tal género que

$$F(x_0, y_0) = 0. \tag{41.78}$$

Sea

$$F_x(x_0, y_0) = F_y(x_0, y_0) = 0 \tag{41.79}$$

Mostremos que incluso si se cumplen estas condiciones, la ecuación (41.77) se resuelve, a veces, en el entorno del punto (x_0, y_0) respecto de una de las variables, de suerte que se obtendrá una función continuamente derivable; no obstante, en el caso general, esto se puede hacer de una manera que no es única. De tal modo, la condición

$$F_x^2(x_0, y_0) + F_y^2(x_0, y_0) \neq 0, \tag{41.80}$$

que en nuestro caso no se cumple (véase (41.79)) y que permite aplicar el teorema 1 sobre las funciones implícitas a una de las variables, puede llamarse, naturalmente, condición de solubilidad unívoca de la ecuación (41.77).

Definición 12. *Un punto (x_0, y_0) cuyas coordenadas satisfacen las condiciones* (41.78) *y* (41.79) *se denomina punto singular de la ecuación* (41.77).

Un punto singular se llama aislado, si existe un entorno suyo en el cual el punto citado es el único punto singular.

En el lenguaje geométrico esto significa que si la ecuación (41.77) es una representación implícita de alguna curva, entonces en el entorno de los puntos singulares de esta ecuación la curva no representa, en el caso general, la gráfica de cierta función unívoca suave (lo que tiene lugar, si se cumplen las condiciones (41.80)); aquí pueden surgir diferentes singularidades que ahora serán el objeto de nuestra consideración.

Introduzcamos, para abreviar, las siguientes designaciones

$$F_{xx}(x_0, y_0) = F^0_{xx}, \quad F_{xy}(x_0, y_0) = F^0_{xy}, \quad F_{yy}(x_0, y_0) = F^0_{yy}.$$

Teorema 9. *Supongamos que la función $F(x, y)$ está definida y es dos veces continuamente derivable en cierto entorno de un punto singular aislado (x_0, y_0) de la ecuación* (41.77) *y sea*

$$F^0_{xx}F^0_{yy} - F^{0^2}_{xy} \neq 0.$$

En este caso, si

$$F^0_{xx}F^0_{yy} - F^{0^2}_{xy} > 0, \tag{41.81}$$

entonces (x_0, y_0) es una solución aislada de la ecuación (41.77), *es decir, existe un entorno del punto (x_0, y_0) tal que ningún punto de él, salvo (x_0, y_0) satisface la ecuación* (41.77); *en cambio, si*

$$F^0_{xx}F^0_{yy} - F^{0^2}_{xy} < 0, \tag{41.82}$$

la ecuación (41.77) *es resoluble en cierto entorno del punto (x_0, y_0) pero no unívocamente: hay dos diferentes funciones derivables que satisfacen la ecuación* (41.77). *Por ello (x_0, y_0) se llama en este caso punto doble.*

Por ejemplo, si

$$F^0_{yy} \neq 0, \tag{41.83}$$

existen dos funciones derivables $f_1(x)$ y $f_2(x)$, definidas en cierto entorno del punto x_0, tales que en dicho entorno $F(x, f_1(x)) = 0$, $F(x, f_2(x)) = 0$, con la particularidad de que $f_1(x_0) = f_2(x_0) = y_0$, mientras que las derivadas de las funciones $f_1(x)$ y $f_2(x)$ en el punto x_0 son diferentes raíces de la ecuación

$$F^0_{xx} + 2F^0_{xy}k + F^0_{yy}k^2 = 0^{*)}. \tag{41.84}$$

DEMOSTRACIÓN. Supongamos que se cumple la condición (41.81). Junto con (41.79) es suficiente para que exista en el punto (x_0, y_0) un extremo estricto de la función $F(x, y)$ (véase el teorema 3 en el p. 40.2). Por eso existe un entorno U del punto (x_0, y_0) tal que para $(x, y) \in U$ y $(x, y) \neq (x_0, y_0)$ o bien siempre $F(x, y) > F(x_0, y_0)$, o bien siempre $F(x, y) < F(x_0, y_0)$ y como $F(x_0, y_0) = 0$, se tiene $F(x, y) \neq 0$ para cualesquiera $(x, y) \in U$, $(x, y) \neq (x_0, y_0)$, es decir, (x_0, y_0) es la solución aislada de la ecuación (41.77)**).

Supongamos ahora que se cumple la condición (41.82). Desarrollemos la función $F(x, y)$, rigiéndonos por la fórmula de Taylor, en el entorno del punto (x_0, y_0)

---

*) Las raíces de esta ecuación son reales y diferentes, en virtud de las condiciones (41.82) y (41.83).

**) Para demostrar esta afirmación se usa no el hecho de que (x_0, y_0) es el punto singular aislado, sino que es sólo un punto simplemente singular en el que se cumple la condición (41.81).

hasta los sumandos de segundo orden. Tomando en consideración las condiciones (41.78) y (41.79), obtendremos

$$F(x, y) = \frac{1}{2}[F^0_{xx}(x - x_0)^2 + 2F^0_{xy}(x - x_0)(y - y_0) + F^0_{yy}(y - y_0)^2] + o(r^2), \tag{41.85}$$

donde $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$. Pongamos $x - x_0 = r \cos \varphi$, $y - y_0 = r \operatorname{sen} \varphi$. Evidentemente, (r, φ) son las coordenadas polares del punto (x, y), con la particularidad de que a título de origen del sistema polar de coordenadas se ha adoptado el punto (x_0, y_0).

En estas coordenadas

$$F(x, y) = \frac{r^2}{2}(F^0_{xx}\cos^2\varphi + 2F^0_{xy}\cos\varphi \operatorname{sen}\varphi + F^0_{yy}\operatorname{sen}^2\varphi) + o(r^2) =$$

$$= \frac{r^2}{2}P(\varphi) + o(r^2), \tag{41.86}$$

donde

$$P(\varphi) = F^0_{xx}\cos^2\varphi + 2F^0_{xy}\cos\varphi\operatorname{sen}\varphi + F^0_{yy}\operatorname{sen}^2\varphi, \tag{41.87}$$

o bien, para $\varphi \neq \frac{\pi}{2}(2k + 1)$, $k = 0, \pm 1, \pm 2, \ldots$,

$$P(\varphi) = \cos^2\varphi(F^0_{xx} + F^0_{xy}\operatorname{tg}\varphi + F^0_{yy}\operatorname{tg}^2\varphi). \tag{41.88}$$

Supongamos ahora que queda cumplida también la condición (41.83). Sean k_1 y k_2 las raíces de la ecuación (41.84) y sean $\varphi_1 = \operatorname{arctg} k_1$ y $\varphi_2 = \operatorname{arctg} k_2$. En este caso

$$\varphi_1 \neq \pm\pi/2, \quad \varphi_2 \neq \pm\pi/2, \tag{41.89}$$

y de (41.88) se infiere que

$$P(\varphi) = \cos^2\varphi(\operatorname{tg}\varphi - \operatorname{tg}\varphi_1)(\operatorname{tg}\varphi - \operatorname{tg}\varphi_2). \tag{41.90}$$

De la fórmula (41.90) se ve que la función $P(\varphi)$ se anula, para $\varphi \neq \frac{\pi}{2}(2k + 1)$, $k = 0, \pm 1, \pm 2, \ldots$, cuando $\varphi = \varphi_1 + k\pi$ y $\varphi = \varphi_2 + k\pi$, $k = 0, \pm 1, \pm 2, \ldots$, con la particularidad de que al pasar el argumento por estos valores, la función cambia de signo. Será cómodo para nosotros interpretar $P(\varphi)$ como función de un punto de la circunferencia C de radio igual a la unidad (este radio se elige a fin de simplificar el caso, para que las longitudes de los arcos coincidan con los ángulos φ) y centro en el punto (x_0, y_0).

Sea $\varepsilon > 0$. Designemos mediante $U_1 = U_1(\varepsilon)$ un ángulo abierto determinado por la desigualdad $\varphi_1 - \varepsilon < \varphi < \varphi_1 + \varepsilon$, es decir,

$$U_1 = \{(r, \varphi): \varphi_1 - \varepsilon < \varphi < \varphi_1 + \varepsilon\},$$

y pongamos correspondientemente

$$U_2 = \{(r, \varphi): \varphi_2 - \varepsilon < \varphi < \varphi_2 + \varepsilon\};$$

eligiendo, además, $\varepsilon > 0$ tan pequeño que U_1 y U_2 no se corten y no contengan en sí los semiejes de las ordenadas y, por consiguiente, las semirrectas verticales en general (lo último siempre puede conseguirse en vista de las condiciones (41.89)).

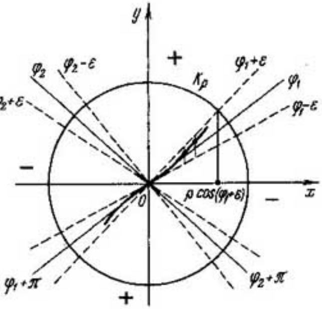

Fig. 166

Sean U_1^* y U_2^* los ángulos centrales simétricos con U_1 y U_2 respecto del punto (x_0, y_0):

$$U_1^* = \{(r, \varphi): \varphi_1 + \pi - \varepsilon < \varphi < \varphi_1 + \pi + \varepsilon\},$$
$$U_2^* = \{(r, \varphi): \varphi_2 + \pi - \varepsilon < \varphi < \varphi_2 + \pi + \varepsilon\}.$$

Por la elección del número ε los conjuntos U_1, U_2, U_1^* y U_2^* no se cortan dos a dos (fig. 166).

Examinemos ahora $P(\varphi)$ como función de un punto de la circunferencia mencionada C. Para simplificar, designemos también mediante φ un punto de la circunferencia C, al cual corresonde el ángulo polar φ. Eliminemos en la circunferencia C los intervalos con centros en los puntos φ_1, φ_2, $\varphi_1 + \pi$ y $\varphi_2 + \pi$ de longitud $2\,\varepsilon$*); por la elección adecuada de $\varepsilon > 0$, estos intervalos no tienen puntos comunes. El conjunto restante, que se designará con B, es limitado y cerrado y, por lo tanto, representa un compacto. En B la función $P(\varphi)$ es continua y no se anula, a consecuencia de lo cual

$$\inf_{\varphi \in B} |P(\varphi)| = \mu > 0. \tag{41.91}$$

Designemos mediante K_ρ un círculo cerrado con centro en el punto (x_0, y_0) y radio ρ:

$$K_\rho = \{(r, \varphi): 0 \leqslant r \leqslant \rho\},$$

y mediante L_ρ, un conjunto que se obtiene por sustracción (en el sentido de la teoría de conjuntos, véase el p. 1.1) de los conjuntos U_1, U_2, U_1^* y U_2^* del círculo K_ρ. Es ob-

---

*) Se llama intervalo de longitud $2\,\varepsilon$ en la circunferencia con centro en un punto, cuyo ángulo polar es igual a φ_0, un conjunto de sus puntos cuyos ángulos polares φ satisfacen la desigualdad $\varphi_0 - \varepsilon < \varphi < \varphi_0 + \varepsilon$.

vio que debido a (41.91) se tiene

$$\inf_{(r,\,\varphi)\in L_\rho} |P(\varphi)| = \mu > 0.$$

Ahora, observando que de (41.86) se deduce

$$F(x, y) = \frac{r^2}{2}[P(\varphi) + \alpha(r, \varphi)], \tag{41.92}$$

donde $\lim\limits_{r\to 0} \alpha(r, \varphi) = 0$, elijamos $\rho > 0$ de modo tal que para $r \leqslant \rho$ se verifique la desigualdad

$$|\alpha(r, \varphi)| < \mu. \tag{41.93}$$

Entonces, de (41.92) se desprende que para todos los puntos $(r, \varphi) \in L_\rho$ la expresión que figura en el segundo miembro de la fórmula (41.92) es del mismo signo que $P(\varphi)$.

El conjunto L_ρ se compone de cuatro sectores cerrados (véase fig. 166), en cada uno de los cuales, a excepción de su centro, la función $P(\varphi)$ y, por ende, debido a la elección adecuada de ρ, la función $F(x, y)$ adquieren valores de un mismo signo, y en los sectores vecinos, de signos contrarios.

Examinemos ahora el ángulo $U_1 = U_1(\varepsilon)$. Sea, para concretar, $0 \leqslant \varphi_1 < \pi/2$. La intersección de la clausura $\overline{U}_1$ del ángulo U_1 con la recta vertical $x = x^*$, $x_0 < x^* \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon)$, representa un segmento en cuyos extremos superior e inferior la función $F(x^*, y)$ toma los valores de signos diferentes. La función $F(x^*, y)$ considerada como función de una sola variable y con x^* fijado, es continua en el segmento mencionado y por eso se anula en cierto punto y^* del segmento, es decir, para todo x^*, donde $x_0 < x^* \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon)$, existe por lo menos un punto y^* tal que

$$F(x^*, y^*) = 0, \quad (x^*, y^*) \in U_1(\varepsilon) \cap K_\rho. \tag{41.94}$$

Definamos $y = f_1(x)$ como una función que pone en correspondencia al número x^* el número y^*:

$$f_1(x^*) = y^*, \quad x_0 < x^* \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon).$$

Mostremos que para ε y ρ suficientemente pequeños la función f_1 está definida unívocamente, es decir, existen tales $\varepsilon > 0$ y $\rho > 0$ que con x^* dado las condiciones (41.94) definen unívocamente y^*. Admitamos lo contrario. Tomemos las sucesiones $\varepsilon_n \to 0$ y $\rho_n \to 0$, para $n \to \infty$. En este caso existen dos sucesiones de los puntos que tienen abscisas iguales x_n y ordenadas distintas y'_n e y''_n, tales que

$$(x_n, y'_n) \in U_1(\varepsilon_n) \cap K_{\rho_n}, \quad F(x_n, y'_n) = 0,$$
$$(x_n, y''_n) \in U_1(\varepsilon_n) \cap K_{\rho_n}, \quad F(x_n, y''_n) = 0.$$

En virtud del teorema de Rolle, en el segmento $[y'_n, y''_n]$ de la recta $x = x_n$ existe tal punto y_n que

$$F_y(x_n, y_n) = 0. \tag{41.95}$$

siendo, además, en este caso $(x_n, y_n) \in U_1(\varepsilon_n) \cap K_{\rho_n}$; por hipótesis (véase (41.79)) teníamos

$$F_y(x_0, y_0) = 0. \tag{41.96}$$

Según la fórmula de incrementos finitos aplicada a la función $F_y(x, y)$ se tiene

$$F_y(x_n, y_n) - F_y(x_0, y_0) = F_{yx}(\xi_n, \eta_n)(x_n - x_0) + F_{yy}(\xi_n, \eta_n)(y_n - y_0),$$
$$(\xi_n, \eta_n) \in U_1(\varepsilon_n) \cap K_{\rho_n},$$

de donde, en virtud de (41.95) y (41.96),

$$F_{xy}(\xi_n, \eta_n) + F_{yy}(\xi_n, \eta_n)\frac{y_n - y_0}{x_n - x_0} = 0. \tag{41.97}$$

Sea $(x_n, y_n) = (r_n, \psi_n)$. Evidentemente, $|\psi_n - \varphi_1| < \varepsilon_n$; y por esta razón de la condición $\varepsilon_n \to 0$ se deduce que $\psi_n \to \varphi_1$ cuando $n \to \infty$, y como $\operatorname{tg}\psi_n = \dfrac{y_n - y_0}{x_n - x_0}$, entonces

$$\lim_{n \to \infty} \frac{y_n - y_0}{x_n - x_0} = \operatorname{tg}\varphi_1 = k_1. \tag{41.98}$$

Pasando al límite en la igualdad (41.97) para $n \to \infty$, en vista de (41.98) tenemos

$$F^0_{xy} + F^0_{yy}k_1 = 0, \quad \text{es decir,} \quad k_1 = -\frac{F^0_{xy}}{F^0_{yy}};$$

sustituyendo este valor de la raíz en la ecuación (41.84), obtenemos

$$F^0_{xx}F^0_{yy} - F^{0\,2}_{xy} = 0,$$

que contradice la condición (41.82).

Así pues, la función $y = f_1(x)$ se define realmente de modo unívoco cuando ε y ρ son suficientemente pequeños. En lo que sigue supondremos que ε y ρ se han elegido precisamente de la manera indicada.

Definamos adicionalmente la función f_1 en el punto x_0, al poner $y_0 = f_1(x_0)$. Obviamente, por la propia definición de la función $f_1(x)$, tenemos

$$F(x, f_1(x)) = 0, \quad x_0 \leqslant x \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon).$$

Mostremos que en el punto x_0 la función $f_1(x)$ tiene una derivada a la derecha y que esta derivada es igual a k_1. Sea $\varepsilon > 0$ arbitrariamente fijado. De lo dicho más arriba se desprende la existencia de tal $\rho = \rho(\varepsilon) > 0$ que la parte correspondiente de la gráfica de la función $f_1(x)$ se dispone íntegramente en $U_1(\varepsilon) \cap K_\rho$: .

$$(x, f_1(x)) \in U_1(\varepsilon) \cap K_\rho, \quad x_0 \leqslant x \leqslant x_0 + \rho\cos(\varphi_1 + \varepsilon) \tag{41.99}$$

Tomemos $\delta = \rho\cos(\varphi_1 + \varepsilon)$ y sea x tal que $0 < x - x_0 < \delta$, $y = f_1(x)$ y $(x, y) = (r, \varphi)$. En virtud de (41.99), tenemos $|\varphi - \varphi_1| < \varepsilon$. Esto significa que $\lim\limits_{x \to x_0 + 0} \varphi = \varphi_1$, y, por lo tanto, $\lim\limits_{x \to x_0 + 0} \operatorname{tg}\varphi = \operatorname{tg}\varphi_1$. Puesto que $\operatorname{tg}\varphi = \dfrac{y - y_0}{x - x_0}$, de lo demostrado se deduce que

$$\lim_{x \to x_0 + 0} \frac{f_1(x) - f_1(x_0)}{x - x_0} = \lim_{x \to x_0 + 0} \frac{y - y_0}{x - x_0} = \operatorname{tg}\varphi_1,$$

es decir, la función $f_1(x)$ tiene derivada a la derecha en el punto x_0 y esta derivada es igual a $\operatorname{tg}\varphi_1 = k_1$.

De esta misma manera se demuestra, analizando el comportamiento de la función $F(x, y)$ en el ángulo U_1^*, que en el segmento $[x_0 - \delta', x_0]$ existe, para cierto $\delta' > 0$, tal función $f_1(x)$ que siendo $x_0 - \delta' \leqslant x \leqslant x_0$, se tiene

$$F(x, f_1(x)) = 0, \quad (x, f_1(x)) \in U^*_1, \quad f_1'(x_0) = k_1$$

(por derivada se entiende en el caso dado la derivada a la izquierda).

Si ρ se toma tan pequeño que en un entorno circular de radio ρ del punto (x_0, y_0) no se contengan otros puntos singulares de la ecuación (41.77), a excepción de (x_0, y_0), la función $f_1(x)$ será derivable en todo punto $x \neq x_0$. Esto se deduce en seguida del teorema sobre las funciones implícitas que fue demostrado anteriormente (véase el teorema 1 en el p. 41.1). De resultas hemos obtenido la función $f_1(x)$ que está definida en cierto entorno del punto x_0 y que posee todas las propiedades requeridas.

Análogamente se demuestra la existencia de la función $f_2(x)$ que también es una solución de la ecuación (41.77) y satisface las condiciones del teorema, con la particularidad de que la gráfica de esta función pasa en los ángulos U_2 y U_2^* y por el punto (x_0, y_0).

Si $F^0_{yy} = 0$ y $F^0_{xx} \neq 0$, entonces todos los razonamientos quedan los mismos; sólo conviene cambiar de lugar los ejes Ox y Oy, de suerte que obtendremos, como resultado, las soluciones de la ecuación (41.77) en forma de las funciones de la variable y: $f_1(y)$ y $f_2(y)$.

Si, por fin, $F^0_{xx} = F^0_{yy} = 0$ y, por lo tanto, $F^0_{xy} \neq 0$, entonces resulta más conveniente realizar el cambio de variables: $x = \xi + \eta$ e $y = \xi - \eta$ (gírense los ejes de coordenadas al ángulo $\pi/4$). En este caso (de lo que es fácil convencerse por derivabilidad directa)

$$F^0_{\xi\xi} = -F^0_{\eta\eta} = 2F^0_{xy} \neq 0, \; F^0_{\xi\eta} = 0,$$

es decir, en el nuevo sistema de coordenadas se obtendrá el caso ya estudiado. En particular, la ecuación (41.84) para los coeficientes angulares de las tangentes en el punto singular tiene, en el sistema de coordenadas ξ, η, la forma siguiente

$$k^2 - 1 = 0,$$

y, por consiguiente, $k_{1,2} = \pm 1$. En otras palabras, las bisectrices de los ángulos coordenados, que representan los ejes en el sistema antiguo de coordenadas x, y, son las tangentes a las gráficas de dos funciones definidas por la ecuación (41.77) en cierto entorno del punto singular que se considera. □

Si la ecuación $F(x, y) = 0$ es una representación implícita de alguna curva, en el punto singular (x_0, y_0) de esta ecuación la curva puede tener (aunque no sea necesariamente) ciertas singularidades, es decir, en el entorno del punto singular de dicha ecuación la curva no es, en el caso general, la gráfica de cierta función unívoca suave.

Cabe recordar también que un conjunto de los puntos cuyas coordenadas satisfacen la ecuación (41.77) no es (en el caso general) siempre una curva en el sentido de la definición de la curva dada paramétricamente, enunciada anteriormente (véase el p. 16.2*).

Ejemplos. 1. Sea dada una ecuación $y^2(x^2 + y^2 + 1) = 0$. Aquí, $F(x, y) = y^2(x^2 + y^2 + 1)$, por lo cual $F_x = 2xy^2$, $F_y = 2x^2y + 4y^3 + 2y$. Las condiciones para la presencia de un punto singular (41.78) y (41.79) nos dan en este caso

$$x_0 = 0, y_0 = 0.$$

De este modo, de punto singular sirve el punto (0, 0). No obstante, en este punto la curva, definida por la ecuación, no tiene singularidad, puesto que la ecuación (el factor $x^2 + y^2 + 1$ nunca se reduce a cero) es equivalente a la $y = 0$ y la curva en consideración representa una gráfica de la función explícita $y = f(x) \equiv 0$. Observemos que en el punto (0, 0) para este caso

$$F_{xx}F_{yy} - F_{xy}^2 = 0, \tag{41.100}$$

de lo que podemos convencernos con facilidad.

2. Para la ecuación

$$(x^2 + y^2)(x^2 + y^2 - 1) = 0, \tag{41.101}$$

las condiciones (41.79) se convierten en el siguiente sistema de ecuaciones:

$$2x^3 + 2xy^2 - x = 0,$$
$$2y^3 + 2x^2y - y = 0.$$

Al sumar y restar estas ecuaciones, obtenemos el sistema

$$(x + y)(2x^2 + 2y^2 - 1) = 0$$
$$(x - y)(2x^2 + 2y^2 - 1) = 0.$$

De aquí, o bien $x = y = 0$, o bien $2x^2 + 2y^2 - 1 = 0$, sin embargo, el punto (x, y) cuyas coordenadas satisfacen la última correlación no es una raíz de la ecuación (41.101) (para este punto $x^2 + y^2 = \frac{1}{2}$, y, por lo tanto, en el primer miembro no hay factor que se reduzca a cero).

De este modo, el único punto singular es (0, 0). Es fácil comprobar que aquí se cumple la condición (41.81) y, por ende, el punto (0, 0) es una raíz aislada de la ecuación (41.101). Desde el punto de vista geométrico, como se ve en seguida, la ecuación (41.101) define la circunferencia unidad y el centro de ésta (0, 0) (este conjunto no porta, evidentemente, ninguna curva dada paramétricamente en el sentido del p. 16.2°).

3. Para la ecuación

$$x^3 + y^3 - 3axy = 0, \tag{41.102}$$

las condiciones (41.79) para la existencia de un punto singular conducen al sistema de ecuaciones

$$x^2 - ay = 0,$$
$$y^2 - ax = 0,$$

de donde o bien $x = y = 0$, y en este caso dicho punto satisface la ecuación (41.102), o bien $x = a, y = a$, pero las coordenadas de este punto no constituyen la solución de la ecuación (41.102). Aquí, otra vez, (0, 0) es el único punto singular.

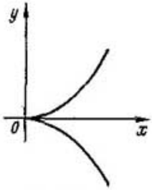

Fig. 167

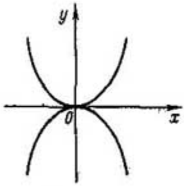

Fig. 168

No es difícil convencerse de que en el caso dado se cumplen las condiciones (41.82) y, por consiguiente, (0, 0) es un punto doble.

Geométricamente, para una curva cuya representación explícita es la ecuación (41.102) (se llama folio de Descartes y con ella ya chocamos en el p. 14.5) el punto (0, 0) es un *punto múltiple* (véase la fig. 70 en el tomo I).

4. Para la ecuación

$$y^2 - x^3 = 0, \tag{41.103}$$

(0, 0) es un punto singular; en dicho punto ya se cumple la condición (41.100) y, consecuentemente, en este caso no se cumplen las condiciones del teorema 6. Geométricamente, una curva expresada por la ecuación (41.103) y llamada parábola semicúbica $y = \pm x^{3/2}$, tiene en el punto (0, 0) una tangente y se dispone en el entorno de este punto por un lado respecto de la normal.

Los puntos de esta índole se llaman *puntos de retroceso* (fig. 167).

5. Para la ecuación

$$y^2 - x^4 = 0 \tag{41.104}$$

el punto (0, 0) es también singular y aquí se cumple nuevamente la condición (41.100). La ecuación (41.104) se descompone, evidentemente, en dos ecuaciones: $y = x^2$ e $y = -x^2$ que definen dos parábolas y estas últimas tienen en el punto (0, 0) una tangente común.

Los puntos singulares, en cierto entorno de los cuales la ecuación (41.77) define dos curvas continuamente derivables que tienen en el punto (x_0, y_0) una tangente común, se llaman *puntos autoadherentes* (fig. 168) de estas dos curvas.

Puede ocurrir que, cumplida la condición (41.100), el punto singular resulte ser la solución aislada de la ecuación (41.77) o su punto doble.

Como conclusión, ofrecemos algunas explicaciones para la ecuación (41.84). Si (x_0, y_0) es un punto singular de la ecuación (41.77), entonces, realizado el traslado paralelo del origen de coordenadas al punto (x_0, y_0), la ecuación (41.77) adquiere la forma

$$F_{xx}^0 x^2 + 2F_{xy}^0 xy + F_{yy}^0 y^2 + o(x^2 + y^2) = 0, \tag{41.105}$$

(aquí, con x e y están designadas las coordenadas del punto en el nuevo sistema de coordenadas y por índice 0, puesto por arriba, se designan los valores de las derivadas parciales en el punto (0,0) de este sistema), de donde, con una exactitud de hasta los infinitésimos de orden superior, nuestra ecuación puede escribirse del modo siguiente:

$$F^0_{xx}x^2 + 2F^0_{xy}xy + F^0_{yy}y^2 = 0. \tag{41.106}$$

Si se cumple la condición (41.82), el primer miembro de la ecuación (41.106) se descompone en dos factores reales, cada uno de los cuales, siendo igualado a cero, proporciona las tangentes a las dos ramas de la curva en el punto (0, 0) (véase (41.84)). Cuando se cumple la condición (41.81), el primer miembro de la ecuación (41.106) se descompone en dos factores complejos: "las tangentes son imaginarias". Esto es natural, puesto que no hay sentido en hablar aquí de una tangente, pues en el caso dado el punto singular es aislado.

Esta observación es de comodidad especial cuando se trata de determinar el carácter del punto singular en el caso de una curva algebraica, es decir, de la curva definida por la ecuación

$$P(x, y) = 0, \tag{41.107}$$

donde $P(x, y)$ es un polinomio de dos variables x e y. Si (0, 0) es un punto singular de esta ecuación, de las condiciones (41.78) y (41.79) se deduce que dicho polinomio no contiene término independiente ni tampoco los términos de primer orden, es decir, la ecuación (41.107) tiene por expresión

$$ax^2 + 2bxy + cy^2 + Q(x, y) = 0,$$

donde $Q(x, y)$ es un polinomio, todos los términos del cual son por lo menos de tercer orden. El comportamiento de las soluciones de esta ecuación se determina por su parte principal, es decir, por la ecuación

$$ax^2 + 2bxy + cy^2 = 0,$$

que para el caso dado es precisamente la ecuación (41.106), pues aquí se ve claramente que

$$a = F^0_{xx}, \quad b = F^0_{xy} \quad \text{y} \quad c = F^0_{yy}.$$

En cambio, si el punto (0, 0) satisface la ecuación (41.107), pero no es un punto singular, entonces (41.107) tiene por expresión

$$Ax + By + R(x, y) = 0, \quad A^2 + B^2 > 0,$$

donde $R(x, y)$ es un polinomio, todos los términos del cual tienen el orden no inferior al segundo. Del teorema sobre las funciones implícitas (véase el teorema 1 en el p. 41.1) se deduce que la ecuación

$$Ax + By = 0$$

es en este caso una ecuación de la tangente en el punto (0, 0) a la gráfica de la solución de la ecuación (41.107).

Ejercicios. Investíguese el comportamiento de cada una de las siguientes curvas en el entorno de sus puntos singulares; hállense las tangentes en el punto singular.

13. $y^2 = x^2 + x^3$. 14. $y^2 = x^2 - x^3$.

15. $y^2 = x^2 - x^4$.
16. $(x^2 - 9)y^2 = x^4$.
17. $y^2 = x(x - 3)^2$.
18. $y(y - 2)^2 = x^2$.
19. $4y^2 = x^5 + 5x^4$.
20. $(x^2 - y^2)y = x^4$.
21. $(y - x^2)^2 = x^5$.

41.10. CAMBIO DE VARIABLES

En las diversas cuestiones del análisis matemático y en las aplicaciones de éste, al estudiar tal o cual fórmula que contiene unas funciones y sus derivadas (ordinarias o parciales), resulta a menudo conveniente pasar a otras variables independientes y, a veces, a otras funciones ligadas por relaciones determinadas con las funciones que figuran en la fórmula que se considera. Todas estas transformaciones se realizan sobre la base de las reglas de derivación de las funciones implícitas y compuestas. He aquí unos ejemplos.

Sea $u = u(x, y)$. Transformemos las expresiones

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 \text{ y } \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$

reduciéndolas a las coordenadas polares r y φ. La primera de estas expresiones es el cuadrado de la longitud del gradiente ∇u de la función u, es decir, es igual a $|\nabla u|^2$, mientras que la segunda expresión tiene una designación especial Δu:

$$|\nabla u|^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2, \tag{41.108}$$

$$\Delta u \overset{\text{def}}{=} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. \tag{41.109}$$

El símbolo Δ, que es indicio de la aplicación a la función u de la operación (41.109), lleva el nombre de *operador de Laplace**).

De las fórmulas que ligan las coordenadas cartesianas con las polares

$$x = r\cos\varphi, y = r\operatorname{sen}\varphi, \tag{41.110}$$

encontramos

$$\frac{\partial x}{\partial r} = \cos\varphi, \frac{\partial x}{\partial \varphi} = -r\operatorname{sen}\varphi, \frac{\partial y}{\partial r} = \operatorname{sen}\varphi, \frac{\partial y}{\partial \varphi} = r\cos\varphi. \tag{41,111}$$

Apliquemos las fórmulas de derivación de una función compuesta:

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial u}{\partial x}\cos\varphi + \frac{\partial u}{\partial y}\operatorname{sen}\varphi,$$

$$\frac{\partial u}{\partial \varphi} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial \varphi} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \varphi} = -\frac{\partial u}{\partial x} r\operatorname{sen}\varphi + \frac{\partial u}{\partial y} r\cos\varphi.$$

*) P. Laplace (1749 — 1827), mecánico y matemático francés.

Resolvamos estas igualdades respecto de $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$:

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}, \frac{\partial u}{\partial y} = \frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r} \tag{41.112}$$

y sustituyamos las expresiones obtenidas en (41.108):

$$|\nabla u|^2 = \left(\frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)^2 + \left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)^2 = \\ = \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \varphi}\right)^2.$$

Pasemos ahora al cálculo de la expresión (41.109). Derivemos las fórmulas (41.110) respecto de x y luego respecto de y:

$$\left.\begin{aligned} 1 &= \cos\varphi\frac{\partial r}{\partial x} - r\operatorname{sen}\varphi\frac{\partial \varphi}{\partial x}, \\ 0 &= \operatorname{sen}\varphi\frac{\partial r}{\partial x} + r\cos\varphi\frac{\partial \varphi}{\partial x}, \end{aligned}\right\} \quad \left.\begin{aligned} 0 &= \cos\varphi\frac{\partial r}{\partial y} - r\operatorname{sen}\varphi\frac{\partial \varphi}{\partial y}, \\ 1 &= \operatorname{sen}\varphi\frac{\partial r}{\partial y} + r\cos\varphi\frac{\partial \varphi}{\partial y}, \end{aligned}\right\}$$

Resolvamos los sistemas obtenidos respecto de $\frac{\partial r}{\partial x}$, $\frac{\partial \varphi}{\partial x}$, $\frac{\partial r}{\partial y}$ y $\frac{\partial \varphi}{\partial y}$;

$$\frac{\partial r}{\partial x} = \cos\varphi, \frac{\partial r}{\partial y} = \operatorname{sen}\varphi, \frac{\partial \varphi}{\partial x} = -\frac{\operatorname{sen}\varphi}{r}, \frac{\partial \varphi}{\partial y} = \frac{\cos\varphi}{r}. \tag{41.113}$$

Derivemos ahora las fórmulas (41.112) respecto de x e y; entonces, empleando (41.113), obtendremos

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial r}\left(\frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)\frac{\partial r}{\partial x} + \frac{\partial}{\partial \varphi}\left(\frac{\partial u}{\partial r}\cos\varphi - \right. \\ \left. - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)\frac{\partial \varphi}{\partial x} = \frac{\partial^2 u}{\partial r^2}\cos^2\varphi - \frac{2\cos\varphi\operatorname{sen}\varphi}{r}\frac{\partial^2 u}{\partial r\partial\varphi} + \\ + \frac{\operatorname{sen}^2\varphi}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\operatorname{sen}^2\varphi}{r}\frac{\partial u}{\partial r} + \frac{2\cos\varphi\operatorname{sen}\varphi}{r^2}\frac{\partial u}{\partial \varphi};$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{\partial}{\partial r}\left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)\frac{\partial r}{\partial y} + \\ + \frac{\partial}{\partial \varphi}\left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)\frac{\partial \varphi}{\partial y} = \frac{\partial^2 u}{\partial r^2}\operatorname{sen}^2\varphi + \\ + \frac{2\cos\varphi\operatorname{sen}\varphi}{r}\frac{\partial^2 u}{\partial r\partial\varphi} + \frac{\cos^2\varphi}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\cos^2\varphi}{r}\frac{\partial u}{\partial r} - \frac{2\cos\varphi\operatorname{sen}\varphi}{r^2}\frac{\partial u}{\partial \varphi}.$$

Al sustituir las expresiones obtenidas en (41.109), tendremos

$$\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{1}{r}\frac{\partial u}{\partial r}.$$

Cuando en la expresión a transformar figuran no una sola, sino varias derivadas de orden dado, resulta más cómodo emplear el método de cálculo de las diferenciales, en lugar de las derivadas. Por ejemplo, considerando x e y como variables independientes, obtendremos las expresiones para las diferenciales dr y $d\varphi$. De las fórmulas (41.110) tenemos

$$dx = \cos\varphi\, dr - r \operatorname{sen}\varphi\, d\varphi,\ dy = \operatorname{sen}\varphi\, dr + r\cos\varphi\, d\varphi,$$

de aquí

$$dr = \cos\varphi dx + \operatorname{sen}\varphi\, dy,\ d\varphi = -\frac{\operatorname{sen}\varphi}{r}dx + \frac{\cos\varphi}{r}dy \tag{41.114}$$

(observemos que de estas fórmulas se deducen inmediatamente también las fórmulas (41.113).

Para la función $u = u(x, y)$ tenemos

$$du = \frac{\partial u}{\partial r}dr + \frac{\partial u}{\partial \varphi}d\varphi =$$

$$= \left(\frac{\partial u}{\partial r}\cos\varphi - \frac{\partial u}{\partial \varphi}\frac{\operatorname{sen}\varphi}{r}\right)dx + \left(\frac{\partial u}{\partial r}\operatorname{sen}\varphi + \frac{\partial u}{\partial \varphi}\frac{\cos\varphi}{r}\right)dy. \tag{41.115}$$

En la expresión para la diferencial du los coeficientes de dx y dy son las derivadas $\frac{\partial u}{\partial x}$ y $\frac{\partial u}{\partial y}$, por lo cual, de (41.115) se obtienen las dos fórmulas (41.112). Hallemos, a continuación, las segundas diferenciales d^2r y $d^2\varphi$, haciendo uso de (41.114):

$$d^2r = -\operatorname{sen}\varphi\, d\varphi\, dx + \cos\varphi\, d\varphi\, dy =$$

$$= \frac{\operatorname{sen}^2\varphi\, dx^2 - 2\cos\varphi \operatorname{sen}\varphi\, dx\, dy + \cos^2\varphi\, dy^2}{r},$$

$$d^2\varphi = -\left(\frac{\cos\varphi}{r}dx + \frac{\operatorname{sen}\varphi}{r}dy\right)d\varphi + \left(\frac{\operatorname{sen}\varphi}{r^2}dx - \frac{\cos\varphi}{r^2}dy\right)dr =$$

$$= \frac{2\cos\varphi\operatorname{sen}\varphi\, dx^2 - 2(\cos^2\varphi - \operatorname{sen}^2\varphi)\, dx\, dy - 2\cos\varphi\operatorname{sen}\varphi\, dy^2}{r^2}.$$

Ahora, de (41.115) obtendremos para d^2u

$$d^2u = \frac{\partial^2 u}{\partial r^2}dr^2 + 2\frac{\partial^2 u}{\partial r\, d\varphi}dr\, d\varphi + \frac{\partial^2 u}{\partial \varphi^2}d\varphi^2 + \frac{\partial u}{\partial r}d^2r + \frac{\partial u}{\partial \varphi}d^2\varphi =$$

$$= \left(\cos^2\varphi\,\frac{\partial^2 u}{\partial r^2} - \frac{2\cos\varphi\,\mathrm{sen}\,\varphi}{r}\,\frac{\partial^2 u}{\partial r\partial\varphi} + \frac{\mathrm{sen}^2\varphi}{r^2}\,\frac{\partial^2 u}{\partial\varphi^2} + \right.$$

$$\left. + \frac{\mathrm{sen}^2\varphi}{r}\,\frac{\partial u}{\partial r} + \frac{2\cos\varphi\,\mathrm{sen}\,\varphi}{r^2}\,\frac{\partial u}{\partial\varphi}\right) dx^2 + 2(\dots)\,dx\,dy + (\dots)\,dy^2.$$

De aquí se obtienen las expresiones para las segundas derivadas $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x\partial y}$ y $\frac{\partial^2 u}{\partial y^2}$ que intervienen como los coeficientes de dx^2, $2dx\,dy$ y dy^2.

Los métodos análogos son aplicables, por supuesto, en el caso en que se realiza algún otro cambio de variables $x = x(u, v)$, $y = y(u, v)$, cuando se tienen derivadas de órdenes superiores y también cuando se trata de las funciones del mayor número de variables.

Ejercicio 22. Transfórmese la expresión $|\nabla u|^2$, donde $u = u(x, y)$ a las coordenadas ortogonales ξ, η, es decir, tales que

$$\frac{\partial x}{\partial\xi}\,\frac{\partial x}{\partial\eta} + \frac{\partial y}{\partial\xi}\,\frac{\partial y}{\partial\eta} = 0.$$

23. Transfórmese la ecuación $y'' - xy'^3 + e^y y'^3 = 0$, considerando y como una nueva variable independiente y x, como una función de y.

24. En la ecuación $\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} = 0$ pásese a las nuevas variables independientes $u = x + y$, $v = x - y$.

25. En la expresión $\frac{1}{2}z\left(\frac{1}{x^2}\frac{\partial^2 z}{\partial x^2} + \frac{1}{y^2}\frac{\partial^2 z}{\partial y^2}\right) + \frac{1}{2}\left[\left(\frac{1}{x}\frac{\partial z}{\partial x}\right)^2 + \left(\frac{1}{y}\frac{\partial z}{\partial y}\right)^2\right] - \frac{1}{2}z\left(\frac{1}{x^3}\frac{\partial z}{\partial x} + \frac{1}{y^3}\,\frac{\partial z}{\partial y}\right)$ pásese a las variables u, v, $w = w(u, v)$, si $u = x^2$, $v = y^2$, $w = z^2$ $\left(\text{respuesta: } \frac{\partial^2 w}{\partial u^2} + \frac{\partial^2 w}{\partial v^2}\right)$.

Problema 27. En un espacio n-dimensional transfórmese la expresión $|\nabla u|^2$, donde $u = u(x_1, \dots, x_n)$, reduciéndola a las coordenadas ortogonales $\xi_1, \dots, \xi_n$, es decir, a tales coordenadas que para $i \neq k$ se verifica la igualdad

$$\sum_{j=1}^{n} \frac{\partial x_j}{\partial\xi_i}\,\frac{\partial x_j}{\partial\xi_k} = 0, \quad i, k = 1, 2, \dots, n.$$

§ 42. DEPENDENCIA DE LAS FUNCIONES

42.1. CONCEPTO DE DEPENDENCIA DE LAS FUNCIONES. CONDICIÓN NECESARIA PARA LA DEPENDENCIA DE LAS FUNCIONES

Definición 1. *Sean dadas en el conjunto abierto $G \subset R^n$ las funciones continuamente derivables*

$$y_i = \varphi_i(x), i = 1, 2, \dots, m, \; x = (x_1, \dots, x_n) \in G. \tag{42.1}$$

Si existen un conjunto abierto D en el espacio $R^{m-1}_{y_1, \dots, y_{m-1}}$ y una función, continuamente derivable en D, $\Phi(y_1, \dots, y_{m-1})$ tales que en todo punto $x \in G$ se cumplen las condiciones $(\varphi_1(x), \dots, \varphi_{m-1}(x)) \in D$ y $\Phi(\varphi_1(x), \dots, \varphi_{m-1}(x)) = \varphi_m(x)$, entonces la función φ_m se llama dependiente en el conjunto G de las funciones $\varphi_1, \dots, \varphi_{m-1}$.

Definición 2. *Si entre las funciones del sistema* (42.1) *hay una función, dependiente de las demás en el conjunto G, dicho sistema se denomina dependiente en el conjunto G.*

Si ninguna función del sistema (42.1) *depende de las restantes en el conjunto G, este sistema se llama independiente en G.*

A veces, para abreviar, en lugar de la expresión "sistema de funciones dependiente (independiente)" diremos simplemente "funciones dependientes (independientes)".

En la cuestión sobre la dependencia del sistema de funciones (42.1) el papel fundamental lo desempeña la matriz de Jacobi de este sistema

$$\left\| \frac{\partial y_i}{\partial x_j} \right\|, \; i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n, \tag{42.2}$$

i es el número de la fila y j, el número de la columna.

Teorema 1 (condición necesaria para la dependencia de las funciones). *Supongamos que $m \leqslant n$ y el sistema de funciones* (42.1) *es dependiente en un conjunto abierto G. Entonces, en todo punto de este conjunto el rango de la matriz de Jacobi* (42.2)*) *del sistema citado es inferior a m.*

DEMOSTRACIÓN. Por hipótesis, el sistema de funciones (42.1) es dependiente en G, puesto que por lo menos una de estas funciones depende de las demás. Sea, para concretar, que φ_m depende de $\varphi_1, \dots, \varphi_{m-1}$:

$$\varphi_m(x) = \Phi(\varphi_1(x), \dots, \varphi_{m-1}(x)), \; x \in G,$$

donde Φ es una función continuamente derivable de $(m-1)$ argumentos $y_1, \dots, y_{m-1}$. De aquí

$$\frac{\partial y_m}{\partial x_j} = \sum_{i=1}^{m-1} \frac{\partial \Phi}{\partial y_i} \frac{\partial y_i}{\partial x_j} \text{ para cualquier } j = 1, 2, \dots, n.$$

Esta fórmula muestra que la m-ésima fila de la matriz de Jacobi (42.2) en todo punto $x \in G$ es una combinación lineal de las filas restantes de esta matriz y, por lo tanto, el rango de la matriz de Jacobi (42.2) es inferior a m en todo punto $x \in G$. □

Corolario. *Supongamos que $m = n$ y el sistema de funciones* (42.1) *es depen-*

---

*) Recordemos que se llama rango de una matriz el número máximo de sus filas linealmente independientes. Este número coincide con el orden máximo del menor de dicha matriz distinto de cero.

diente en G. En este caso su jacobiano $\dfrac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}$ *es igual a cero en todos los puntos del conjunto G.*

Corolario 2 (condiciones suficientes para la independencia de las funciones). *Supongamos que $m \leqslant n$ y el rango de la matriz de Jacobi* (42.2) *por lo menos en un punto del conjunto abierto G es igual a m. En este caso el sistema* (42.1) *es independiente en el conjunto G.*

El corolario 1 se obtiene inmediatamente del teorema demostrado cuando $m = n$.

El corolario 2 se demuestra con facilidad por reducción al absurdo.

Por cuanto las filas de la matriz de Jacobi (42.2) son coordenadas de los gradientes de la función (42.1), el teorema 1 puede parafrasearse del modo siguiente.

Si el sistema de funciones (42.1) *es dependiente en la región G, los gradientes $\nabla\varphi_1, \dots, \nabla\varphi_m$ de estas funciones son linealmente dependientes en todo punto de G.*

42.2. CONDICIONES SUFICIENTES PARA LA DEPENDENCIA DE LAS FUNCIONES

En este punto conservamos intactas las designaciones del párrafo anterior y supondremos, como hasta ahora, que las funciones (42.1) son continuamente derivables en el conjunto abierto $G \subset R^n$.

Teorema 2 (condiciones suficientes para la dependencia local de las funciones). *Supongamos que el rango de la matriz de Jacobi* (42.2) *del sistema de funciones* (42.1) *en cualquier punto del conjunto abierto G no es superior al número r, $r < m \leqslant \leqslant n$, y en cierto punto $x^{(0)} \in G$ es igual a r, en otras palabras, existen tales variables $x_{j_1}, \dots, x_{j_r}$ y las funciones $y_{i_1} = \varphi_{i_1}(x), \dots, y_{i_r} = \varphi_{i_r}(x)$ que*

$$\left.\frac{\partial(y_{i_1}, \dots, y_{i_r})}{\partial(x_{j_1}, \dots, x_{j_r})}\right|_{x^{(0)}} \neq 0. \tag{42.3}$$

En este caso todas r funciones, que figuran en la condición (42.3), *son independientes en el conjunto G y existe un entorno del punto $x^{(0)}$ tal que cualquiera de $m - r$ funciones restantes depende en dicho entorno de las r funciones mencionadas.*

DEMOSTRACIÓN. Supongamos, para simplificar la anotación, que la condición (42.3) tiene por expresión

$$\left.\frac{\partial(y_1, \dots, y_r)}{\partial(x_1, \dots, x_r)}\right|_{x^{(0)}} \neq 0 \tag{42.4}$$

(esto siempre puede lograrse, numerando, si se desea, las funciones y los argumentos del sistema (42.1) en el orden debido). De acuerdo con el corolario 2 del teorema 1 en el p. 42.1, las funciones $y_1, \dots, y_r$ son independientes en G.

Mostremos que cada una de las funciones restantes depende de ellas en cierto entorno del punto $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$. Sea $y_i^{(0)} = \varphi_i(x^{(0)})$, $i = 1, 2, \dots, m$. Examinemos el sistema de las primeras r funciones del sistema (42.1):

$$\begin{matrix} y_1 = \varphi_1(x_1, \dots, x_n) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ y_r = \varphi_r(x_1, \dots, y_n)\,. \end{matrix} \tag{42.5}$$

Elijamos ante todo tal η_0 que todo punto $x = (x_1, \dots, x_n)$ perteneciente a η_0-entorno cúbico del punto $x^{(0)}$, es decir, todo punto x, para el cual $|x_i - x_i^{(0)}| < \eta_0$, $i = 1, 2, \dots, n$, pertenezca al conjunto G: $x \in G$. Esto es siempre posible, pues el conjunto mencionado es abierto.

Luego, en virtud de la condición (42.4) y el teorema sobre las funciones implícitas (véase el p. 41.3), el sistema (42.5) es resoluble respecto de las variables $x_1, \dots, x_r$ en cierto entorno del punto $(x^{(0)}, y^{(0)})$:

$$\begin{aligned} x_1 &= f_1(y_1, \dots, y_r, x_{r+1}, \dots, x_n),\\ &\dots\dots\dots\dots\dots\dots\dots\dots\\ x_r &= f_r(y_1, \dots, y_r, x_{r+1}, \dots, x_n). \end{aligned} \tag{42.6}$$

Las funciones $f_1, \dots, f_r$ en este caso están definidas y son continuamente derivables en cierto entorno del punto $(y_1^{(0)}, \dots, y_r^{(0)}, x_{r+1}^{(0)}, \dots, x_n^{(0)})$.

En la forma más amplia (si a título de entorno se toman los entornos cúbicos) esto significa lo siguiente: pueden escogerse tales números $\delta > 0$ y $\eta > 0$ (con el fin de comodidad, pueden ser incluso inferiores a η_0: $\delta < \eta_0$, $\eta < \eta_0$) que si designamos con U el entorno cúbico del punto $(y_1^{(0)}, \dots, y_r^{(0)}, x_{r+1}^{(0)}, \dots, x_n^{(0)})$ prefijado por las desigualdades

$$|y_i - y_i^{(0)}| < \delta, \quad i = 1, 2, \dots, r, \quad |x_j - x_j^{(0)}| < \delta, \quad j = r + 1, \dots, n$$

entonces

1) en el entorno U las funciones f_k, $k = 1, 2, \dots, r$, están definidas y son continuamente derivables;

2) para todos los puntos $(y_1, \dots, y_r, x_{r+1}, \dots, x_n) \in U$ son válidas las desigualdades

$$|f_k(y_1, \dots, y_r, x_{r+1}, \dots, x_n) - x_k^{(0)}| < \eta, \quad k = 1, 2, \dots, r;$$

3) en el entorno U se verifican las igualdades

$$\varphi_i(f_1, \dots, f_r, x_{r+1}, \dots, x_n) = y_i, \quad i = 1, 2, \dots, r,$$

donde por f_k, $k = 1, \dots, r$, se entienden los segundos miembros de las igualdades (42.6).

Consideremos una composición de funciones (42.6) y $\varphi_{r+1}(x_1, \dots, x_n)$, es decir, una función

$$y_{r+1} = \varphi_{r+1}(f_1, \dots, f_r, x_{r+1}, \dots, x_n), \tag{42.7}$$

en la que $f_k = f_k(y_1, \dots, y_r, x_{r+1}, \dots, x_n)$, $k = 1, \dots, r$. Esta función compuesta está definida a ciencia cierta y es continuamente derivable en el entorno cúbico U, mencionado arriba, del punto

$$(y_1^{(0)}, \dots, y_r^{(0)}, x_{r+1}^{(0)}, \dots, x_n^{(0)}).$$

Probemos que, realmente, la función (42.7) en este entorno U no depende de las variables $x_{r+1}, \dots x_n$, es decir, no varía cuando cambian las variables citadas y, consecuentemente, es, de hecho, sólo función de las variables $y_1, \dots, y_r$. Con este objeto basta mostrar que para la función (42.7) en el entorno U se verifica la igualdad

$$\frac{\partial y_{r+1}}{\partial x_j} = 0, j = r + 1, \dots, n \tag{42.8}$$

(véase el p. 20.4 o la fórmula de los incrementos finitos de Lagrange en el p. 39.2, de la cual se deduce directamente la suficiencia de la condición (42.8) para que una función sea independiente de las variables $x_{r+1}, \dots, x_n$ en una región convexa y, por consiguiente, en un entorno cúbico).

Para demostrar la igualdad (42.8), fijemos uno de los valores $j (j = r + 1, \dots, n)$ y las coordenadas x_k cuyos índices k toman los valores $r + 1, \dots, j - 1, j + 1, \dots, n$; designémoslas mediante x_k^* y elijamos x_k^* de una manera tal que sea $|x_k^* - x_k^{(0)}| < \delta$, $k = r + 1, \dots, j - 1, j + 1, \dots, n$.

Consideraremos una aplicación

$$\begin{gathered} y_1 = y_1, \\ \dots\dots\dots \\ y_r = y_r, \\ y_{r+1} = \varphi_{r+1}(f_1^*, \dots, f_r^*, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*), \end{gathered} \tag{42.9}$$

donde $f_k^* = f_k(y_1, \dots, y_r, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*)$, del entorno cúbico $U^{(j)}$ del punto $(y_1^{(0)}, \dots, y_r^{(0)}, x_j^{(0)})$ que se define por las desigualdades

$$|y_k - y_k^{(0)}| < \delta, \ k = 1, 2, \dots, r, \ |x_j - x_j^{(0)}| < \delta.$$

Para subrayar cuáles son precisamente las variables que cambian, representemos simbólicamente la aplicación (42.9) en la forma

$$(y_1, \dots, y_r, x_j) \to (y_1, \dots, y_r, y_{r+1}).$$

Esta aplicación es continuamente derivable en $U^{(j)}$; su matriz de Jacobi tiene por expresión

$$\begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial y_{r+1}}{\partial y_1} & \dfrac{\partial y_{r+1}}{\partial y_2} & \dfrac{\partial y_{r+1}}{\partial y_r} & \dfrac{\partial y_{r+1}}{\partial x_j} \end{Vmatrix}$$

y, por esta razón

$$\frac{\partial(y_1, \dots, y_r, y_{r+1})}{\partial(y_1, \dots, y_r, x_j)} = \frac{\partial y_{r+1}}{\partial x_j}, \tag{42.10}$$

es decir, el jacobiano de la aplicación en consideración es igual a la derivada que nos interesa.

En el entorno $U^{(j)}$ esta aplicación puede representarse como composición de dos aplicaciones: de la aplicación continuamente derivable

$$\begin{gathered} x_1 = f_1(y_1, \dots, y_r, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_r = f_r(y_1, \dots, y_r, x_{r+1}^*, \dots, x_{j-1}^*, x_j, x_{j+1}^*, \dots, x_n^*), \\ x_j = x_j \end{gathered}$$

del entorno $U^{(j)}$ y la aplicación continuamente derivable

$$y_1 = \varphi_1(x_1, \dots, x_r, x^*_{r+1}, \dots, x^*_{j-1}, x_j, x^*_{j+1}, \dots, x^*_n),$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$y_r = \varphi_r(x_1, \dots, x_r, x^*_{r+1}, \dots, x^*_{j-1}, x_j, x^*_{j+1}, \dots, x^*_n),$$
$$y_{r+1} = \varphi_{r+1}(x_1, \dots, x_r, x^*_{r+1}, \dots, x^*_{j-1}, x_j, x^*_{j+1}, \dots, x^*_n)$$

del entorno del punto $(x_1^{(0)}, \dots, x_r^{(0)}, x_j^{(0)})$ definido mediante las desigualdades

$$|x_i - x_i^{(0)}| < \eta,\ i = 1, 2, \dots, r,\ |x_j - x_j^{(0)}| < \delta.$$

Por ser adecuadamente elegidos los números δ y η, la composición de estas aplicaciones la que con los fines ilustrativos puede ser representada en la forma

$$(y_1, \dots, y_r, x_j) \to (x_1, \dots, x_r, x_j) \to (y_1, \dots, y_r, y_{r+1}).$$

está definida y es continuamente derivable en el entorno $U^{(j)}$. La primera de estas aplicaciones es continuamente derivable en el entorno $U^{(j)}$ del punto $(y_1^{(0)}, \dots, y_r^{(0)}, x_j^{(0)})$ y la segunda es continuamente derivable en el entorno correspondiente del punto $(x_1^{(0)}, \dots, x_r^{(0)}, x_j^{(0)})$. Por ello, de (42.10) y de las propiedades de los jacobianos de aplicaciones (véase el p. 41.7) tenemos

$$\frac{\partial y_{r+1}}{\partial x_j} = \frac{\partial(y_1, \dots, y_r, y_{r+1})}{\partial(y_1, \dots, y_r, x_j)} = \frac{\partial(y_1, \dots, y_r, y_{r+1})}{\partial(x_1, \dots, x_r, x_j)}\,\frac{\partial(x_1, \dots, x_r, x_j)}{\partial(y_1, \dots, y_r, x_j)}. \quad (42.11)$$

De acuerdo con la condición del teorema, el rango de la matriz de Jacobi en el conjunto G es igual o inferior a r, por consiguiente,

$$\frac{\partial(y_1, \dots, y_r, y_{r+1})}{\partial(x_1, \dots, x_r, x_j)} = 0$$

en todo punto de G. Por ello, de (42.11) se desprende en seguida que para cualquier punto $(y_1, \dots, y_r, x_j) \in U^{(j)}$ y, por ende, para todo punto

$$(y_1, \dots, y_r, x^*_{r+1}, \dots, x^*_{j-1}, x_j, x^*_{j+1}, \dots, x^*_n) \in U$$

se verifica la igualdad (42.8). Puesto que las coordenadas x^*_k se han fijado arbitrariamente, con tal de que sea $|x^*_k - x_k^{(0)}| < \delta$, $k = r+1, \dots, j-1, j+1, \dots, n$, esto quiere decir que la igualdad (42.8) es justa en todo el entorno U.

De este modo, la función (42.7) depende sólo de las variables $y_1, \dots, y_r$. Al designarla por el símbolo Φ, obtendremos

$$\varphi_{r+1}(f_1, \dots, f_r, x_{r+1}, \dots, x_n) = \Phi(y_1, \dots, y_r).$$

Elijamos ahora δ_0, $\delta_0 < \delta$ y $\delta_0 < \eta$, de una manera tal que para $|x_i - x_i^{(0)}| < \delta_0$, $i = 1, 2, \dots, n$, se cumplan las desigualdades

$$|y_i - y_i^{(0)}| < \delta, j = 1, 2, \dots, r.$$

Esto es posible, puesto que las funciones $y_j = \varphi_j(x_1, \dots, x_n)$, $j = 1, 2, \dots, r$, del sistema (42.5) son continuas en el punto $x^{(0)}$.

En virtud de lo demostrado, para cualquier punto x del entorno δ_0-cúbico del punto $x^{(0)}$, es decir, para todo punto $x = (x_1, \dots, x_n)$ tal que

$$|x_i - x_i^{(0)}| < \delta_0, \; i = 1, 2, \dots, n,$$

será válida la identidad

$$\varphi_{r+1}(x) = \Phi(\varphi_1(x), \dots, \varphi_r(x)),$$

es decir, en el entorno citado del punto $x^{(0)}$ las funciones $\varphi_1, \dots, \varphi_r, \varphi_{r+1}$ son dependientes.

Análogamente se demuestra también la dependencia de cada una de las funciones $\varphi_{r+2}, \dots, \varphi_m$ de $\varphi_1, \dots, \varphi_r$ en cierto entorno del punto $x^{(0)}$. □

Al igual que la condición necesaria para la dependencia de las funciones, las condiciones suficientes pueden enunciarse también en términos de los gradientes. Para simplificar, limitémonos al caso en que $r = m - 1$.

Si los gradientes $\nabla\varphi_1, \dots, \nabla\varphi_m$ son linealmente dependientes en todo punto de la región G, entonces, cualquiera que sea el punto $x \in G$ en el que $m - 1$ de los gradientes mencionados son linealmente independientes, existe un entorno suyo en el cual las funciones $\varphi_1, \dots, \varphi_m$ son dependientes. Además, si, por ejemplo, los gradientes $\nabla\varphi_1, \dots, \nabla\varphi_{m-1}$ son linealmente independientes en el punto dado y, por lo tanto, el gradiente $\nabla\varphi_m$ en él es una combinación lineal de ellos, entonces en el entorno mencionado la función φ_m depende de las funciones $\varphi_1, \dots, \varphi_{m-1}$.

Conviene fijar la atención en que las condiciones suficientes para la dependencia de las funciones, establecidas aquí, tienen un carácter local, a diferencia de los resultados del punto anterior que son de carácter global. Esto significa lo siguiente: si un sistema de m funciones continuamente derivables (42.1) es dependiente en el conjunto abierto $G \subset R^n$, entonces, de acuerdo con el teorema 1 del p. 42.1, en *todo* punto de dicho conjunto el rango de la matriz de Jacobi de este sistema es inferior a m (correspondientemente, si por lo menos en un punto del cunjunto G el rango de la matriz que se considera es igual a m, el sistema es independiente en *todo* el conjunto G). En cuanto al teorema 2 de este punto, se debe decir que éste sólo afirma que si en algún punto $x^{(0)} \in G$ se cumplen sus condiciones, el sistema dado de funciones será dependiente sólo en *cierto entorno* del punto citado (y no en todo el conjunto G). De este modo, efectivamente, la afirmación del teorema 2 tiene un carácter local.

Diremos, además, que si en *todo* punto $x^{(0)}$ del conjunto abierto G se cumplen las condiciones del teorema 2, entonces, naturalmente, en este caso, el sistema de funciones en consideración será dependiente en cierto entorno de todo punto. No obstante, el teorema 2 no garantiza que dicha dependencia será la misma en todos los entornos indicados, es decir, del teorema 2 no se infiere que en diferentes puntos unas mismas funciones serán dependientes de las otras y que las funciones Φ que "cristalizan en realidad" la dependencia de unas mismas funciones, consideradas en distintos entornos, serán coincidentes en los puntos de intersección de estos entornos. Por consiguiente, del teorema 2 no se infiere ni mucho menos que el sistema de funciones que satisface las condiciones de este teorema en todos los puntos $x^{(0)}$ del conjunto G será dependiente en todo el conjunto G en su total en el único sentido, es decir, en el sentido de la definición 1. Esto precisamente sirve de testimonio de que el teorema 2 no lleva el carácter global.

Observemos que existe un acceso algo más general al concepto de dependencia de las funciones que permite construir la teoría global de este problema, sin embargo no nos detendremos en ello.

Ejemplo. Examinemos un sistema de funciones

$$u = \operatorname{sen}(x + y)$$
$$v = \cos(x + y). \tag{42.12}$$

El jacobiano de este sistema es nulo en todo el plano

$$\begin{vmatrix} \cos(x+y) & \cos(x+y) \\ -\operatorname{sen}(x+y) & -\operatorname{sen}(x+y) \end{vmatrix} = 0,$$

y, como es fácil de ver, el rango de la matriz de Jacobi de este sistema es igual a uno en todos los puntos del plano.

Según el teorema 2, las funciones (42.12) son dependientes en el entorno de cada punto del plano. En el caso dado la dependencia de las funciones se halla fácilmente en la forma explícita, por ejemplo, en el conjunto abierto de los puntos (x, y), para los cuales $\cos(x + y) > 0$, ella puede ser dada por la fórmula $v = \sqrt{1 - u^2}$.

Ejercicios. 1. Sean $u = x^2 + y^2 + z^2$, $v = xy + yz + zx$, $w = x + y + z$. Demuéstrese que las funciones u, v, w son dependientes y hállese la ecuación que expresa su dependencia.

2. Investíguese la cuestión sobre la dependencia de las funciones $u = \xi^3 + \eta^3 + \zeta^3$, $v = \xi\eta\zeta$, $w = \xi^2 + \eta^2 + \zeta^2$, $z = \xi\eta + \eta\zeta + \zeta\xi$.

Problema 28. Una función $u = u(x, y)$ se denomina armónica en una región plana, si en todos los puntos de esta región satisface la ecuación $\Delta u = 0$ (véase (41.109)). Demuéstrese que dos funciones armónicas son dependientes en una región plana cuando, y sólo cuando son linealmente dependientes.

§ 43. EXTREMO CONDICIONADO

43.1. CONCEPTO DE EXTREMO CONDICIONADO

Supongamos que en un conjunto abierto $G \subset R^n$ vienen definidas las funciones

$$y_i = f_i(x), i = 1, 2, \dots, m, \tag{43.1}$$

$x = (x_1, \dots, x_n) \in G$. Designemos mediante X un conjunto de los puntos $x \in G$, donde todas las funciones f_i, $i = 1, 2, \dots, m$ se reducen a cero:

$$X = \{x: f_i(x) = 0, i = 1, 2, \dots, m, x \in G\}. \tag{43.2}$$

Las ecuaciones

$$f_i(x) = 0, i = 1, 2, \dots, m, \tag{43.3}$$

se llamarán *ecuaciones de conexión (enlace)*.

Definición 1. *Sea $y = f_0(x)$ una función definida en G. El punto $x^{(0)} \in X$ se llama punto de extremo condicionado**) *de la funcion $f_0(x)$ respecto de las ecuaciones de enlace* (43.3) (*o si se cumplen dichas ecuaciones*), *siempre que es un punto de extremo ordinario de esta función al considerarse ésta sólo en el conjunto X* (véase el p. 40.1).

En otras palabras, el valor de la función $f_0(x)$ en el punto $x^{(0)}$ se compara aquí no con todos los valores suyos en un entorno suficientemente pequeño de este punto, sino sólo con los valores en los puntos pertenecientes, a la vez, al mencionado entorno suficiente pequeño y al conjunto X. Al igual que en el caso de los extremos ordinarios, podemos, desde luego, considerar los puntos de extremo simplemente condicionado y los de extremo estrictamente condicionado.

Ejemplos. 1. Consideraremos una función

$$f(x, y) = x^2 + y^2 \tag{43.4}$$

y la ecuación de enlace

$$x + y - 1 = 0. \tag{43.5}$$

Hallemos el extremo condicionado de la función (43.4) para el caso en que se cumple la ecuación de enlace (43.5). De (43.5) tenemos $y = 1 - x$, de donde

$$f(x, 1 - x) = 2x^2 - 2x + 1.$$

De este modo, cuando se cumple la condición de enlace, la función (43.4) es una función de una sola variable. Su extremo se halla de una manera elemental: igualando a cero su derivada (la condición necesaria de un extremo), obtenemos $2x - 1 = 0$, de donde $x = \frac{1}{2}$. En el punto dado la función que se considera, tiene, evidentemente, un mínimo (la función es un polinomio de segundo grado cuyo término mayor tiene coeficiente positivo). Al valor $x = \frac{1}{2}$, de conformidad con la ecuación de enlace (43.5), le corresponde $y = \frac{1}{2}$.

Por consiguiente, en el punto (1/2, 1/2) la función (43.4) alcanza su mínimo respecto de la ecuación de enlace (43.5). En el lenguaje geométrico esto significa que el punto del paraboloide $z = x^2 + y^2$, que se proyecta en el punto (1/2, 1/2), es el más inferior de todos sus puntos dispuestos por encima de la recta (43.5) (fig. 169). Este ejemplo muestra que un punto en el que la función alcanza su extremo condicionado, no es, en el caso general, punto de extremo de esta función.

2. Consideraremos la función $f(x, y) = y^2 - x^2$ y la ecuación de enlace $y = 2x$.

Tenemos $f(x, 2x) = 3x^2$, es decir, si se cumplen las ecuaciones de enlace, la función en consideración es también una función de una sola variable y, evidentemente, alcanza su mínimo cuando $x = 0$ (fig. 170). De conformidad con la ecuación de enlace, al valor $x = 0$ corresponde el valor de $y = 0$, y, por lo tanto, la función $f(x, y) = y^2 - x^2$ tiene en el punto (0, 0) un mínimo condicionado respecto de la ecuación de enlace $y = 2x$.

*) Se ha adoptado también el término "extremo relativo".

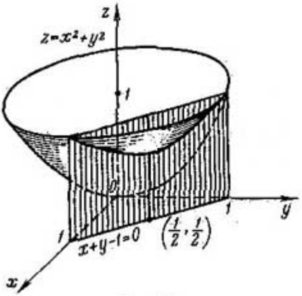

Fig. 169

Cabe destacar que en este caso la propia función $f(x, y)$ no tiene máximo ni mínimo en ningún punto del plano. De este modo, el ejemplo considerado muestra que la función puede no tener un extremo, pero, para ciertas ecuaciones de enlace, puede tener un extremo condicionado.

En lo sucesivo supondremos que

1) todas las funciones $f_0, f_1, \dots, f_m$ son continuamente derivables en el conjunto abierto G;

2) en el punto $x^{(0)}$ que se considera, los vectores $\nabla f_1, \dots, \nabla f_m$ son linealmente independientes, es decir, el rango de la matriz de Jacobi

$$\left(\frac{\partial f_j}{\partial x_i}\right), j = 1, 2, \dots, m, i = 1, 2, \dots, n,$$

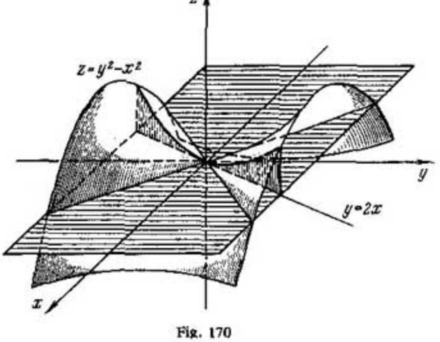

Fig. 170

es igual a m, o sea, al número de sus filas (las filas de la matriz de Jacobi son componentes de los gradientes $\nabla f_1, \dots, \nabla f_m$).

De acuerdo con los resultados del párrafo anterior, esto significa que las funciones del sistema (43.1) son independientes en cierto entorno del punto $x^{(0)}$. Por cuanto en un espacio n-dimensional no pueden haber más de n vectores linealmente independientes y el rango de la matriz no puede exceder del número de columnas, entonces, de la condición 2) proviene que $m \leqslant n$.

Conforme a la condición 2), en el punto $x^{(0)}$ al menos uno de los determinantes del tipo

$$\frac{\partial(f_1, \dots, f_m)}{\partial(x_{i_1}, \dots, x_{i_m})}$$

es distinto de cero. Supongamos, para concretar, que en el punto $x^{(0)}$

$$\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_m)} \neq 0. \tag{43.6}$$

Entonces, en virtud del teorema sobre las funciones implícitas (véase el p. 41.3), el sistema de ecuaciones (43.3) en cierto entorno del punto $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ puede resolverse respecto de las variables $x_1, \dots, x_m$:

$$\begin{gathered} x_1 = \varphi_1(x_{m+1}, \dots, x_n) \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ x_m = \varphi_m(x_{m+1}, \dots, x_n) \end{gathered} \tag{43.7}$$

Al sustituir los valores de $x_1, \dots, x_m$, proporcionados por las fórmulas (43.7), en $y = f_0(x)$, es decir, al examinar la composición de las funciones f_0 y $\varphi_1, \dots, \varphi_m$, obtendremos una función

$$y = f_0(\varphi_1(x_{m+1}, \dots x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) \overset{\text{def}}{=} g(x_{m+1}, \dots, x_n) \tag{43.8}$$

de $n - m$ variables $x_{m+1}, \dots, x_n$ que está definida y es continuamente derivable en cierto entorno del punto $\tilde{x}^{(0)} = (x_{m+1}^{(0)}, \dots, x_n^{(0)})$ en el espacio $(n-m)$-dimensional R^{n-m}

Como, de acuerdo con el teorema sobre las funciones implícitas, las condiciones (43.3) y (43.7) son equivalentes, resulta válida la siguiente afirmación.

El punto $x^{(0)}$ es un punto de extremc condicionado (estricto) para la función $f_0(x)$ respecto de las ecuaciones de enlace (43.3) *cuando, y sólo cuando, $\tilde{x}^{(0)}$ sea un punto de extremo ordinario (estricto) de la función* (43.8).

Si $\tilde{x}(0)$ es un punto de extremo ordinario de la función g, será el punto estacionario de esta función (véase el p. 40.1):

$$dg(\tilde{x}^{(0)}) = 0. \tag{43.9}$$

Recordemos que la diferencial es una función homogénea lineal y el hecho de que ella es nula significa que es nula dicha función, cualesquiera que sean los valores de sus argumentos, en el caso dado, para cualesquiera $dx_{m+1}, dx_{m+2}, \dots, dx_n$. Esto es factible, evidentemente, cuando y sólo cuando, todos los coeficientes de los argumentos citados, es decir, las derivadas $\dfrac{\partial g}{\partial x_{m+k}}$, $k = 1, 2, \ldots\ n - m$, se anulan en el punto $x^{(0)}$. La condición (43.9) es necesaria para que en el punto $x^{(0)}$ haya un extremo condicionado.

De este modo, el método, basado en la resolución del sistema de ecuaciones (43.3), permite reducir el problema de búsqueda del extremo condicionado al problema, ya estudiado, sobre un extremo ordinario. Precisamente este procedimiento se ha empleado en los ejemplos considerados más arriba. Sin embargo, resulta a menudo imposible o muy embarazoso expresar la solución del sistema (43.3) en términos de las funciones elementales, razón por la cual es deseable disponer de un método que permita hallar el extremo condicionado sin resolver el sistema (43.3). Tal método va expuesto más abajo.

43.2. MÉTODO DE LOS MULTIPLICADORES DE LAGRANGE PARA BUSCAR LOS PUNTOS DE EXTREMO CONDICIONADO

En este punto se supondrá que todas las funciones $f_0, f_1, \ldots, f_m$ son continuamente derivables en el conjunto abierto G.

Teorema 1. *Sea $x^{(0)}$ un punto de extremo condicionado de la función f_0, cumpliéndose las ecuaciones de enlace* (43.3). *En este punto los gradientes ∇f_0, $\nabla f_1, \ldots, \nabla f_m$ son linealmente dependientes, es decir, existen tales números λ_0, $\lambda_1, \ldots, \lambda_m$, de los cuales no todos son nulos, que*

$$\lambda_0 \nabla f_0 + \lambda_1 \nabla f_1 + \ldots + \lambda_m \nabla f_m = 0. \qquad (43.10)$$

Corolario. *Si en el punto $x^{(0)}$ de extremo condicionado de la función f_0 respecto de las ecuaciones de enlace* (43.3) *los gradientes $\nabla f_1, \ldots, \nabla f_m$ son linealmente independientes, es decir, el rango de la matriz de Jacobi*

$$\left(\frac{\partial f_j}{\partial x_i} \right), j = 1, 2, \ldots, m \quad i = 1, 2, \ldots, n,$$

es igual a m, entonces existen tales $\lambda_1, \ldots, \lambda_m$ que en dicho punto

$$\nabla f_0 + \sum_{j=1}^{m} \lambda_j \nabla f_j = 0, \qquad (43.11)$$

es decir, ∇f_0 es una combinación lineal de los gradientes $\nabla f_1, \ldots, \nabla f_m$.

En la forma coordenada esta condición tiene por expresión: para cualquier $i = 1, 2, \ldots, n$ en el punto $x^{(0)}$

$$\frac{\partial f_0}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial f_j}{\partial x_i} = 0. \qquad (43.12)$$

La función

$$F(x) \stackrel{\text{def}}{=} f_0(x) + \sum_{j=1}^{m} \lambda_j f_j(x), \qquad (43.13)$$

donde los números $\lambda_1, \dots, \lambda_m$ satisfacen la condición (43.12), se llama *función de Lagrange* del problema en consideración y los propios números, *multiplicadores de Lagrange*.

La condición (43.12) significa que si $x^{(0)}$ es un punto de extremo condicionado de la función f_0 respecto de la ecuación de enlace (43.3), será punto estacionario para la función de Lagrange, es decir,

$$\frac{\partial F(x^{(0)})}{\partial x_i} = 0, i = 1, 2, \dots, n. \tag{43.14}$$

Antes de demostrar el teorema aclaremos su sentido y mostremos cómo usarlo para la obtención de los puntos de extremo condicionado. Ante todo fijemos la atención en que en la función del tipo (43.13), para los números arbitrarios $\lambda_1, \dots, \lambda_m$, cualquier punto de su extremo condicionado es a la vez el punto de extremo condicionado de la función de partida f_0, y viceversa. Elijamos tales valores de $\lambda_1, \dots, \lambda_n$ que se cumplan las condiciones (43.12), es decir, que el punto dado de extremo condicionado resulte ser también el punto estacionario de la función (43.11).

Para la obtención de los puntos de extremo condicionado, conviene considerar un sistema de $n + m$ ecuaciones (43.3) y (43.10) respecto de las incógnitas $x_1^{(0)}, \dots, x_n^{(0)}, \lambda_1, \dots, \lambda_m$ y resolverlo (si esto es factible), hallando $x_1^{(0)}, \dots, x_n^{(0)}$ y eliminando, cuanto sea posible, $\lambda_1, \dots, \lambda_m$. El teorema enunciado afirma que todos los puntos de extremo condicionado estarán entre los puntos $(x_1^{(0)}, \dots, x_n^{(0)})$, determinados de este modo. La cuestión de qué puntos entre los mencionados serán realmente puntos de extremo condicionado requiere unas investigaciones adicionales; algunos razonamientos referentes a esta cuestión serán aducidos en el p. 43.5*.

DEMOSTRACIÓN DEL TEOREMA. Demostremos una afirmación que es equivalente al teorema: si en el punto $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ que satisface las ecuaciones de enlace

$$f_k(x^{(0)}) = 0, k = 1, 2, \dots, m, \tag{43.15}$$

los gradientes $\nabla f_0, \nabla f_1, \dots, \nabla f_m$ son linealmente independientes, entonces $x^{(0)}$ no es el punto de extremo condicionado.

Así pues, sean $\nabla f_0, \nabla f_1, \dots, \nabla f_m$ linealmente independientes y, por consiguiente, el rango de la matriz de Jacobi $\left(\frac{\partial f_j}{\partial x_i}\right), j = 0, 1, \dots, m, i = 1, 2, \dots, n$, es igual a $m + 1$. Entonces, en dicha matriz existe un menor de orden $m + 1$ distinto de cero. Para concretar, convengamos en considerar que el menor se ha formado por las primeras $m + 1$ columnas, es decir,

$$\left.\frac{\partial(f_0, f_1, \dots, f_m)}{\partial(x_1, x_2, \dots, x_{m+1})}\right|_{x = x^{(0)}} \neq 0. \tag{43.16}$$

El conjunto G es abierto, por lo cual existe tal $\delta_0 > 0$, que para todo δ, $0 < \delta < \delta_0$, el cubo

$$Q^n_\delta = \{x : |x_i - x_i^{(0)}| < \delta, i = 1, 2, \dots, n\}$$

se dispone dentro de G y, por ende, en este cubo están definidas todas las funciones $f_0, f_1, \dots, f_m$.

Fijemos $x_{m+2} = x_{m+2}^{(0)}, \dots, x_n = x_n^{(0)}$ e introduzcamos las designaciones

$$x^* = (x_1, \dots, x_{m+1})$$

$$Q_\delta^{m+1} = \{x^* : |x_i - x_i^{(0)}| < \delta, i = 1, 2, \dots, m+1\}.$$

Obviamente, las funciones $f_j(x_1, \dots, x_{m+1}, x_{m+2}^{(0)}, \dots, x_n^{(0)}), j = 0, 1, \dots, m$, están definidas y son continuamente derivables en todo punto de Q_δ^{m+1}. Consideraremos una aplicación $\Phi: Q_\delta^{m+1} \to R^{m+1}$, definida por las fórmulas

$$\begin{aligned} y_1 &= f_0(x_1, \dots, x_{m+1}, x_{m+2}^{(0)}, \dots, x_n^{(0)}) \\ y_2 &= f_1(x_1, \dots, x_{m+1}, x_{m+2}^{(0)}, \dots, x_n^{(0)}) \\ &\dots\dots\dots\dots\dots\dots \\ y_{m+1} &= f_{m+1}(x_1, \dots, x_{m+1}, x_{m+2}^{(0)}, \dots, x_n^{(0)}). \end{aligned} \tag{43.17}$$

En vista de (43.16), para el punto $x^{*(0)} = (x_1^{(0)}, \dots, x_{m+1}^{(0)})$ tenemos

$$\left.\frac{\partial(y_1, \dots, y_{m+1})}{\partial(x_1, \dots, x_{m+1})}\right|_{x^* = x^{*(0)}} = \left.\frac{\partial(f_0, f_1, \dots, f_m)}{\partial(x_1, x_2, \dots, x_{m+1})}\right|_{x = x^{(0)}} \neq 0,$$

y debido a (43.15),

$$\Phi(x^{*(0)}) = (f_0(x^{(0)}), 0, \dots, 0).$$

Por ello, (véase el teorema 7 en el p. 41.8 sobre la invertibilidad local de la aplicación continuamente derivable en un punto, donde su jacobiano es distinto de cero) existe tal $\varepsilon > 0$ que en el entorno

$$V = \{y = (y_1, \dots, y_{m+1}) : |y_1 - f_0(x^{(0)})| < \varepsilon, |y_j| < \varepsilon, j = 2, 3, \dots, m+1\}$$

(fig. 171) está definida una aplicación inversa respecto de Φ, y, por lo tanto, en cualquier punto de este entorno se aplica algún punto de Q_δ^{m+1}.

En particular, puesto que para cualquier η, $0 < \eta < \varepsilon$, tiene lugar la inclusión $(f(x^{(0)}) \pm \eta, 0, \dots, 0) \in V$, entonces en el cubo Q_δ^{m+1} existen los puntos $x'^* = (x_1', \dots, x_{m+1}')$ y $x''^* = (x_1'', \dots, x_{m+1}'')$ que, realizándose la aplicación Φ, se aplican en los puntos indicados del entorno V:

$$\Phi(x'^*) = (f(x^{(0)}) + \eta, 0, \dots, 0), \Phi(x''^*) = (f(x^{(0)}) - \eta, 0, \dots, 0).$$

Si ponemos, para abreviar, $x' = (x_1', \dots, x_{m+1}', x_{m+2}^{(0)}, \dots, x_n^{(0)})$ y $x'' = (x_1'', \dots, x_{m+1}'', x_{m+2}^{(0)}, \dots, x_n^{(0)})$, entonces en la inscripción coordenada obtendremos (véase (43.17)):

$$f_0(x') = f_0(x^{(0)}) + \eta > f(x^{(0)}),$$
$$f_k(x') = 0, k = 1, 2, \dots, m, x' \in Q_\delta^n$$

y

$$f_0(x'') = f_0(x^{(0)}) - \eta < f(x^{(0)}),$$
$$f_k(x'') = 0, k = 1, 2, \dots, m, x'' \in Q_\delta^n.$$

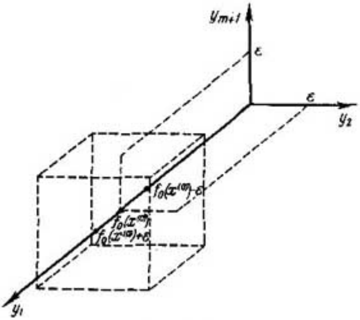

Fig. 171

Puesto que $\delta > 0$, $0 < \delta < \delta_0$, es arbitrario, esto es indicio de que $x^{(0)}$ no es el punto de extremo condicionado. □

DEMOSTRACIÓN DEL COROLARIO. Si los vectores $\nabla f_1, \ldots, \nabla f_m$ son linealmente independientes, en la igualdad (43.10) tenemos $\lambda_0 \neq 0$, puesto que si fuera $\lambda_0 = 0$, los vectores mencionados serían, en virtud de (43.10), linealmente dependientes. Al dividir ambos miembros de (43.10) por λ_0, obtendremos una igualdad del tipo (43.11). □

43.3*. INTERPRETACIÓN GEOMÉTRICA DEL MÉTODO DE LAGRANGE

Daremos ahora algunas explicaciones geométricas al teorema 1. Consideraremos, para no complicar, el caso de un extremo condicionado de la función de dos variables $z = f(x, y)$ verificándose la ecuación de enlace $\varphi(x, y) = 0$.

Supongamos que las funciones f y φ son continuamente derivables en el entorno del punto (x_0, y_0), $\nabla\varphi(x_0, y_0) = \left(\frac{\partial\varphi(x_0, y_0)}{\partial x}, \frac{\partial\varphi(x_0, y_0)}{\partial y}\right) \neq 0$ y $\varphi(x_0, y_0) = 0$. En vista de la condición $\nabla\varphi(x_0, y_0) \neq 0$, de acuerdo con el teorema sobre las funciones implícitas, la ecuación $\varphi(x, y) = 0$ en el entorno del punto (x_0, y_0) define cierta curva suave cuya representación explícita tiene por expresión o bien $y = y(x)$, o bien $x = x(y)$. Por cuanto nos interesan sólo los puntos suficientemente próximos a (x_0, y_0), la curva mencionada se llamará simplemente curva $\varphi(x, y) = 0$ (es decir, en otras palabras, en lo que sigue se considerará siempre la restricción de las funciones f y φ al entorno indicado del punto (x_0, y_0))

El gradiente $\nabla\varphi(x_0, y_0)$ es una normal a la curva $\varphi(x, y) = 0$ en el punto (x_0, y_0) (p. 20.6). Designemos con τ el vector unidad tangente a la curva $\varphi(x, y)$ en el punto (x_0, y_0). Supongamos, para concretar, que la curva en consideración se da por la

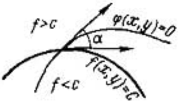

Fig. 172

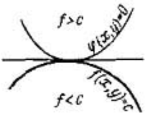

Fig. 173

ecuación $y = y(x)$. Si el punto (x_0, y_0) es un punto de extremo condicionado, entonces x_0 es un punto de extremo ordinario para la funcion $g(x) \stackrel{\text{def}}{=} f(x, g(x))$ (véase el p. 43.1), y, por lo tanto, $g'(x) = 0$, es decir, la derivada de la función f en el punto (x_0, y_0) según la dirección de la curva $\varphi(x, y) = 0$, o bien, que es lo mismo (vease el p. 20.7), según la dirección del vector τ, es igual a cero

$$\frac{\partial f(x_0, y_0)}{\partial \tau} = (\nabla f(x_0, y_0), \tau) = 0.$$

Esto significa la ortogonalidad del gradiente $\nabla f(x_0, y_0)$ y del vector tangente τ, lo que es equivalente al carácter colineal de los vectores $\nabla f(x_0, y_0)$ y $\nabla \varphi(x_0, y_0)$:

$$\nabla f(x_0, y_0) = \lambda \nabla \varphi(x_0, y_0),$$

es decir, se cumple la condición (43.11). El cumplimiento de esta condición en el punto de extremo condicionado puede explicarse también de otro modo. Sea $f(x_0, y_0) \stackrel{\text{def}}{=} c$. Si en el punto (x_0, y_0) no se cumple la condición (43.11), es decir, cuando los gradientes ∇f y $\nabla \varphi$ no son colineales, esto significa que en dicho punto $\nabla f \neq 0$ y la línea de nivel $f(x, y) = c$ corta en este punto la curva $\varphi(x, y) = 0$, haciendo con ésta cierto ángulo α, distinto de 0 y π (fig. 172). Por eso en todo entorno suficientemente pequeño del punto (x_0, y_0) una parte de la curva $\varphi(x, y) = 0$ resultará ser dispuesta en el dominio $f < c$ ("en el dominio de los valores menores") y otra parte, en el dominio $f > c$ ("en el dominio de los valores mayores"). Esto significa que en el punto (x_0, y_0) no hay extremo condicionado de que se trata.

En el caso en que los vectores ∇f y $\nabla \varphi$ sean colineales, $\nabla f = \lambda \nabla \varphi$, una parte de la curva $\varphi(x, y) = 0$ puede pertenecer a cierto entorno del punto (x_0, y_0), integramente dispuesto en el dominio de valores menores $f < c$ (fig. 173) o bien en el dominio de valores mayores $f > c$. En este caso en el punto (x_0, y_0) se alcanza un extremo condicionado.

Sin embargo, siendo colineales los vectores ∇f y $\nabla \varphi$ la curva $\varphi(x, y) = 0$ puede también disponerse en cualquier entorno suficientemente pequeño del punto (x_0, y_0), con la particularidad de que una parte de dicha curva se dispondrá en el dominio de valores menores y la otra parte, en el de valores mayores de la función f (fig. 174); en este caso en el punto (x_0, y_0) tampoco habrá extremo condicionado. Una situación semejante aparece, por ejemplo, cuando las curvas $f(x, y) = c$ y $\varphi(x, y) = 0$ tienen en el punto (x_0, y_0) una tangente común, con la particularidad de que la curva $f(x, y) = c$ está dispuesta en un entorno suficientemente pequeño del punto (x_0, y_0) por un lado de la tangente mencionada, mientras que la curva $\varphi(x, y) = 0$ sufre inflexión en dicho punto, pasando de un lado de la tangente al otro lado.

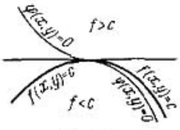

Fig. 174

Lo dicho aclara aquella circunstancia que (43.10) es una *condición necesaria* para que haya extremo condicionado, *pero no suficiente*.

Los razonamientos geométricos aducidos referentes a la cuestión acerca del extremo condicionado quedan válidos también en el caso multidimensional.

43.4*. PUNTOS ESTACIONARIOS DE LA FUNCIÓN DE LAGRANGE

Se dará aquí la descripción de los puntos estacionarios de la función de Lagrange (43.13) por medio de la función $g(x_{m+1}, \dots, x_n)$ introducida en el p. 43.1 (véase (43.8)). Demostremos previamente un lema sencillo conocido del curso de álgebra lineal.

Sea dado un sistema de ecuaciones homogéneas lineales

$$a_{i1}x_1 + \dots + a_{in}x_n = 0, i = 1, 2, \dots, m, \tag{43.18}$$

y una ecuación homogénea lineal más

$$b_1x_1 + \dots + b_nx_n = 0. \tag{43.19}$$

El sistema de ecuaciones obtenido por adición de la ecuación (43.19) al sistema (43.18) se denominará *sistema ampliado* (43.18)—(43.19).

Lema. *Para que el sistema ampliado* (43.18) — (43.19) *sea equivalente al sistema básico* (43.18), *es necesario y suficiente que la ecuación* (43.19) *sea una combinación lineal del sistema de ecuaciones* (43.18).

Corolario. *Para que la ecuación* (43.19) *sea una combinación lineal de las ecuaciones* (43.18) *o bien, que es lo mismo, para que el vector*

$$b \overset{\text{def}}{=} (b_1, \dots, b_n) \tag{43.20}$$

constituya una combinación lineal de los vectores

$$a_i \overset{\text{def}}{=} (a_{i1}, \dots, a_{in}), i = 1, 2, \dots, m, \tag{43.21}$$

es necesario y suficiente que toda solución del sistema (43.18) *sea la solución de la ecuación* (43.19).

DEMOSTRACIÓN DEL LEMA. Supongamos que el rango de la matriz (a_{ij}) de los coeficientes del sistema (43.18) es igual a m_0. Es evidente que $m_0 \leqslant m$. Si $m_0 < m$, entonces $m - m_0$ ecuaciones del sistema (43.18) son combinaciones lineales de las demás. Al desechar aquellas $m - m_0$ ecuaciones lineales que son combinaciones lineales de las restantes, obtendremos un sistema de m_0 ecuaciones linealmente independientes equivalente al sistema (43.18), con la particularidad de que la ecuación

(43.19) constituye una combinación lineal de las ecuaciones del sistema (43.18) cuando, y sólo cuando, es una combinación lineal del sistema mencionado formado por m_0 ecuaciones restantes. Por eso, desde el mismo principio convengamos en considerar que $m = m_0$, es decir, que el rango de la matriz (a_{ij}) de los coeficientes del sistema (43.18) es igual a m, es decir, al número de ecuaciones de este sistema.

Supongamos que los sistemas (43.18) y (43.18) — (43.19) son equivalentes. Esto significa que los espacios de sus soluciones coinciden. Por cuanto todas las ecuaciones del sistema básico (43.18) figuran en el sistema ampliado (43.18) — (43.19), toda solución del sistema ampliado será también la solución del sistema básico, es decir, el espacio de soluciones del sistema ampliado está contenido en el espacio de soluciones del sistema básico. Por consiguiente, la coincidencia de estos espacios equivale a la igualdad de sus dimensiones.

La dimensión s del espacio de soluciones de un sistema de ecuaciones homogéneas lineales es igual, como se sabe, al número de las incógnitas n de dicho sistema, del cual se ha sustraído el rango r de la matriz de los coeficientes del sistema: $s = n - r$. De aquí se deduce que la equivalencia de los sistemas (43.18) y (43.18)—(43.19) implica la igualdad de los rangos de sus matrices. El rango de la matriz de los coeficientes del sistema (43.18) es igual, por hipótesis, a m, es decir, los vectores (43.21) son linealmente independientes.

El rango de la matriz de los coeficientes del sistema ampliado (48.18) — (43.19), de acuerdo con lo expuesto en nuestras condiciones, es también igual a m. Por ello, los vectores (véanse (43.20) y (43.21))

$$b, a_1, \dots, a_m \tag{43.22}$$

son linealmente dependientes. Esto es indicio de que b es una combinación lineal de los vectores $a_1, \dots, a_m$.

Efectivamente, la dependencia lineal de los vectores (43.22) significa que existen tales números $\mu_0, \mu_1, \dots, \mu_m$, no todos nulos, que

$$\mu_0 b + \mu_1 a_1 + \dots + \mu_m a_m = 0. \tag{43.23}$$

Aquí, a ciencia cierta, $\mu_0 \neq 0$, puesto que de lo contrario los vectores $a_1, \dots, a_m$ resultarían ser linealmente dependientes. Al dividir la igualdad (43.23) por μ_0, llegamos a que b es una combinación lineal de los vectores $a_1, \dots, a_m$.

Viceversa, si b es una combinación lineal de los vectores (43.21), entonces los sistemas de vectores (43.21) y (43.22) tienen cada uno exactamente m vectores linealmente independientes, es decir, los rangos de las matrices de los coeficientes de los sistemas de ecuaciones (43.18) y (43.18) — (43.19) son iguales.

Así pues, la condición de que el vector b es una combinación lineal de los vectores (43.21):

$$b = \lambda_1 a_1 + \dots + \lambda_m a_m$$

es equivalente a la igualdad de rangos de las matrices de los coeficientes de los sistemas básico y ampliado en consideración y, por consiguiente, significa su equivalencia. □

La afirmación del corolario proviene seguidamente del lema, ya que los sistemas (43.18) y (43.18) — (43.19) son, evidentemente equivalentes cuando, y sólo cuando, toda solución del sistema (43.18) es, a la vez, la solución del sistema (43.19); las demás ecuaciones de estos sistemas son simplemente coincidentes. □

OBSERVACIÓN 1. El lema demostrado y sus corolarios tienen una sencilla interpretación geométrica en el espacio vectorial *euclídeo* n-dimensional R^n, es decir, en un espacio n-dimensional provisto de un producto escalar. Haciendo uso de la designación del producto escalar, el sistema (43.18) puede escribirse en la forma

$$(a_i, x) = 0, i = 1, 2, \dots, m, \tag{43.24}$$

y la ecuación (43.19), en la forma

$$(b, x) = 0, \tag{43.25}$$

donde los vectores $a_1, \dots, a_m$ y b están definidos en (43.20) y (43.21), mientras que $x = (x_1, \dots, x_n)$.

El conjunto de toda una serie de combinaciones lineales de los vectores $a_1, \dots$ $\dots, a_m$ forma un subespacio del espacio R^n y se denomina *subespacio tendido sobre dichos vectores*. Designémoslo $\mathscr{L}(a_1, \dots, a_m)$.,

El conjunto de soluciones del sistema (43.24) se compone de todos los vectores x ortogonales al subespacio $\mathscr{L}(a_1, \dots, a_m)$. Designemos este conjunto de soluciones con T. Es también un subespacio del espacio R^n.

Los subespacios $L \overset{\text{def}}{=} \mathscr{L}(a_1, \dots, a_m)$ y T se denominan *complementos ortogonales* uno al otro en el *espacio* R^n.

Dado que $L = \mathscr{L}(a_1, \dots, a_m)$, la representación del vector b en forma de una combinación lineal de los vectores $a_1, \dots, a_m$ es equivalente a su pertenencia al subespacio L del espacio R^n: $b \in L$. Esta condición es equivalente, a su vez, a la ortogonalidad del vector b al subespacio T: $b \perp T$, la que significa que para todo $x \in T$ tiene lugar la igualdad $(b, x) = 0$, es decir, que toda solución x del sistema (43.24) es la solución de la ecuación (43.25). Esto es precisamente la afirmación del corolario del lema.

OBSERVACIÓN 2. Recordemos un método por medio del cual pueden obtenerse todas las soluciones de un sistema homogéneo de ecuaciones lineales. Supongamos que el sistema (43.18) se compone de las ecuaciones linealmente independientes. En este caso el rango de la matriz de sus coeficientes es igual a m. Esto significa que existe un menor de dicha matriz de orden m, distinto de cero. Sea, para concretar,

$$\begin{vmatrix} a_{11} \dots a_{1m} \\ a_{m1} \dots a_{mm} \end{vmatrix} \neq 0. \tag{43.26}$$

En este caso todas las soluciones del sistema (43.18) pueden obtenerse, fijando arbitrariamente las últimas $n - m$ coordenadas del vector $(x_1, \dots, x_n)$. Las coordenadas restantes se obtienen unívocamente del sistema de ecuaciones (43.18). En efecto, tomemos una solución arbitraria $(x_1^{(0)}, \dots, x_n^{(0)})$ del sistema (43.18). Realizada la sustitución $x_{m+1} = x_{m+1}^{(0)}, \dots, x_n = x_n^{(0)}$ en (43.18), se obtendrá un sistema de m ecuaciones lineales (con m incógnitas $x_1, \dots, x_m$) tal que la matriz de los coeficientes de este sistema es regular, en virtud de la condición (43.26). Por esta razón existen los únicos valores $x_1, \dots, x_m$ que satisfacen el sistema obtenido. Por cuanto $(x_1^{(0)}, \dots, x_n^{(0)})$ fue también la solución del sistema (43.18), se tiene $x_1 = x_1^{(0)}, \dots,$ $x_m = x_m^{(0)}$.

Pasemos ahora al análisis de los puntos estacionarios de la función de Lagrange.

Teorema 2. *Supongamos que las funciones $f_0, f_1, \ldots, f_m$ son continuamente derivables en la región $G \subset R^n$, $x^{(0)} \in G$,*

$$f_i(x^{(0)}) = 0,\ i = 1, 2, \ldots, m,$$

y el rango de la matriz de Jacobi de las funciones $f_1, \ldots, f_m$ en el punto $x^{(0)}$ es igual a m. Para que en el punto $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ el gradiente ∇f_0 sea una combinación lineal de los gradientes $\nabla f_1, \ldots, \nabla f_m$ es necesario y suficiente que el punto $x^{(0)} = (x_{m+1}^{(0)}, \ldots, x_n^{(0)})$ sea estacionario para la funcion $g(x) = g(x_{m+1}, \ldots, x_n)$ (véase (43.8)).

Recordemos que si en el punto $x^{(0)}$ el gradiente ∇f_0 es una combinación lineal

$$\nabla f_0 = \lambda_1 \nabla f_1 + \ldots + \lambda_m \nabla f_m, \tag{43.27}$$

de los gradientes $\nabla f_1, \ldots, \nabla f_m$, esto es equivalente a que existe una función de Lagrange

$$F = f_0 - \lambda_1 f_1 - \ldots - \lambda_m f_m, \tag{43.28}$$

para la cual el punto $x^{(0)}$ es estacionario:

$$\frac{\partial F(x^{(0)})}{\partial x_i} = 0,\ i = 1, 2, \ldots, n. \tag{43.29}$$

Esta es simplemente una inscripción coordenada de la condición (43.27), pues, en virtud de (43.28),

$$\frac{\partial F}{\partial x_i} = \frac{\partial f_0}{\partial x_i} - \lambda_1 \frac{\partial f_1}{\partial x_i} - \ldots - \lambda_m \frac{\partial f_m}{\partial x_i},\ i = 1, \ldots, m.$$

DEMOSTRACIÓN. Por hipótesis, el rango de la matriz de Jacobi del sistema de funciones $f_1, \ldots, f_m$ en el punto $x^{(0)}$ es igual a m. Convengamos en considerar, para concretar, como en el p. 43.1, que

$$\left.\frac{\partial(f_1, \ldots, f_m)}{\partial(x_1, \ldots, x_m)}\right|_{x^{(0)}} \neq 0. \tag{43.30}$$

Sustituyamos en la ecuación de enlace (43.3) las funciones (43.7) que representan las soluciones de estas ecuaciones y derivemos las identidades obtenidas respecto de las variables $x_{m+1}, \ldots, x_n$. Obtendremos para el punto $x^{(0)}$ las igualdades $df_i(x^{(0)}) = 0$, $i = 1, 2, \ldots, m$, válidas para cualesquiera incrementos $dx_{m+1}, \ldots, dx_n$ de las variables independientes $x_{m+1}, \ldots, x_n$ (recordemos que la diferencial es una función lineal definida en todo el espacio). Al hacer uso de la invariación de la forma de la primera diferencial respecto de la elección de las variables, llegamos a que en el punto $x^{(0)}$ se cumplen las igualdades

$$\frac{\partial f_i}{\partial x_1} dx_1 + \ldots + \frac{\partial f_i}{\partial x_m} dx_m + \frac{\partial f_i}{\partial x_{m+1}} dx_{m+1} + \ldots$$
$$\ldots + \frac{\partial f_i}{\partial x_n} dx_n = 0,\ i = 1, 2, \ldots, m, \tag{43.31}$$

donde $dx_{m+1}, \dots dx_n$ son arbitrarios, mientras que $dx_1, \dots, dx_m$ se hallan a partir de las fórmulas (43.7). De este modo, el vector

$$dx = (dx_1, \dots, dx_m, dx_{m+1}, \dots, dx_n) \tag{43.32}$$

es la solución del sistema homogéneo lineal (43.31).

Cabe notar que en virtud de la condición (43.30), los valores de $dx_1, \dots, dx_m$, para $dx_{m+1}, \dots, dx_n$ dados, se hallan también unívocamente del sistema (43.31). De la observación 2 se infiere, además, que mediante el método indicado se obtienen todas las soluciones del sistema (43.31).

El carácter estacionario del punto $\tilde{x}^{(0)}$ para la función $g(\tilde{x}) = g(x_{m+1}, \dots, x_n)$ significa que $dg(\tilde{x}^{(0)}) = 0$. Por ser invariante la forma de la primera diferencial, esta igualdad puede escribirse más detalladamente como sigue

$$\frac{\partial f_0}{\partial x_1} dx_1 + \dots + \frac{\partial f_0}{\partial x_m} dx_m + \frac{\partial f_0}{\partial x_{m+1}} dx_{m+1} + \dots + \frac{\partial f_0}{\partial x_n} dx_n = 0, \tag{43.33}$$

donde $dx_{m+1}, \dots, dx_n$ pueden prefijarse de modo arbitrario, mientras que $dx_1, \dots, dx_m$ se deben hallar de las fórmulas (43.7), o bien, que proporciona el mismo resultado, de las fórmulas (43.31). En otras palabras, cualquier solución del sistema de ecuaciones (43.31) es, a la vez, la solución de la ecuación (43.33). Conforme al corolario del lema, esto resulta posible cuando, y sólo cuando, la ecuación (43.33) es una combinación lineal de las ecuaciones del sistema (43.31), es decir, cuando existen tales números $\lambda_1, \dots, \lambda_m$ que

$$\nabla f_0 = \lambda_1 \nabla f_1 + \dots + \lambda_m \nabla f_m. \ \square$$

OBSERVACIÓN 3. De acuerdo con la observación 2, la totalidad de todas las soluciones del sistema de ecuaciones (43.31) forma un subespacio T del espacio R^n que es un complemento ortogonal al subespacio $L = \mathcal{L}(\nabla f_1, \dots, \nabla f_m)$. Todo vector $y \in T$ es ortogonal respecto de cualquier gradiente ∇f_i, por lo cual será natural llamarlo *vector tangente* en el punto $x^{(0)}$ a la hipersuperficie $f_i(x) = 0$, la cual representa un conjunto de nivel (véase § 19) de la función f_i, $i = 1, \dots, m$.

De este modo, el espacio de soluciones T del sistema (43.31) se compone de los vectores tangentes simultáneamente a todas las hipersuperficies $f_i(x) = 0$, $i = 1, \dots, m$, y por esta razón lo llaman espacio tangente de la intersección de todas las hipersuperficies $f_i(x) = 0$, $i = 1, 2, \dots, m$. Recordemos que los vectores del espacio tangente T, es decir, las soluciones del sistema (43.31) han sido designados mediante dx (43.32).

Puesto que, de acuerdo con el teorema 2, en el punto de extremo condicionado tiene lugar la inclusión

$$\nabla f_0 \in L = \mathcal{L}(\nabla f_1, \dots, \nabla f_m),$$

se tiene

$$\nabla f_0 \perp T.$$

En otras palabras, el gradiente ∇f_0 es ortogonal simultáneamente a todas las tangentes dx a las superficies $f_i(x) = 0$, $i = 1, 2, \dots, m$:

$$(\nabla f_0, dx) = 0$$

(ésta es otra notación de la ecuación (43.33)), es decir, el gradiente ∇f_0 es perpendicular al espacio tangente T en el punto $x^{(0)}$. Pero el conjunto de todos los vectores ortogonales a ∇f_0 forma un subespacio $(n - 1)$-dimensional T_0, denominado espacio tangente a la hipersuperficie $f_0(x) = f_0(x^{(0)})$. En virtud de lo dicho anteriormente, cada vector de T, siendo ortogonal al gradiente ∇f_0, pertenece a T_0, es decir, $T \subset T_0$.

Así pues, *si $x^{(0)}$ es un punto de extremo condicionado, entonces $T \subset T_0$, es decir, el espacio tangente en el punto $x^{(0)}$ de intersección de todas las hipersuperficies, definidas por las ecuaciones de enlace, está contenido en el espacio tangente en el mismo punto de la hipersuperficie $f_0(x) = f_0(x^{(0)})$.*

OBSERVACIÓN 4. Del teorema 2 se deduce una vez más el corolario del teorema 1. Efectivamente, si $x^{(0)}$ es un punto de extremo condicionado, entonces $x^{(0)}$ será el punto de extremo ordinario para la función g (véase el p. 43.1) y, por consiguiente, su punto estacionario. Por ello, de acuerdo con el teorema 2, $x^{(0)}$ es el punto estacionario para la función de Lagrange, es decir, se cumple la condición (43.29).

43.5*. CONDICIONES SUFICIENTES PARA LOS PUNTOS DE EXTREMO CONDICIONADO

En este punto se supondrán también vigentes todas las suposiciones impuestas en las funciones f_0 y f_i, $i = 1, 2, \ldots, m$, en el p. 43.1. Sea

$$F = f_0 + \sum_{i=1}^{m} \lambda_i f_i$$

una función de Lagrange (véase (43.13)) para la función f_0 y las ecuaciones de enlace (43.3). Supongamos que $x^{(0)} \in G$ satisface las ecuaciones de enlace (43.3) y es un punto estacionario de la función de Lagrange, es decir, un punto cuyas coordenadas satisfacen el sistema de ecuaciones (43.12) y (43.3). Nuestro objeto es la obtención de un método, por cuyo intermedio se pueden establecer las condiciones suficientes para que $x^{(0)}$ sea punto de extremo condicionado del problema en consideración.

Observemos ante todo que si el punto $x \in G$ satisface las ecuaciones de enlace (43.3), entonces

$$\Delta f_0 = f_0(x) - f_0(x^{(0)}) = F(x) - F(x^{(0)}) = \Delta F. \tag{43.34}$$

De aquí se ve en seguida que si $x^{(0)}$ es un punto de extremo habitual para la función F, es decir, si ΔF no cambia de signo en cierto entorno del punto $x^{(0)}$, entonces $x^{(0)}$ es el punto de extremo condicionado para la función f_0.

En efecto, de (43.34) se deduce en este caso que el incremento Δf_0 para los valores admisibles de x, es decir, para aquellos que satisfacen las ecuaciones de enlace, tampoco cambia de signo. Esta condición suficiente impone, sin embargo, una restricción demasiado fuerte sobre el comportamiento de la función de Lagrange $F(x)$ en el punto considerado: ésta debe tener extremo ordinario, lo que hace muy estrecho el dominio de la aplicación eventual de la condición indicada en la resolución de los problemas. Por esta razón parece conveniente obtener un criterio suficiente más general para un extremo condicionado.

Supongamos que $x^{(0)} = (x_1^{(0)}, \dots, x_n^{(0)})$ satisface las ecuaciones de enlace (43.3). Volvamos a considerar la función (43.8), es decir, una función $g(\tilde{x}) = g(x_{m+1}, \dots \dots, x_n)$, $\tilde{x} = (x_{m+1}, \dots, x_n)$, que se obtiene de $f_0(x) = f_0(x_1, \dots, x_n)$ a condición de que $x_1, \dots, x_m$ son funciones de las variables $x_{m+1}, \dots, x_n$, determinadas por las ecuaciones de enlace (43.3) en cierto entorno del punto $x^{(0)}$. Supondremos adicionalmente que $f_0(x)$ y $f_i(x)$, $i = 1, 2, \dots, m$, son dos veces continuamente derivables en el punto $x^{(0)}$.

Se ha notado más arriba (véase el p. 43.1) que $x^{(0)}$ es un punto de extremo condicionado (estricto) para la función $f_0(x)$ respecto de las ecuaciones de enlace (43.3) cuando, y sólo cuando, $\tilde{x}^{(0)} = (x_{m+1}^{(0)}, \dots, x_n^{(0)})$ es un punto de extremo ordinario (estricto) para la función $g(x)$. Por eso, si, por ejemplo, en el punto $x^{(0)}$ la función $g(x)$ satisface las condiciones suficientes para la existencia de un extremo estricto, en este punto la función $f_0(x)$ tendrá extremo estricto condicionado respecto de las ecuaciones de enlace (43.3). Las condiciones suficientes para un extremo estricto ordinario se han obtenido anteriormente (véase el teorema 2 en el p. 40.2). Para nuestro caso tienen por expresión

$$1)\ \frac{\partial g(\tilde{x}^{(0)})}{\partial x_i} = 0, i = m + 1, \dots, n; \tag{43.35}$$

2) la segunda diferencial

$$d^2 g(\tilde{x}^{(0)}) = \sum_{i,j=m+1}^{n} \frac{\partial^2 g(\tilde{x}^{(0)})}{\partial x_i \partial x_j} dx_i dx_j \tag{43.36}$$

es una forma cuadrática definida positiva o definida negativa.

Si se cumplen estas condiciones, $\tilde{x}^{(0)}$ es un punto de mínimo estricto o de máximo estricto para la función $g(x)$. En virtud de lo expuesto anteriormente, las condiciones mencionadas son también suficientes para que $x^{(0)}$ sea un punto de mínimo (máximo) estricto condicionado para la función $f_0(x)$ respecto de las ecuaciones de enlace (43.3). No obstante, no son cómodas para el empleo práctico, pues requieren que se conozca la función $g(\tilde{x})$. Por ello, partiendo de las condiciones suficientes obtenidas de un extremo estricto condicionado, que se expresan mediante la función $g(\tilde{x})$, llegaremos a las condiciones suficientes del mismo extremo, pero expresadas esta vez sólo en términos de la función de Lagrange y las ecuaciones de enlace.

Observemos ante todo que en virtud de la condición (43.6), el sistema (43.31) es resoluble y, además, unívocamente respecto de $dx_1, \dots, dx_m$, siendo fijadas arbitrariamente $dx_{m+1}, \dots, dx_n$. El sistema (43.31) que expresa la igualdad a cero de las diferenciales de las funciones $f_i(x)$ en el punto $x^{(0)}$:

$$df_i = 0, i = 1, 2, \dots, m,$$

cuando se cumplen las condiciones (43.3), se escribirá en la forma breve así:

$$df = 0, \tag{43.37}$$

donde $f = (f_1, \dots, f_m)$.

Sea $x^{(0)}$ un punto estacionario para la función de Lagrange $F(x)$ (véase (43.13)). Esto significa que $dF(x^{(0)}) = 0$, es decir, que en este punto $\nabla f_0 + \sum_{i=1}^{m} \nabla f_i = 0$.

En el teorema 2, p. 43.4*, se ha mostrado que en este caso $\hat{x}^{(0)}$ es un punto estacionario para la función $g(\hat{x})$, es decir,

$$dg(\hat{x}^{(0)}) = 0. \tag{43.38}$$

Expliquemos una vez más la deducción de esta fórmula y mostremos que

$$d^2g(\hat{x}^{(0)}) = d^2F(x^{(0)})|_{df = 0}. \tag{43.39}$$

Esta igualdad debe entenderse como una igualdad de las funciones de $n - m$ variables $dx_{m+1}, \dots, dx_n$. En el segundo miembro de la igualdad (43.39) las variables restantes $dx_1, \dots, dx_m$, que figuran en las expresiones de las diferenciales escritas, se determinan a partir del sistema de ecuaciones (43.37) o bien, que es lo mismo (véanse las fórmulas (43.7)),

$$dx_k = d\varphi_k(x_1, \dots, x_{n-m}), k = 1, 2, \dots, m.$$

Aprovechando la invariación de la forma de la primera diferencial respecto de la elección de las variables y la fórmula (43.8), tenemos

$$dg(\hat{x}^{(0)}) = \sum_{j=1}^{n} \frac{\partial f_0(x^{(0)})}{\partial x_j} dx_j.$$

Sumemos a esta igualdad la suma (igual a cero) de los primeros miembros de las identidades (43.31) multiplicadas, respectivamente, por las constantes λ_i que integran la función de Lagrange $F(x)$ (con más precisión, la i-ésima igualdad (43.31) se multiplica por la constante λ_i). Luego, al hacer uso de la condición (43.13), obtendremos

$$dg(\hat{x}^{(0)}) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left[f_0(x) + \sum_{i=1}^{n} \lambda_i f_i(x)\right] dx_j\Big|_{x = x^{(0)}} =$$

$$= \sum_{j=1}^{n} \frac{\partial F(x^{(0)})}{\partial x_j} dx_j = 0.$$

La afirmación (43.38) queda demostrada.

La igualdad (43.39) se demuestra del modo análogo. Escribamos, ante todo, la segunda diferencial para la función $g(\hat{x})$ en el punto $\hat{x}^{(0)}$:

$$d^2g(\hat{x}^{(0)}) = \sum_{j,k=1}^{n} \frac{\partial^2 f_0(x^{(0)})}{\partial x_j \partial x_k} dx_j dx_k + \sum_{j=1}^{n} \frac{\partial f_0(x^{(0)})}{\partial x_j} d^2x_j. \tag{43.40}$$

Luego, al haber derivado las identidades que se obtienen como resultado de derivar las ecuaciones de enlace (43.3), es decir, las identidades

$$\frac{\partial f_i}{\partial x_1} dx_1 + \dots + \frac{\partial f_i}{\partial x_n} dx_n = 0, i = 1, 2, \dots, n,$$

tendremos en el punto $x^{(0)}$:

$$\sum_{j,k=1}^{n} \frac{\partial^2 f_i(x^{(0)})}{\partial x_j \partial x_k} dx_j dx_k + \sum_{j=1}^{n} \frac{\partial f_i(x^{(0)})}{\partial x_j} d^2x_j = 0, i = 1, 2, \dots, m. \tag{43.41}$$

Al multiplicar la i-ésima igualdad (43.41) por la constante λ_i, que figura en la función de Lagrange $F(x)$, sumemos las expresiones obtenidas al segundo miembro de la igualdad (43.40); en este caso tendremos

$$d^2g(\hat{x}^{(0)}) = \sum_{j,k=1}^{n} \frac{\partial^2 F(x^{(0)})}{\partial x_j \partial x_k} dx_j dx_k + \sum_{j=1}^{n} \frac{\partial F(x^{(0)})}{\partial x_j} d^2x_j,$$

donde dx_i, $i = 1, 2, \ldots, n$, satisface el sistema de ecuaciones (43.37). Por cuanto el punto $x^{(0)}$ es estacionario para la función de Lagrange, el segundo miembro de la igualdad obtenida se reduce a cero y de este modo la fórmula (43.39) queda demostrada.

Diremos que la forma cuadrática $d^2F(x^{(0)})$ es cuadrática definida positiva (negativa) de las variables dx_i, $i = 1, 2, \ldots, n$, a condición de que estas variables satisfacen el sistema de ecuaciones (43.37) siempre que para cualesquiera dx_i, $i = 1, 2, \ldots, n$, que satisfacen este sistema de ecuaciones y son tales que $\sum_{i=1}^{n} (dx_i)^2 > 0$, se verifica la desigualdad $d^2F(x^{(0)}) > 0$ ($d^2F(x^{(0)}) < 0$, respectivamente).

Supongamos que el punto $x^{(0)}$ satisface las ecuaciones de enlace (43.3) y es estacionario para la función de Lagrange (43.13) y supongamos también que la segunda diferencial de la función de Lagrange en este punto es una forma cuadrática definida positiva (negativa) de las variables $dx_1, \ldots, dx_n$ a condición de que ellas satisfacen el sistema de ecuaciones (43.37). Entonces, de (43.38) y (43.39) proviene que $\hat{x}^{(0)}$ es un punto estacionario para la función $g(\hat{x})$ y que la segunda diferencial de esta función en el punto $\hat{x}^{(0)}$ es una forma cuadrática definida positiva (negativa) de las variables $dx_{m+1}, \ldots, dx_n$, y, por consiguiente, la función $g(\hat{x})$ tiene en el punto $\hat{x}^{(0)}$ un mínimo (máximo) estricto; quiere decir que la función $f_0(x)$ admite en el punto $x^{(0)}$ un mínimo (máximo) estricto condicionado respecto de las ecuaciones de enlace (43.3).

Enunciemos el resultado obtenido en forma del teorema.

Teorema 3. *Si $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$ satisface las ecuaciones de enlace* (43.3) *y es un punto estacionario para la función de Lagrange* (43.13) *y si la segunda diferencial de la función de Lagrange en este punto es la forma cuadrática definida positiva (negativa) de las variables $dx_1, \ldots, dx_n$ a condición de que éstas satisfacen el sistema de ecuaciones* (43.31), *entonces $x^{(0)}$ es un punto de mínimo (máximo) estricto condicionado para la función f_0 respecto de las ecuaciones de enlace* (43.3).

De este modo, con el fin de investigar el extremo condicionado de un punto estacionario de la función de Lagrange (43.13), se debe analizar si es o no definida la forma cuadrática (43.39), es decir, examinar la segunda diferencial de la función de Lagrange en este punto, siendo cumplidas las condiciones de enlace (43.3) (cuando las diferenciales dx_i, $i = 1, 2, \ldots, n$, están ligadas entre sí por las relaciones (43.31)). En este caso hemos de tener en cuenta que si la segunda diferencial de la función de Lagrange en el punto que se considera resulta ser definida positiva (negativa) sin que se cumplan las condiciones de enlace, lo será, por supuesto, cuando dichas condiciones se cumplan.

Supongamos, por ejemplo, que se requiere hallar los puntos de extremo de la función $f(x, y) = xy$, cuando el punto (x, y) se dispone en la recta $x - y = 0$. La función de Lagrange será, en este caso, $F(x, y) = xy - \lambda(x - y)$, y, como $\frac{\partial F}{\partial x} = y - \tilde{\lambda}$, $\frac{\partial F}{\partial y} = x + \lambda$, entonces para la determinación de los puntos estacionarios de la función $F(x, y)$ que satisfagan las condiciones de enlace tenemos un sistema de ecuaciones

$$\begin{aligned} x - y &= 0, \\ y - \lambda &= 0, \\ x + \lambda &= 0, \end{aligned}$$

de las cuales se desprende que $x = y = \lambda = 0$.

Investiguemos en el punto $(0, 0)$ la segunda diferencial de la función $F(x, y)$, siendo cumplidas las condiciones de enlace, es decir, cuando $dx - dy = 0$. Tenemos

$$d^2F = 2dx\,dy, \tag{43.42}$$

y, por consiguiente, si se cumplen las condiciones de enlace,

$$d^2F = dx^2 \geqslant 0, \tag{43.43}$$

es decir, la segunda diferencial (43.42), representando una forma cuadrática indefinida, se convierte, siempre que se cumplan las condiciones de enlace, en una forma cuadrática definida positiva (43.43). Por eso $(0, 0)$ es el punto de mínimo estricto condicionado para el problema que acabamos de considerar. Esto, además, se ve claramente de lo siguiente: a lo largo de la recta $x - y = 0$ la función $f(x, y) = xy$ tomará la forma $f(x, x) = x^2$, admitiendo, evidentemente, en el punto $x = 0$ un mínimo estricto.

Ejercicios: Hállence los puntos de extremo condicionado de las funciones con las ecuaciones indicadas de enlace:

1. $z = \frac{x}{a} + \frac{y}{b}$, $x^2 + y^2 = 1$.
2. $z = x^2 + y^2$, $\frac{x}{a} + \frac{y}{b} = 1$.
3. $u = xyz$, $x^2 + y^2 + z^2 = 1$, $x + y + z = 0$.
4. $u = x^2 + y^2 + z^2$, $Ax + By + Cz + D = 0$.
5. $u = 2x + y - z + 1$, $x^2 + y^2 - 2z^2 - 22 = 0$.
6. Hállese el valor máximo de la función $z = xy$ en el círculo unitario $x^2 + y^2 \leqslant 1$.
7. Inscríbase en un círculo de radio dado un n-ágano de área máxima.
8. Represéntese el número $a > 0$ en forma de una suma de los sumandos $x_1, \ldots, x_n$ de un modo tal que el producto $x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$ $(\alpha_i > 0, i = 1, \ldots, n)$ asuma el valor máximo.