Ньютона — Лейбница для абстрактных функций (см. формулу п. 6). Получим

$$F(x+h)-F(x)=\int_0^1 F'(x+th)\,h\,dt=$$

$$=\int_0^1\left\{F'(x)+tF''(x)\,h+\frac{t^2}{2!}F'''(x)(h,h)+\ldots+\right.$$

$$\left.+\frac{1}{(n-1)!}t^{n-1}F^{(n)}(x)(h,h,\ldots,h)\right\}h\,dt+R_n,$$

где

$$R_n=\int_0^1 r_1(x,th)\,h\,dt.$$

Здесь мы воспользовались тем, что по определению интеграла от абстрактных функций справедливо равенство $\int_x^{x+h} F'(x+h)\,dh=$ $=\int_0^1 F'(x+th)\,h\,dt$, а также формулой для отображения $F'(x+th)$.

Таким образом *,

$$F(x+h)-F(x)=F'(x)\cdot h+\frac{1}{2!}F''(x)(h,h)+\ldots+$$

$$+\frac{1}{n!}F^{(n)}(x)(h,h,\ldots,h)+R_n,$$

где для R_n в силу свойств 3 интеграла (см. п. 5) справедлива оценка

$$\|R_n\|\leqslant\left\|\int_0^1 r_1(x,th)\,h\,dt\right\|\leqslant\int_0^1\|r_1(x,th)\|\,\|h\|\,dt=o(\|h\|^n).$$

Следовательно, формула Тейлора для отображения $F(x)$ одного нормированного пространства N_1 в другое нормированное пространство N_2 установлена. Теорема доказана.

---

* Ниже мы пользуемся тем, что выражение $F''(x)h$, примененное к h, есть квадратичное выражение $F''(x)(h,h)$, $F''(x)\in(N_1^2\to N_2)$, и действует на пару векторов (h,h), переводя их в элемент пространства N_2; аналогично $F^{(n)}(x)\underbrace{(h,h,\ldots,h)}_{n-1}$ примененное к h, есть уже n-линейное отображение вида $F^{(n)}(x)\underbrace{(h,h,\ldots,h)}_{n}$.

Исследование на экстремум функционалов в нормированных пространствах

В главе, посвященной дифференциальному исчислению функции одной переменной, и в § 6 гл. 12 мы уже занимались задачей об отыскании экстремумов функции одной или нескольких переменных. Однако ряд задач, имеющих важное практическое применение, приводит к отысканию экстремумов функционалов, определенных на нормированном пространстве. Так, например, еще в 1696 г. Иоганн Бернулли опубликовал письмо, в котором предлагал вниманию математиков задачу о линии наиболее быстрого ската — *брахистохроне*. В этой задаче требуется найти линию, соединяющую две точки A_1 и A_2, не лежащие на одной вертикальной прямой, и обладающую тем свойством, что материальная точка скатится по этой линии из точки A_1 в точку A_2 в кратчайшее время. Эта задача приводит к исследованию на экстремум функционала $F(x)$, определенного на нормированном пространстве $C_1[a, b]$ непрерывно дифференцируемых функций (см. дополнение 1) вида

$$F(x)=\int_a^b \Phi(x(t), x'(t))\,dt,$$

где $x(t)$ — элемент пространства $C_1[a, b]$. Оказалось, что линией наиболее быстрого ската является циклоида (см. § 7 гл. 5)

$$t=A(\tau-\sin\tau),$$
$$x=A(1-\cos\tau).$$

Основным принципом в механике является *принцип стационарного действия Остроградского — Гамильтона*, утверждающий, что среди возможных, т. е. совместимых со связями, движений системы материальных точек в действительности осуществляется движение, которое можно определить, исследовав на экстремум функционал

$$F=\int_a^b (T-U)\,dt,$$

где T — кинетическая, а U — потенциальная энергия системы.

Изучение таких задач, в которых надо исследовать на экстремум определенный функционал, составляет содержание так называемого вариационного исчисления. В настоящем пункте мы коснемся лишь некоторых основных фактов этой дисциплины.

1. Необходимое условие экстремума. Пусть $F(x)$ — некоторый действительный функционал, определенный на банаховом пространстве B, т. е. $F(x)$ — отображение полного нормированного пространства B в действительную числовую ось.

Определение. *Функционал $F(x)$ достигает в точке x_0 локального минимума (максимума), если для всех x, находящихся в достаточно малой окрестности точки x_0 (т. е. для всех x, для которых норма разности $x-x_0$ достаточно мала), выполнено неравенство $F(x)-F(x_0)\geqslant 0\,(\leqslant 0)$. Локальный минимум и локальный максимум объединяются общим наименованием «локальный экстремум».*

Для функций m переменных мы установили в § 6 следующее необходимое условие экстремума: если функция $f(x_1, x_2, \ldots, x_m)$ дифференцируема в точке $\overset{\circ}{x}_0=(\overset{\circ}{x}_1, \overset{\circ}{x}_2, \ldots, \overset{\circ}{x}_m)$ и имеет в этой точке экстремум, то в этой точке либо $df=0$, либо, что то же,

$$\frac{\partial f}{\partial x_1}=\frac{\partial f}{\partial x_2}=\ldots=\frac{\partial f}{\partial x_m}=0.$$

Для функционалов на произвольном нормированном пространстве справедлива аналогичная теорема.

Теорема. *Для того чтобы дифференцируемый на нормированном пространстве функционал F достигал в точке x_0 экстремума, необходимо, чтобы его дифференциал в этой точке равнялся нулю тождественно относительно h:*

$$dF=F'(x_0)h\equiv 0,$$

т. е. необходимо, чтобы $F'(x_0)=0$.

Доказательство. Поскольку функционал F дифференцируем, то по определению дифференцируемости

$$F(x_0+h)-F(x_0)=F'(x_0)h+o(h).$$

Если $F'(x_0)h\neq 0$ для некоторого элемента h, то для всех достаточно малых действительных чисел λ знак всего выражения $F'(x_0)(\lambda h)+o(\lambda h)$ совпадает со знаком его главного члена $F'(x_0)(\lambda h)$, так как если λ мало, то величина $\|\lambda h\|$ также мала и $o(\lambda h)$ есть величина более высокого порядка малости по сравнению с $F'(x_0)(\lambda h)$. Функционал $F'(x_0)$ линейный, поэтому $F'(x_0)(\lambda h)=\lambda F'(x_0)h$. Следовательно, если $F'(x_0)h\neq 0$, то приращение

$$F(x_0+\lambda h)-F(x_0)=F'(x_0)(\lambda h)+o(\lambda h)$$

(при данном h, для которого $F'(x_0)h\neq 0$) может принимать при сколь угодно малых по норме элементах λh как положительные, так и отрицательные значения, т. е. экстремума в точке x_0 быть не может. Теорема доказана.

Пример. Рассмотрим функционал

$$F(x)=\int_a^b f(t, x(t), x'(t))\,dt,$$

заданный на пространстве $C_1[a, b]$ непрерывно дифференцируе-

мых на отрезке $[a, b]$ функций. Здесь $x'(t)=\frac{dx(t)}{dt}$, $f(t, x(t), x'(t))$ — дважды дифференцируемая функция своих аргументов на множестве $a\leqslant t\leqslant b$, $-\infty<x<+\infty$, $-\infty<x'<\infty$. Этот функционал играет важную роль в вопросах вариационного исчисления.

Согласно предыдущей теореме, для того чтобы найти те точки, в которых возможен экстремум, т. е. найти стационарные точки, необходимо найти сильный дифференциал этого функционала и приравнять его нулю. По формуле Тейлора с остаточным членом в форме Пеано для функции $f(t, x, x')$ получим, что

$$F(x+h)-F(x)=\int_a^b [f(t, x+h, x'+h')-f(t, x, x')]\,dt=$$

$$=\int_a^b [f'_x(t, x, x')\,h(t)+f'_x(t, x, x')\,h'(t)+o(\rho)]\,dt,$$

где $\rho=[h^2+h'^2]^{1/2}$. Напомним, что норма элемента h в пространстве $C_1[a, b]$ определяется по правилу

$$\|h\|_{C_1[a,b]}=\max\{\max_{a\leqslant t\leqslant b}|h(t)|,\ \max_{a\leqslant t\leqslant b}|h'(t)|\}.$$

Поэтому

$$\left|\int_a^b o(\rho)\,dt\right|\leqslant(b-a)\cdot o(\|h\|_{C_1[a,b]})=o(\|h\|_{C_1[a,b]}).$$

Итак, необходимое условие экстремума для данного функционала имеет вид ($h(t)$ — произвольное приращение аргумента $F(x)$)

$$dF=\int_a^b [f'_x(t, x, x')\,h(t)+f'_{x'}(t, x, x')\,h'(t)]\,dt=0.$$

На самом деле, можно убедиться, что это условие эквивалентно следующим:

$$f'_x-\frac{d}{dt}f'_{x'}=0,\quad f'_{x'}|_{t=a}=0,\quad f'_{x'}|_{t=b}=0.$$

Это так называемая *краевая задача для функции* $x(t)$ определяет некоторую кривую $x(t)$, на которой возможно достижение функционалом экстремума. Указанная краевая задача состоит из дифференциального уравнения для функции $x(t)$ (уравнения, содержащего функцию $x(t)$ и ее производные) вида $f'_x(t, x(t), x'(t))-\frac{d}{dt}f'_{x'}(t, x(t), x'(t))=0$, которое называется **урав-**

нением Эйлера для функционала $F=\int_a^b f(t, x, x')dt$ и краевых условий вида $f'_{x'}(t, x(t), x'(t))|_{t=a}=0$, $f'_{x'}(t, x(t), x'(t))|_{t=b}=0$.

2. Достаточные условия экстремума. Напомним, что для функции m переменных $f(x_1, x_2, ..., x_m)$ справедливы следующие утверждения.

Если функция $f(x_1, x_2, ..., x_m)$ имеет в точке $(\mathring{x}_1, \mathring{x}_2, ..., \mathring{x}_m)$ локальный минимум (максимум), то в этой точке $d^2f \geqslant 0 (\leqslant 0)$.

Если в точке $(\mathring{x}_1, \mathring{x}_2, ..., \mathring{x}_m)$ выполнены условия:

$$1)\ df=0 \text{ и } 2)\ d^2f=\sum_{i,k=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_k} dx_i dx_k > 0 \ (d^2f < 0),$$

когда не все $dx_i=0$, то в этой точке $f(x)$ имеет локальный минимум (максимум).

Посмотрим, в какой мере эти факты переносятся на функционалы, определенные на банаховом пространстве.

Теорема. *Пусть $F(x)$ — функционал, принимающий действительные значения и заданный в банаховом пространстве B. Пусть $F(x)$ имеет в некоторой окрестности точки x_0 непрерывную вторую производную. Если этот функционал $F(x)$ достигает в точке x_0 локального минимума (максимума), то*

$$d^2F(x_0) \geqslant 0 \quad (d^2F(x_0) \leqslant 0).$$

Доказательство. Заметим прежде всего, что $d^2F(x_0)=F'(x_0)(h, h)$. Поэтому неравенство $d^2F(x_0) \geqslant 0$ означает, что $F'(x_0)(h, h) \geqslant 0$ для всех $h \in B$. По формуле Тейлора получаем

$$F(x_0+h)-F(x_0)=F'(x_0)h+\frac{1}{2}F''(x_0)(h, h)+o(\|h\|^2).$$

Допустим для определенности, что в точке x_0 функционал $F(x)$ имеет локальный минимум. Тогда $F'(x_0)=0$ и, следовательно,

$$F(x_0+h)-F(x_0)=\frac{1}{2}F''(x_0)(h, h)+o(\|h\|^2).$$

Если при каком-либо h выполнено противоположное неравенство $F''(x_0)(h, h)<0$, то, поскольку $F''(x_0)(\lambda h, \lambda h)=\lambda^2 F''(x_0)(h, h)$ при любых действительных λ, существуют сколь угодно малые по норме элементы $g=\lambda h$ *, для которых $F''(x_0)(g, g)<0$. Но при достаточно малых $\|g\|$ знак выражения $F(x_0+g)-F(x_0)=$

* Действительно, $\|g\|=|\lambda|\|h\|$, и при малых по модулю числах λ и фиксированном элементе h норма элемента g мала.

$=\frac{1}{2}F''(x_0)(g, g)+o(\|g\|^2)$ определяется знаком главного члена $\frac{1}{2}F''(x_0)(g, g)$, и мы получаем, что

$$F(x_0+g)-F(x_0)=\frac{1}{2}F''(x_0)(g, g)+o(\|g\|^2)<0,$$

т. е. минимума в точке x_0 нет.

Аналогичные рассуждения проводятся и для того случая, когда в точке x_0 функционал имеет локальный максимум. Теорема доказана.

Как видим, эта теорема есть прямое обобщение соответствующего утверждения о необходимом условии экстремума для случая функции m переменных.

Однако если речь идет о достаточных условиях экстремума функционалов, то полная аналогия со случаем функции m переменных исчезает.

Условие $d^2f>0$ положительной определенности второго дифференциала, достаточное для минимума в стационарной точке в случае функции $f(x_1, x_2, ..., x_m)$ m переменных, *не является достаточным для минимума в стационарной точке функционалов, определенных на произвольных банаховых пространствах.*

Для примера рассмотрим пространство l_2 всевозможных последовательностей $(x_1, x_2, ...)$ действительных чисел $x_1, x_2, ...$ таких, что $\sum_{k=1}^{\infty}|x_k|^2=\lim_{n\to\infty}\sum_{k=1}^{n}|x_k|^2<\infty$. Сумма двух последовательностей $x=(x_1, x_2, ...)$ и $y=(y_1, y_2, ...)$ определяется так: $x+y=(x_1+y_1, x_2+y_2, ...)$; произведение последовательности $x=(x_1, x_2, ...)$ на число λ определяется по правилу: $\lambda x=(\lambda x_1, \lambda x_2, ...)$. Из элементарного неравенства $|x_k\pm y_k|^2\leqslant 2|x_k|^2+2|y_k|^2$ следует, что элемент $x+y$ принадлежит нашему пространству, если x и y ему принадлежат, поскольку $\sum_{k=1}^{\infty}|x_k+y_k|^2\leqslant 2\sum_{k=1}^{\infty}|x_k|^2+2\sum_{k=1}^{\infty}|y_k|^2<\infty$. Таким образом, введенное пространство *линейно.*

Если ввести в этом линейном пространстве *норму*, положив по определению, что

$$\|x\|^2=\sum_{k=1}^{\infty}|x_k|^2,$$

то введенное нами пространство l_2 становится нормированным. Аксиомы нормы проверяются здесь без труда. В частности, ак-

сиома треугольника следует из того, что при каждом натуральном n справедливо неравенство

$$\left(\sum_{k=1}^{n} |x_k \pm y_k|^2\right)^{1/2} \leqslant \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2} + \left(\sum_{k=1}^{n} |y_k|^2\right)^{1/2}$$

(см. дополнение 2). Переходя здесь к пределу при $n \to \infty$, получаем неравенство треугольника.

Рассмотрим в данном нормированном пространстве l_2 функционал

$$F(x) = \sum_{k=1}^{\infty} \frac{x_k^2}{k^3} - \sum_{k=1}^{\infty} x_k^4.$$

Вычислим первый и второй дифференциалы этого функционала в точке 0. Имеем

$$F(x+h) - F(x) = \sum_{k=1}^{\infty} \frac{(x_k + h_k)^2}{k^3} - \sum_{k=1}^{\infty} (x_k + h_k)^4 - \sum_{k=1}^{\infty} \frac{x_k^2}{k^3} -$$

$$- \sum_{k=1}^{\infty} x_k^4 = 2 \sum_{k=1}^{\infty} \left(\frac{x_k}{k^3} - 2x_k^3\right) h_k + \ldots,$$

где многоточием обозначены члены, не являющиеся линейными по $h = (h_1, h_2, \ldots)$. Таким образом, в точке $x = 0$ первый дифференциал равен нулю. Вычисляя, аналогично находим, что второй дифференциал в точке 0 $d^2F = 2 \sum_{k=1}^{\infty} \frac{h_k^2}{k^3} > 0$ положительно определен. Тем не менее в точке 0 минимума нет. Действительно, $F(0) = 0$, $F\left(0, \ldots, 0, \frac{1}{k}, 0, \ldots\right) = \frac{1}{k^5} - \frac{1}{k^4} < 0$, и поэтому $F\left(0, 0, \ldots, 0, \frac{1}{k}, 0, \ldots\right) - F(0, 0, \ldots) < 0$. Поскольку элемент $x = \left(0, 0, \ldots, 0, \frac{1}{k}, 0, \ldots\right)$ может быть по норме l_2 сделан сколь угодно малым $\left(\|x\|^2 = \frac{1}{k^2}\right)$, где k — произвольно большое натуральное число), то в любой сколь угодно малой окрестности точки 0 существуют точки x, в которых $F(x) < F(0)$. Следовательно, в точке 0 минимума нет.

Введем следующее определение.

Определение. *Квадратичный функционал $B(x, x)$ называется сильно положительным, если существует такое*

положительное число γ, *что* $B(x, x) \geqslant \gamma \|x\|^2$ *для всех* x. *Если* $B(x, x) > 0$ *для всех ненулевых* x, *то функционал называется* *положительно определенным.*

Очевидно, что *сильно положительный квадратичный функционал* $B(x, x)$ *является положительно определенным. Обратное, вообще говоря, неверно.*

Теорема. *Если дважды дифференцируемый функционал* $F(x)$, *определенный в банаховом пространстве* B, *удовлетворяет в точке* $x_0 \in B$ *условиям:* $dF(x_0) = 0$ *и* $d^2F(x_0)$ — *сильно положительный квадратичный функционал, то* $F(x)$ *имеет в точке* x_0 *локальный минимум.*

Доказательство. Пусть $F''(x_0)(h, h) \geqslant \gamma \|h\|^2$. Выберем $\varepsilon > 0$ настолько малым, чтобы при $\|h\| < \varepsilon$ величина $o(\|h\|^2)$ в равенстве $F(x_0 + h) - F(x_0) = \frac{1}{2} F''(x_0)(h, h) + o(\|h\|^2)$ удовлетворяла условию $\|o(\|h\|^2)\| < \frac{\gamma}{4} \|h\|^2$. Тогда

$$F(x_0 + h) - F(x_0) = \frac{1}{2} F''(x_0)(h, h) + o(\|h\|^2) > \frac{\gamma}{4} \|h\|^2 > 0$$

при $\|h\| < \varepsilon$. Следовательно, в точке x_0 имеется локальный минимум функционала $F(x)$. Теорема доказана.

Эта теорема вместе с приведенным выше примером функционала в l_2 показывает, в частности, что *сильная положительность есть более сильное условие, чем положительная определенность квадратичного функционала.* Заметим, что *в конечномерном пространстве сильная положительность квадратичной формы эквивалентна ее положительной определенности.*

Заканчивая дополнение 3, подчеркнем, что предыдущая теорема устанавливает достаточное условие минимума, которое является весьма общим и вместе с тем трудно проверяемым в ряде задач. Поэтому в вариационном исчислении, исходя из конкретного вида функционала, устанавливают другие более тонкие достаточные условия экстремумов функционала. На этом вопросе в данном курсе мы останавливаться не будем.

Глава 13

НЕЯВНЫЕ ФУНКЦИИ

Во многих вопросах естествознания приходится сталкиваться с такой ситуацией, когда переменная u, являющаяся по смыслу функцией аргументов $x_1, x_2, \ldots$, задается посредством функционального уравнения $F(u, x_1, x_2, \ldots) = 0$.

В этом случае говорят, что u как функция аргументов $x_1, x_2, \ldots$ задана неявно. Естественно, возникает вопрос о том, при каких условиях функциональное уравнение $F(u, x_1, x_2, \ldots) = 0$ однозначно разрешимо относительно u, т. е. однозначно определяет явную функцию $u = \varphi(x_1, x_2, \ldots)$, и более тонкий вопрос о том, при каких условиях эта явная функция непрерывна и дифференцируема.

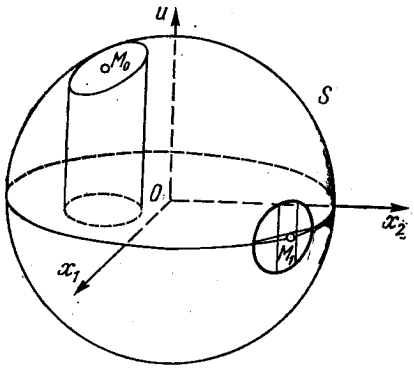

Рис. 13.1

Указанные вопросы не являются простыми. В качестве примера обратимся к функциональному уравнению $F(u, x_1, x_2) = u^2 + x_1^2 + x_2^2 - 1 = 0$, определяющему в пространстве переменных (u, x_1, x_2) сферу S радиуса 1 с центром в начале координат (рис. 13.1). Очевидно, указанное функциональное уравнение определяет в круге $x_1^2 + x_2^2 \leqslant 1$ бесконечно много явных функций. Таковыми являются функция $u = +\sqrt{1 - x_1^2 - x_2^2}$, функция $u = -\sqrt{1 - x_1^2 - x_2^2}$ и любая функция, равная $+\sqrt{1 - x_1^2 - x_2^2}$ для одних точек круга $x_1^2 + x_2^2 \leqslant 1$ и равная $-\sqrt{1 - x_1^2 - x_2^2}$ для других точек этого круга.

При каких же условиях существует единственная явная функция, удовлетворяющая уравнению $F(u, x_1, x_2) = u^2 + x_1^2 + x_2^2 - 1 = 0$? Фиксируем на сфере S произвольную точку $M_0(\overset{\circ}{u}, \overset{\circ}{x}_1, \overset{\circ}{x}_2)$, не лежащую в плоскости Ox_1x_2, т. е. такую, что $\overset{\circ}{u} \neq 0$. Очевидно, что часть сферы S, лежащая в достаточно малой окрестности точки M_0, однозначно проектируется на плоскость Ox_1x_2 (см. рис. 13.1). Аналитически это означает, что если рассматривать функцию $F(u, x_1, x_2) = u^2 + x_1^2 + x_2^2 - 1$ только в указанной достаточ-

но малой окрестности точки M_0, то уравнение $F(u, x_1, x_2)=u^2+x_1^2+x_2^2-1=0$ однозначно разрешимо относительно u и определяет единственную явную функцию $u=+\sqrt{1-x_1^2-x_2^2}$ при $\mathring{u}>0$ и соответственно $u=-\sqrt{1-x_1^2-x_2^2}$ при $\mathring{u}<0$.

Если же на сфере S взять точку $M_1(0, x_1, x_2)$, лежащую в плоскости Ox_1x_2 (см. рис. 13.1), то очевидно, что часть сферы S, лежащая в любой окрестности M_1, неоднозначно проектируется на плоскость Ox_1x_2. Аналитически это означает, что если рассматривать функцию $F(u, x_1, x_2)=u^2+x_1^2+x_2^2-1$ в любой окрестности точки M_1, то уравнение $F(u, x_1, x_2)=u^2+x_1^2+x_2^2-1=0$ не является однозначно разрешимым относительно u. Обратим внимание на то, что частная производная $\frac{\partial F}{\partial u}=2u$ функции $F(u, x_1, x_2)=u^2+x_1^2+x_2^2-1$ не обращается в нуль в точке M_0 и обращается в нуль в точке M_1. Ниже мы установим, что для однозначной разрешимости в окрестности точки M_0 общего функционального уравнения $F(u, x_1, x_2, \ldots)=0$ относительно u принципиальную роль играет необращение в нуль в точке M_0 частной производной $\frac{\partial F}{\partial u}$. Попутно мы установим условия, при которых явная функция, представляющая собой единственное решение этого уравнения является непрерывной и дифференцируемой.

Полученные нами результаты будут перенесены на случай неявных функций, определяемых системой функциональных уравнений, а также на случай абстрактных функций, осуществляющих отображения произвольных банаховых пространств.

§ 1. СУЩЕСТВОВАНИЕ И ДИФФЕРЕНЦИРУЕМОСТЬ НЕЯВНО ЗАДАННОЙ ФУНКЦИИ

1. Теорема о существовании и дифференцируемости неявной функции. В дальнейшем мы договоримся обозначать пространство переменных $(u, x_1, x_2, \ldots)$ символом R, а пространство переменных $(x_1, x_2, \ldots)$ символом R'. Для сокращения записи и для удобства геометрической иллюстрации будем рассматривать две переменные x_1 и x_2.

Теорема 13.1. *Пусть функция $F(u, x_1, x_2)$ дифференцируема в некоторой окрестности точки $M_0(\mathring{u}, \mathring{x}_1, \mathring{x}_2)$ пространства R, причем частная производная $\partial F/\partial u$ непрерывна в точке M_0. Тогда если в точке M_0 функция F обращается в нуль, а частная производная $\partial F/\partial u$ не обращается в нуль, то для любого достаточно малого положительного числа ε найдется такая окрестность точки $M_0'(\mathring{x}_1, \mathring{x}_2)$ пространства R', что в пределах этой окрестности существует единственная функция $u=\varphi(x_1, x_2)$, которая удовлетво-*

ряет условию $|u-\mathring{u}|<\varepsilon$ *и является решением уравнения*

$$F(u, x_1, x_2)=0, \tag{13.1}$$

причем эта функция $u=\varphi(x_1, x_2)$ *непрерывна и дифференцируема в указанной окрестности точки* M_0'.

Замечание 1. В условиях теоремы 13.1 можно опустить требование непрерывности частной производной $\partial F/\partial u$ в точке M_0, но тогда придется дополнительно потребовать, чтобы эта производная не обращалась в нуль не только в самой точке M_0, но и в некоторой ее окрестности и сохраняла определенный знак в этой окрестности.

Доказательство теоремы. 13.1. 1. Прежде всего докажем, что *для достаточно малого* $\varepsilon>0$ *в окрестности точки* $M_0'(x_1, x_2)$ *существует единственная функция* $u=\varphi(x_1, x_2)$, *удовлетворяющая условию* $|u-\mathring{u}|<\varepsilon$ *и являющаяся решением уравнения* (13.1). Чтобы сделать доказательство более наглядным, будем сопровождать его геометрической иллюстрацией. Из аналитической геометрии известно, что уравнение (13.1) определяет в пространстве R некоторую поверхность S (рис. 13.2), причем в силу условия $F(M_0)=0$ точка M_0 лежит на этой поверхности. С геометрической точки зрения однозначная разрешимость уравнения (13.1) относительно u означает, что часть поверхности S, лежащая в непосредственной близости к точке M_0, может быть однозначно спроектирована на плоскость Ox_1x_2.

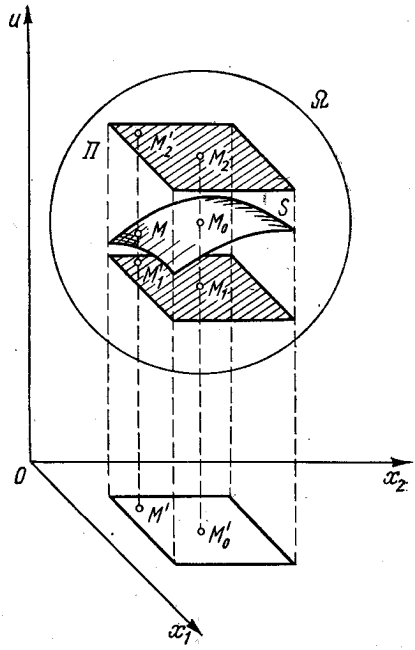

Рис. 13.2

Ради определенности будем считать, что частная производная $\partial F/\partial u$ положительна в точке M_0. Тогда из непрерывности указанной производной в M_0 и из теоремы об устойчивости знака непрерывной функции вытекает, что *найдется такая окрестность точки* M_0 *всюду в пределах которой* $\partial F/\partial u$ *положительна*. Эту окрестность мы можем взять в виде шара Ω достаточно малого радиуса с центром в точке M_0. Фиксируем далее положительное число ε настолько малым, чтобы каждая из точек $M_1(\mathring{u}-\varepsilon, \mathring{x}_1, \mathring{x}_2)$ и

$M_2(\mathring{u}+\varepsilon, \mathring{x}_1, \mathring{x}_2)$ лежала внутри шара Ω (для этого достаточно взять ε меньшим радиуса шара Ω). Подчеркнем, что при этом снизу ε ограничено лишь нулем, и мы можем брать его как угодно малым — это будет использовано нами ниже.

Рассмотрим функцию $F(u, \mathring{x}_1, \mathring{x}_2)$ одной переменной u на сегменте $\mathring{u}-\varepsilon \leqslant u \leqslant \mathring{u}+\varepsilon$. С геометрической точки зрения это означает, что мы рассматриваем функцию трех переменных $F(u, x_1, x_2)$ вдоль отрезка M_1M_2 (см. рис. 13.2). Так как производная $\frac{\partial F}{\partial u}(u, \mathring{x}_1, \mathring{x}_2)$ положительна на сегменте $\mathring{u}-\varepsilon \leqslant u \leqslant \mathring{u}+\varepsilon$, то функция $F(u, \mathring{x}_1, \mathring{x}_2)$ возрастает на этом сегменте. Но тогда, поскольку эта функция равна нулю в середине указанного сегмента (т. е. при $u=\mathring{u}$), то $F(u, \mathring{x}_1, \mathring{x}_2)$ имеет отрицательное значение на левом конце и положительное значение на правом конце указанного сегмента, т. е. $F(M_1)<0$, $F(M_2)>0$. Далее рассмотрим функции $F(\mathring{u}-\varepsilon, x_1, x_2)$ и $F(\mathring{u}+\varepsilon, x_1, x_2)$ двух переменных x_1 и x_2, т. е., выражаясь геометрическим языком, рассмотрим функцию $F(u, x_1, x_2)$ на двух плоскостях, параллельных координатной плоскости Ox_1x_2, первая из которых проходит через точку M_1, а вторая — через точку M_2. Поскольку $F(M_1)<0$, $F(M_2)>0$ и функция $F(u, x_1, x_2)$ непрерывна всюду в шаре Ω, то по теореме об устойчивости знака непрерывной функции на указанных плоскостях найдутся такие окрестности точек M_1 и M_2, в пределах которых функция F сохраняет те же знаки, что и в точках M_1 и M_2. Эти окрестности мы можем взять в виде открытых квадратов с центрами в точках M_1 и M_2 и с достаточно малой стороной 2δ (на рис. 13.2 указанные квадраты заштрихованы). Аналитически тот факт, что функция $F(u, x_1, x_2)$ сохраняет постоянный знак на указанных квадратах, выражается неравенствами

$$\left.\begin{aligned} F(\mathring{u}-\varepsilon, x_1, x_2) &< 0 \\ F(\mathring{u}+\varepsilon, x_1, x_2) &> 0 \end{aligned}\right\} \text{ при } |x_1-\mathring{x}_1|<\delta, \quad |x_2-\mathring{x}_2|<\delta. \tag{13.2}$$

Выборы стороны указанных квадратов мы подчиним и еще одному условию: возьмем δ столь малым, чтобы оба указанных квадрата лежали внутри шара Ω (это заведомо можно сделать, ибо центры квадратов M_1 и M_2 являются внутренними точками шара Ω). При таком выборе δ любая точка пространства (u, x_1, x_2), координаты которой удовлетворяют неравенствам

$$|x_1-\mathring{x}_1|<\delta, \ |x_2-\mathring{x}_2|<\delta, \ |u-\mathring{u}|<\delta, \tag{13.3}$$

будет лежать внутри шара Ω. С геометрической точки зрения неравенства (13.3) определяют открытый прямоугольный параллелепипед с центром в точке M_0 и со сторонами, параллельными осям координат u, x_1, x_2 и соответственно равными 2ε, 2δ и 2δ.

Этот параллелепипед мы будем обозначать символом П. Так как параллелепипед П лежит внутри шара Ω, то *всюду в параллелепипеде* П* *производная* $\partial F/\partial u$ *положительна.* Кроме того, в силу (13.2) *функция* $F(u, x_1, x_2)$ *отрицательна на нижнем основании и положительна на верхнем основании* П.

Докажем теперь, что уравнение (13.1) однозначно разрешимо относительно u, если функцию $F(u, x_1, x_2)$ рассматривать лишь для значений u, x_1, x_2, лежащих внутри параллелепипеда П. Пусть $M'(x_1, x_2)$ — любая точка пространства R', координаты которой удовлетворяют неравенствам

$$|x_1-\mathring{x}_1|<\delta, \quad |x_2-\mathring{x}_2|<\delta. \tag{13.4}$$

Иначе говоря, пусть $M'(x_1, x_2)$ — любая точка плоскости Ox_1x_2, лежащая внутри квадрата с центром в точке $M_0'(\mathring{x}_1, \mathring{x}_2)$ и со сторонами, равными 2δ. Требуется доказать, что для координат x_1, x_2 точки M' найдется, и притом единственное, число u из интервала $\mathring{u}-\varepsilon<u<\mathring{u}+\varepsilon$ такое, что $F(u, x_1, x_2)=0$. (С геометрической точки зрения это означает, что любая прямая, параллельная оси u и пересекающая параллелепипед П, пересекает поверхность S внутри параллелепипеда П в одной и только в одной точке.)

Зафиксировав значения x_1 и x_2, удовлетворяющие неравенствам (13.4), рассмотрим функцию $F(u, x_1, x_2)$ аргумента u на сегменте $\mathring{u}-\varepsilon\leqslant u\leqslant\mathring{u}+\varepsilon$, т. е. рассмотрим функцию $F(u, x_1, x_2)$ на отрезке $M_1'M_2'$, где M_1' и M_2' — точки пересечения прямой, проходящей через точку $M'(x_1, x_2)$ и параллельной оси Ou, с основаниями параллелепипеда П (см. рис. 13.2). Так как производная $\dfrac{\partial F}{\partial u}(u, x_1, x_2)$ положительна на сегменте $\mathring{u}-\varepsilon\leqslant u\leqslant\mathring{u}+\varepsilon$, то функция $F(u, x_1, x_2)$ возрастает на этом сегменте (или, что то же самое, возрастает на отрезке $M_1'M_2'$). Но тогда из условий $F(M_1')<0$, $F(M_2')>0$ вытекает, что внутри сегмента $\mathring{u}-\varepsilon\leqslant u\leqslant\mathring{u}+\varepsilon$ найдется одно-единственное значение u такое, что $F(u, x_1, x_2)=0$ (или, выражаясь геометрически, внутри отрезка $M_1'M_2'$ найдется единственная точка M, лежащая на S).

Пусть теперь функция $u=\varphi(x_1, x_2)$ символизирует то правило, посредством которого каждой точке $M'(x_1, x_2)$ из окрестности (13.4) ставится в соответствие единственное число u из интервала $\mathring{u}-\varepsilon<u<\mathring{u}+\varepsilon$, для которого $F(u, x_1, x_2)=0$. Мы доказали, что в окрестности (13.4) существует единственная функция $u=\varphi(x_1, x_2)$, удовлетворяющая условию $|u-\mathring{u}|<\varepsilon$ и являющаяся решением уравнения (13.1).

2. Докажем теперь, что *функция* $u=\varphi(x_1, x_2)$ *непрерывна в любой точке* $M'(x_1, x_2)$ *окрестности* (13.4). Так как для любой точки

* Включая открытые квадраты, лежащие в его основаниях.

$M'(x_1, x_2)$ из окрестности (13.4) выполнены те же условия *, что и для точки $M_0'(\mathring{x}_1, \mathring{x}_2)$, то достаточно доказать непрерывность функции $u=\varphi(x_1, x_2)$ лишь в точке $M_0'(\mathring{x}_1, \mathring{x}_2)$. Требуется доказать, что для любого достаточно малого положительного ε существует положительное число δ такое, что для любых x_1 и x_2, удовлетворяющих неравенствам $|x_1-\mathring{x}_1|<\delta$, $|x_2-\mathring{x}_2|<\delta$, справедливо неравенство $|u-\mathring{u}|<\varepsilon$, где $u=\varphi(x_1, x_2)$, $\mathring{u}=\varphi(\mathring{x}_1, \mathring{x}_2)$. Если взять в качестве ε то число, которое выбрано выше при рассмотрении п. 1, то существование ε обеспечивается неравенствами (13.3). Остается заметить, что в рассуждениях п. 1 положительное число ε может быть взято как угодно малым (это отмечалось в п. 1).

Тем самым непрерывность функции $u=\varphi(x_1, x_2)$ установлена. Запишем условие непрерывности функции $u=\varphi(x_1, x_2)$ в точке $M_0'(\mathring{x}_1, \mathring{x}_2)$ в разностной форме. Обозначая через Δu полное приращение функции $u=\varphi(x_1, x_2)$ в точке $M_0'(x_1, x_2)$, соответствующее приращениям аргументов Δx_1 и Δx_2, мы получим, что

$$\Delta u \to 0 \text{ при } \begin{cases} \Delta x_1 \to 0, \\ \Delta x_2 \to 0. \end{cases}$$

3. Остается доказать *дифференцируемость функции* $u=\varphi(x_1, x_2)$ *в любой точке* $M'(x_1, x_2)$ *окрестности* (13.4). В силу замечания, сделанного в п. 2, достаточно доказать дифференцируемость функции $u=\varphi(x_1, x_2)$ в самой точке $M_0'(\mathring{x}_1, \mathring{x}_2)$. Чтобы это сделать, вычислим полное приращение Δu функции $u=\varphi(x_1, x_2)$ в точке $M_0'(\mathring{x}_1, \mathring{x}_2)$, соответствующее приращениям аргументов Δx_1 и Δx_2. Поскольку $F(\mathring{u}, \mathring{x}_1, \mathring{x}_2)=0$ и $F(\mathring{u}+\Delta u, \mathring{x}_1+\Delta x_1, \mathring{x}_2+\Delta x_2)=0$, то полное приращение ΔF функции $F(u, x_1, x_2)$ в точке $M_0(\mathring{u}, \mathring{x}_1, \mathring{x}_2)$, соответствующее приращениям аргументов Δu, Δx_1 и Δx_2, равно нулю. Но в силу условия дифференцируемости функции $F(u, x_1, x_2)$ в точке $M_0(\mathring{u}, \mathring{x}_1, \mathring{x}_2)$ это полное приращение имеет вид

$$\Delta F = \left(\frac{\partial F}{\partial u} + \gamma\right)\Delta u + \left(\frac{\partial F}{\partial x_1} + \alpha\right)\Delta x_1 + \left(\frac{\partial F}{\partial x_2} + \beta\right)\Delta x_2. \quad (13.5)$$

Здесь *все частные производные* $\partial F/\partial u$, $\partial F/\partial x_1$ и $\partial F/\partial x_2$ *берутся в точке* $M_0(\mathring{u}, \mathring{x}_1, \mathring{x}_2)$;

$$\alpha, \beta \text{ и } \gamma \to 0 \text{ при } \begin{cases} \Delta u \to 0, \\ \Delta x_1 \to 0, \\ \Delta x_2 \to 0. \end{cases}$$

* Именно, любой точке $M'(x_1, x_2)$ из окрестности (13.4) соответствует точка $M(u, x_1, x_2)$ пространства R такая, что функция $F(u, x_1, x_2)$ обращается в нуль в точке M, дифференцируема в некоторой окрестности точки M и имеет в этой окрестности отличную от нуля частную производную $\partial F/\partial u$.

Из соотношения (13.5), учитывая, что $\Delta F=0$, получаем

$$\left(\frac{\partial F}{\partial u}+\gamma\right)\Delta u+\left(\frac{\partial F}{\partial x_1}+\alpha\right)\Delta x_1+\left(\frac{\partial F}{\partial x_2}+\beta\right)\Delta x_2=0. \qquad (13.6)$$

Согласно разностной форме условия непрерывности функции $u=\varphi(x_1, x_2)$ в точке $M_0'(\mathring{x}_1, \mathring{x}_2)$ можно утверждать, что $\Delta u\to 0$ при $\Delta x_1\to 0$, $\Delta x_2\to 0$. Таким образом, можно утверждать, что

$$\alpha,\ \beta \text{ и } \gamma\to 0 \text{ лишь при условии} \begin{cases}\Delta x_1\to 0,\\ \Delta x_2\to 0.\end{cases}$$

По условию теоремы частная производная $\partial F/\partial u$ *отлична от нуля в точке* M_0. Поскольку $\gamma\to 0$ при $\begin{cases}\Delta x_1\to 0,\\ \Delta x_2\to 0,\end{cases}$ то *при достаточно малых* Δx_1 *и* Δx_2 *выражение* $(\partial F/\partial u+\gamma)$ *не обращается в нуль.* В таком случае формулу (13.6) можно поделить на $(\partial F/\partial u+\gamma)$, в результате чего мы получим

$$\Delta u=\left(-\frac{\partial F/\partial x_1+\alpha}{\partial F/\partial u+\gamma}\right)\Delta x_1+\left(-\frac{\partial F/\partial x_2+\beta}{\partial F/\partial u+\gamma}\right)\Delta x_2. \qquad (13.7)$$

По теореме о пределе частного двух функций можем утверждать, что

$$-\frac{\partial F/\partial x_1+\alpha}{\partial F/\partial u+\gamma}=-\frac{\partial F/\partial x_1}{\partial F/\partial u}+\mu,\quad -\frac{\partial F/\partial x_2+\beta}{\partial F/\partial u+\gamma}=-\frac{\partial F/\partial x_2}{\partial F/\partial u}+\nu, \qquad (13.8)$$

где μ и $\nu\to 0$ при $\begin{cases}\Delta x_1\to 0,\\ \Delta x_2\to 0,\end{cases}$

Сопоставляя формулы (13.7) и (13.8), окончательно получим

$$\Delta u=\left(-\frac{\partial F/\partial x_1}{\partial F/\partial u}\right)\Delta x_1+\left(-\frac{\partial F/\partial x_2}{\partial F/\partial u}\right)\Delta x_2+\mu\Delta x_1+\nu\Delta x_2. \qquad (13.9)$$

Формула (13.9) доказывает дифференцируемость функции $u=\varphi(x_1, x_2)$ в точке $M_0'(\mathring{x}_1, \mathring{x}_2)$. Тем самым теорема 13.1 полностью доказана.

Замечание 2. Приведенное доказательство без всяких затруднений переносится на случай неявной функции, зависящей не от двух, а от любого конечного числа аргументов $x_1, x_2, \ldots, x_n$*. Случай двух аргументов x_1 и x_2 имеет лишь то преимущество, что допускает наглядную геометрическую иллюстрацию в пространстве (u, x_1, x_2).

2. Вычисление частных производных неявно заданной функции. Остановимся на вычислении частных производных функции, неявно заданной посредством уравнения (13.1). Пусть выполнены условия теоремы 13.1. Тогда для полного приращения функции $u=$

* И, в частности, от одного аргумента.

$=\varphi(x_1, x_2)$ справедливо представление (13.9). Это представление и теорема 13.9 позволяют утверждать, что частные производные функции $u=\varphi(x_1, x_2)$ определяются формулами

$$\frac{\partial u}{\partial x_1}=-\frac{\partial F/\partial x_1}{\partial F/\partial u}, \quad \frac{\partial u}{\partial x_2}=-\frac{\partial F/\partial x_2}{\partial F/\partial u}. \tag{13.10}$$

Аналогичные формулы справедливы и для случая, когда неявно заданная функция зависит не от двух, а от любого конечного числа аргументов $x_1, x_2, \ldots, x_n$. В этом случае

$$\frac{\partial u}{\partial x_k}=-\frac{\partial F/\partial x_k}{\partial F/\partial u} \quad (k=1, 2, \ldots, m). \tag{13.11}$$

Если мы хотим обеспечить существование у неявно заданной функции $u=\varphi(x_1, x_2)$ частных производных второго порядка, то, естественно, приходится усилить требования, наложенные на функцию $F(u, x_1, x_2)$ в теореме 13.1, именно приходится дополнительно требовать, чтобы функция $F(u, x_1, x_2)$ была два раза дифференцируема в рассматриваемой точке. В этих предположениях остановимся на вычислении частных производных второго порядка.

Введем полезное в дальнейшем понятие полной частной производной функции.

Предположим, что нам дана дифференцируемая функция трех аргументов $\Phi(u, x_1, x_2)$, причем один из этих аргументов u сам является дифференцируемой функцией двух других аргументов x_1 и x_2. Тогда функцию $\Phi(u, x_1, x_2)$ можно рассматривать как сложную функцию двух аргументов x_1, x_2. Частные производные этой сложной функции по x_1 и x_2 будем называть полными частными производными функции $\Phi(u, x_1, x_2)$ по x_1 и x_2 и обозначать символами $D\Phi/Dx_1$ и $D\Phi/Dx_2$. По правилу дифференцирования сложной функции мы получим следующие формулы для указанных полных частных производных:

$$\frac{D\Phi}{Dx_1}=\frac{\partial\Phi}{\partial u}\frac{\partial u}{\partial x_1}+\frac{\partial\Phi}{\partial x_1}, \quad \frac{D\Phi}{Dx_2}=\frac{\partial\Phi}{\partial u}\frac{\partial u}{\partial x_2}+\frac{\partial\Phi}{\partial x_2}.$$

Переходим к вычислению частных производных второго порядка неявно заданной функции. Ради определенности вычислим производную $\partial^2 u/\partial x_2 \partial x_1$. Дифференцируя первую из формул (13.10) по x_2 и принимая во внимание, что каждая из частных производных $\partial F/\partial x_1$ и $\partial F/\partial u$ зависит от трех аргументов u, x_1, x_2, первый из которых сам является функцией x_1 и x_2, будем иметь

$$\frac{\partial^2 u}{\partial x_2 \partial x_1}=\frac{D\left[-\dfrac{\partial F/\partial x_1}{\partial F/\partial u}\right]}{Dx_2}=\frac{-\dfrac{\partial F}{\partial u}\dfrac{D[\partial F/\partial x_1]}{Dx_2}+\dfrac{\partial F}{\partial x_1}\dfrac{D[\partial F/\partial u]}{Dx_2}}{(\partial F/\partial u)^2}=$$

$$= \frac{-\frac{\partial F}{\partial u}\left(\frac{\partial^2 F}{\partial x_1 \partial u}\frac{\partial u}{\partial x_2}+\frac{\partial^2 F}{\partial x_1 \partial x_2}\right)+\frac{\partial F}{\partial x_1}\left(\frac{\partial^2 F}{\partial u^2}\frac{\partial u}{\partial x_2}+\frac{\partial^2 F}{\partial u \partial x_2}\right)}{\left(\frac{\partial F}{\partial u}\right)^2}.$$

Вставляя в полученную формулу выражение $\partial u/\partial x_2$, определяемое второй из формул (13.10), окончательно будем иметь

$$\frac{\partial^2 u}{\partial x_2 \partial x_1} = \frac{\frac{\partial^2 F}{\partial x_1 \partial u}\frac{\partial F}{\partial x_2}\frac{\partial F}{\partial u}-\frac{\partial^2 F}{\partial x_1 \partial x_2}\left(\frac{\partial F}{\partial u}\right)^2-\frac{\partial^2 F}{\partial u^2}\frac{\partial F}{\partial x_1}\frac{\partial F}{\partial x_2}+\frac{\partial^2 F}{\partial x_2 \partial u}\frac{\partial F}{\partial x_1}\frac{\partial F}{\partial u}}{\left(\frac{\partial F}{\partial u}\right)^3}. \tag{13.12}$$

Совершенно аналогично вычисляются частные производные $\partial^2 u/\partial x_1^2$ и $\partial^2 u/\partial x_2^2$. Аналогичным методом могут быть вычислены и частные производные третьего и последующих порядков.

Пример. Вычислить частную производную $\partial^2 u/\partial x_2 \partial x_1$ функции, неявно заданной уравнением

$$u^2+x_1^4+x_2^4-1=0.$$

Используя равенства (13.10), получим, что

$$\frac{\partial u}{\partial x_1} = -\frac{2x_1^3}{u}, \quad \frac{\partial u}{\partial x_2} = -\frac{2x_2^3}{u}.$$

Далее получим, что

$$\frac{\partial^2 u}{\partial x_2 \partial x_1} = \frac{D(\partial u/\partial x_1)}{Dx_2} = \frac{\partial}{\partial u}\left(-\frac{2x_1^3}{u}\right)\frac{\partial u}{\partial x_2} = \frac{2x_1^3}{u^2}\frac{\partial u}{\partial x_2} = -\frac{4x_1^3 x_2^3}{u^3}.$$

3. Особые точки поверхности и плоской кривой. Рассмотрим некоторую поверхность S [плоскую кривую L], определяемую в заданной декартовой прямоугольной системе координат уравнением $F(x, y, z)=0$ [$F(x, y)=0$]. Относительно функции $F(x, y, z)$ [$F(x, y)$] предположим, что она имеет непрерывные частные производные первого порядка по всем аргументам всюду в некоторой окрестности любой точки поверхности S [кривой L]. Будем называть данную точку поверхности S [кривой L] особой, если в этой точке обращаются в нуль все частные производные первого порядка функции $F(x, y, z)$ [$F(x, y)$]. В окрестности особой точки нельзя применить к уравнению $F(x, y, z)=0$ [$F(x, y)=0$] теорему 13.1, т. е. нельзя утверждать, что это уравнение разрешимо хотя бы относительно одной из переменных x, y, z [x, y]. Таким образом, участок поверхности S [кривой L], прилегающей к особой точке, может не допускать однозначного проектирования ни на одну из координатных плоскостей [ни на одну из осей координат]. Структура поверхности S [кривой L] в окрестности особой точки

может быть сложной и требует дополнительного исследования.

Точки поверхности S [кривой L], не являющиеся особыми, принято называть обыкновенными. В окрестности обыкновенной точки действует теорема 13.1, так что прилегающий к обыкновенной точке участок поверхности S [кривой L] допускает однозначное проектирование хотя бы на одну из координатных плоскостей [хотя бы на одну из осей координат], что существенно облегчает исследование этого участка.

Примеры. 1) Найти особые точки кругового конуса $x^2+y_2-z^2=0$. Поскольку $F(x, y, z)=x^2+y^2-z^2$, то $\frac{\partial F}{\partial x}=2x$, $\frac{\partial F}{\partial y}=2y$, $\frac{\partial F}{\partial z}=-2z$. Единственной особой точкой является начало координат. Хорошо известно, что в окрестности этой точки поверхность конуса не может быть однозначно спроектирована ни на одну из координатных плоскостей (рис. 13.3).

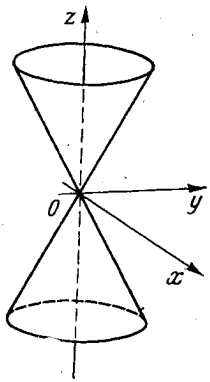

Рис. 13.3

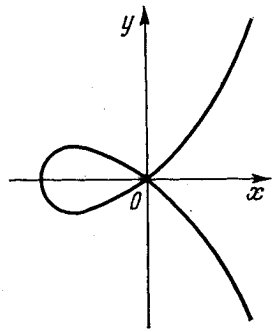

Рис. 13.4

2) Найдем особые точки плоской кривой $x^2-y^2+x^3=0$. Так как $\partial F/\partial x=2x+3x^2=x(2+3x)$, $\partial F/\partial y=-2y$, то обе частные производные $\partial F/\partial x$ и $\partial F/\partial y$ обращаются в нуль в двух точках $(0, 0)$ и $(-2/3, 0)$. Из этих двух точек только первая принадлежит рассматриваемой кривой, т. е. является особой. Построив кривую $x^2-y^2+x^3=0$ в окрестности точки $(0, 0)$, мы убедимся в том, что эта точка является точкой самопересечения графика (рис. 13.4). Ясно, что в окрестности этой точки кривую нельзя однозначно спроектировать ни на ось Ox, ни на ось Oy.

4. Условия, обеспечивающие существование для функции $y=f(x)$ обратной функции. Применим теорему 13.1 для выяснения условий, при выполнении которых функция $y=f(x)$ имеет в некоторой окрестности точки x_0 обратную функцию $x=f^{-1}(y)$, опреде-

ленную в некоторой окрестности точки y_0, где $y_0=f(x_0)$. Будем рассматривать функцию $y=f(x)$ как функцию, определяемую функциональным уравнением вида $F(x, y)=f(x)-y=0$.

Тогда вопрос о существовании обратной функции совпадает с вопросом о разрешимости относительно x указанного функционального уравнения. Как следствие теоремы 13.1 и замечания 1 перед доказательством этой теоремы, мы получим следующее

Утверждение. *Если функция $y=f(x)$ имеет отличную от нуля и сохраняющую определенный знак производную в некоторой окрестности точки x_0, то для этой функции в окрестности x_0 существует обратная функция $x=f^{-1}(y)$, определенная и дифференцируемая в некоторой окрестности точки y_0, где $y_0=f(x_0)$. Производная указанной обратной функции в точке y_0 в силу второй из формул* (13.10) *равна* $[f'(x_0)]^{-1}$.

§ 2. НЕЯВНЫЕ ФУНКЦИИ, ОПРЕДЕЛЯЕМЫЕ СИСТЕМОЙ ФУНКЦИОНАЛЬНЫХ УРАВНЕНИЙ

1. Теорема о разрешимости системы функциональных уравнений. В предыдущем параграфе мы рассматривали вопрос о существовании и дифференцируемости неявной функции, определяемой посредством одного функционального уравнения. В этом параграфе мы рассмотрим аналогичный вопрос для совокупности m (m — любое натуральное число) неявных функций, определяемых посредством системы функциональных уравнений.

Итак, предположим, что m функций

$$\left.\begin{aligned} u_1 &= \varphi_1(x_1, x_2, \ldots, x_n), \\ u_2 &= \varphi_2(x_1, x_2, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots \\ u_m &= \varphi_m(x_1, x_2, \ldots, x_n) \end{aligned}\right\} \tag{13.13}$$

ищутся как решение системы m функциональных уравнений

$$\left.\begin{aligned} F_1(u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n) &= 0, \\ F_2(u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n) &= 0, \\ &\ldots\ldots\ldots\ldots \\ F_m(u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n) &= 0. \end{aligned}\right\} \tag{13.14}$$

Изучим вопрос о разрешимости системы функциональных уравнений (13.14) относительно $u_1, u_2, \ldots, u_m$. Под термином «решение системы (13.14)» мы в дальнейшем будем понимать совокупность m функций (13.13) таких, что при подстановке этих функций в систему (13.14) все уравнения этой системы обращаются в тождества. Это решение мы будем называть непрерывным и дифференцируемым в некоторой области D изменения переменных $x_1, x_2, \ldots, x_n$, если каждая из функций (13.13) непрерывна и

дифференцируема в области D. Договоримся обозначать символом R пространство $m+n$ переменных $u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n$, а символом R' пространство n переменных $x_1, x_2, \ldots, x_n$.

Рассмотрим m функций $F_1, F_2, \ldots, F_m$, стоящих в левых частях системы (13.14), и составим из частных производных этих функций следующий определитель:

$$\Delta = \begin{vmatrix} \frac{\partial F_1}{\partial u_1} & \frac{\partial F_1}{\partial u_2} & \cdots & \frac{\partial F_1}{\partial u_m} \\ \frac{\partial F_2}{\partial u_1} & \frac{\partial F_2}{\partial u_2} & \cdots & \frac{\partial F_2}{\partial u_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F_m}{\partial u_1} & \frac{\partial F_m}{\partial u_2} & \cdots & \frac{\partial F_m}{\partial u_m} \end{vmatrix}. \tag{13.15}$$

Будем называть определитель вида (13.15) определителем Якоби* (или кратко якобианом) функций $F_1, F_2, \ldots, F_m$ по переменным $u_1, u_2, \ldots, u_m$ и кратко обозначать символом

$$\frac{D(F_1, F_2, \ldots, F_m)}{D(u_1, u_2, \ldots, u_m)}.$$

Имеет место следующее замечательное утверждение.

Теорема 13.2 (обобщение теоремы 13.1). *Пусть m функций*

$$\left.\begin{aligned} &F_1(u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n), \\ &F_2(u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &F_m(u_1, u_2, \ldots, u_m, x_1, x_2, \ldots, x_n) \end{aligned}\right\} \tag{13.16}$$

дифференцируемы в некоторой окрестности точки $M_0(\mathring{u}_1, \mathring{u}_2, \ldots, \mathring{u}_m, \mathring{x}_1, \mathring{x}_2, \ldots, \mathring{x}_n)$ пространства R, причем частные производные этих функций по переменным $u_1, u_2, \ldots, u_m$ непрерывны в точке M_0. Тогда если в точке M_0 все функции (13.16) *обращаются в нуль, а якобиан $\frac{D(F_1, F_2, \ldots, F_m)}{D(u_1, u_2, \ldots, u_m)}$ отличен от нуля в M_0, то для достаточно малых положительных чисел $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ найдется такая окрестность точки $M_0'(\mathring{x}_1, \ldots, \mathring{x}_n)$ пространства R', что в пределах этой окрестности существуют единственные m функций* (13.13), *которые удовлетворяют условиям $|u_1-\mathring{u}_1|<\varepsilon_1$, $|u_2-\mathring{u}_2|<\varepsilon_2, \ldots, |u_m\mathring{u}_m-|<\varepsilon_m$ и являются решением системы уравнений* (13.14), *причем это решение непрерывно и дифференцируемо в указанной окрестности точки M_0'.*

---

* Карл Густав Яков Якоби — немецкий математик (1804—1851).

Замечание. При $m=1$ теорема 13.2 переходит в доказанную выше* теорему 13.1, ибо в этом случае якобиан (13.15) обращается в частную производную $\partial F_1/\partial u_1$.

Доказательство теоремы 13.2 проведем методом математической индукции. При $m=1$ теорема уже доказана. Поэтому достаточно, предположив теорему 13.2 справедливой для системы $m-1$ функциональных уравнений, доказать справедливость этой теоремы и для системы m функциональных уравнений. Поскольку, по предположению, якобиан

$$\Delta = \frac{D(F_1, F_2, \ldots, F_m)}{D(u_1, u_2, \ldots, u_m)} \begin{vmatrix} \frac{\partial F_1}{\partial u_1} & \cdots & \frac{\partial F_1}{\partial u_{m-1}} & \frac{\partial F_1}{\partial u_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial F_{m-1}}{\partial u_1} & \cdots & \frac{\partial F_{m-1}}{\partial u_{m-1}} & \frac{\partial F_{m-1}}{\partial u_m} \\ \frac{\partial F_m}{\partial u_1} & \cdots & \frac{\partial F_m}{\partial u_{m-1}} & \frac{\partial F_m}{\partial u_m} \end{vmatrix} \tag{13.17}$$

отличен от нуля в точке M_0, то *хотя бы один из миноров* $(m-1)$*-го порядка*** *этого якобиана отличен от нуля в точке* M_0. Не ограничивая общности, будем считать, что в точке M_0 отличен от нуля обведенный пунктиром минор, стоящий в левом верхнем углу. Тогда в силу предположения индукции первые $m-1$ уравнений системы (13.14) разрешимы относительно $u_1, u_2, \ldots, u_{m-1}$. Точнее, для достаточно малых положительных чисел $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{m-1}$ найдется такая окрестность точки $M_0''(\mathring{u}_m, \mathring{x}_1, \ldots, \mathring{x}_n)$ пространства R'' переменных $(u_m, x_1, \ldots, x_n)$, что в пределах этой окрестности определены $m-1$ функций

$$\left.\begin{aligned} u_1 &= \Phi_1(u_m, x_1, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ u_{m-1} &= \Phi_{m-1}(u_m, x_1, \ldots, x_n), \end{aligned}\right\} \tag{13.18}$$

которые удовлетворяют условиям $|u_1-\mathring{u}_1|<\varepsilon_1, \ldots, |u_{m-1}-\mathring{u}_{m-1}|<\varepsilon_{m-1}$ и являются при наличии этих условий единственным и дифференцируемым решением системы первых $m-1$ уравнений (13.14).

Подставим найденные функции (13.18) в левую часть последнего из уравнений (13.14). При этом левая часть последнего из уравнений (13.14) превращается в функцию, зависящую только от $u_m, x_1, \ldots, x_n$:

$$F_m(u_1, \ldots, u_{m-1}, u_m, x_1, \ldots, x_n) = F_m[\Phi_1(u_m, x_1, \ldots, x_n), \ldots, \Phi_{m-1}(u_m, x_1, \ldots, x_n), u_m, x_1, \ldots, x_n] = \psi(u_m, x_1, \ldots, x_n) \tag{13.19}$$

---

* При этом следует учесть замечание 2 к теореме 13.1.

** Напомним, что минором $(m-1)$-го порядка данного определителя m-го порядка называется определитель $(m-1)$-го порядка, полученный из данного определителя m-го порядка вычеркиванием одной строки и одного столбца.

(эту функцию мы обозначили символом ψ). Таким образом, последнее из уравнений системы (13.14) приводит нас к уравнению

$$\psi(u_m, x_1, \ldots, x_n) = 0. \tag{13.20}$$

В силу равенства (13.19) $\psi(u_m, x_1, \ldots, x_n)$ можно рассматривать как сложную функцию своих аргументов. Тогда, применяя теорему о дифференцируемости сложной функции, мы можем утверждать, что *функция* $\psi(u_m, x_1, \ldots, x_n)$ *дифференцируема в некоторой окрестности точки* $M_0''(\mathring{u}_m, \mathring{x}_1, \ldots, \mathring{x}_n)$ *пространства* R''. Равенство (13.19) и последнее из уравнений (13.14) позволяют утверждать, что $\psi(\mathring{u}_m, \mathring{x}_1, \ldots, \mathring{x}_n) = 0$. Поэтому, для того чтобы доказать, что к уравнению (13.20) применима теорема 13.1 и это уравнение разрешимо относительно u_m, достаточно установить, что частная производная $\partial\psi/\partial u_m$ непрерывна и отлична от нуля в точке M_0''. Для того чтобы сделать это, вычислим указанную частную производную. Подставим в первые $m-1$ уравнений системы (13.14) функции (13.18), являющиеся решением этих уравнений, и продифференцируем полученные при этом тождества по u_m. Получим

$$\frac{\partial F_1}{\partial u_1}\frac{\partial \Phi_1}{\partial u_m} + \ldots + \frac{\partial F_1}{\partial u_{m-1}}\frac{\partial \Phi_{m-1}}{\partial u_m} + \frac{\partial F_1}{\partial u_m} = 0, \tag{13.21^1}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial F_{m-1}}{\partial u_1}\frac{\partial \Phi_1}{\partial u_m} + \ldots + \frac{\partial F_{m-1}}{\partial u_{m-1}}\frac{\partial \Phi_{m-1}}{\partial u_m} + \frac{\partial F_{m-1}}{\partial u_m} = 0. \tag{13.21^{m-1}}$$

Далее продифференцируем по u_m равенство (13.19). Получим

$$\frac{\partial F_m}{\partial u_1}\frac{\partial \Phi_1}{\partial u_m} + \ldots + \frac{\partial F_m}{\partial u_{m-1}}\frac{\partial \Phi_{m-1}}{\partial u_m} + \frac{\partial F_m}{\partial u_m} = \frac{\partial \psi}{\partial u_m}. \tag{13.21^m}$$

Умножим теперь равенства (13.21^1)—(13.21^m) на соответствующие алгебраические дополнения $\Delta_1, \Delta_2, \ldots, \Delta_m$ элементов последнего столбца якобиана (13.17) и после этого сложим эти равенства. Получим

$$\sum_{k=1}^{m-1} \frac{\partial \Phi_k}{\partial u_m}\left[\Delta_1 \frac{\partial F_1}{\partial u_k} + \Delta_2 \frac{\partial F_2}{\partial u_k} + \ldots + \Delta_m \frac{\partial F_m}{\partial u_k}\right] +$$

$$+ \left(\Delta_1 \frac{\partial F_1}{\partial u_m} + \Delta_2 \frac{\partial F_2}{\partial u_m} + \ldots + \Delta_m \frac{\partial F_m}{\partial u_m}\right) = \Delta_m \frac{\partial \psi}{\partial u_m}.$$

Так как сумма произведений элементов данного столбца определителя на соответствующие алгебраические дополнения элементов этого (другого) столбца равна определителю (нулю), то каждая квадратная скобка равна нулю, а круглая скобка равна якобиану (13.17).

Таким образом, мы получим, что

$$\Delta = \Delta_m \frac{\partial \psi}{\partial u_m}. \tag{13.22}$$

Здесь символом Δ обозначен якобиан (13.17), а Δ_m — алгебраическое дополнение последнего элемента последнего столбца, которое совпадает с минором, обведенным пунктиром и, по предположению, отличным от нуля в точке M_0. Поделив равенство (13.22) на Δ_m, окончательно найдем

$$\frac{\partial \psi}{\partial u_m} = \frac{\Delta}{\Delta_m}. \tag{13.23}$$

Формула (13.23), справедливая в точке M_0'', доказывает непрерывность частной производной $\partial\psi/\partial u_m$ в точке M_0'' (ибо Δ и Δ_m состоят из частных производных функций (13.16) по $u_1, u_2, \ldots, u_m$, непрерывных в точке M_0). Кроме того, из формулы (13.23) вытекает, что $\partial\psi/\partial u_m$ в точке M_0'' отлична от нуля (ибо якобиан Δ отличен от нуля в точке M_0). Тем самым мы доказали, что к уравнению (13.20) можно применить теорему 13.1.

Согласно этой теореме для достаточно малого положительного числа ε_m найдется такая окрестность точки $M_0'(\mathring{x}_1, \mathring{x}_2, \ldots, \mathring{x}_n)$ пространства R', что всюду в пределах этой окрестности определена функция

$$u_m = \varphi_m(x_1, x_2, \ldots, x_n), \tag{13.24}$$

которая удовлетворяет условию $|u_m - \mathring{u}_m| < \varepsilon_m$ и является при наличии этого условия единственным, непрерывным и дифференцируемым решением уравнения (13.20). Имея в виду, что функции (13.18) являются решениями первых $m-1$ уравнений (13.14) при любых $u_m, x_1, \ldots, x_n$ из окрестности точки M_0'', и вставляя найденную функцию (13.24) в (13.18), мы получим функции, зависящие только от переменных $x_1, \ldots, x_n$:

$$u_1 = \Phi_1[\varphi_m(x_1, \ldots, x_n), x_1, \ldots, x_n] = \varphi_1(x_1, \ldots, x_n),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$u_{m-1} = \Phi_{m-1}[\varphi_m(x_1, \ldots, x_n), x_1, \ldots, x_n] = \varphi_{m-1}(x_1, \ldots, x_n).$$

(Эти функции мы обозначили символами $\varphi_1, \ldots, \varphi_{m-1}$.) Теорема о дифференцируемости сложной функции дает право утверждать, что каждая из функций $\varphi_1, \ldots, \varphi_{m-1}$ дифференцируема в окрестности точки $M_0'(\mathring{x}_1, \ldots, \mathring{x}_n)$. Таким образом, мы доказали, что m функций

$$\left.\begin{aligned} u_1 &= \varphi_1(x_1, x_2, \ldots, x_n), \\ &\cdots\cdots\cdots\cdots \\ u_m &= \varphi_m(x_1, x_2, \ldots, x_n) \end{aligned}\right\} \tag{13.25}$$

удовлетворяют в окрестности точки M_0' условиям $|u_1 - \mathring{u}_1| < \varepsilon_1, \ldots$

..., $|u_m - \mathring{u}_m| < \varepsilon_m$ и представляют собой при наличии этих условий непрерывное и дифференцируемое в некоторой окрестности точки $M_0'(\mathring{x}_1, ..., \mathring{x}_n)$ решение системы (13.14).

Остается доказать, что функции (13.25) представляют собой единственное решение системы (13.14), удовлетворяющее условиям $|u_1 - \mathring{u}_1| < \varepsilon_1, ..., |u_m - \mathring{u}_m| < \varepsilon_m$ (при достаточно малых положительных $\varepsilon_1, ..., \varepsilon_m$).

Предположим что кроме функций (13.25) существуют еще m функций

$$\left.\begin{aligned} \hat{u}_1 &= \hat{\varphi}_1(x_1, x_2, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots \\ \hat{u}_m &= \hat{\varphi}_m(x_1, x_2, \ldots, x_n), \end{aligned}\right\} \tag{13.25}$$

также являющихся решением системы (13.14) и удовлетворяющих условиям $|\hat{u}_1 - \mathring{u}_1| < \varepsilon_1, ..., |\hat{u}_m - \mathring{u}_m| < \varepsilon_m$.

Тогда в силу предположения индукции первые $m-1$ функций (13.25) представляют собой при заданном $u_m = \hat{u}_m$ единственное и дифференцируемое решение системы первых $m-1$ уравнений (13.14). Но при заданном u_m единственное решение системы первых $m-1$ уравнений (13.14) дается равенствами (13.18). Таким образом, справедливы соотношения

$$\left.\begin{aligned} \hat{u}_1 &= \Phi_1(\hat{u}_m, x_1, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots \\ \hat{u}_{m-1} &= \Phi_{m-1}(\hat{u}_m, x_1, \ldots, x_n), \end{aligned}\right\} \tag{13.18'}$$

в которых $\Phi_1, ..., \Phi_{m-1}$ — те же функции, что и (13.18).

В таком случае последнее уравнение (13.14) и соотношение (13.19) позволяют нам утверждать, что $\hat{u}_m$ является единственным решением уравнения (13.20), т. е. $\hat{u}_m = u_m$.

При наличии равенства $\hat{u}_m = u_m$ из соотношений (13.18′) и (13.18) сразу же вытекает, что $\hat{u}_1 = u_1, ..., \hat{u}_{m-1} = u_{m-1}$.

Теорема 13.2 полностью доказана.

2. Вычисление частных производных функций, неявно определяемых посредством системы функциональных уравнений. В этом пункте мы предположим, что выполнены условия теоремы 13.2, и займемся вычислением частных производных функций (13.25). Подставим функции (13.25) в систему уравнений (13.14), решением которой эти функции являются, и продифференцируем получившиеся тождества по x_l ($l = 1, 2, ..., n$). Получим

$$\left.\begin{aligned} &\frac{\partial F_1}{\partial u_1}\frac{\partial u_1}{\partial x_l} + \ldots + \frac{\partial F_1}{\partial u_m}\frac{\partial u_m}{\partial x_l} + \frac{\partial F_1}{\partial x_l} = 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\frac{\partial F_m}{\partial u_1}\frac{\partial u_1}{\partial x_l} + \ldots + \frac{\partial F_m}{\partial u_m}\frac{\partial u_m}{\partial x_l} + \frac{\partial F_m}{\partial x_l} = 0. \end{aligned}\right\} \tag{13.26}$$

Равенства (13.26) представляют собой линейную систему уравнений относительно m неизвестных $\frac{\partial u_1}{\partial x_l}, \ldots, \frac{\partial u_m}{\partial x_l}$. Определитель этой системы якобиан (13.17) отличен от нуля в окрестности точки M_0. Значит, система (13.26) имеет единственное решение, определяемое формулами Крамера:

$$\frac{\partial u_k}{\partial x_l} = -\frac{\dfrac{D(F_1, F_2, \ldots, F_m)}{D(u_1, \ldots, u_{k-1}, x_l, u_{k+1}, \ldots, u_m)}}{\dfrac{D(F_1, F_2, \ldots, F_m)}{D(u_1, u_2, \ldots, u_m)}}.$$

Выражения для частных производных второго и последующих порядков * можно получить посредством дифференцирования этих формул.

3. Взаимно однозначное отображение двух множеств m-мерного пространства. Рассмотрим в некоторой окрестности точки $M_0(\overset{\circ}{x}_1, \ldots, \overset{\circ}{x}_m)$ функции

$$\left.\begin{aligned} u_1 &= \varphi_1(x_1, \ldots, x_m), \\ &\ldots\ldots\ldots\ldots \\ u_m &= \varphi_m(x_1, \ldots, x_m). \end{aligned}\right\} \qquad (13.27)$$

Рассматриваемые m функций осуществляют отображение указанной окрестности точки M_0 на некоторое множество $\{N\}$ m-мерного пространства переменных $u_1, u_2, \ldots, u_m$. Это отображение называется взаимно однозначным, если каждой точке из указанной окрестности точки M_0 соответствует только одна точка множества $\{N\}$, так что при этом каждая точка множества $\{N\}$ соответствует только одной точке указанной окрестности точки M_0.

Из теоремы 13.2 непосредственно вытекает следующее

Утверждение. *Если функции* (13.27) *дифференцируемы в окрестности точки M_0, причем все частные производные первого порядка непрерывны в самой точке M_0, а якобиан* $\frac{D(\varphi_1, \ldots, \varphi_m)}{D(x_1, \ldots, x_m)}$ *отличен от нуля в этой точке, то функции* (13.27) *осуществляют взаимно однозначное отображение некоторой окрестности точки $M_0(\overset{\circ}{x}_1, \ldots, \overset{\circ}{x}_m)$ на некоторую окрестность точки $N_0(\mathring{u}_1, \ldots, \mathring{u}_m)$, где $u_i = \varphi_i(x_1, \ldots, x_m)$ $(i = 1, 2, \ldots, m)$.*

В самом деле, соотношения (13.27) можно рассматривать как систему функциональных уравнений относительно $x_1, x_2, \ldots, x_m$, для которой выполнены условия теоремы 13.2. Но тогда эта система всюду в некоторой окрестности точки $N_0(\mathring{u}_1, \ldots, \mathring{u}_m)$ имеет единственное решение:

---

* Существование этих частных производных обеспечивается дополнительными ограничениями на функции (13.16).

$$\left.\begin{array}{l} x_1=\psi_1(u_1, \ldots, u_m), \\ \cdots\cdots\cdots\cdots \\ x_m=\psi_m(u_1, \ldots, u_m). \end{array}\right\} \qquad (13.27')$$

Очевидно, что функции (13.27′) осуществляют обратное отображение.

Заметим, что в условиях сформулированного утверждения как функции (13.27), осуществляющие прямое отображение, так и функции (13.27′), осуществляющие обратное отображение, являются непрерывными. Взаимно однозначное отображение, обладающее таким свойством, называется гомеоморфизмом.

Более того, в условиях сформулированного утверждения как функции (13.27), осуществляющие прямое отображение, так и функции (13.27′), осуществляющие обратное отображение, являются дифференцируемыми. Взаимно однозначное отображение, обладающее таким свойством, принято называть диффеоморфизмом.

§ 3. ЗАВИСИМОСТЬ ФУНКЦИЙ

1. Понятие зависимости функций. Достаточное условие независимости. Пусть m функций от одних и тех же n переменных

$$\left.\begin{array}{l} u_1=\varphi_1(x_1, x_2, \ldots, x_n), \\ \cdots\cdots\cdots\cdots \\ u_m=\varphi_m(x_1, x_2, \ldots, x_n) \end{array}\right\} \qquad (13.28)$$

определены и дифференцируемы в некоторой открытой n-мерной области * D.

Будем говорить, что *одна из этих функций, например u_k, зависит в области D от остальных функций, если сразу для всех точек $(x_1, x_2, \ldots, x_n)$ области D*

$$u_k=\Phi(u_1, \ldots, u_{k-1}, u_{k+1}, \ldots, u_m), \qquad (13.29)$$

где Φ — некоторая функция, определенная и дифференцируемая в соответствующей области изменения своих аргументов. Функции $u_1, u_2, \ldots, u_m$ будем называть зависимыми в области D, если одна из этих функций (все равно какая) зависит в области D от остальных.

Если же не существует дифференцируемой функции Φ такой, что сразу для всех точек области D справедливо тождество вида (13.29) хотя бы для одного k ($k=1, \ldots, m$), то мы будем называть функции $u_1, u_2, \ldots, u_m$ *независимыми в области D.*

---

* В частности, в качестве области D можно взять некоторую окрестность фиксированной точки M_0 n-мерного пространства.

Примеры. 1) Легко убедиться в том, что три функции четырех переменных

$$u_1 = x_1^2 + x_2^2 + x_3^2 + x_4^2,$$

$$u_2 = x_1 + x_2 + x_3 + x_4,$$

$$u_3 = 2x_1x_2 + 2x_1x_3 + 2x_1x_4 + 2x_2x_3 + 2x_3x_4$$

зависимы в любой области D четырехмерного пространства, ибо для всех точек (x_1, x_2, x_3, x_4) этой области

$$u_1 = u_2^2 - u_3.$$

2) Покажем теперь, что две функции двух переменных $u_1 = x + y$ и $u_2 = x - y$ независимы в любой области D плоскости x, y, содержащей начало координат. Ясно, что функция u_1 сохраняет постоянное значение нуль на прямой $x + y = 0$, проходящей через начало координат (рис. 13.5). Но на этой прямой функция u_2 имеет переменное значение $u_2 = 2x$. Поэтому на том участке этой прямой, который лежит внутри D, u_2 заведомо не зависит от u_1. Совершенно аналогично доказывается, что на лежащем внутри области D участке прямой $x - y = 0$, $u_2 = 0$, $u_1 = 2x$ и, значит, u_1 не зависит от u_2.

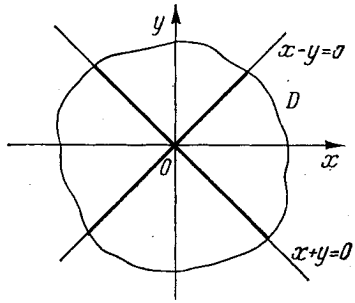

Рис. 13.5

Замечание. В курсе линейной алгебры вводится понятие линейной зависимости функций: m функций $u_1, u_2, \ldots, u_m$ называются линейно зависимыми в области D, если для всех точек области D одна из этих функций выражается в виде линейной функции от остальных. Ясно, что линейная зависимость функций является частным случаем зависимости этих функций, ибо если функции $u_1, u_2, \ldots, u_m$ линейно зависимы в области D, то они зависимы в этой области, но существуют функции, зависимые в области D, но не являющиеся в D линейно зависимыми (например, функции, выписанные в примере 1).

Теорема 13.3 (достаточное условие независимости функций). *Пусть m функций от $n \geqslant m$ переменных*

$$\left.\begin{array}{l} u_1 = \varphi_1(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ u_m = \varphi_m(x_1, \ldots, x_m, x_{m+1}, \ldots, x_n) \end{array}\right\}$$

определены и дифференцируемы в окрестности точки $M_0(\mathring{x}_1, \ldots, \mathring{x}_m, \mathring{x}_{m+1}, \ldots, \mathring{x}_n)$. Тогда если якобиан из этих функций по каким-либо m переменным отличен от нуля в точке M_0, то эти функции независимы в некоторой окрестности точки M_0.

Доказательство. Не ограничивая общности, будем считать, что в точке M_0 отличен от нуля якобиан

$$\frac{D(u_1, u_2, \ldots, u_m)}{D(x_1, x_2, \ldots, x_m)}. \tag{13.30}$$

Докажем теорему от противного. Предположим, что функции u_1, $u_2, \ldots, u_m$ зависимы в некоторой окрестности точки M_0, т. е. одна из этих функций, например u_k, для всех точек этой окрестности выражается в виде $u_k = \Phi(u_1, \ldots, u_{k-1}, u_{k+1}, \ldots, u_m)$, где Φ — некоторая дифференцируемая функция. Пользуясь правилом дифференцирования сложной функции, вычислим производную функции u_k по любой из переменных x_l $(l = 1, 2, \ldots, m)$. Будем иметь

$$\frac{\partial u_k}{\partial x_l} = \frac{\partial \Phi}{\partial u_1}\frac{\partial u_1}{\partial x_l} + \ldots + \frac{\partial \Phi}{\partial u_{k-1}}\frac{\partial u_{k-1}}{\partial x_l} +$$

$$+ \frac{\partial \Phi}{\partial u_{k+1}}\frac{\partial u_{k+1}}{\partial x_l} + \ldots + \frac{\partial \Phi}{\partial u_m}\frac{\partial u_m}{\partial x_l} \quad (l = 1, 2, \ldots, m). \tag{13.31}$$

Формулы (13.31), если их взять для любого значения $l = 1, 2, \ldots, m$ в точке M_0, говорят о том, что k-я строка якобиана (13.30) представляет собой линейную комбинацию остальных строк с коэффициентами, соответственно равными $\frac{\partial \Phi}{\partial u_1}, \frac{\partial \Phi}{\partial u_2}, \ldots, \frac{\partial \Phi}{\partial u_{k-1}}, \frac{\partial \Phi}{\partial u_{k+1}}, \ldots, \frac{\partial \Phi}{\partial u_m}$. Но в этом случае якобиан (13.30) равен нулю в точке M_0, что противоречит условию теоремы.

Пример. Уже рассмотренные выше две функции $u_1 = x + y$ и $u_2 = x - y$ независимы в окрестности любой точки $M(x, y)$, ибо якобиан $\frac{D(u_1, u_2)}{D(x, y)} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 \neq 0$ всюду.

2. Функциональные матрицы и их приложения. Снова рассмотрим m функций от n переменных (13.28):

$$\left.\begin{aligned} u_1 &= \varphi_1(x_1, x_2, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots \\ u_m &= \varphi_m(x_1, x_2, \ldots, x_n). \end{aligned}\right\} \tag{13.28}$$

На этот раз предположим, что *функции* (13.28) *определены и дифференцируемы в некоторой окрестности точки* $M_0(\mathring{x}_1, \ldots, \mathring{x}_n)$, *причем все частные производные первого порядка этих функций непрерывны в самой точке* M_0.

Составим из частных производных функций (13.28) следующую функциональную матрицу:

$$\left\| \begin{matrix} \frac{\partial\varphi_1}{\partial x_1} & \frac{\partial\varphi_1}{\partial x_2} & \dots & \frac{\partial\varphi_1}{\partial x_n} \\ \frac{\partial\varphi_2}{\partial x_1} & \frac{\partial\varphi_2}{\partial x_2} & \dots & \frac{\partial\varphi_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial\varphi_m}{\partial x_1} & \frac{\partial\varphi_m}{\partial x_2} & \dots & \frac{\partial\varphi_m}{\partial x_n} \end{matrix} \right\|, \tag{13.32}$$

содержащую m строк и n столбцов.

Имеет место следующее утверждение.

Теорема 13.4. *Пусть у функциональной матрицы* (13.32): 1) *некоторый минор r-го порядка* отличен от нуля в точке $M_0(\mathring{x}_1, \mathring{x}_2, \dots, \mathring{x}_n)$*; 2) *все миноры $(r+1)$-го порядка равны нулю в некоторой окрестности точки M_0**. Тогда r функций, представленных в указанном миноре r-го порядка, независимы в окрестности точки M_0, а каждая из остальных функций зависит в этой окрестности от указанных r функций.*

Доказательство. Не ограничивая общности, будем считать, что в точке M_0 отличен от нуля минор, стоящий в левом верхнем углу матрицы (13.32), т. е. отличен от нуля определитель

$$\begin{vmatrix} \frac{\partial\varphi_1}{\partial x_1} & \dots & \frac{\partial\varphi_1}{\partial x_r} \\ \frac{\partial\varphi_r}{\partial x_1} & \dots & \frac{\partial\varphi_r}{\partial x_r} \end{vmatrix}. \tag{13.33}$$

Тогда независимость в окрестности точки M_0 *функций* $u_1, u_2, \dots, u_r$ сразу вытекает из теоремы 13.3. Остается доказать, что любая из функций $u_{r+1}, \dots, u_m$*** зависит в окрестности M_0 от $u_1, u_2, \dots, u_r$. Докажем, например, что u_{r+1} зависит в окрестности точки M_0 от $u_1, u_2, \dots, u_r$. Сосредоточим свое внимание на первых r функциях (13.28). Если обозначить через $\mathring{u}_1, \dots, \mathring{u}_r$ числа вида $\mathring{u}_1 = \varphi_1(\mathring{x}_1, \mathring{x}_2, \dots, \mathring{x}_n), \dots, \mathring{u}_r = \varphi_r(\mathring{x}_1, \mathring{x}_2, \dots, \mathring{x}_n)$, то всюду в некоторой окрестности точки N_0 $(\mathring{u}_1, \dots, \mathring{u}_r, \mathring{x}_1, \dots, \mathring{x}_n)$ $(n+r)$-мерного пространства переменных $(u_1, \dots, u_r, x_1, \dots, x_n)$ первые r функций (13.28) представляют собой единственное и дифференцируемое решение следующей системы уравнений****:

---

* Напомним, что минором r-го порядка данной матрицы называется определитель, составленный из элементов, стоящих на пересечении каких-либо r столбцов и r строк матрицы.

** В случае, если $r = \min(m, n)$, требование 2) следует опустить.

*** Конечно, при этом предполагается, что $m > r$.

**** В самом деле, в указанной точке N_0 все функции $F_1, \dots, F_r$ обращаются в нуль, а якобиан $\frac{D(F_1, \dots, F_r)}{D(u_1, \dots, u_r)} = (-1) \neq 0$, так что выполнены условия теоремы 13.2.

$$\left.\begin{array}{c} F_1(u_1, \ldots, u_r, x_1, \ldots, x_n) \equiv \varphi_1(x_1, \ldots, x_n) - u_1 = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ F_r(u_1, \ldots, u_r, x_1, \ldots, x_n) \equiv \varphi_r(x_1, \ldots, x_n) - u_r = 0. \end{array}\right\} \quad (13.34)$$

С другой стороны, поскольку якобиан $\frac{D(F_1, \ldots, F_r)}{D(x_1, \ldots, x_r)}$, совпадающий с минором (13.33), отличен от нуля в точке N_0, то систему (13.34) можно в окрестности этой точки однозначно разрешить относительно $x_1, \ldots, x_r$. Иными словами, всюду в достаточно малой окрестности точки N_0 система (13.34) имеет единственное и дифференцируемое решение

$$\left.\begin{array}{c} x_1 = \psi_1(u_1, \ldots, u_r, x_{r+1}, \ldots, x_n), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_r = \psi_r(u_1, \ldots, u_r, x_{r+1}, \ldots, x_n). \end{array}\right\} \quad (13.35)$$

Подчеркнем, что равенства (13.35) и первые r равенств (13.28) полностью эквивалентны в окрестности точки N_0. В частности, если подставить $x_1, x_2, \ldots, x_r$, определяемые уравнениями (13.35), в первые r равенств (13.28), то указанные равенства обратятся в тождества относительно $x_{r+1}, \ldots, x_n, u_1, \ldots, u_r$. Дифференцируя эти тождества по переменной x_l $(l = r+1, \ldots, n)$ и замечая, что $u_1, \ldots, u_r$ не зависят от $x_{r+1}, \ldots, x_n$ будем иметь

$$\left.\begin{array}{ll} \frac{d\varphi_1}{\partial x_1}\frac{\partial\psi_1}{\partial x_l} + \ldots + \frac{\partial\varphi_1}{\partial x_r}\frac{\partial\psi_r}{\partial x_l} + \frac{\partial\varphi_1}{\partial x_l} = 0, & (13.36^1) \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ \frac{\partial\varphi_r}{\partial x_1}\frac{\partial\psi_1}{\partial x_l} + \ldots + \frac{\partial\varphi_r}{\partial x_r}\frac{\partial\psi_r}{\partial x_l} + \frac{\partial\varphi_r}{\partial x_l} = 0. & (13.36^r) \end{array}\right\}$$

Заметим, что равенства (13.36^1)—$13.36^r)$ справедливы для всех значений переменных $x_1, \ldots, x_r, x_{r+1}, \ldots, x_n$ из некоторой окрестности точки M_0.

Для того, чтобы убедиться в том, что функция u_{r+1} зависит в некоторой окрестности точки M_0 от $u_1, \ldots, u_r$, подставим значения $x_1, \ldots, x_r$, определяемые уравнениями (13.35), в $(r+1)$-е равенство (13.28). При этом u_{r+1} превращается в функцию аргументов $u_1, \ldots, u_r, x_{r+1}, \ldots, x_n$, ибо $u_{r+1} = \varphi_{r+1}(x_1, \ldots, x_r, x_{r+1}, \ldots, x_n) = \varphi_{r+1}[\psi_1(u_1, \ldots, u_r, x_{r+1}, \ldots, x_n), \ldots, \psi_r(u_1, \ldots, u_r, x_{r+1}, \ldots, x_n), x_{r+1}, \ldots, x_n] = \Phi(u_1, \ldots, u_r, x_{r+1}, \ldots, x_n)$ (эту функцию мы обозначили символом Φ). Остается доказать, что для всех значений переменных $x_1, \ldots, x_r, x_{r+1}, \ldots, x_n$, лежащих в достаточно малой окрестности точки M_0, функция Φ не зависит от $x_{r+1}, \ldots, x_n$. Для этого достаточно доказать, что для всех $x_1, \ldots, x_n$ из достаточно малой окрестности точки M_0 справедливы равенства

$$\frac{\partial\Phi}{\partial x_l} = 0 \quad (l = r+1, \ldots, n). \quad (13.37)$$

Продифференцируем функцию Φ по переменной x_l $(l=r+1, \dots n)$ как сложную функцию. При этом получим

$$\frac{\partial\varphi_{r+1}}{\partial x_1}\frac{\partial\psi_1}{\partial x_l}+\dots+\frac{\partial\varphi_{r+1}}{\partial x_r}\frac{\partial\psi_r}{\partial x_l}+\frac{\partial\varphi_{r+1}}{\partial x_l}=\frac{\partial\Phi}{\partial x_l}. \quad (13.36^{r+1})$$

Рассмотрим теперь следующий минор $(r+1)$-го порядка матрицы (13.32):

$$\Delta=\begin{vmatrix} \frac{\partial\varphi_1}{\partial x_1} & \dots & \frac{\partial\varphi_1}{\partial x_r} & \frac{\partial\varphi_1}{\partial x_l} \\ \dots & \dots & \dots & \dots \\ \frac{\partial\varphi_r}{\partial x_1} & \dots & \frac{\partial\varphi_r}{\partial x_r} & \frac{\partial\varphi_r}{\partial x_l} \\ \frac{\partial\varphi_{r+1}}{\partial x_1} & \dots & \frac{\partial\varphi_{r+1}}{\partial x_r} & \frac{\partial\varphi_{r+1}}{\partial x_l} \end{vmatrix}. \quad (13.38)$$

По условию теоремы этот минор равен нулю всюду в окрестности точки M_0. Умножим равенства (13.36^1)—(13.36^{r+1}) на соответствующие алгебраические дополнения $\Delta_1, \dots, \Delta_r$, Δ_{r+1} элементов последнего столбца минора (13.38) и после этого сложим все эти равенства. В силу теоремы о том, что сумма произведений элементов данного столбца на соответствующие алгебраические дополнения элементов этого (другого) столбца равна определителю (нулю), получим *

$$\Delta=\frac{\partial\Phi}{\partial x_l}\Delta_{r+1}. \quad (13.39)$$

В равенстве (13.39) символ Δ обозначает минор (13.38), равный нулю всюду в окрестности точки M_0, а алгебраическое дополнение Δ_{r+1} совпадает с минором (13.33), отличным от нуля в точке M_0, а значит, и в некоторой окрестности этой точки **. Из равенства (13.39) заключаем, что всюду в некоторой окрестности точки M_0 справедливы равенства (13.37). Теорема доказана.

Пример. Вернемся к исследованию зависимости функций

$$u_1=x_1^2+x_2^2+x_3^2+x_4^2,$$
$$u_2=x_1+x_2+x_3+x_4,$$
$$u_3=2x_1x_2+2x_1x_3+2x_1x_4+2x_2x_3+2x_3x_4.$$

---

* При этом мы повторяем рассуждения, подробно описанные при доказательстве теоремы 13.2.

** Поскольку все частные производные, входящие в минор (13.33), непрерывны в точке M_0, то и сам минор (13.33) непрерывен в точке M_0. Но тогда по теореме об устойчивости знака непрерывной функции этот минор отличен от нуля не только в самой точке M_0, но и в некоторой ее окрестности.

Функциональная матрица (13.32) имеет вид

$$\begin{Vmatrix} 2x_1 & 2x_2 & 2x_3 & 2x_4 \\ 1 & 1 & 1 & 1 \\ 2(x_2+x_3+x_4) & 2(x_1+x_3+x_4) & 2(x_1+x_2+x_4) & 2(x_1+x_2+x_3) \end{Vmatrix}.$$

Легко убедиться в том, что все определители третьего порядка тождественно равны нулю, причем в любой точке пространства (x_1, x_2, x_3, x_4), у которой не все четыре координаты x_1, x_2, x_3 и x_4 совпадают, хотя бы один из определителей второго порядка

$$\begin{vmatrix} 2x_1 & 2x_2 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2x_1 & 2x_3 \\ 1 & 1 \end{vmatrix}, \quad \begin{vmatrix} 2x_1 & 2x_4 \\ 1 & 1 \end{vmatrix}$$

отличен от нуля. Значит, в окрестности любой указанной точки u_1 и u_2 независимы, а u_3 зависит от u_1 и u_2.

§ 4. УСЛОВНЫЙ ЭКСТРЕМУМ

1. Понятие условного экстремума. В § 6 гл. 12 мы занимались отысканием локальных экстремумов функции, аргументы которой не связаны никакими дополнительными условиями. Вместе с тем в математике и в ее приложениях часто встречается задача об отыскании экстремумов функции, аргументы которой удовлетворяют дополнительным условиям связи. Экстремумы такого рода мы будем называть условными, чтобы отличить их от (безусловных) экстремумов, изученных в § 6 гл. 12.

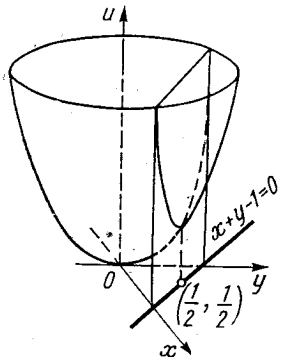

Рис. 13.6

Приведем пример задачи об отыскании условного экстремума. Пусть требуется найти экстремумы функции $u=x^2+y^2$ при условии, что аргументы этой функции удовлетворяют условию связи $x+y-1=0$. Таким образом, *экстремумы функции $u=x^2+y^2$ ищутся не на всей плоскости Oxy, а лишь на прямой $x+y-1=0$*. Для решения поставленной задачи подставим в выражение функции $u=x^2+y^2$ значение y, определяемое из условия связи $x+y-1=0$. Таким путем мы сведем поставленную задачу к задаче об отыскании безусловного экстремума функции $u=2x^2-2x+1$.

Последний экстремум находится без труда: поскольку $u'=4(x-1/2)$, $u^{(2)}=4$, то функция $u=2x^2-2x+1$ имеет минимум $u=1/2$ при $x=1/2$. Таким образом, функция $u=x^2+y^2$ с условием связи $x+y-1=0$ имеет условный минимум $u=1/2$ в точке $(1/2, 1/2)$. Отметим, что безусловный минимум функции $u=x^2+y^2$ дос-

тигается в точке (0, 0) и равен $u=0$. Впрочем, даже из наглядных соображений (рис. 13.6) очевидно, что минимум функции $u=x^2+y^2$ (графиком которой служит параболоид вращения) на всей плоскости Oxy не совпадает с ее минимумом на прямой $x+y-1=0$.

Переходим к общей постановке задачи об отыскании условного экстремума. Пусть требуется найти экстремум функции $m+n$ переменных

$$u=f(x_1, \ldots, x_n, y_1, \ldots, y_m) \tag{13.40}$$

при наличии m условий связи

$$\left.\begin{aligned} F_1(x_1, \ldots, x_n, y_1, \ldots, y_m)&=0,\\ F_2(x_1, \ldots, x_n, y_1, \ldots, y_m)&=0,\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ F_m(x_1, \ldots, x_n, y_1, \ldots, y_m)&=0. \end{aligned}\right\} \tag{13.41}$$

Прежде всего уточним само понятие условного экстремума функции (13.40) при наличии связей (13.41). Будем говорить, что *функция* (13.40) *при наличии связей* (13.41) *имеет условный максимум (минимум) в точке* $M_0(\mathring{x}_1, \ldots, \mathring{x}_n, \mathring{y}_1, \ldots \mathring{y}_m)$, *координаты которой удовлетворяют условиям связи* (13.41), *если найдется такая окрестность точки* M_0, *в пределах которой значение функции* (13.40) *в точке* M_0 *является наибольшим (наименьшим) среди ее значений во всех точках, координаты которых удовлетворяют условиям связи* (13.41).

Для нахождения условного экстремума функции (13.40) при наличии связей (13.41) предположим, что функции, стоящие в левых частях равенств (13.41), дифференцируемы в некоторой окрестности рассматриваемой точки M_0, причем в самой точке M_0 частные производные указанных функций по $y_1, \ldots, y_m$ непрерывны, а якобиан

$$\frac{D(F_1, \ldots, F_m)}{D(y_1, \ldots, y_m)} \tag{13.42}$$

отличен от нуля.

В таком случае в силу теоремы 13.2 для достаточно малых положительных чисел $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ найдется такая окрестность точки $M'_0(\mathring{x}_1, \ldots, \mathring{x}_n)$ пространства переменных $(x_1, \ldots, x_n)$, что всюду в пределах этой окрестности определены m функций

$$\left.\begin{aligned} y_1&=\varphi_1(x_1, \ldots, x_n),\\ y_2&=\varphi_2(x_1, \ldots, x_n),\\ \ldots\ldots\ldots\\ y_m&=\varphi_m(x_1, \ldots, x_n), \end{aligned}\right\} \tag{13.43}$$

удовлетворяющих условиям $|y_1-\mathring{y}_1|<\varepsilon_1, \ldots, |y_m-\mathring{y}_m|<\varepsilon_m$ и являющихся при наличии этих условий единственным и дифференци-

руемым решением системы уравнений (13.41). Подставляя найденные функции (13.43) в (13.40), мы сведем вопрос о существовании условного экстремума в точке M_0 у функции (13.40) при наличии связей (13.41) к вопросу о существовании безусловного экстремума в точке M'_0 у сложной функции аргументов $x_1, x_2, \ldots, x_n$.

$$u=f[x_1, \ldots, x_n, \varphi_1(x_1, \ldots, x_n), \ldots, \varphi_m(x_1, \ldots, x_n)]= \\ =\Phi(x_1, \ldots, x_n). \tag{13.44}$$

Вопрос о существовании безусловного экстремума функции (13.44) может быть решен методами, указанными в § 6 гл. 12*. Изложенная нами общая схема сведения условного экстремума к безусловному была реализована в рассмотренном выше частном примере. Постараемся теперь, не прибегая к решению системы (13.41), установить по крайней мере необходимые условия существования условного экстремума в точке M_0. Итак, пусть функция (13.40) дифференцируема в точке M_0 и имеет в этой точке условный экстремум при наличии связей (13.41) или, что то же самое, функция (13.44) имеет в точке M_0' безусловный экстремум. Согласно установленному в § 6 гл. 12 необходимым условием безусловного экстремума функции $u=\Phi(x_1, \ldots, x_n)$ в точке M_0' является равенство нулю в этой точке дифференциала этой функции

$$du=\frac{\partial\Phi}{\partial x_1}dx_1+\ldots+\frac{\partial\Phi}{\partial x_n}dx_n, \tag{13.45}$$

тождественное относительно $dx_1, \ldots, dx_n$. В силу инвариантности формы первого дифференциала и равенства (13.44) формулу (13.45) можно переписать в виде

$$du=\frac{\partial f}{\partial x_1}dx_1+\ldots+\frac{\partial f}{\partial x_n}dx_n+\frac{\partial f}{\partial y_1}dy_1+\ldots+\frac{\partial f}{\partial y_m}dy_m=0. \tag{13.46}$$

(В этой формуле все частные производные берутся в точке M_0.) Подчеркнем, однако, что в равенстве (13.46) $dy_1, \ldots, dy_m$ представляют собой дифференциалы функций (13.43), так что равенство (13.46) не является тождеством относительно $dy_1, \ldots, dy_m$. Предположим, что в уравнения связи (13.41) мы подставили функции (13.43), являющиеся решением системы (13.41). При этом уравнения (13.41) обратятся в тождества, и мы получим, дифференцируя эти тождества,

* При этом, конечно, придется подчинить функцию (13.40) некоторым условиям.

$$\left.\begin{array}{l}\frac{\partial F_1}{\partial x_1}dx_1+\ldots+\frac{\partial F_1}{\partial x_n}dx_n+\frac{\partial F_1}{\partial y_1}dy_1+\ldots+\frac{\partial F_1}{\partial y_m}dy_m=0,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\partial F_m}{\partial x_1}dx_1+\ldots+\frac{\partial F_m}{\partial x_n}dx_n+\frac{\partial F_m}{\partial y_1}dy_1+\ldots+\frac{\partial F_m}{\partial y_m}dy_m=0.\end{array}\right\}\quad(13.47)$$

Так как якобиан (13.42), по предположению, отличен от нуля в точке M_0, то из линейной системы (13.47) $dy_1, \ldots, dy_m$ могут быть выражены как линейные функции $dx_1, \ldots, dx_n$. Если найти эти выражения и подставить их в (13.46), то, собирая в полученном равенстве члены, содержащие $dx_1, \ldots, dx_n$, мы будем иметь

$$A_1dx_1+\ldots+A_ndx_n=0, \quad (13.48)$$

где через $A_1, \ldots, A_n$ обозначены некоторые рациональные функции частных производных f, $F_1, \ldots, F_m$ в точке M_0. Так как в равенстве (13.48) фигурируют *лишь дифференциалы независимых переменных*, то из этого равенства заключаем, что $A_1=0, \ldots, A_n=0$. Присоединяя к указанным равенствам m условий связи (13.41), мы получим необходимые условия существования условного экстремума функции (13.40) при наличии связей (13.41) в виде

$$A_1=0, \ldots, A_n=0, \quad F_1=0, \ldots, F_m=0. \quad (13.49)$$

Равенства (13.49) представляют собой систему $m+n$ уравнений для определения $m+n$ координат точки возможного экстремума.

2. Метод неопределенных множителей Лагранжа. При изложенном выше методе отыскания точек возможного условного экстремума мы нарушили симметрию в отношении переменных $x_1, \ldots, x_n$, $y_1, \ldots, y_m$. Часть из этих переменных $x_1, \ldots, x_n$ мы рассматривали как независимые, остальные — как функции этих переменных. В ряде случаев это приводит к усложнению выкладок. Лагранжем предложен метод, симметризирующий роль переменных. Изложению этого метода и посвящен настоящий пункт. Умножим равенства (13.47) соответственно на произвольные (и пока еще неопределенные) постоянные множители $\lambda_1, \ldots, \lambda_m$. Полученные после умножения равенства сложим почленно с равенством (13.46). В результате получим следующее равенство:

$$\frac{\partial\psi}{\partial x_1}dx_1+\ldots+\frac{\partial\psi}{\partial x_n}dx_n+\frac{\partial\psi}{\partial y_1}dy_1+\ldots+\frac{\partial\psi}{\partial y_m}dy_m=0, \quad (13.50)$$

где символом $\psi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ обозначена следующая функция:

$$\psi=f+\lambda_1F_1+\ldots+\lambda_mF_m. \quad (13.51)$$

Эту функцию мы в дальнейшем будем называть функцией Лагранжа. Считая, что для функций (13.41) выполнены усло-

вия, сформулированные в предыдущем пункте, и что функция (13.40) дифференцируема, выберем множители $\lambda_1, \ldots, \lambda_m$ так, чтобы выполнялись равенства

$$\frac{\partial\psi}{\partial y_1}=0, \ \ldots \ , \frac{\partial\psi}{\partial y_m}=0. \tag{13.52}$$

Это заведомо можно сделать, ибо равенства (13.52) приводят к линейной системе

$$\begin{cases} \dfrac{\partial f}{\partial y_1}+\lambda_1\dfrac{\partial F_1}{\partial y_1}+\ldots+\lambda_m\dfrac{\partial F_m}{\partial y_1}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{\partial f}{\partial y_1}+\lambda_1\dfrac{\partial F_1}{\partial y_m}+\ldots+\lambda_m\dfrac{\partial F_m}{\partial y_m}=0, \end{cases}$$

определитель которой (якобиан (13.42)) отличен от нуля. В силу равенства (13.52) равенство (13.50) принимает вид

$$\frac{\partial\psi}{\partial x_1}dx_1+\ldots+\frac{\partial\psi}{\partial x_n}dx_n=0. \tag{13.53}$$

Поскольку при сделанных выше предположениях переменные $x_1, \ldots, x_n$ являются независимыми, то из равенства (13.53) заключаем, что

$$\frac{\partial\psi}{\partial x_1}=0, \ \ldots \ , \frac{\partial\psi}{\partial x_n}=0. \tag{13.54}$$

Присоединяя к уравнениям (13.52) и (13.54) условия связи (13.41), мы получим систему $n+2m$ уравнений

$$\frac{\partial\psi}{\partial x_1}=0, \ \ldots \ , \frac{\partial\psi}{\partial x_n}=0, \ \frac{\partial\psi}{\partial y_1}=0, \ \ldots \ , \frac{\partial\psi}{\partial y_m}=0, \tag{13.55}$$
$$F_1=0, \ \ldots \ , F_m=0$$

для определения $n+m$ координат точек возможного условного экстремума и m множителей $\lambda_1, \ldots, \lambda_m$. Практически при реализации этого метода поступают следующим образом. Составляют функцию Лагранжа (13.51) и для этой функции находят точки возможного безусловного экстремума. Для исключения множителей $\lambda_1, \ldots, \lambda_m$ привлекают условия связи (13.41). Такой путь отыскания точек возможного условного экстремума является законным, ибо он приводит нас как раз к системе $n+2m$ уравнений (13.55). Пример применения метода множителей Лагранжа будет рассмотрен в п. 4.

3. Достаточные условия. В этом пункте мы рассмотрим один из путей дополнительного исследования точек возможного условного экстремума. Предположим, что в точке M_0 выполнены необходимые условия экстремума (13.55). Кроме того, дополнительно потребуем двукратной дифференцируемости функций (13.40) и

(13.41) в окрестности точки M_0 и непрерывности всех частных производных второго порядка в самой точке M_0. Из конструкции функции Лагранжа (13.51) очевидно, что при наличии связей (13.41) экстремумы функций (13.40) и Лагранжа совпадают *. Но тогда из результатов § 6 гл. 12 вытекает, что для получения достаточного условия экстремума в точке M_0 у функции (13.40) при наличии связей (13.41) *следует присоединить к условиям* (13.55) *требование знакоопределенности в этой точке* $d^2\psi$. При этом в соответствии с результатами § 6 гл. 12 мы можем констатировать наличие в точке M_0 минимума, если при наличии связей (13.41) $d^2\psi/M_0>0$, и максимума, если $d^2\psi/M_0<0$. Сделаем еще несколько замечаний практического характера. Прежде всего отметим, что *второй дифференциал* $d^2\psi$ *можно в данной точке* M_0 *возможного экстремума вычислять так, как если бы все переменные* $x_1, \ldots, x_n, y_1, \ldots, y_m$ *были независимыми.* В самом деле, в общем случае второй дифференциал $d^2\psi$ функции ψ не обладает свойством инвариантности формы и должен был бы с учетом зависимости $y_1, \ldots, y_m$ от $x_1, \ldots, x_n$ определяться равенством

$$d^2\psi=\left(dx_1\frac{\partial}{\partial x_1}+\ldots+dx_n\frac{\partial}{\partial x_n}+dy_1\frac{\partial}{\partial y_1}+\ldots+dy_m\frac{\partial}{\partial y_m}\right)^2\psi+$$
$$+\frac{\partial\psi}{\partial y_1}d^2y_1+\ldots+\frac{\partial\psi}{\partial y_m}d^2y_m.$$

Но в точке возможного экстремума справедливы равенства

$$\frac{\partial\psi}{\partial y_1}=0,\ \ldots,\ \frac{\partial\psi}{\partial y_m}=0,$$

так что $d^2\psi$ определяется той же формулой

$$d^2\psi=\left(dx_1\frac{\partial}{\partial x_1}+\ldots+dx_n\frac{\partial}{\partial x_n}+dy_1\frac{\partial}{\partial y_1}+\ldots+dy_m\frac{\partial}{\partial y_m}\right)^2\psi, \tag{13.56}$$

что и в случае, когда все переменные $x_1, \ldots, x_n, y_1, \ldots, y_m$ независимы. Далее, заметим, что поскольку нам требуется установить знакоопределенность $d^2\psi$ лишь при наличии связей (13.41), то при проведении вычислений следует в формулу (13.56) для $d^2\psi$ подставить вместо $dy_1, \ldots, dy_m$ их значения, определяемые из системы (13.47). После этого следует изучить вопрос о знакоопределенности $d^2\psi$ в данной точке M_0.

4. Пример. Найдем экстремальные значения функции переменных

$$u=x_1^2+x_2^2+\ldots+x_m^2 \tag{13.57}$$

* Это вытекает из того, что при наличии связей (13.41) разность $f(M)-f(M_0)$ совпадает с разностью $\psi(M)-\psi(M_0)$.

при наличии связи

$$x_1+x_2+\ldots+x_m+1=0. \tag{13.58}$$

Составим функцию Лагранжа

$$\psi=x_1^2+\ldots+x^2_m+\lambda(x_1+\ldots+x_m+1) \tag{13.59}$$

и для нее изучим вопрос о точках безусловного экстремума.

Так как для любого номера i, равного $1, 2, \ldots, m$, справедливо соотношение $\frac{\partial\psi}{\partial x_i}=2x_i+\lambda$, то единственной стационарной точкой является точка M_0 с координатами

$$M_0\left(-\frac{\lambda}{2}, -\frac{\lambda}{2}, \ldots, -\frac{\lambda}{2}\right).$$

Для определения λ используем условие связи (13.58), из которого получим, что

$$\left(-\frac{\lambda}{2}\right)m+1=0.$$

Таким образом, $\lambda=2/m$ и единственная стационарная точка имеет координаты

$$M_0(-1/m, -1/m, \ldots, -1/m).$$

Поскольку второй дифференциал функции Лагранжа (13.59), равный

$$d^2\psi=2[(dx_1)^2+\ldots+(dx_m)^2],$$

всегда положительно определен, то функция (13.57) при наличии связи (13.58) имеет в точке $M_0(-1/m, -1/m, \ldots, -1/m)$ условный минимум.

Подставляя координаты точки M_0 в (13.57), мы получим, что минимальное значение функции (13.57) при наличии связи (13.58) равно $u_{\min}=\frac{1}{m}$.

ДОПОЛНЕНИЕ

Отображения банаховых пространств. Аналог теоремы о неявной функции

1. Теорема о существовании и дифференцируемости неявной функции. В § 1 гл. 13 мы установили теорему 13.1 существования и дифференцируемости неявной функции, зависящей от числовых аргументов x, y, причем эта функция $u=\varphi(x, y)$ находилась как решение уравнения $F(u, x, y)=0$ в окрестности некоторой точки $M_0(\mathring{u}, \mathring{x}, \mathring{y})$.

Ниже мы убедимся, что указанная теорема 13.1 без больших изменений переносится с числовых функций на отображения произвольных банаховых пространств.

Предпошлем формулировке и доказательству этой теоремы некоторые обозначения.

Символом B мы обозначим некоторое банахово (т. е. полное нормированное) пространство.

Символом N будем обозначать нормированное пространство, через $X \times Y$ обозначим прямое произведение множеств X и Y, т. е. множество всевозможных упорядоченных пар (x, y), для которых $x \in X$, $y \in Y$.

Всюду ниже под производной некоторого отображения мы понимаем сильную производную, введенную в дополнении 3 к гл. 12.

Докажем теорему об отображениях нормированных пространств, обобщающую теорему 13.1 о неявной функции.

Теорема. *Пусть N_1, N_2 и B — линейные нормированные пространства, причем B — банахово пространство, Σ — окрестность точки (x_0, y_0), принадлежащей произведению пространств N_1 и B, а $F(x, y)$ — отображение окрестности Σ в пространство N_2, обладающее следующими свойствами:*

а) *$F(x, y)$ непрерывно в точке (x_0, y_0), $x, x_0 \in N_1$, $y, y_0 \in B$;*

б) *$F(x_0, y_0) = 0$;*

в) *частная производная $F_y'(x, y)$, т. е. производная по аргументу y при фиксированном элементе x*, существует в Σ и непрерывна в точке (x_0, y_0), а оператор $F_y'(x_0, y_0)$ имеет ограниченный обратный оператор** $[F_y'(x_0, y_0)]^{-1}$;*

г) *частная производная $F_x'(x, y)$ существует в каждой точке окрестности Σ и непрерывна в самой точке (x_0, y_0).*

При этих условиях в некоторой окрестности Σ_0 точки x_0 пространства N_1 определено отображение $y = \varphi(x)$: $\Sigma_0 \to B$ такое, что:

а°. *$F(x, \varphi(x)) = 0$ в Σ_0 и $y(x) = \varphi(x)$ непрерывно в точке x_0;*

б°. *$y_0 = \varphi(x_0)$;*

в°. *если $\varphi_1(x)$ — какое-либо отображение, определенное в некоторой окрестности Σ_1 точки x_0, обладающее свойствами а° и б°, то $\varphi_1(x) \equiv \varphi(x)$ в некоторой окрестности $\widetilde{\Sigma} \subset \Sigma_1 \cap \Sigma_0$ точки x_0;*

---

* См. по этому поводу также п. 2.

** Пусть A — оператор, действующий из B в N, D_A — область определения (т. е. множество, на котором определено отображение A), а R_A — образ множества D_A, т. е. множество элементов вида Ax, $x \in D_A$. Оператор A называется обратимым, если для любого $z \in R_A$ уравнение $Ax = z$ имеет единственное решение. Если A обратим, то каждому $z \in R_A$ можно поставить в соответствие единственный элемент $x \in D_A$, являющийся решением уравнения $Ax = z$. Оператор, осуществляющий это соответствие, называется обратным к A и обозначается символом A^{-1}.

г°. *отображение* $y=\varphi(x)$ *дифференцируемо в точке* x_0 *и*

$$y'=\varphi'(x_0)=-[F_y'(x_0, y_0)]^{-1}F_x'(x_0, y_0).$$

Доказательство. Без ограничения общности можно считать, что $x_0=0$, $y_0=0$ (в противном случае можно сделать замену $x_1=x-x_0$, $y_1=y-y_0$, и утверждение сведется к указанному случаю $x_0=0$, $y_0=0$).

Согласно условию теоремы окрестность Σ точки $(0,0)$ принадлежит произведению $N_1\times B$ пространств N_1 и B. Очевидно, что окрестность Σ содержит такую окрестность точки $(0,0)$, которая является прямым произведением окрестности $\|x\|<\delta$ точки $0\in N_1$ и окрестности $\|y\|<\varepsilon$ точки $0\in B$.

Итак, можно считать, что $x_0=0$, $y_0=0$ и

$$\Sigma=\{(x, y)\in N_1\times B : \|x\|<\delta, \|y\|<\varepsilon\}.$$

Рассмотрим при *фиксированном* x отображение

$$A_xy=y-[F_y'(0,0)]^{-1}F(x, y)$$

окрестности точки $0\in B$ в пространство B. Подчеркнем, что в силу условий теоремы отображение $F(x, y)$ определено при $\|x\|<\delta$, $\|y\|<\varepsilon$ и его значения лежат в пространстве N_2. Отображение $F_y'(x, y)$ определено и непрерывно в точке $(0,0)$. По условию в) отображение $F_y'(0,0) : B\to N_2$ имеет обратное $[F_y'(0,0)]^{-1} : N_2\to B$. Таким образом, последовательное применение отображений $F(x, y)$ и $[F_y'(0,0)]^{-1}$, т. е. отображение $[F_y'(0,0)]^{-1}F(x, y)$ определено и его значения лежат в B.

Заметим, что точка $y=\varphi(x)$ является неподвижной точкой отображения A_x (см. дополнение 2 к гл. 12) тогда и только тогда, когда $F(x, \varphi(x))=0$. Таким образом, отыскание и исследование функции $y=\varphi(x)$ сводится к отысканию и исследованию неподвижных точек отображения A_x.

Для того чтобы отыскать неподвижные точки отображения A_x, применим принцип сжимающих отображений.

Заметим для этого, что при любом фиксированном элементе $x\in N_1$ таком, что $\|x\|<\delta$, отображение A_xy дифференцируемо по y в области $\|y\|<\varepsilon$, $y\in B$, что следует из свойства производной сложной функции (см. дополнение 3 к гл. 12, п. 1, свойство 3), причем

$$(A_xy)_y'=(Ey)_y'-[F_y'(0,0)]^{-1}F_y'(x, y);$$

здесь E — единичный оператор. Поскольку, согласно свойству 2 (дополнение 3 к гл. 12), производная ограниченного линейного оператора есть сам этот оператор, т. е. $(Ey)'=E$, то мы получим, что

$$(A_xy)_y'=E_y-[F_y'(0,0)]^{-1}F_y'(x, y)=$$
$$=[F_y'(0,0)]^{-1}[F_y'(0,0)-F_y'(x, y)].$$

Отображение $F_y'(x, y)$ по условию а) теоремы непрерывно в точке (0,0), поэтому всегда можно найти такую окрестность $\|x\|<\beta<\delta$, $\|y\|<\beta<\varepsilon$ точки $(0, 0)\in N_1\times B$, в которой *

$$\|F_y'(0, 0)-F_y'(x, y)\|<(2\|[F_y'(0, 0)]^{-1}\|)^{-1}.$$

Следовательно, при $\|x\|<\beta$, $\|y\|<\beta$ получим, что

$$\|(A_xy)_y'\|\leqslant\|[F_y'(0,0)]^{-1}\|\,\|F_y'(0, 0)-F_y'(x, y)\|<2^{-1}.$$

Всюду ниже будем считать, что $\|x\|<\beta$, $\|y\|<\beta$, и поэтому $\|(A_xy)_y'\|<2^{-1}$.

По формуле конечных приращений для отображений при любом фиксированном x, $\|x\|<\beta$, и любых y_1 и y_2 таких, что $\|y_1\|<\beta$, $\|y_2\|<\beta$, получим (см. дополнение 3 к гл. 12, п. 2)

$$\|A_xy_1-A_xy_2\|\leqslant\sup_{0<\theta<1}\|A_x'y_1+\theta(y_2-y_1)\|\|y_1-y_2\|<\frac{1}{2}\|y_1-y_2\|^{**}.$$

Здесь мы воспользовались тем, что $\|y_1+\theta(y_2-y_1)\|<\beta$ при $\|y_1\|<\beta$, $\|y_2\|<\beta$ и при любом $0<\theta<1$

$$\|y_1+\theta(y_2-y_1)\|\leqslant\|y_1(1-\theta)\|+\|\theta y_2\|=(1-\theta)\|y_1\|+\theta\|y_2\|<\beta.$$

Таким образом, A_x — семейство сжимающих отображений окрестности $\|y\|<\beta$ точки $0\in B$ в пространстве B, причем коэффициент сжатия 1/2 не зависит от параметра $x\in N_1$, $\|x\|<\beta$.

Для того чтобы применить принцип сжимающих отображений к отображению A_x, следует указать то полное метрическое или нормированное пространство, которое при этом отображении переходит в себя.

Укажем это пространство. Зафиксируем произвольное число ε_1 такое, что $0<\varepsilon_1<\beta$. Покажем, что существует такое число $\delta_1(\varepsilon_1)$, $0<\delta_1(\varepsilon_1)<\beta$, что при любом $x\in N_1$ и $\|x\|<\delta_1(\varepsilon_1)$ (эту окрестность в пространстве N_1 обозначим через Σ_0) отображение A_x преобразует замкнутый шар $K(0, \varepsilon_1)$ с центром в точке 0 и радиуса ε_1 в пространстве B в себя.

Действительно, в силу того, что отображение $F(x, y)$ непрерывно в точке $(0, 0)$, и того, что $F(0, 0)=0$, для указанного ε_1 найдется такое положительное число $\delta_1(\varepsilon_1)$, что при $\|x\|<\delta_1(\varepsilon_1)<\beta$

$$\|A_x0\|=\|[F_y'(0,0)]^{-1}F(x, 0)\|\leqslant\|[F_y'(0, 0)]^{-1}\|\,\|F(x, 0)\|<\frac{1}{2}\varepsilon_1.$$

---

* Напомним, что по условию в) теоремы оператор $[F_y'(0, 0)]^{-1}$ непрерывен, а следовательно, ограничен по норме; норма этого оператора отлична от нуля, так как область его значений, совпадая с областью определения оператора $F_y'(0, 0)$, не является лишь нулевым элементом.

** Здесь мы обозначили $(A_x(\cdot))_y'$ через $A_x'(\cdot)$.

Далее, если $\|x\|<\delta_1(\varepsilon_1)$, $\|y\|\leqslant\varepsilon_1$, то в силу того, что отображение A_x сжимающее, и полученной оценки для $\|A_x0\|$ имеем

$$\|A_xy\|=\|A_xy-A_x0+A_x0\|\leqslant\|A_xy-A_x0\|+\|A_x0\|<\frac{1}{2}\|y\|+$$

$$+\frac{1}{2}\varepsilon_1\leqslant\varepsilon_1.$$

Следовательно, A_x преобразует замкнутый шар $K(0, \varepsilon_1)$ пространства B в себя. Элемент x при этом фиксирован и $\|x\|<\delta_1(\varepsilon_1)$.

Замкнутый шар $K(0, \varepsilon_1)$ является метрическим пространством, причем (в силу замкнутости) полным. В самом деле, B — полное метрическое пространство. Всякая фундаментальная последовательность точек $y_n\in K(0, \varepsilon_1)\subset B$ сходится к точке $y_0\in B$, которая в силу замкнутости $K(0, \varepsilon_1)$ принадлежит $K(0, \varepsilon_1)$. Поэтому $K(0, \varepsilon_1)$ — полное метрическое пространство и A_x — сжимающее отображение, определенное на нем. Согласно принципу неподвижной точки (см. дополнение 2 к гл. 12) получаем, что при каждом $x\in\Sigma_0$ найдется единственная точка $y=\varphi(x)$ в шаре $K(0, \varepsilon_1)$, неподвижная относительно отображения A_x. Для этой точки справедливо соотношение

$$A_x\varphi(x)=\varphi(x)=\varphi(x)-[F_y'(0, 0)]^{-1}F(x, \varphi(x)),$$

или

$$F(x, \varphi(x))\equiv0,$$

если

$$\|x\|<\delta_1(\varepsilon_1).$$

Следовательно, функция $y=\varphi(x)$ удовлетворяет уравнению $F(x, \varphi(x))=0$. Эта функция непрерывна в нуле, так как по любому положительному $\varepsilon_1<\beta$, мы можем найти такое $\delta_1(\varepsilon_1)$, что при $\|x\|<\delta_1(\varepsilon_1)$ отображение A_x переводит шар $K(0, \varepsilon_1)$ в себя, т. е. единственная неподвижная точка $y=\varphi(x)$ этого отображения при $x\in\Sigma_0$ принадлежит шару $K(0, \varepsilon_1)$, т. е. удовлетворяет условию $\|\varphi(x)\|\leqslant\varepsilon_1$.

Утверждение а° полностью доказано.

Далее, поскольку

$$A_00=0-[F_y'(0, 0)]^{-1}F(0, 0)=0,$$

то в силу единственности неподвижной точки имеем при $x=0$ $y_0=\varphi(0)=0$. Тем самым установлено утверждение б°.

Далее, если $\varphi_1(x)$ — отображение, удовлетворяющее условиям а° и б° в некоторой окрестности Σ_1 точки $0\in N_1$, то найдется такое $\delta_2(\varepsilon_1)$, что $\|\varphi_1(x)\|\leqslant\varepsilon_1$ при $\|x\|<\delta_2(\varepsilon_1)$. При $\|x\|<\min(\delta_1(\varepsilon_1), \delta_2(\varepsilon_1))$ одновременно будут выполнены соотношения $\|\varphi(x)\|\leqslant\varepsilon_1$ и $\|\varphi_1(x)\|\leqslant\varepsilon_1$. В силу того, что $F(x, \varphi_1(x))=0$, будет справедливо равенство

$$A_x\varphi_1(x)=\varphi_1(x)-[F_y'(0,0)]^{-1}F(x,\ \varphi_1(x))=\varphi_1(x)$$

при любом x таком, что $\|x\|<\min(\delta_1(\varepsilon_1),\ \delta_2(\varepsilon_1))$.

Следовательно, $\varphi_1(x)$ — также неподвижная точка отображения A_x, принадлежащая шару $K(0,\ \varepsilon_1)$. Однако неподвижная точка может быть только одна, поэтому $\varphi_1(x)\equiv\varphi(x)$ в окрестности

$$\widetilde{\Sigma}=\{x:\|x\|<\min\ (\delta_1(\varepsilon_1),\ \delta_2(\varepsilon_1))\}\subset\Sigma_1\cap\Sigma_0.$$

Тем самым доказано утверждение в°.

Осталось установить утверждение г° о дифференцируемости неявной функции.

Обозначим через L линейный ограниченный оператор, действующий из N_1 в B по правилу

$$L=\varphi'(0)=-[F_y'(0,0)]^{-1}F_x'(0,0).$$

Проверим, что этот оператор L является производной отображения $y=\varphi(x)$ в точке 0. Напомним, что для этого необходимо существование для каждого $\varepsilon>0$ такого $\delta>0$, что для любого x, удовлетворяющего условию $\|x\|<\delta$, выполнено неравенство

$$\|\varphi(x)-\varphi(0)-Lx\|\leqslant\varepsilon\|x\|.$$

Учитывая, что $\varphi(0)=0$, и используя выражение для L запишем соотношение

$$\varphi(x)=\varphi(0)-Lx=\varphi(x)+[F_y'(0,0)]^{-1}F_x'(0,0)x=$$
$$=[F_y'(0,0)]^{-1}[F_x'(0,0)x+F_y'(0,0)\varphi(x)].$$

Поскольку * $F(x,\ \varphi(x))=F(0,0)=0$, то с помощью формулы конечных приращений для отображений получим, обозначая буквой y функцию $\varphi(x)$:

$$\|y-Lx\|\leqslant\|[F_y'(0,0)]^{-1}\|\|F(x,y)-F(0,\ 0)-F_x'(0,0)x-$$
$$-F_y'(0,0)y\|\leqslant\|[F_y'(0,0)]^{-1}\|[\sup_{0<\theta,\ \theta_1<1}\|F_x'(\theta x,\theta_1 y)-F_x'(0,0)\|\|x\|+$$
$$+\sup_{0<\theta,\ \theta_1<1}\|F_y'(\theta x,\theta_1 y)-F_y'(0,0)\|\|y\|]\leqslant\eta(\|x\|+\|y\|),$$

где величина η может быть сделана сколь угодно малой (в силу непрерывности в точке $(0,0)$ производных F_x' и F_y'), если величина δ достаточно мала. Таким образом, $\|y-Lx\|\leqslant\eta(\|x\|+\|y\|)\leqslant\eta(\|x\|+\|Lx\|+\|y-Lx\|)$. Отсюда при достаточно малом η получаем ** $\|y-Lx\|\leqslant\eta(1-\eta)^{-1}(1+\|L\|)\ \|x\|$. Достаточно выбрать η так, чтобы было выполнено неравенство

$$\eta(1-\eta)^{-1}(1+\|L\|)\leqslant\varepsilon.$$

* Все рассмотрения ведутся в окрестности Σ_0 точки 0, в которой существует единственная неявная функция $y=\varphi(x)$.

** Ниже мы используем неравенство $\|Lx\|\leqslant\|L\|\|x\|$.

Дифференцируемость отображения $y=\varphi(x)$ в точке 0 доказана, и утверждение г° установлено.

Таким образом, теорема полностью доказана *.

2. Случай конечномерных пространств. Рассмотрим важный частный случай, когда нормированное пространство N_1 совпадает с пространством E^m, а пространства B и N_2 совпадают с E^n **. В качестве следствия теоремы получим теорему о неявной функции для этого частного случая.

Выберем в пространствах E^m и E^n базисы и запишем отображение $F(x, y)$ в координатной форме:

$$F_1=F_1(x_1, \ldots, x_m, y_1, \ldots, y_n),$$
$$\cdots\cdots\cdots\cdots$$
$$F_n=F_n(x_1, \ldots, x_m, y_1, \ldots, y_n),$$

где $(x_1, \ldots, x_m)=x$ — точка пространства E^m, $(y_1, \ldots, y_n)=y$ — точка пространства E^n, причем точка $(F_1, \ldots, F_n)$ принадлежит также E^n.

Подчеркнем, что отображение $F(x, y)$ определено на прямом произведении пространств $N_1\times B$ или в данном частном случае на прямом произведении $E^m\times E^n$. Если фиксировать переменную y, положив ее равной y_0, то мы получим функцию, определенную *** на некотором множестве $E^m_{y_0}$ — сечении $E^m\times E^n$. Отображение $F(x, y_0)$ представляет собой частичное отображение по переменной x (по отношению к исходному отображению $F(x, y)$).

Если в некоторой точке $x_0\in E^m_{y_0}$ это частичное отображение $F(x, y_0)$ дифференцируемо, то его производная в этой точке x_0 называется **частной производной** *исходного отображения $F(x, y)$ в точке (x_0, y_0) (или частным производным отображением) по переменной x. Обозначается эта частная производная символом $F'_x(x_0, y_0)$ или $\frac{\partial F(x_0, y_0)}{\partial x}$ (иногда $D_x F(x_0, y_0)$).*

Ясно, что наши рассмотрения пригодны и в *общем случае нормированных пространств* N_1, B, а не только в случае конечномерных пространств E^m и E^n.

---

* Можно показать, что если в окрестности Σ точки (x_0, y_0) существуют и непрерывны частные производные отображения $F: F'_x$ и F'_y то и в некоторой окрестности точки x_0 функция $y(x)$ непрерывно дифференцируема и ее производное отображение вычисляется по формуле

$$y'(x)=-[F'_y(x, y(x))]^{-1}[F'_x(x, y(x))].$$

** Мы не предполагаем какой-либо конкретной реализации элементов конечномерных пространств E^m, E^n, т. е. элементы этих пространств не обязательно упорядоченные совокупности чисел (ср. п. 1 § 1 гл. 13).

*** Все рассмотрения проводятся аналогично, если отображение $F(x, y)$ определено не на всем произведении $E^m\times E^n$, а лишь на некотором, например открытом, множестве Σ, принадлежащем $E^m\times E^n$.

Замечание. Нетрудно, так же как и в случае функций многих переменных, доказать, что *если отображение* $F(x, y)$: $N_1 \times B \to N_2$ *дифференцируемо в точке* (x_0, y_0), *то в этой точке у него есть все частные производные отображения и справедливо соотношение* $dF(x_0, y_0) = \frac{\partial F(x_0, y_0)}{\partial x} dx + \frac{\partial F(x_0, y_0)}{\partial y} dy$. *Здесь* $dF(x_0, y_0)$ — *дифференциал Фреше от функции* $F(x, y)$ *в точке* (x_0, y_0); $\frac{\partial F(x_0, y_0)}{\partial x}$ и $\frac{\partial F(x_0, y_0)}{\partial y}$ — *частные производные отображения* $F(x, y)$ *в точке* (x_0, y_0), *а* (dx, dy) — *элемент касательного пространства* $K_{(x_0, y_0)} N_1 \times B^*$.

Вернемся снова к изучению отображения $F(x, y)$: $E^m \times E^n \to E^n$. Согласно рассмотрениям п. 8 дополнения 3 к гл. 12 частные производные отображения $F_x' : N_1 \to N_2$, $F_y' : B \to N_2$ или, что то же самое, $F_x' : E^m \to E^n$, $F_y' : E^n \to E^n$ задаются матрицами

$$F_x' = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_m} \\ \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_m} \end{pmatrix}, \quad F_y' = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_n} \\ \dots & \dots & \dots \\ \frac{\partial F_n}{\partial y_1} & \dots & \frac{\partial F_n}{\partial y_n} \end{pmatrix}.$$

Непрерывность отображений F_x' и F_y' (операторов в конечномерных пространствах, задаваемых матрицами), как известно из курса линейной алгебры, равносильна непрерывности всех элементов соответствующей матрицы. Обратимость линейного отображения $F_y'(x_0, y_0) : E^n \to E^n$ равносильна невырожденности матрицы, задающей это преобразование.

Теперь мы подготовлены к тому, чтобы сформулировать теорему в случае отображения $F(x, y) : E^m \times E^n \to E^n$.

Теорема. *Пусть* E^m *и* E^n — *два конечномерных пространства размерностей* m *и* n *соответственно,* Σ — *окрестность точки* (x_0, y_0), $x_0 = (\mathring{x}_1, \dots, \mathring{x}_m)$, $y_0 = (\mathring{y}_1, \dots, \mathring{y}_n)$, *лежащей в произведении пространств* E^m *и* E^n, *и* $F(x, y)$ — *отображение* $\Sigma \subset E^m \times E^n$ *в пространство* E^n, *обладающее свойствами*:

а) $F(x, y)$ *непрерывно в точке* (x_0, y_0); $x \in E^m$, $y, y_0 \in E^n$ (*т. е. в точке* (x_0, y_0) *непрерывны все координаты* F_i, $i = 1, \dots, n$);

---

* Если x — фиксированная точка, то $K_x N$ можно отождествить с пространством N. Пусть $F(x)$ — дифференцируемое отображение нормированного пространства N_1 в нормированное пространство N_2, а $L_x h$ — дифференциал Фреше (см. дополнение 3 к гл. 12), тогда можно сказать, что L_x — линейное отображение касательного пространства $K_x N_1$ к N_1 в точке x в касательное пространство $K_x N_2$ к N_2 в точке $F(x)$. Часто элементы пространства $K_x N_1$ обозначают через dx, элементы пространства $K_x N_2$ — через dF, поэтому $dF = h_x dx$. Касательным пространством к N в точке x называют пару $K_x N = (x, N)$. Понятие касательного пространства формализует такие понятия, как «вектор, приложенный в точке», «вектор с началом в точке» и т. д. В рассматриваемом случае касательное пространство $K_{(x_0, y_0)} N_1 \times B$ есть произведение касательных пространств $K_{x_0} N_1 \times K_{y_0} N_2$.

б) $F(x_0, y_0)=0$ (*т. е.* $F_1(\overset{\circ}{x}_1, \ldots, \overset{\circ}{x}_m, \overset{\circ}{y}_1, \ldots, \overset{\circ}{y}_n)=0, \ldots, F_n(\overset{\circ}{x}_1, \ldots$ $\ldots, \overset{\circ}{x}_m, \overset{\circ}{y}_1, \ldots, \overset{\circ}{y}_n)=0$), *где*

$$x_0=(\overset{\circ}{x}_1, \ldots, \overset{\circ}{x}_m), \quad y_0=(\overset{\circ}{y}_1, \ldots, \overset{\circ}{y}_n);$$

в) *частная производная*

$$F'_y(x, y)=\begin{pmatrix} \dfrac{\partial F_1(x, y)}{\partial y_1} & \ldots & \dfrac{\partial F_1(x, y)}{\partial y_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial F_n(x, y)}{\partial y_1} & \ldots & \dfrac{\partial F_n(x, y)}{\partial y_n} \end{pmatrix}$$

существует в каждой точке окрестности Σ *и непрерывна в точке* (x_0, y_0) (*т. е. в точке* (x_0, y_0) *непрерывны все элементы указанной матрицы*), *отображение*

$$F'_y(x_0, y_0)=\begin{pmatrix} \dfrac{\partial F_1(x_0, y_0)}{\partial y_1} & \ldots & \dfrac{\partial F_1(x_0, y_0)}{\partial y_n} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial F_n(x_0, y_0)}{\partial y_1} & \ldots & \dfrac{\partial F_n(x_0, y_0)}{\partial y_n} \end{pmatrix}$$

не вырождено (*т. е. определитель указанной матрицы Якоби отличен от нуля*);

г) *частная производная*

$$F'_x(x, y)=\begin{pmatrix} \dfrac{\partial F_1(x, y)}{\partial x_1} & \ldots & \dfrac{\partial F_1(x, y)}{\partial x_m} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial F_n(x, y)}{\partial x_1} & \ldots & \dfrac{\partial F_n(x, y)}{\partial x_m} \end{pmatrix}$$

существует в каждой точке окрестности Σ *и непрерывна в точке* (x_0, y_0) (*т. е. в точке* (x_0, y_0) *непрерывны все элементы указанной матрицы*).

При этих условиях в некоторой окрестности Σ_0 *точки* $x_0 \in E^m$ *определено отображение* $y=\varphi(x)$ (*т. е. определены функции* $y_k=\varphi_k(x_1, \ldots, x_m)$, $k=1, \ldots, n$) *такое, что*:

а°. $F(x, \varphi(x))=0$ ($F_i(x_1, \ldots, x_m, y_1, \ldots, y_n)\equiv 0$, $i=1, \ldots, n$ *и* $y=\varphi(x)$ *непрерывно в точке* x_0 (*т. е. координаты* $\varphi_1(x), \ldots, \varphi_n(x)$ *отображения* $\varphi(x)$, $x=(x_1, \ldots, x_m)$ *непрерывны в точке* x_0);

б°. $y_0=\varphi(x_0)$ (*т. е.* $\overset{\circ}{y}_k=\varphi_k(\overset{\circ}{x}_1, \ldots, \overset{\circ}{x}_m)$, $k=1, \ldots, n$);

в°. *если* $\varphi_1(x)$ — *другое отображение, определенное в некоторой окрестности* Σ_1 *точки* x_0, *обладающее свойствами* а° *и* б°, *то* $\varphi_1(x)\equiv\varphi(x)$ *в некоторой окрестности* $\widetilde{\Sigma}\subset\Sigma_1\cap\Sigma_0$ *точки* x_0;

г°. *отображение* $y=\varphi(x)$ *дифференцируемо в точке* x_0 *и*

$$y'=\varphi'(x_0)=-[F'_y(x_0, y_0)]^{-1}F'_x(x_0, y_0)=$$

$$=-\begin{pmatrix} \frac{\partial F_1(x_0, y_0)}{\partial y_1} & \dots & \frac{\partial F_1(x_0, y_0)}{\partial y_n} \\ \dots & \dots & \dots \\ \frac{\partial F_n(x_0, y_0)}{\partial y_1} & \dots & \frac{\partial F_n(x_0, y_0)}{\partial y_n} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1(x_0, y_0)}{\partial x_1} & \dots & \frac{\partial F_1(x_0, y_0)}{\partial x_m} \\ \dots & \dots & \dots \\ \frac{\partial F_n(x_0, y_0)}{\partial x_1} & \dots & \frac{\partial F_n(x_0, y_0)}{\partial x_m} \end{pmatrix}.$$

Замечание. В случае, если E^n одномерно, т. е. E^n совпадает с E^1, мы получим всего одно скалярное уравнение $F(x_1, \dots, x_m, y)=0$, где $x=(x_1, \dots, x_m)$ — точка из E^m, $y \in E^1$ — числовой аргумент. В этом случае отображение F_y' состоит только из одного элемента $\frac{\partial F}{\partial y}$, и поэтому $(F_y')^{-1} = 1 \Big/ \frac{\partial F}{\partial y}$. Поскольку в этом случае $\frac{\partial F}{\partial x} = F_x' = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_m}\right)$, то согласно утверждению г° теоремы получаем, что

$$y' = \left(\frac{\partial y}{\partial x_1}, \dots, \frac{\partial y}{\partial x_m}\right) = -\left(\frac{\partial F}{\partial y}\right)^{-1} \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_m}\right),$$

или, приравнивая координаты, получаем

$$\frac{\partial y}{\partial x_i} = -\frac{\partial F/\partial x_i}{\partial F/\partial y}, \quad i = 1, \dots, m.$$

Если и пространство E^m одномерно, т. е. $E^m = E^1$, то точка x — число и в этом случае $\frac{\partial y}{\partial x} = -\frac{\partial F}{\partial x} \Big/ \frac{\partial F}{\partial y}$. Эти формулы были нами установлены в § 1 и 2.

3. Особые точки поверхности в пространстве n измерений. Обратное отображение. В п. 3 § 1 настоящей главы мы рассматривали понятия особой и обыкновенной точек поверхности в трехмерном пространстве E^3. Рассмотрим теперь обобщения этих понятий на случай пространства любого конечного числа измерений.

Пусть $y=f(x)$ — отображение окрестности точки x_0 пространства E^m в пространство E^n, дифференцируемое в точке x_0. Если в пространствах E^m и E^n выбрать базисы, то отображение $y=f(x)$ может быть записано в координатной форме $[x = (x_1, x_2, \dots, x_m)]$:

$$y_1 = f_1(x_1, \dots, x_m),$$
$$\dots\dots\dots$$
$$y_n = f_n(x_1, \dots, x_m).$$

Рангом отображения $y=f(x)$ в точке $x_0 = (\mathring{x}_1, \dots, \mathring{x}_m)$ называется число, равное рангу матрицы

$$\begin{pmatrix} \frac{\partial f_1(x_0)}{\partial x_1} & \dots & \frac{\partial f_1(x_0)}{\partial x_m} \\ \frac{\partial f_n(x_0)}{\partial x_1} & \dots & \frac{\partial f_n(x_0)}{\partial x_m} \end{pmatrix}.$$

Обычно ранг отображения $f(x)$ в точке x_0 обозначается символом $\operatorname{rang} f(x_0)$. Ясно, что $\operatorname{rang} f(x_0) \leqslant \min(m, n)$.

Определение. *Точка x_0 пространства E^m называется особой (критической, сингулярной) точкой отображения $f(x)$ некоторой окрестности этой точки в пространство E^n, если ранг этого отображения в точке x_0 меньше наименьшего из чисел m и n, т. е. если*

$$\operatorname{rang} f(x_0) < \min(m, n).$$

Если в точке x_0 выполнено соотношение $\operatorname{rang} f(x_0) = \min(m, n)$, то точка x_0 называется обыкновенной или неособой для отображения $y = f(x)$.

Пусть $x_0 \in E^m$ и $x_0 = (\mathring{x}_1, \ldots, \mathring{x}_m)$. Рассмотрим соотношение $F(\mathring{x}_1, \ldots, \mathring{x}_m)$, где F — числовая непрерывно дифференцируемая функция, $F: E^m \to E^n$. Предположим, что точка $x_0 = (\mathring{x}_1, \ldots, \mathring{x}_m)$ не является особой для функции F. Поскольку в этом случае матрица, определяющая ранг отображения, состоит из одной строки

$$\left(\frac{\partial F(x_0)}{\partial x_1}, \ldots, \frac{\partial F(x_0)}{\partial x_m}\right),$$

то сформулированное выше условие того, что точка x_0 не является особой для отображения $F(x)$ (т. е. условие $\operatorname{rang} F = \min(m, 1) = 1$), означает, что хотя бы одна из частных производных F в этой точке отлична от нуля. Пусть ради определенности $\frac{\partial F(x_0)}{\partial x_m} \neq 0$. Поскольку $F(x_0) = 0$, то по теореме о неявном отображении переменную x_m в окрестности точки $x_0 = (\mathring{x}_1, \ldots, \mathring{x}_m)$ можно выразить в виде функциональной зависимости

$$x_m = f(x_1, \ldots, x_{m-1}).$$

Таким образом, в окрестности неособой точки мы разрешили функциональную систему уравнений, состоящую из одного уравнения, зависящего от m переменных.

Пусть теперь в некоторой окрестности точки $x_0 = (\mathring{x}_1, \ldots, \mathring{x}_{n+m})$ пространства E^{n+m} задано обращающееся в нуль в этой точке отображение $F: E^{n+m} \to E^m$, $m > 1$, указанной окрестности в пространство E^m, непрерывное в этой окрестности и обладающее в ней частными производными первого порядка по всем переменным, непрерывным в самой точке x_0. Тогда если точка x_0 не является особой, то $\operatorname{rang} F(x_0) = m$, т. е. найдется хотя бы один минор порядка m, отличный от нуля. Пусть это будет, например, минор $D(F_1, \ldots, F_m)/D(x_1, \ldots, x_m)$. Тогда по теореме о неявной функции уравнение $F(x) = 0$, т. е. система

$$\left.\begin{array}{l} F_1(x_1, \ldots, x_{n+m}) = 0, \\ \cdots\cdots\cdots\cdots \\ F_m(x_1, \ldots, x_{n+m}) = 0 \end{array}\right\} \qquad (*)$$

разрешима в указанной окрестности точки x_0 относительно переменных $x_1, \ldots, x_m$, т. е. существуют единственные функции

$$x_1 = \varphi_1(x_{m+1}, \ldots, x_{m+n}),$$
$$\cdots\cdots\cdots\cdots$$
$$x_m = \varphi_m(x_{m+1}, \ldots, x_{m+n}),$$

являющиеся решением системы функциональных уравнений $(*)$; эти функции будут непрерывны в некоторой окрестности точки $x_0' = (\mathring{x}_{m+1}, \ldots, \mathring{x}_{m+n})$ и дифференцируемы в этой точке. (Если предположить, что частные производные $\frac{\partial F_i}{\partial x_j}$ непрерывны и в некоторой окрестности точки x_0, то и функции $\varphi_1, \ldots, \varphi_m$ будут дифференцируемы в некоторой окрестности точки $(\mathring{x}_{m+1}, \ldots, \mathring{x}_{m+n})$; см. сноску п. 2.)

Подчеркнем, что условие неравенства нулю якобиана

$$\frac{D(F_1, \ldots, F_m)}{D(x_1, \ldots, x_m)}$$

означает при принятых нами определениях, что отображение $F(x', x_0') : E^m \to E^m$ (x_0' — фиксированная точка) имеет ранг в точке $(\mathring{x}_1, \ldots, \mathring{x}_m) = \mathring{x}'$, равный m: $\operatorname{rang} F(\mathring{x}', x_0') = \min(m, m) = m$, т. е. означает, что точка $\mathring{x}'$ является неособой для отображения $F(x', x_0') : E^m \to E^m$; выше $x_0' = (\mathring{x}_{m+1}, \ldots, \mathring{x}_{m+n})$, $x' = (x_1, \ldots, x_m)$.

В качестве еще одного следствия теоремы о неявном отображении является следующая теорема об обратном отображении (об обратной функции) (см. также п. 4 § 1 настоящей главы).

Теорема. *Пусть $F : \Sigma \to N$ — отображение открытой окрестности Σ точки x_0 банахова пространства B в нормированное пространство N. Пусть отображение $y = F(x)$ дифференцируемо в Σ, $F'(x)$ непрерывно в точке x_0 и оператор $F'(x_0)$ — обратимый.*

Тогда найдутся окрестность Σ_1 точки x_0, $\Sigma_1 \subset B$, и окрестность Σ_2 точки $y_0 = F(x_0)$, $\Sigma_2 \subset N$, такие, что отображение $F : \Sigma_1 \to \Sigma_2$ взаимно однозначно отображает Σ_1 на Σ_2, а обратное отображение F^{-1}: $\Sigma_2 \to \Sigma_1$ непрерывно в Σ_2 и дифференцируемо в точке x_0, причем $(F^{-1})'(y_0) = [F'(x_0)]^{-1}$.

Для доказательства этой теоремы достаточно применить теорему о неявной функции (о неявном отображении) к отображению

$$G(y, x) = F(x) - y,$$

определенному на множестве $N \times \Sigma$, $\Sigma \subset B$, и принимающему значения в пространстве N.

Действительно, в этом случае для $G(y, x)$ выполнены все условия теоремы п. 1: $G(y_0, x_0)=0$, $G(y, x)$ непрерывно в точке (y_0, x_0), $G_x'(y, x)=F'(x)$, $G_y'(y, x)=-E$, E — единичный оператор, поэтому G_y' и G_x' определены в $N\times\Sigma$ и непрерывны в точке (y_0, x_0); наконец, $G_x'(y_0, x_0)=F'(x_0)$ — обратимый оператор. Поэтому, применяя теорему о неявном отображении к отображению $G(y, x)$, мы можем утверждать, что в некоторой окрестности Σ_2 точки $y_0\in N$ существует функция $x=F^{-1}(y)$, которая непрерывна в этой окрестности Σ_2; справедливо соотношение $x_0=F^{-1}(y_0)$; функция $x(y)=F^{-1}(y)$ дифференцируема в точке y_0, причем

$$x'(y_0)=-[G_x'(y_0, x_0)]^{-1}G_y'(y_0, x_0)=[F'(x_0)]^{-1}E=[F'(x_0)]^{-1}.$$

Отображение $x=F^{-1}(y)$ осуществляет взаимно однозначное отображение окрестности Σ_2 на Σ_1 и является обратным по отношению к отображению $y=F(x)$. Если предположить, что отображение $y=F(x)$ непрерывно дифференцируемо в некоторой окрестности точки x_0, то отображение $x=F^{-1}(y)$ будет диффеоморфизмом, т. е. будет взаимно однозначным и дифференцируемым отображением Σ_2 на Σ_1.

Геометрически это утверждение можно истолковать так: множество $G(y, x)=F(x)-y=0$, заданное как график отображения $y=F(x)$, в некоторой окрестности (y_0, x_0) точки из пространства $N\times B$ можно изобразить как график отображения $x=F^{-1}(y)$, заданного в некоторой окрестности точки $y_0\in N$.

В случае, если пространства B и N совпадают с пространством E^m, т. е. $B=N=E^m$, теорема об обратном отображении утверждает, что *если задано отображение $y=F(x)$, или в координатной форме*

$$y_1=F_1(x_1, \ldots, x_m),$$
$$\cdots\cdots\cdots$$
$$y_m=F_m(x_1, \ldots, x_m),$$

причем функции F_i, $i=1, \ldots, m$ дифференцируемы в окрестности точки $x_0=(\mathring{x}_1, \ldots, \mathring{x}_m)$ и их частные производные непрерывны в $y_0=F(x_0)$; якобиан $\dfrac{D(F_1, \ldots, F_m)}{D(x_1, \ldots, x_m)}$ в точке x_0 отличен от нуля, то существуют окрестности $\Sigma_1\subset E^m$, $\Sigma_2\subset E^m$ точек $x_0\in\Sigma_1$ и $y_0\in\Sigma_2$ соответственно, которые взаимно однозначно соответствуют друг другу при помощи отображения $y=F(x)$: $\Sigma_1\to\Sigma_2$, причем существует гомеоморфное отображение $x=F^{-1}(y)$: $\Sigma_2\to\Sigma_1$ окрестности Σ_2 на окрестность Σ_1, обратное по отношению к F, дифференцируемое в точке $y_0=(\mathring{y}_1, \ldots, \mathring{y}_m)$ и*

* Взаимно однозначное непрерывное отображение, обратное к которому также непрерывно, называется гомеоморфным отображением или гомеоморфизмом.