

## Глава 12. ТЕОРЕМА О НЕЯВНО ЗАДАННОЙ ФУНКЦИИ И ЕЕ ПРИЛОЖЕНИЯ

### 12.1. неявно заданные функции

Начнем с примера. Пусть задано *одно* уравнение с *двумя вещественными* переменными

$$x^2 + y^2 - 1 = 0.$$

Будем задавать произвольные значения  $x$  и решать получившееся уравнение относительно  $y$  (напоминаем, что нас интересуют только *вещественные* решения!). Возможны три ситуации:

1. При данном  $x$  уравнение *не имеет* решений. Например, при  $x = 2$  получаем уравнение  $y^2 = -3$ .
2. При данном  $x$  уравнение имеет *несколько* решений. Например, при  $x = 0$  получаем уравнение  $y^2 = 1$ , т.е.  $y_1 = 1, y_2 = -1$ .
3. При данном  $x$  уравнение имеет *ровно одно* решение. Например, при  $x = 1$  получаем уравнение  $y^2 = 0$ , т.е.  $y = 0$ .

**Определение.** Пусть на открытом прямоугольнике  $]a, b[ \times ]c, d[$  задана *вещественная* функция  $F$ . Если для всякого  $x \in ]a, b[$  уравнение  $F(x, y) = 0$  имеет *ровно одно* решение  $y \in ]c, d[$ , то говорят, что это уравнение  *неявно задает* на интервале  $]a, b[$  функцию со значениями из интервала  $]c, d[$ . Если обозначить эту функцию  $f$ , то при всех  $x \in ]a, b[$   $F(x, f(x)) = 0$ .

**Замечание.** Отличительной особенностью неявного задания функции является наличие в алгоритме вычисления ее значения операции "решить уравнение" (естественно, без указания способа фактического выполнения этой операции).

Начнем с *линейного* уравнения

$$a \cdot (x - x_0) + b \cdot (y - y_0) = 0, \quad (12.1.1)$$

где  $a, b, x_0, y_0$  – заданные числа и  $b \neq 0$ .

Очевидно, что это уравнение при любом  $x$  однозначно разрешимо относительно  $y$ , т.е. задает  *неявно* функцию  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Нетрудно получить явное выражение для этой функции:

$$y = f(x) = y_0 - b^{-1}a \cdot (x - x_0); \quad (\text{заметьте, что } f(x_0) = y_0).$$

Теперь перейдем к уравнению общего вида  $F(x, y) = 0$  и предположим, что пара чисел  $(x_0, y_0)$  есть решение этого уравнения, т.е.  $F(x_0, y_0) = 0$ .

Используя для функции  $F$  формулу Тейлора<sup>1</sup> (11.7.5) и полагая

$$a = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}, \quad h = \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}, \quad a + h = \begin{bmatrix} x \\ y \end{bmatrix},$$

перепишем наше уравнение в виде

$$F(x, y) = F(x_0, y_0) + F'(x_0, y_0) \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \alpha(x, y) = 0$$

или

$$D_x F(x_0, y_0) \cdot (x - x_0) + D_y F(x_0, y_0) \cdot (y - y_0) + \alpha(x, y) = 0, \quad (12.1.2)$$

где  $\alpha(x, y) = \frac{1}{2} \cdot h^T \cdot F''(a + \xi e) \cdot h$  – остаточный член формулы Тейлора.

Мы видим, что уравнение (12.1.2) отличается от уравнения (12.1.1) только остаточным членом формулы Тейлора. Поскольку при  $F'(x_0, y_0) \neq \theta$  и  $(x, y)$  близких к  $(x_0, y_0)$ , остаточный член  $\alpha(x, y)$  мал по сравнению с первыми слагаемыми, можно ожидать, что при  $D_y F(x_0, y_0) \neq 0$  уравнение (12.1.2), как и (12.1.1) будет однозначно разрешимо относительно  $y$  при любом  $x$  хотя бы в малой окрестности точки  $(x_0, y_0)$ .

Действительно, можно показать, что справедлива

**Теорема.** Пусть на открытом прямоугольнике  $J = ]a, b[ \times ]c, d[$  задана непрерывно дифференцируемая функция  $F : J \rightarrow \mathbb{R}$ . Пусть  $(x_0, y_0) \in J$ , причем  $F(x_0, y_0) = 0$ , а  $D_y F(x_0, y_0) \neq 0$ . Тогда существуют такие  $J_x \subset ]a, b[$  (окрестность точки  $x_0$ ) и  $J_y \subset ]c, d[$  (окрестность точки  $y_0$ ), что

- 1) при любом  $x \in J_x$  уравнение  $F(x, y) = 0$  имеет единственное решение  $y \in J_y$  (обратите внимание, не вообще единственное решение, а единственное решение, *лежащее в  $J_y$* !). Таким образом, уравнение  $F(x, y) = 0$  неявно задает функцию  $f : J_x \rightarrow J_y$ ;

<sup>1</sup>В соответствии с *Важным соглашением* мы считаем, что функция имеет все нужные нам производные.

- 2)  $F(x, f(x)) = 0$  при всех  $x \in J_x$ ;
- 3)  $f(x_0) = y_0$ ;
- 4)  $f$  непрерывно дифференцируема на  $J_x$ .

Геометрический смысл этой теоремы очень прост: в некоторой окрестности точки  $(x_0, y_0)$  множество решений уравнения  $F(x, y) = 0$  есть график гладкой функции  $f$ .

Примеры. 1. (рис. 12.1a) Пусть  $F(x, y) = x^2 + y^2 - 1$ , и  $(x_0, y_0)$  – некоторая точка, лежащая на верхней дуге единичной окружности, т.е.  $F(x_0, y_0) = 0$ ,  $y_0 > 0$ . Тогда  $D_y F(x_0, y_0) = 2y_0 \neq 0$ . Из рисунка видно, что в некоторой окрестности точки  $(x_0, y_0)$  множество решений уравнения  $x^2 + y^2 - 1 = 0$  представляет собой график функции  $y = f(x)$ , т.е. для каждого  $x$ , достаточно близкого к  $x_0$ , существует ровно один  $y$ , близкий к  $y_0$ , такой, что  $x^2 + y^2 - 1 = 0$ . Фрагмент графика функции  $y = f(x)$  изображен на рис.12.1.b. Видно, что если не требовать близости  $y$  к  $y_0$ , то нашлось бы не одно, а два решения уравнения:  $y$  и  $-y$ .

К сожалению, теорема не дает сведений об окрестностях  $J_x$  и  $J_y$ . Очевидно, что на рисунке они взяты не наибольшими из возможных.

Заметим, что в нашем примере нетрудно найти *явное* выражение функции  $f$ :  $f(x) = +\sqrt{1-x^2}$  (здесь учтено, что  $y_0 > 0$ ).

В точке же  $(-1, 0)$   $D_y F = 0$ . Из рисунка видно, что в любой окрестности этой точки множество решений уравнения не является графиком функции.

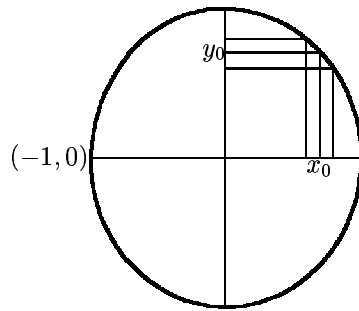


Рис. 12.1a

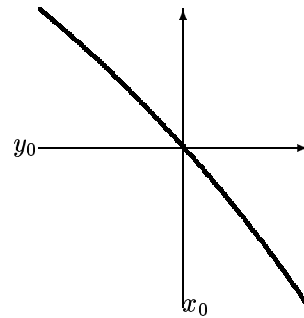


Рис. 12.1b

2.  $F(x, y) = x^7 - xy + y^6 - 1$ ,  $x_0 = 0$ ,  $y_0 = 1$ .  $F(0, 1) = 0$ ,  $DF_y(0, 1) \neq 0$ . В силу сформулированной теоремы уравнение  $x^7 - xy + y^6 - 1 = 0$  определяет в некоторой окрестности точки  $(0, 1)$  функцию  $y = f(x)$ , однако нетрудно убедиться, что ее явное задание получить не удастся. Фрагмент графика этой функции изображен на рис 12.2.

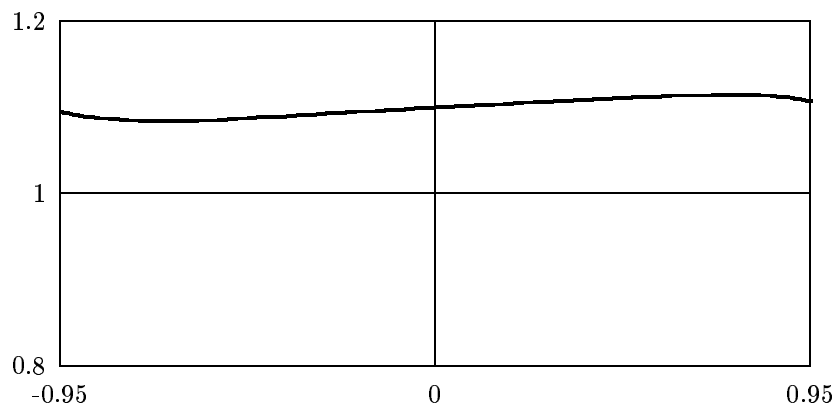


Рис. 12.2

Замечание. В случае  $D_y F(x_0, y_0) = 0$  теорема не дает никаких заключений о структуре множества решений уравнения в окрестности точки  $(x_0, y_0)$ . Приведем несколько простых примеров, в которых  $x_0 = y_0 = 0$  и  $D_y F(0, 0) = 0$ .

1.  $x - y^2 = 0$  (рис 12.3). Здесь при  $x > 0$  уравнение имеет два решения, при  $x = 0$  — одно, а при  $x < 0$  — ни одного.

2.  $x - y^3 = 0$  (рис 12.4). Здесь при всех  $x \in \mathbb{R}$  уравнение имеет одно решение.

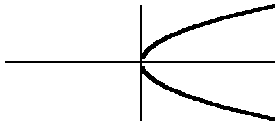


Рис. 12.3

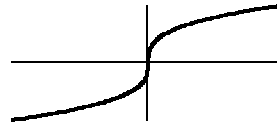


Рис. 12.4

3.  $x^2 - y^2 = 0$  (рис 12.5). Здесь при всех  $x \neq 0$  уравнение имеет два решения.

4.  $x^2 + y^2 = 0$  (рис 12.6). Здесь при любом  $x \neq 0$  уравнение не имеет ни одного решения.

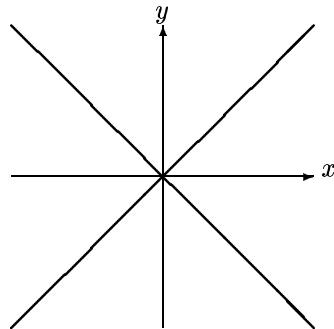


Рис.12.5

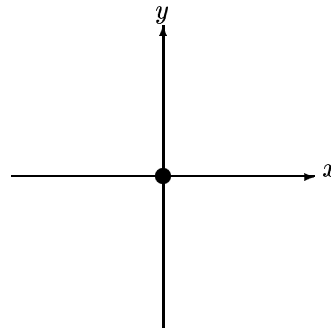


Рис.12.6

Обратите внимание на то, что множества решений уравнений в примерах 1, 3, 4 не являются графиками функций ни в каком открытом прямоугольнике, содержащем начало координат, в то время как в примере 2 уравнение задает неявно функцию без каких-либо ограничений на переменные (нетрудно найти ее явное выражение:  $f(x) = \text{sign}(x) \cdot |x|^{1/3}$ ).

Пусть теперь условия теоремы выполнены и уравнение  $F(x, y) = 0$  задает неявно в некоторой окрестности точки  $x_0$  функцию  $y = f(x)$  со значениями из окрестности точки  $y_0$ . Получим формулу для вычисления производной неявно заданной функции  $f$  (существование этой производной гарантировано теоремой).

Подставив в уравнение  $F(x, y) = 0$  его единственное решение  $y = f(x)$ , получим тождество

$$(F \circ \varphi)(u) = F(u, f(u)) \equiv 0, \quad (12.1.3)$$

где отображение  $\varphi$  (см. рис. 12.7) определено на  $J_x$ .

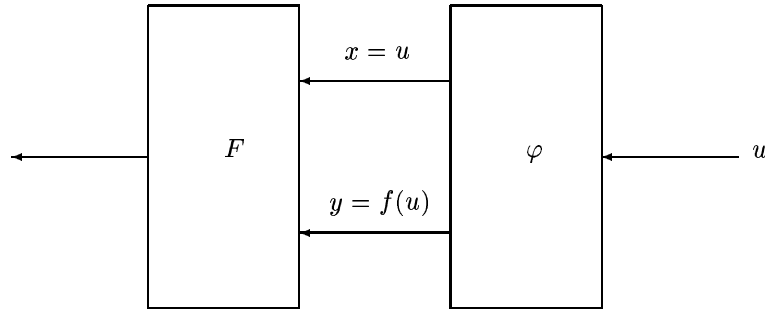


Рис. 12.7

Дифференцируем *тождество* (12.1.3)

$$(F' \circ \varphi)(u) \cdot \varphi'(u) \equiv 0.$$

Подставив сюда

$$F'(x, y) = [D_x F(x, y), D_y F(x, y)], \quad \varphi'(u) = \begin{bmatrix} 1 \\ f'(u) \end{bmatrix},$$

получим

$$D_x F(u, f(u)) + D_y F(u, f(u)) \cdot f'(u) \equiv 0, \quad \text{откуда}$$

$$f'(u) = -\left(D_y F(u, f(u))\right)^{-1} \cdot D_x F(u, f(u)) = -\frac{D_x F(u, f(u))}{D_y F(u, f(u))}. \quad (12.1.4)$$

**Замечания.** 1. В формуле (12.1.4) деление на  $D_y F(u, f(u))$  возможно, так как по условию теоремы  $D_y F$  отлична от нуля в точке  $(x_0, y_0)$ , а следовательно, по непрерывности, и в некоторой ее окрестности (вне которой эта формула, естественно, не работает).

2. Для фактического вычисления значения производной неявно заданной функции  $f$  по формуле (12.1.4), необходимо предварительно найти значение этой функции, т.е. решить (например, численно) уравнение  $F(x, y) = 0$ .

Вернемся к нашим примерам. Если  $F(x, y) = x^2 + y^2 - 1$ ,  $x > 0$ ,  $y > 0$ , то  $D_x F(x, y) = 2x$ ,  $D_y F(x, y) = 2y$ . Поэтому  $f'(x) = -\frac{2x}{2y} = -\frac{x}{f(x)}$ . Здесь можно найти явное выражение для функции:  $f(x) = +\sqrt{1-x^2}$ . Поэтому  $f'(x) = -\frac{x}{\sqrt{1-x^2}}$ , что легко проверить непосредственно.

Если же  $F(x, y) = x^7 - xy + y^6 - 1$ , то  $D_x F(x, y) = 7x^6 - y$ ,  $D_y F(x, y) = -x + 6y^5$ ,  $f'(x) = -\frac{7x^6 - y}{6y^5 - x}$ . Здесь получить явное выражение функции невозможно, однако при  $x = 0$  имеем  $y = 1$  и

$$f'(0) = -\frac{7x^6 - y}{6y^5 - x} \Big|_{x=0, y=1} = \frac{1}{6}.$$

Обратимся теперь к неявному заданию *векторного поля*. Пусть задана система из  $n$  уравнений с  $n + m$  переменными:

$$F(x, y) = \theta_n. \quad (12.1.5)$$

где  $x \in \mathbb{R}^m$ ,  $y \in \mathbb{R}^n$ ,  $F(x, y) = [F_1(x, y), \dots, F_n(x, y)]^T$  — функция из  $\mathbb{R}^m \times \mathbb{R}^n$  в  $\mathbb{R}^n$ .

Если существуют такие множества  $U \subset \mathbb{R}^m$  и  $V \subset \mathbb{R}^n$ , что для всякого  $x \in U$  система (12.1.5) имеет *ровно одно* решение  $y \in V$ , то говорят, что эта система задает *неявно* функцию  $f : U \rightarrow V$ . Очевидно, что при всех  $x \in U$   $F(x, f(x)) = \theta_n$ .

Вновь начнем с рассмотрения *линейной* системы

$$A \cdot (x - x^{(0)}) + B \cdot (y - y^{(0)}) = \theta_n, \quad (12.1.6)$$

где  $A$  и  $B$  — матрицы размеров  $n \times m$  и  $n \times n$  соответственно.

Очевидно, что при  $\det(B) \neq 0$  система (12.1.6) однозначно разрешима относительно  $y$  и задает неявно функцию  $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ . Нетрудно получить явное задание этой функции:

$$y = f(x) = y^{(0)} - B^{-1}A \cdot (x - x^{(0)}); \quad f(x^{(0)}) = y^{(0)}.$$

Пусть теперь дана система общего вида (12.1.5) и известно ее решение – упорядоченная пара векторов  $(x^{(0)}, y^{(0)})$ :  $F(x^{(0)}, y^{(0)}) = \theta_n$ .

Используя формулу Тейлора (11.7.7), запишем систему (12.1.5) в виде

$$F'(x^{(0)}, y^{(0)}) \cdot \begin{bmatrix} x - x^{(0)} \\ y - y^{(0)} \end{bmatrix} + \alpha(x, y) = \theta_n, \quad (12.1.7)$$

где  $F'(x^{(0)}, y^{(0)}) = [D_x F(x^{(0)}, y^{(0)}), D_y F(x^{(0)}, y^{(0)})]$  – матрица Якоби размера  $n \times (m + n)$  (ее блоки – матрицы  $D_x F$  и  $D_y F$ , составленные из частных производных компонент  $F$  по координатам векторов  $x$  и  $y$  соответственно, – имеют размеры  $n \times m$  и  $n \times n$ );

$\begin{bmatrix} x - x^{(0)} \\ y - y^{(0)} \end{bmatrix}$  – столбец высоты  $m + n$ ;

$\alpha(x, y)$  – остаточный член формулы Тейлора – столбец высоты  $n$ .

Расписав (12.1.7) подробнее, получим уравнение

$$D_x F(x^{(0)}, y^{(0)}) \cdot (x - x^{(0)}) + D_y F(x^{(0)}, y^{(0)}) \cdot (y - y^{(0)}) + \alpha(x, y) = \theta_n,$$

отличающееся от (12.1.6) только остаточным членом  $\alpha(x, y)$ . Поэтому можно ожидать, что при

$$\det(B) = \det(D_y F(x^{(0)}, y^{(0)})) \neq 0$$

система при заданном значении  $x$  будет однозначно разрешима относительно  $y$  в некоторой окрестности точки  $(x^{(0)}, y^{(0)})$ .

Сформулируем соответствующую теорему.

**Теорема.** Пусть

$J_m = ]a_1, b_1[ \times ]a_2, b_2[ \times \dots \times ]a_m, b_m[$  – параллелепипед в  $\mathbb{R}^m$ ;

$J_n = ]c_1, d_1[ \times ]c_2, d_2[ \times \dots \times ]c_n, d_n[$  – параллелепипед в  $\mathbb{R}^n$ ;

$J_{m+n} = J_m \times J_n$  – параллелепипед в  $\mathbb{R}^{m+n}$ ;

$F: J_{m+n} \rightarrow \mathbb{R}^n$  – непрерывно дифференцируемая вектор-функция;

$(x^{(0)}, y^{(0)}) \in J_{m+n}$ ,  $F(x^{(0)}, y^{(0)}) = \theta_n$ ,  $\det(D_y F(x^{(0)}, y^{(0)})) \neq 0$ .

Тогда существуют такие  $U \in J_m$  – окрестность точки  $x^{(0)}$  и  $V \in J_n$  – окрестность точки  $y^{(0)}$ , что

1) для каждого  $x \in U$  существует единственный  $y \in V$  – решение уравнения  $F(x, y) = \theta_n$ , т.е. это уравнение неявно задает функцию  $f: U \rightarrow V$ ;

2)  $F(x, f(x)) = \theta_n$ ;

3)  $f(x^{(0)}) = y^{(0)}$ ;

4)  $f$  непрерывно дифференцируема на  $U$ .

Для вычисления матрицы Якоби неявно заданной функции  $f$  рассмотрим вновь рис. 12.7 (отображение  $\varphi$  при этом определено на  $U$ ). Дифференцируя тождество  $(F \circ \varphi)(u) = F(u, f(u)) \equiv \theta_n$ , выполняющееся на  $U$ , получим

$$(F \circ \varphi)'(u) = (F' \circ \varphi)(u) \cdot \varphi'(u) = \Theta_{n \times m}$$

( $\Theta_{n \times m}$  – нуль-матрица размера  $n \times m$ ).

Подставляя сюда

$$\varphi'(u) = \begin{bmatrix} u \\ f(u) \end{bmatrix}' = \begin{bmatrix} I_m \\ f'(u) \end{bmatrix}$$

( $I_m$  – единичная матрица порядка  $m$ ), получим

$$D_x F(u, f(u)) + D_y F(u, f(u)) \cdot f'(u) = \Theta_{n \times m}$$

(проверьте согласованность размеров матриц!), откуда

$$f'(u) = -\left(D_y F(u, f(u))\right)^{-1} \cdot D_x F(u, f(u)). \quad (12.1.8)$$

Пример. Пусть

$$m = n, \quad x, y \in \mathbb{R}^n, \quad g: \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad F(x, y) = x - g(y).$$

Тогда уравнением  $x - g(y) = \theta_n$  задается неявно функция  $g^{-1}$ , обратная функции  $g$ .

Поскольку  $D_y F = -g'$ , теорему о неявно заданной функции можно переформулировать в теорему об обратной функции так:

Пусть

$$J_n = ]c_1, d_1[ \times ]c_2, d_2[ \times \dots \times ]c_n, d_n[ \text{ — параллелепипед в } \mathbb{R}^n;$$

$$g: J_n \rightarrow \mathbb{R}^n, \quad y^{(0)} \in J_n, \quad \det(g'(y^{(0)})) \neq 0.$$

Тогда в некоторой окрестности точки  $y^{(0)}$  функция  $g$  обратима, обратная ей функция  $g^{-1}$  непрерывно дифференцируема и

$$(g^{-1})' = [g' \circ g^{-1}]^{-1}. \quad (12.1.9)$$

В случае  $n = 1$  эта формула дает известный результат:  $(g^{-1})' = \frac{1}{g' \circ g^{-1}}$ .

Замечания. 1. Поскольку  $\det(D_y F)$  не равен нулю в точке  $(x_0, y_0)$ , то по непрерывности он отличен от нуля и в некоторой ее окрестности, в которой, следовательно, обратима матрица  $D_y F$  и применима формула (12.1.9). Остается также в силе замечание 2 на стр.77.

2. В реальной задаче переменные, конечно, не поделены изначально на независимые ("иксы") и зависимые ("игреки"). Постановщик задачи сам выбирает те переменные относительно которых он хочет разрешить систему, и проверяет выполнение условий теоремы.

3. В формуле (12.1.9) имеет место "коллизия символов":  $g^{-1}$  обозначает функцию, обратную функции  $g$ , а  $[g' \circ g^{-1}]^{-1}$  — матрицу, обратную матрице  $[g' \circ g^{-1}]$ .

## 12.2 Задание гладкой поверхности в $\mathbb{R}^3$ уравнением

Рассмотрим уравнение

$$F(x) = c, \quad (12.2.1)$$

где  $F: \mathbb{R}^3 \rightarrow \mathbb{R}$  — непрерывно дифференцируемый функционал, а  $c$  — число. Напомним, что множество решений этого уравнения называется *множеством уровня* функционала  $F$ .

Предположим, что точка  $a$  удовлетворяет уравнению (12.2.1) и  $\nabla F(a) \neq \theta$ . Тогда, не умаляя общности, можно считать, например, что  $D_3 F(a) \neq 0$ . Применяя к уравнению теорему о неявно заданной функции (здесь  $m = 2, n = 1$ ), получаем, что в некоторой окрестности точки  $a$  уравнение можно *однозначно* разрешить относительно  $x_3$ , т.е. множество уровня функционала в этой окрестности представляет собой график функции  $x_3 = f(x_1, x_2)$  и

$$F(x_1, x_2, f(x_1, x_2)) - c \equiv 0, \quad f(a_1, a_2) = a_3.$$

Функция  $f$  непрерывно дифференцируема, и, следовательно (см. п.11.3), ее график есть гладкая поверхность. Проведем в точке  $a$  касательную плоскость к этой поверхности.

Из формулы (12.1.8) получаем

$$f'(a_1, a_2) = -D_3^{-1} F(a_1, a_2, a_3) \cdot [D_1 F(a_1, a_2, a_3), D_2 F(a_1, a_2, a_3)].$$

Отсюда

$$D_1 f(a_1, a_2) = -\frac{D_1 F(a)}{D_3 F(a)}, \quad D_2 f(a_1, a_2) = -\frac{D_2 F(a)}{D_3 F(a)}.$$

Подставляя результат в формулу (11.3.2), получим

$$x_3 - a_3 = \frac{D_1 F(a)}{D_3 F(a)} \cdot (x_1 - a_1) + \frac{D_2 F(a)}{D_3 F(a)} \cdot (x_2 - a_2)$$

или, после умножения на  $D_3 F(a) \neq 0$ ,

$$D_1 F(a) \cdot (x_1 - a_1) + D_2 F(a) \cdot (x_2 - a_2) + D_3 F(a) \cdot (x_3 - a_3) = 0$$

$$\text{или } F'(a) \cdot (x - a) = 0, \quad \text{или, наконец, } \langle \nabla F(a), (x - a) \rangle = 0. \quad (12.2.2)$$

В силу симметрии этого уравнения относительно координат оно будет выполняться, если хотя бы одна из трех частных производных отлична от нуля, т.е. когда отличен от нуля градиент функционала  $F$ .

Итак, если  $F(a) = c$  и  $\nabla F(a) \neq \theta$ , то в некоторой окрестности точки  $a$  множество уровня функционала есть гладкая поверхность. Если же градиент отличен от нуля во всех точках множества уровня функционала, то естественно сказать, что все это множество есть гладкая поверхность (ее называют *поверхностью уровня функционала*  $F$ .)

Примеры. 1.  $F(x_1, x_2, x_3) \equiv x_1^2 + x_2^2 - x_3^2 = -1$ .

$\nabla F(x) = 2[x_1, x_2, -x_3]^T$ . Градиент, очевидно, не обращается в нуль ни в одной точке *множества уровня*. Как известно из курса линейной алгебры, это множество – двухполостный гиперболоид.

2.  $F(x_1, x_2, x_3) \equiv x_1^2 + x_2^2 - x_3^2 = 0$ . Это – другое множество уровня того же функционала.

$\nabla F(x) = 2[x_1, x_2, -x_3]^T$ . Градиент обращается в нуль в одной точке *множества уровня* (в начале координат). Теорема о неявно заданной функции не работает в окрестности этой точки. Как известно из курса линейной алгебры, рассматриваемое множество уровня – конус, "негладкий" в вершине, но гладкий во всех остальных своих точках.

Единичный направленный отрезок, перпендикулярный касательной плоскости, называется *нормалью* к поверхности в точке касания. Из (12.2.2) видно, что  $\nabla F(a)$  перпендикулярен касательной плоскости, т.е. коллинеарен нормали. Говорят, что *градиент функционала ортогонален поверхности уровня этого функционала*.

### 12.3. Условный экстремум

В приложениях часто возникают задачи о нахождении экстремума функционала при некоторых дополнительных условиях. Одну из таких задач мы рассмотрим в этом пункте.

Пусть  $F, G : \mathbb{R}^3 \rightarrow \mathbb{R}$  – непрерывно дифференцируемые ффункционалы. Требуется найти локальные экстремумы функционала  $F$  на гладкой поверхности уровня функционала  $G$ , заданной уравнением  $G(x) = c$ .

Пусть точка  $a$  лежит на этой поверхности уровня. Вследствие гладкости поверхности  $\nabla G(a) \neq \theta$ , и мы можем разрешить уравнение  $G(x) = c$  в окрестности точки  $a$  относительно хотя бы одной из координат. Пусть, например,  $D_3 G(a) \neq 0$ . Тогда поверхность  $G(x) = c$  в окрестности точки  $a$  представляет собой график функции  $x_3 = f(x_1, x_2)$ . Поэтому наличие в точке  $a$  локального экстремума у функционала  $F$  на поверхности уровня функционала  $G$  равносильно наличию локального экстремума у функционала

$$\phi(x_1, x_2) = F(x_1, x_2, f(x_1, x_2))$$

в точке  $(a_1, a_2) \in \mathbb{R}^2$ .

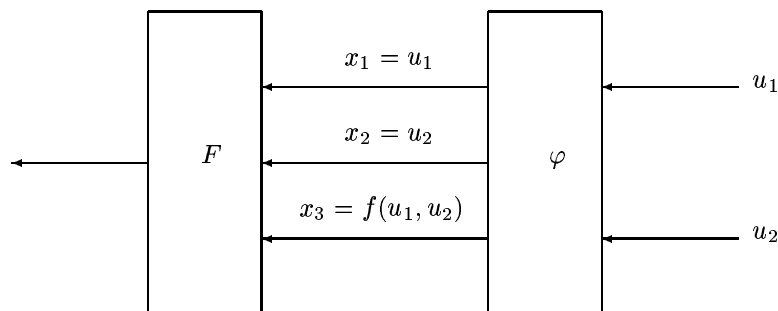


Рис. 12.8

Как известно, точка локального экстремума гладкого функционала  $\phi$  обязана быть его стационарной точкой. Таким образом  $\nabla \phi(a_1, a_2) = \theta$ .

Из рис.12.8 видно, что  $\phi' = (F \circ \varphi)' = (F' \circ \varphi) \cdot \varphi'$ .

Поскольку  $(F' \circ \varphi)(u_1, u_2) = F'(u_1, u_2, f(u_1, u_2))$

и

$$\varphi'(u) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ D_1 f(u) & D_2 f(u) \end{bmatrix},$$

то

$$\nabla \phi(a_1, a_2) = (\phi'(a_1, a_2))^T =$$

$$= \begin{bmatrix} D_1 F(a) \cdot 1 + D_2 F(a) \cdot 0 + D_3 F(a) \cdot D_1 f(a_1, a_2) \\ D_1 F(a) \cdot 0 + D_2 F(a) \cdot 1 + D_3 F(a) \cdot D_2 f(a_1, a_2) \end{bmatrix}.$$

Отсюда

$$D_1 F(a) + D_3 F(a) \cdot D_1 f(a_1, a_2) = D_2 F(a) + D_3 F(a) \cdot D_2 f(a_1, a_2) = 0$$

Вспоминая формулы (12.2.2), получаем

$$D_1 F(a) - D_3 F(a) \cdot \frac{D_1 G(a)}{D_3 G(a)} = D_2 F(a) - D_3 F(a) \cdot \frac{D_2 G(a)}{D_3 G(a)} = 0$$

или

$$[D_1 F(a), D_2 F(a), D_3 F(a)] = \frac{D_3 F(a)}{D_3 G(a)} \cdot [D_1 G(a), D_2 G(a), D_3 G(a)],$$

или

$$\nabla F(a) = \lambda \cdot \nabla G(a), \quad (12.3.1)$$

где  $\lambda = \frac{D_3 F(a)}{D_3 G(a)}$  — число.

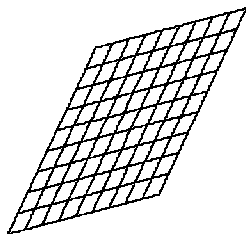
Если предположить, что отлична от нуля другая компонента вектора  $\nabla G(a)$ , то после аналогичных преобразований придем опять к соотношению (12.3.1).

Таким образом, если векторы  $\nabla F(a)$  и  $\nabla G(a)$  *линейно независимы*, то в точке  $a$  функционал  $F$  *не может иметь локального экстремума на поверхности уровня функционала  $G$* . Точки поверхности  $G(x) = c$ , в которых выполнено условие (12.3.1) называются *стационарными точками* функционала  $F$  на этой поверхности.

Можно показать, что аналогичный результат имеет место в пространстве любой размерности: если  $F, G : \mathbb{R}^n \rightarrow \mathbb{R}$  — непрерывно дифференцируемые функционалы, а уравнение  $G(x) = c$  задает гладкую „поверхность“, то в точках этой поверхности, где градиенты  $\nabla F$  и  $\nabla G$  линейно независимы, функционал  $F$  *не может иметь локального экстремума на поверхности уровня функционала  $G$* .

Это утверждение называется *теоремой Эйлера*.

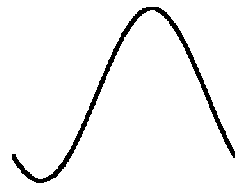
По историческим причинам рассмотренную задачу обычно называют задачей об экстремуме функционала  $F$  при дополнительном условии  $G(x) = c$  или задачей об *условном экстремуме*. Важно понимать, что это просто задача об экстремуме нового функционала (изменяется *область определения* функционала  $F$ )!



(a)



(b)



(c)

Рис.12.9

Пример (см.рис.12.9). Известно, что линейный функционал

$$F : \mathbb{R}^2 \rightarrow \mathbb{R} \quad F(x_1, x_2) = x_1 + x_2$$

(фрагмент его графика — наклонной плоскости — изображен на рис.12.9а) не имеет локальных экстремумов. Однако если рассматривать поведение этого функционала не на всей координатной плоскости, а на окружности  $G(x) = x_1^2 + x_2^2 = 1$ , то экстремумы появляются.

Действительно, параметрические уравнения нашей окружности

$$x_1 = \cos(t), \quad x_2 = \sin(t); \quad t \in [-\pi, \pi],$$

график сужения  $F$  на нее (пространственная кривая) изображен на рис.12.9б), а его развертка — на рис.12.9с.



Теорема Эйлера имеет в  $\mathbb{R}^2$  и  $\mathbb{R}^3$  простую геометрическую интерпретацию. Ранее (п. 11.7) было показано, что функционал  $F$  не может иметь локального экстремума в точке, где его градиент отличен от нуля, так как  $F$  возрастает в направлении градиента и убывает в направлении антиградиента. В случае условного экстремума ситуация осложняется тем, что движение из точки, лежащей на поверхности уровня функционала  $G$  может происходить не в произвольном направлении, а только по этой поверхности. Рассмотрим ситуации, схематически изображенные на рис.12.10.

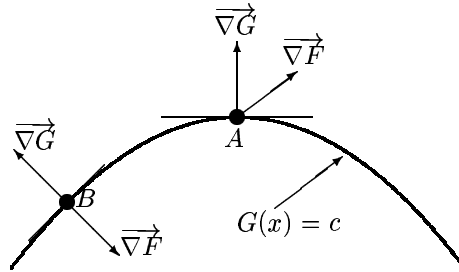


Рис.12.10

Известно, что в точке, лежащей на поверхности уровня функционала  $G$ , направленный отрезок  $\overrightarrow{\nabla G}$  перпендикулярен этой поверхности. Если направленный отрезок  $\overrightarrow{\nabla F}$  не коллинеарен  $\overrightarrow{\nabla G}$  (точка  $A$  на рис 12.10), т.е. не перпендикулярен поверхности уровня функционала  $G$ , то его проекция на касательную плоскость (производная по направлению) отлична от нуля. Следовательно, в окрестности точки  $A$  функционал принимает *на поверхности уровня* значения и большие и меньшие, чем  $F(A)$ , т.е. не имеет в этой точке экстремума. В точке  $B$  направленные отрезки  $\overrightarrow{\nabla F}$  и  $\overrightarrow{\nabla G}$  коллинеарны, проекция  $\overrightarrow{\nabla F}$  на касательную плоскость к поверхности уровня равна нулю – в этой точке условный экстремум *может быть*.

Соотношение (12.3.1) может быть переписано так:

$$\nabla(F - \lambda G) = \theta,$$

т.е. стационарные точки функционала  $F$  на поверхности уровня функционала  $G$  – это в точности стационарные точки функционала  $F - \lambda G$  (с дополнительной переменной  $\lambda$ , которая называется *множителем Лагранжа*).

**Замечание.** Для поиска этих стационарных точек следует решить систему из  $n + 1$  уравнения

$$\begin{cases} \nabla(F - \lambda G)(x) &= \theta_n \\ G(x) &= c \end{cases} \quad (12.3.2)$$

с  $n + 1$  переменными:  $x_1, \dots, x_n; \lambda$ .

**Примеры.** 1. Найти расстояние от точки  $x^{(0)} \in \mathbb{R}^3$  до плоскости  $\Pi : \langle x, a \rangle = c$ ,  $a \neq \theta_3$ , не проходящей через эту точку.

По определению,  $\text{dist}(x^{(0)}, \Pi) = \min_{x \in \Pi} \|x^{(0)} - x\|$ . Минимизируя, как обычно, квадрат нормы, получим задачу на условный экстремум: найти минимум функционала  $\|x^{(0)} - x\|^2$  при условии  $\langle x, a \rangle = c$ .

Составим уравнение (12.3.1)

$$\nabla \|x^{(0)} - x\|^2 = \lambda \cdot \nabla \langle x, a \rangle \quad \text{или} \quad 2 \cdot (x - x^{(0)}) = \lambda \cdot a. \quad (12.3.3)$$

Как известно,  $\vec{a} \perp \Pi$ . Отсюда и из (12.3.3) следует, что  $x - x^{(0)} \perp \Pi$ , и мы получили известный результат: из всех отрезков, соединяющих данную точку с точками плоскости, наименьшую длину имеет перпендикуляр. (На самом деле мы доказали только, что *никакой другой отрезок не может дать минимума*; то, что перпендикуляр дает минимум, требует еще проверки!)

2. Пусть  $A$  – вещественная симметричная матрица порядка  $n$ . Найти экстремумы квадратичной формы  $\langle Ax, x \rangle$  на сфере  $\|x\|^2 = r^2$ .

---

Поскольку (см. пример в п.11.6)  $\nabla \langle Ax, x \rangle = 2Ax$ ,  $\nabla \|x\|^2 = 2x$ , система (12.3.2) приобретает вид

$$\begin{cases} (A - \lambda I_n)x &= \theta_n \\ \|x\|^2 &= r^2 \end{cases}.$$

Таким образом, стационарные точки квадратичной формы *на сфере* – собственные векторы матрицы этой квадратичной формы, а множители Лагранжа – соответствующие собственные числа. Из курса линейной алгебры известно, что *глобальные* экстремумы квадратичной формы на сфере действительно достигаются на собственных векторах, соответствующих наибольшему и наименьшему собственным числам. Можно показать, что собственные векторы, соответствующие остальным собственным числам, не дают квадратичной форме даже локальных экстремумов.