

# Математический анализ, Коллоквиум 4

Балюк Игорь  
[@lodthe](#), [GitHub](#)

Материалы предоставил Егор Косов.

Дата изменения: 2020.05.24 в 08:14

## Содержание

1	Метрические и нормированные пространства.	2
2	Компакты в метрических пространствах	4
3	Дифференцируемость отображений	5
4	Градиент и достаточное условие дифференцируемости	6
5	Частные производные высокого порядка	7
6	Правила дифференцирования	8
7	Дифференциалы высоких порядков.	8
8	Дифференцирование сложной функции.	9
9	Дифференциал обратного отображения	10

Исходный код предоставил Егор Косов. В данном файле я попытался исправить опечатки и облегчить некоторые моменты для понимания.

[Оригинальный список вопросов](#)

## 1 Метрические и нормированные пространства.

[Оригинальный конспект.](#)

**Определение.** Пусть  $X$  — множество. Функция  $d : X \times X \rightarrow [0; +\infty)$  называется метрикой, если

1.  $d(x, y) = 0 \iff x = y$ ;
2.  $d(x, y) = d(y, x) \forall x, y \in X$ ;
3.  $\|\lambda x\| = |\lambda| \|x\|, \forall x \in X$ ;
4.  $\|x + y\| \leq \|x\| + \|y\| \forall x, y \in X$ .

Пара  $(X, \|\cdot\|)$  называется нормированным пространством.

Нормой является привычная нам длина вектора. Аналогично матрике, мы будем часто работать с Евклидовой нормой: пусть  $x \in \mathbb{R}^n$ , тогда  $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$ .

Всякое нормированное пространство является метрическим с метрикой  $d(x, y) = \|x - y\|$ .

**Определение.** Пусть  $X$  — линейное пространство. Функция  $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{R}$  называется скалярным произведением, если для всех  $x, y, z \in X$  и всех  $a, b \in \mathbb{R}$  выполнены следующие условия:

1.  $\langle x, x \rangle \geq 0$  и  $\langle x, x \rangle = 0 \iff x = 0$ ;
2.  $\langle x, y \rangle = \langle y, x \rangle$ ;
3.  $\langle ax + by, z \rangle = a\langle x, z \rangle + b\langle y, z \rangle$ .

Линейное пространство  $X$  со скалярным произведением называется Евклидовым.

Мы будем часто работать со следующим скалярным произведением: пусть  $x, y \in \mathbb{R}^n$ , тогда  $\langle x, y \rangle = x_1 \cdot y_1 + \dots + x_n \cdot y_n$ .

**Лемма.** (Неравенство Коши-Буняковского) Пусть  $\langle \cdot, \cdot \rangle$  скалярное произведение на линейном пространстве  $X$ , тогда  $\forall x, y \in X$

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \cdot \sqrt{\langle y, y \rangle}.$$

*Доказательство.* Заметим, что для  $\lambda \in \mathbb{R}$  выполнено

$$0 \leq \langle x + \lambda y, x + \lambda y \rangle = \lambda^2 \langle y, y \rangle + 2\lambda \langle x, y \rangle + \langle x, x \rangle.$$

Не ограничивая общности, считаем, что  $\langle y, y \rangle > 0$  (иначе  $y$  — нулевой вектор, доказательство тривиально). Это означает, что ветви параболы смотрят вверх. Но парабола не касается оси  $Ox$ , поэтому дискриминант этого трехчлена неположителен, т.е.  $4|\langle x, y \rangle|^2 - 4\langle y, y \rangle \langle x, x \rangle \leq 0$ . ■

**Следствие.** На евклидовом пространстве функция  $\|x\| := \sqrt{\langle x, x \rangle}$  является нормой.

*Доказательство.* Первые два свойства следуют из определения скалярного произведения. Неравенство треугольника следует из неравенства Коши-Буняковского:

$$\|x + y\|^2 = \langle x + y, x + y \rangle \leq \|x\|^2 + 2 \cdot |\langle x, y \rangle| + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

■

**Пример.** На линейном пространстве  $\mathbb{R}^k$  всех упорядоченных наборов  $(x_1, \dots, x_k)$  задано скалярное произведение  $\langle x, y \rangle := \sum_{j=1}^k x_j y_j$ . Тем самым, на  $\mathbb{R}^k$  задана естественная евклидова метрика  $\|x - y\| := \sqrt{|x_1 - y_1|^2 + \dots + |x_k - y_k|^2}$ .

**Определение.** Пусть  $(X, d)$  метрическое пространство.

1. Множество

$$B_r(x_0) := \{x \in X \mid d(x, x_0) < r\}$$

называется **открытым шаром** радиуса  $r$ .

2. Множество

$$\overline{B}_r(x_0) := \{x \in X \mid d(x, x_0) \leq r\}$$

называется **замкнутым шаром** радиуса  $r$ .

3. Последовательность точек  $x_n \in X$  называется **сходящейся к точке**  $x$ , если для всякого  $\varepsilon > 0$  найдется такой номер  $N(\varepsilon)$ , что  $d(x, x_n) < \varepsilon$  для каждого  $n \geq N(\varepsilon)$ .

4. Последовательность точек  $x_n \in X$  называется **фундаментальной**, если для всякого  $\varepsilon > 0$  найдется такой номер  $N(\varepsilon)$ , что  $d(x_k, x_n) < \varepsilon$  для всех  $k, n \geq N(\varepsilon)$ .

5. Точка  $x$  называется **предельной** для множества  $M \subset X$ , если для всякого  $\varepsilon > 0$  выполнено  $B_\varepsilon(x) \cap (M \setminus \{x\}) \neq \emptyset$ .

6. Множество  $U \subset X$  называется **открытым**, если для всякого  $x \in U$  найдется такое  $\varepsilon > 0$ , что  $B_\varepsilon(x) \subset U$ .

7. Множество  $F \subset X$  называется **замкнутым**, если множество  $X \setminus F$  открыто.

**Лемма.** Пусть  $(X, d)$  метрическое пространство. Тогда

1. если  $x_n \rightarrow x, y_n \rightarrow y$ , то  $d(x_n, y_n) \rightarrow d(x, y)$ ;
2. предел сходящейся последовательности единственный;
3. любой открытый шар является открытым множеством;
4. множество  $F$  замкнуто тогда и только тогда, когда множество  $F$  содержит все свои предельные точки.

*Доказательство.*

1. Следует из оценки

$$|d(x_n, y_n) - d(x, y)| \leq |d(x_n, y_n) - d(x_n, y)| + |d(x_n, y) - d(x, y)| \leq d(y_n, y) + d(x_n, x).$$

2. Следует из пункта 1).

3. Если  $x \in B_r(x_0)$ , то по неравенству треугольника  $B_\varepsilon(x) \subset B_r(x_0)$  при  $\varepsilon + d(x, x_0) < r$ .

4. Множество  $F$  замкнуто тогда и только тогда, когда  $\forall x \notin F \exists \varepsilon > 0 : B_\varepsilon \cap F = \emptyset \iff$  всякая точка  $x \notin F$  — не предельная для  $F$ .

■

**Определение.** Метрическое пространство называется **полным**, если каждая фундаментальная последовательность в нем сходится.

**Замечание.** На  $\mathbb{R}^k$  справедливы соотношения

$$\max_{1 \leq j \leq k} |x_j| \leq \|x\| \leq \sqrt{k} \cdot \max_{1 \leq j \leq k} |x_j|$$

для векторов  $x = (x_1, \dots, x_k)$ . Тем самым, последовательность  $x_n \rightarrow x$  в  $\mathbb{R}^k$  тогда и только тогда, когда  $(x_n)_j \rightarrow x_j$ .

**Пример.** Пространство  $\mathbb{R}^k$  со стандартной евклидовой метрикой полное. Действительно, если последовательность векторов  $x_n \in \mathbb{R}^k$  фундаментальна, то фундаментальны и последовательности координат  $\{(x_n)_j\}_{j=1}^\infty$  для всякого  $j \in \{1, \dots, k\}$ .

Тем самым, у  $j$ -ой координаты есть предел  $x_j$  для каждого  $j \in \{1, \dots, k\}$ . То есть  $|(x_n)_j - x_j| \rightarrow 0$ . Значит,  $x_n \rightarrow x := (x_1, \dots, x_k)$ .

**Пример.** Пусть  $X = [0; \pi/2)$ . Пространство  $X$  не является полным с метрикой  $d_1(x, y) = |x - y|$ , но является полным с метрикой  $d_2(x, y) = |\operatorname{tg} x - \operatorname{tg} y|$ .

**Определение.** Пусть  $(X, d_X)$  и  $(Y, d_Y)$  — два метрических пространства. Отображение  $f : X \rightarrow Y$  называется непрерывным в точке  $x_0 \in X$ , если для всякой последовательности  $x_n \rightarrow x_0$  выполнено  $f(x_n) \rightarrow f(x_0)$ .

**Лемма.** Пусть  $(X, d_X)$  и  $(Y, d_Y)$  — два метрических пространства.

1. Отображение  $f : X \rightarrow Y$  является непрерывным в точке  $x \in X$  тогда и только тогда, когда для всякого  $\varepsilon > 0$  найдется  $\delta > 0$  такое, что  $d_Y(f(x), f(x_0)) < \varepsilon$ , если  $d_X(x, x_0) < \delta$ .
2. Отображение  $f : X \rightarrow Y$  является непрерывным в каждой точке  $x \in X$  тогда и только тогда, когда прообраз каждого открытого множества в  $Y$  будет открытым множеством в  $X$  (такие отображения будем называть просто непрерывными).

*Доказательство.*

1. Отображение  $f$  разрывно в точке  $x_0 \iff$  найдется последовательность  $x_n \rightarrow x_0$ , для которой  $f(x_n)$  не сходится к  $f(x_0) \iff$  найдется число  $\varepsilon > 0$  и последовательность  $x'_n \rightarrow x_0$ , для которой  $d_Y(f(x'_n), f(x_0)) \geq \varepsilon \iff$  найдется такое число  $\varepsilon > 0$ , что для произвольного  $\delta > 0$  существует  $x_\delta \in B_\delta(x_0)$ , для которого  $d_Y(f(x_\delta), f(x_0)) \geq \varepsilon$ .
2. Если прообраз любого открытого множества открыт, то для произвольного  $\varepsilon > 0$  найдется такое  $\delta > 0$ , что  $f^{-1}(B_\varepsilon(f(x_0))) \supset B_\delta(x_0)$ , и значит отображение  $f$  непрерывно в точке  $x_0$ . Наоборот: пусть  $U$  — открыто в  $Y$  и  $x_0 \in f^{-1}(U)$ . Тогда в силу открытости найдется  $\varepsilon > 0$ , для которого  $B_\varepsilon(f(x_0)) \subset U$ . Из-за непрерывности в точке  $x_0$  найдется такое  $\delta > 0$ , что  $f^{-1}(B_\varepsilon(f(x_0))) \supset B_\delta(x_0)$ , что дает открытость множества  $f^{-1}(U)$ . ■

**Предложение.** Пусть  $f : X \rightarrow Y$  непрерывна в точке  $a \in X$ ,  $g : Y \rightarrow Z$  непрерывна в точке  $f(a) \in Y$ . Тогда композиция  $g \circ f : X \rightarrow Z$  непрерывна в точке  $a$ .

*Доказательство.* Следует из определения непрерывности. ■

**Следствие.** Пусть  $f, g : \mathbb{R}^k \rightarrow \mathbb{R}^m$  — непрерывные в точке  $a$  функции. Тогда  $f + g$  и  $f \cdot g$  — непрерывны в точке  $a$ .

*Доказательство.* Следует из того, что отображение  $(x_1, x_2) \rightarrow x_1 + x_2$  и  $(x_1, x_2) \rightarrow x_1 \cdot x_2$  непрерывны на  $\mathbb{R}^2$ . ■

**Определение.** Пусть  $(X, d_X)$  и  $(Y, d_Y)$  — метрические пространства и пусть  $x_0$  — предельная точка в  $X$ . Скажем, что предел функции  $f : X \rightarrow Y$  в точке  $x_0$  равен  $y_0$ , если функция  $g$ , определенная соотношением  $g(x) = f(x)$  при  $x \neq x_0$  и  $g(x_0) = y_0$  иначе, непрерывна в точке  $x_0$ .

## 2 Компакты в метрических пространствах

**Определение.** Множество  $K$  в метрическом пространстве называется компактным тогда и только тогда, когда из произвольной последовательности  $\{x_n\}_{n=1}^\infty \subset K$  можно выделить сходящуюся подпоследовательность  $x_{n_k} \rightarrow x \in K$ .

**Лемма.** Пусть  $K$  — компакт. Тогда

1.  $K$  — ограниченное множество;
2.  $K$  — замкнутое множество;
3. образ  $K$  при непрерывном отображении компактен.

*Доказательство.*

1. Зафиксируем произвольную точку  $x_0 \in K$ . Если  $K$  — неограниченное множество, то найдется последовательность  $x_n \in K$ ,  $d(x_n, x_0) \rightarrow \infty$ . Переходя к подпоследовательности, имеем  $x_{n_k} \rightarrow x$ ,  $d(x_{n_k}, x_0) \rightarrow d(x, x_0)$ . Противоречие.
2. Если  $x_n \in K$ ,  $x_n \rightarrow x_0$ , то переходя к подпоследовательности  $x_{n_k} \rightarrow x \in K$ , в силу единственности предела  $x_0 = x \in K$ .

3. Рассмотрим последовательность  $\{f(x_n)\}_{n=1}^\infty$ ,  $x_n \in K$ . Переходя к подпоследовательности имеем  $x_{n_k} \rightarrow x \in K$ . Так как  $f$  — непрерывное отображение, то  $f(x_{n_k}) \rightarrow f(x) \in f(K)$ . ■

**Предложение.** Множество  $K$  в  $\mathbb{R}^k$  со стандартной евклидовой метрикой компактно тогда и только тогда, когда оно замкнуто и ограничено.

*Доказательство.* Необходимость этого условия следует из предыдущей леммы. Проверим достаточность. Пусть множество  $K$  — замкнуто и ограничено, и пусть  $x_n \in K$ . В силу ограниченности  $K$  ограниченными будут и все координаты  $(x_n)_j$  последовательности  $x_n$ . Тогда найдется сходящаяся подпоследовательность первых координат  $(x_{n_m})_1$ . Далее, из последовательности  $(x_{n_m})_2$  можно также извлечь сходящуюся подпоследовательность. Повторяя процедуру, получим подпоследовательность  $x'_n$ , у которой каждая координата сходится, то есть  $(x'_n)_j \rightarrow x_j$  для некоторого  $x_j$ . Тем самым,  $x'_n \rightarrow x = (x_1, \dots, x_k)$ . В силу замкнутости  $K$ , вектор  $x \in K$ . ■

**Следствие.** Пусть  $K$  — компакт,  $f : K \rightarrow \mathbb{R}$  — непрерывная функция. Тогда образ  $f(K)$  — ограниченное множество и найдутся точки  $x_m, x_M \in K$ , для которых  $f(x_m) = \inf_{x \in K} f(x)$ ,  $f(x_M) = \sup_{x \in K} f(x)$ .

### 3 Дифференцируемость отображений

**Определение.** Отображение  $f : \mathbb{R}^k \rightarrow \mathbb{R}^m$  называется дифференцируемым в точке  $x$ , если для каждого  $h \in \mathbb{R}^k$

$$f(x+h) = f(x) + Lh + \alpha(h) \|h\|,$$

где  $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$  — линейное отображение,  $\lim_{\|h\| \rightarrow 0} \|\alpha(h)\| = 0$ . Линейное отображение  $L$  называют дифференциалом  $f$  в точке  $x$  и обозначают  $df$ .

**Замечание.** Напомним, что отображение  $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$  называется линейным, если

$$L(a_1 h_1 + a_2 h_2) = a_1 Lh_1 + a_2 Lh_2$$

для произвольных векторов  $h_1, h_2 \in \mathbb{R}^k$  и произвольных чисел  $a_1, a_2 \in \mathbb{R}$ .

Если в  $\mathbb{R}^k$  фиксирован базис  $e := \{e_1, \dots, e_k\}$ , а в  $\mathbb{R}^m$  фиксирован  $e' := \{e'_1, \dots, e'_m\}$ , то линейное отображение  $L$  представимо в виде  $L(h) = L(e_1)h_1 + \dots + L(e_k)h_k$ , где  $h = (h_1, \dots, h_k)$  в базисе  $e$ , а векторы  $L(e_j) = (e_{1,j}, \dots, e_{m,j})$  в базисе  $e'$ .

В частности, каждое линейное отображение при фиксированных базисах  $e$  и  $e'$  в  $\mathbb{R}^k$  и  $\mathbb{R}^m$  соответственно записывается с помощью матрицы  $A = (a_{ij})$ . Кроме того,

$$\|Lh\| \leq (\|L(e_1)\| + \dots + \|L(e_k)\|) \cdot \max_{1 \leq j \leq k} |h_j| \leq C \|h\|$$

и каждое линейное отображение непрерывно на  $\mathbb{R}^k$ .

**Следствие.** Если отображение  $f : \mathbb{R}^k \rightarrow \mathbb{R}^m$  дифференцируемо в точке  $x$ , то оно непрерывно в точке  $x$ .

*Доказательство.* Действительно,  $\|f(x+h) - f(x)\| = \|df(h) + \alpha(h) \|h\|\| \leq C \|h\|$  при  $h$  из некоторой окрестности нуля. ■

**Замечание.** Так как дифференцируемость  $f$  в точке  $x$  равносильна тому, что

$$\lim_{h \rightarrow 0} \frac{\|f(x+h) - f(x) - Lh\|}{\|h\|} = 0,$$

и так как сходимость по норме равносильна покоординатной сходимости, то при фиксированном базисе  $e' := \{e'_1, \dots, e'_m\}$  в  $\mathbb{R}^m$  дифференцируемость отображения  $f$  равносильна дифференцируемости каждой координаты  $f_j$  в точке  $x$ . В этом случае  $Lh = (L_1 h, \dots, L_m h)$  в базисе  $e'$ , где  $L_j = df_j$  — дифференциал  $j$ -ой координаты.

**Лемма.** Если отображение  $f : \mathbb{R}^k \rightarrow \mathbb{R}^m$  дифференцируемо в точке  $x$ , то для каждого вектора  $h \in \mathbb{R}^k$  функция  $t \rightarrow f(x+th)$  дифференцируема в точке 0 и  $\left. \frac{d}{dt} f(x+th) \right|_{t=0} = df(h)$ .

*Доказательство.* По определению

$$f(x + th) - f(x) = t df(h) + t\alpha(th) \|h\|.$$

Разделив на  $t$  и перейдя к пределу при  $t \rightarrow 0$ , получаем требуемое соотношение. ■

**Определение.** Производная  $\frac{\partial f}{\partial h}(x) := \frac{d}{dt} f(x + th) \Big|_{t=0}$  называется производной вдоль вектора  $h$  и может существовать и в случае, когда сама функция  $f$  не дифференцируема в точке  $x$ .

Как мы уже поняли, для дифференцируемости отображения достаточно исследовать дифференцируемость его координат, то есть дифференцируемость функции  $f : \mathbb{R}^k \rightarrow \mathbb{R}$ . Зафиксировав базис  $e := \{e_1, \dots, e_k\}$  в  $\mathbb{R}^n$ , условие дифференцируемости в точке  $x = (x_1, \dots, x_k)$  переписывается в виде

$$f(x_1 + h_1, \dots, x_k + h_k) = f(x_1, \dots, x_k) + c_1 h_1 + \dots + c_k h_k + \bar{o}(\|h\|),$$

то есть  $df(h) = c_1 h_1 + \dots + c_k h_k$ . Из уже доказанного ясно, что  $\frac{\partial f}{\partial e_j}(x) = df(e_j) = c_j$ .

**Определение.** Частной производной  $\frac{\partial f}{\partial x_j}$  функции  $f : \mathbb{R}^k \rightarrow \mathbb{R}$  в точке  $x = (x_1, \dots, x_k)$  называется производная вдоль вектора  $e_j$ , то есть

$$\frac{\partial f}{\partial x_j}(x) = \frac{d}{dt} f(x_1, \dots, x_{j-1}, t, x_{j+1}, \dots, x_k) \Big|_{t=x_j}.$$

**Замечание.** При фиксированном базисе  $e = \{e_1, \dots, e_k\}$  в  $\mathbb{R}^k$  линейные функционалы  $dx_1, \dots, dx_k$  оказываются сопряженным базисом к  $e$ . То есть  $dx_i(e_j) = \delta_{i,j}$ . Таким образом,  $df = \frac{\partial f}{\partial x_1} dx_1 + \dots + \frac{\partial f}{\partial x_k} dx_k$ .

**Замечание.** В случае отображения  $f : \mathbb{R}^k \rightarrow \mathbb{R}^m$  при фиксированных базисах  $e$  и  $e'$  в  $\mathbb{R}^k$  и в  $\mathbb{R}^m$  соответственно, компоненты матрицы дифференциала  $df$  имеют вид  $a_{i,j} = \frac{\partial f_i}{\partial x_j}(x)$ , то есть по строкам написаны градиенты  $\nabla f_i(x)$ .

**Определение.** При фиксированных базисах  $e$  в  $\mathbb{R}^k$  и  $e'$  в  $\mathbb{R}^m$  матрицу, соответствующую линейному отображению  $df$ , называют матрицей Якоби отображения  $f$  в точке  $x$  и обозначают  $J_f(x)$ .

## 4 Градиент и достаточное условие дифференцируемости

**Определение.** Градиентом функции  $f$  называется вектор  $\nabla f := \left( \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$ .

**Лемма.** Если  $f$  дифференцируема в точке  $x$  и  $df \neq 0$ , то наибольшее значение производной вдоль единичного вектора  $v$  (т.е.  $\|v\| = 1$ ) достигается на векторе  $\|\nabla f(x)\|^{-1} \nabla f(x)$ .

*Доказательство.* Так как  $\frac{\partial f}{\partial v}(x) = df(v) = \langle \nabla f(x), v \rangle$ , то по неравенству Коши–Буняковского  $\left| \frac{\partial f}{\partial v}(x) \right| \leq \|\nabla f(x)\| \|v\| = \|\nabla f(x)\|$ . Если  $v = \|\nabla f(x)\|^{-1} \nabla f(x)$ , то в неравенстве достигается равенство. ■

Заметим, что наличия частных производных в точке недостаточно для дифференцируемости функции в этой точке.

**Пример.** Пусть

$$f(x, y) = \begin{cases} \frac{2xy}{x^2 + y^2}, & (x, y) \neq (0, 0), \\ 0, & (x, y) = (0, 0). \end{cases}$$

Функция  $f$  разрывна в нуле, а значит не дифференцируема, но в точке  $(0, 0)$  существуют обе частных производных. Действительно, если  $x = r \cos \varphi, y = r \sin \varphi$ , то функция  $f(x, y) = \sin 2\varphi$ . Таким образом,  $f(x, y)$  в любой окрестности точки  $(0, 0)$  принимает значения  $\pm 1$ , но  $\frac{\partial f}{\partial x}(0, 0) = \frac{d}{dx} f(x, 0) = 0$ . Аналогично  $\frac{\partial f}{\partial y}(0, 0) = 0$ .

В следующей теореме сформулировано достаточное условие дифференцируемости.

**Теорема.** Если все частные производные  $\frac{\partial f}{\partial x_j}$  существуют в окрестности точки  $x_0$  и непрерывны в этой точке, то  $f$  — дифференцируема в точке  $x_0$ .

*Доказательство.* Для сокращения выкладок докажем теорему в случае  $k = 2$ . Заметим, что

$$\begin{aligned} f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2) &= f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2 + h_2) + f(x_1, x_2 + h_2) - f(x_1, x_2) \\ &= \frac{\partial f}{\partial x_1}(\xi_1, x_2 + h_2)h_1 + \frac{\partial f}{\partial x_2}(x_1, \xi_2)h_2, \end{aligned}$$

где  $\xi_1$  принадлежит интервалу с концами  $x_1, x_1 + h_1$ , а  $\xi_2$  — с концами  $x_2, x_2 + h_2$ . Запишем теперь последнюю сумму в виде

$$\frac{\partial f}{\partial x_1}(x_1, x_2)h_1 + \frac{\partial f}{\partial x_2}(x_1, x_2)h_2 + \alpha(h)\|h\|,$$

где

$$\alpha(h) = \left( \frac{\partial f}{\partial x_1}(\xi_1, x_2 + h_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right) \frac{h_1}{\|h\|} + \left( \frac{\partial f}{\partial x_2}(x_1, \xi_2) - \frac{\partial f}{\partial x_2}(x_1, x_2) \right) \frac{h_2}{\|h\|}.$$

При малой  $\|h\|$  выражения в скобках будут малы в силу непрерывности частных производных. Тем самым,  $\lim_{h \rightarrow 0} \|\alpha(h)\| = 0$ . ■

## 5 Частные производные высокого порядка

**Определение.** Пусть  $f: \mathbb{R}^k \rightarrow \mathbb{R}$  и предположим, что в некоторой окрестности  $B_r(x_0)$  точки  $x_0$  существует частная производная  $\frac{\partial f}{\partial x_j}$ . Если функция  $x \mapsto \frac{\partial f}{\partial x_j}(x)$  в точке  $x_0$  имеет частную производную по переменной  $x_i$ , то эта частная производная  $\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right)(x_0)$  называется *частной производной второго порядка* по переменным  $x_j$  и  $x_i$  и обозначается  $\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$ .

**Замечание.** Заметим, что частные производные  $\frac{\partial^2 f}{\partial x_i \partial x_j}$  и  $\frac{\partial^2 f}{\partial x_j \partial x_i}$  являются разными объектами и, вообще говоря, не совпадают (пример будет в рамках семинарских задач).

О совпадении смешанных частных производных позволяют судить следующие две теоремы, которые мы для простоты сформулируем в двумерном случае (общий случай, по сути, ничем не отличается).

**Теорема 1 (Шварц).** Пусть смешанные частные производные  $\frac{\partial^2 f}{\partial x \partial y}$  и  $\frac{\partial^2 f}{\partial y \partial x}$  существуют в окрестности точки  $(x_0, y_0)$  и непрерывны в этой точке. Тогда их значения в точке  $(x_0, y_0)$  совпадают.

**Теорема 2 (Юнг).** Пусть  $f$  — дифференцируема в окрестности точки  $(x_0, y_0)$ , а ее частные производные  $\frac{\partial f}{\partial x}$  и  $\frac{\partial f}{\partial y}$  дифференцируемы в точке  $(x_0, y_0)$ . Тогда смешанные частные производные  $\frac{\partial^2 f}{\partial x \partial y}$  и  $\frac{\partial^2 f}{\partial y \partial x}$  в точке  $(x_0, y_0)$  совпадают.

Приведем доказательство второй из этих теорем.

*Доказательство.* Не ограничивая общности, будем считать, что  $(x_0, y_0) = (0, 0)$ . Рассмотрим функцию

$$F(t, t) = f(t, t) - f(0, t) - f(t, 0) + f(0, 0).$$

Применяя теорему Лагранжа к функции  $g(u) = f(t, u) - f(0, u)$ , получаем

$$F(t, t) = g(t) - g(0) = g'(\xi)t = \left( \frac{\partial f}{\partial y}(t, \xi) - \frac{\partial f}{\partial y}(0, \xi) \right)t.$$

Дифференцируемость  $\frac{\partial f}{\partial y}$  в точке  $(0, 0)$  означает, что

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial f}{\partial y}(0, 0) + \frac{\partial^2 f}{\partial x \partial y}(0, 0)x + \frac{\partial^2 f}{\partial y^2}(0, 0)y + \bar{o}(\sqrt{x^2 + y^2}).$$

Таким образом,

$$\frac{\partial f}{\partial y}(t, \xi) = \frac{\partial f}{\partial y}(0, 0) + \frac{\partial^2 f}{\partial x \partial y}(0, 0)t + \frac{\partial^2 f}{\partial y^2}(0, 0)\xi + \bar{o}(\sqrt{t^2 + \xi^2}).$$

и

$$\frac{\partial f}{\partial y}(0, \xi) = \frac{\partial f}{\partial y}(0, 0) + \frac{\partial^2 f}{\partial y^2}(0, 0)\xi + \bar{o}(\xi).$$

Т.к.  $\xi \leq t$ , то  $\bar{o}(\sqrt{t^2 + \xi^2}) = \bar{o}(t)$  и  $\bar{o}(\xi) = \bar{o}(t)$ . Таким образом,

$$F(t, t) = \frac{\partial^2 f}{\partial x \partial y}(0, 0)t^2 + \bar{o}(t^2).$$

Аналогично,

$$F(t, t) = \frac{\partial^2 f}{\partial y \partial x}(0, 0)t^2 + \bar{o}(t^2).$$

Приравняв полученные выражения, поделив на  $t^2$  и устремив  $t$  к нулю, получаем

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

■

## 6 Правила дифференцирования

**Теорема.** Пусть функции  $f, g: \mathbb{R}^k \rightarrow \mathbb{R}$  дифференцируемы в некоторой точке  $x$ . Тогда, для произвольных чисел  $a, b \in \mathbb{R}$ , функции  $af + bg$  и  $fg$  дифференцируемы в точке  $x$  и  $d(af + bg) = a df + b dg$  и  $d(fg) = f dg + g df$ .

*Доказательство.* Заметим, что

$$\begin{aligned} (af + bg)(x + h) - (af + bg)(x) &= a(f(x + h) - f(x)) + b(g(x + h) - g(x)) \\ &= a(df(h) + \bar{o}(\|h\|)) + b(dg(h) + \bar{o}(\|h\|)) = a df(h) + b dg(h) + \bar{o}(\|h\|). \end{aligned}$$

Таким образом,  $d(af + bg) = a df + b dg$ .

Для доказательства второго равенства заметим, что

$$\begin{aligned} (fg)(x + h) - (fg)(x) &= (f(x + h) - f(x))g(x + h) + f(x)(g(x + h) - g(x)) \\ &= (df(h) + \bar{o}(\|h\|))(g(x) + \bar{o}(1)) + f(x)(dg(h) + \bar{o}(\|h\|)) \\ &= g(x) df(h) + f(x) dg(h) + (df(h))\bar{o}(1) + g(x)\bar{o}(\|h\|) + f(x)\bar{o}(\|h\|) + \bar{o}(\|h\|). \end{aligned}$$

Мы использовали непрерывность функции  $g$ , т.е.  $g(x + h) - g(x) = \bar{o}(1)$  при  $\|h\| \rightarrow 0$ , в силу ее дифференцируемости в точке  $x$ .

Так как  $df$  — линейное отображение, то для некоторого числа  $C > 0$  выполнено  $|df(h)| \leq C\|h\|$ , а значит  $(df(h))\bar{o}(1) = \bar{o}(\|h\|)$ . Т.к.  $f(x)$  и  $g(x)$  просто числа, то  $g(x)\bar{o}(\|h\|) + f(x)\bar{o}(\|h\|) = \bar{o}(\|h\|)$ . Таким образом, теорема доказана. ■

## 7 Дифференциалы высоких порядков.

Предположим, что  $f: \mathbb{R}^k \rightarrow \mathbb{R}$  — дифференцируема в окрестности точки  $a$  и предположим, что ее частные производные  $\frac{\partial f}{\partial x_j}$  дифференцируемы в точке  $a$ . Тогда при каждом  $h \in \mathbb{R}^k$  возникает функция

$$x \mapsto df|_x(h) = \frac{\partial f}{\partial x_1}(x)h_1 + \dots + \frac{\partial f}{\partial x_k}(x)h_k, \text{ дифференцируемая в точке } a.$$

$$\text{Ее дифференциал } d(df(h))|_a(q) = \left( \sum_{j=1}^k \frac{\partial^2 f}{\partial x_j \partial x_1}(a)q_j \right) h_1 + \dots + \left( \sum_{j=1}^k \frac{\partial^2 f}{\partial x_j \partial x_k}(a)q_j \right) h_k.$$



То есть получена билинейная форма  $d(df(h))|_a(q) = \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_j \partial x_i}(a) q_j h_i$ . Эта билинейная форма называется симметричной по теореме Юнга, а т.к. симметричная билинейная форма однозначно задается своей квадратичной формой  $d(df(h))|_a(h) = \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_j \partial x_i}(a) h_j h_i = \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_j \partial x_i}(a) dx_j(h) dx_i(h)$ , то эту квадратичную форму  $d^2 f := \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_j \partial x_i}(a) dx_j dx_i$  и называют **вторым дифференциалом** функции  $f$ .

Аналогично определяется дифференциал  $n$ -го порядка:

**Определение.** Если  $f$  —  $n$  раз дифференцируема в точке  $a$ , то

$$d^n f|_a := \sum_{1 \leq j_1, \dots, j_n \leq k} \frac{\partial^n f}{\partial x_{j_1} \dots \partial x_{j_n}}(a) dx_{j_1} \dots dx_{j_n}.$$

Последняя запись означает лишь то, что при вычислении  $n$ -го дифференциала на векторе  $h \in \mathbb{R}^k$  надо воспользоваться линейностью, а  $[dx_{j_1} \dots dx_{j_n}](h) := dx_{j_1}(h) \dots dx_{j_n}(h) = h_{j_1} \dots h_{j_n}$ .

## 8 Дифференцирование сложной функции.

**Теорема.** Пусть  $f: \mathbb{R}^k \rightarrow \mathbb{R}^m$ ,  $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ , причем отображение  $f$  дифференцируемо в точке  $a$ , отображение  $g$  дифференцируемо в точке  $f(a)$ . Тогда отображение  $g \circ f$  дифференцируемо в точке  $a$  и  $d(g \circ f)|_a = dg|_{f(a)} \circ df|_a$ .

**Замечание.** Поясним запись  $d(g \circ f)|_a = dg|_{f(a)} \circ df|_a$ . Здесь  $df|_a: \mathbb{R}^k \rightarrow \mathbb{R}^m$  есть линейное отображение и  $dg|_{f(a)}: \mathbb{R}^m \rightarrow \mathbb{R}^n$  есть линейное отображение. Тогда их композиция  $dg|_{f(a)} \circ df|_a: \mathbb{R}^k \rightarrow \mathbb{R}^n$  есть линейное отображение, действующее по правилу

$$dg|_{f(a)} \circ df|_a(h) = dg|_{f(a)}(df|_a(h)).$$

*Доказательство.* По условию  $f(a+h) - f(a) = df(h) + \alpha(h)\|h\|$ , где  $\lim_{\|h\| \rightarrow 0} \|\alpha(h)\| = 0$  и  $g(f(a)+q) - g(f(a)) = dg(q) + \beta(q)\|q\|$ , где  $\lim_{\|q\| \rightarrow 0} \|\beta(q)\| = 0$ . Мы также доопределим  $\alpha$  и  $\beta$  в точке нуль нулем (т.е. считаем  $\alpha(0) = 0$  и  $\beta(0) = 0$ ). Тогда

$$\begin{aligned} g(f(a+h)) - g(f(a)) &= g(f(a) + [f(a+h) - f(a)]) - g(f(a)) \\ &= dg[f(a+h) - f(a)] + \beta(f(a+h) - f(a))\|f(a+h) - f(a)\| \\ &= dg[df(h) + \alpha(h)\|h\|] + \beta(f(a+h) - f(a))\|df(h) + \alpha(h)\|h\|. \end{aligned}$$

Тем самым,

$$g(f(a+h)) - g(f(a)) = dg[df(h)] + \gamma(h)\|h\|,$$

где

$$\begin{aligned} \|\gamma(h)\| &= \|dg[\alpha(h)] + \beta(f(a+h) - f(a))\| \|df(h/\|h\|) + \alpha(h)\| \\ &\leq \|dg[\alpha(h)]\| + \|\beta(f(a+h) - f(a))\| (\|df(h/\|h\|)\| + \|\alpha(h)\|). \end{aligned}$$

Напомним, что для линейных отображений  $dg$  и  $df$  существуют такие постоянные  $A$  и  $B$ , что  $\|df(h)\| \leq A\|h\|$  и  $\|dg(q)\| \leq B\|q\|$ , поэтому  $\|df(h/\|h\|)\| + \|\alpha(h)\| \leq A + \|\alpha(h)\|$  и  $\|dg[\alpha(h)]\| \leq B\|\alpha(h)\|$ .

Так как  $\|\beta(f(a+h) - f(a))\| \rightarrow 0$  при  $\|h\| \rightarrow 0$ , получаем, что  $\lim_{\|h\| \rightarrow 0} \|\gamma(h)\| = 0$ . ■

**Замечание.** При фиксированных базисах  $e = \{e_1, \dots, e_k\}$ ,  $e' = \{e'_1, \dots, e'_m\}$ ,  $e'' = \{e''_1, \dots, e''_n\}$  в  $\mathbb{R}^k$ ,  $\mathbb{R}^m$  и  $\mathbb{R}^n$  соответственно, матрица композиции линейных отображений есть произведение матриц этих линейных отображений.

Таким образом, в нашем случае для композиции функций  $g \circ f$ , где  $f: \mathbb{R}^k \rightarrow \mathbb{R}^m$  и  $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ , по предыдущей теореме выполнено

$$\begin{pmatrix} \frac{\partial(g \circ f)_1}{\partial y_1}(a) & \dots & \frac{\partial(g \circ f)_1}{\partial y_k}(a) \\ \vdots & & \vdots \\ \frac{\partial(g \circ f)_n}{\partial y_1}(a) & \dots & \frac{\partial(g \circ f)_n}{\partial y_k}(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(f(a)) & \dots & \frac{\partial g_1}{\partial x_m}(f(a)) \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1}(f(a)) & \dots & \frac{\partial g_n}{\partial x_m}(f(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a) & \dots & \frac{\partial f_1}{\partial y_k}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(a) & \dots & \frac{\partial f_m}{\partial y_k}(a) \end{pmatrix}$$

В частности, в случае, когда  $n = 1$ , для функции  $g(x_1, \dots, x_m)$  и отображения

$$f(y_1, \dots, y_k) = (f_1(y_1, \dots, y_k), \dots, f_m(y_1, \dots, y_k)),$$

выполнено:

$$\begin{pmatrix} \frac{\partial(g \circ f)}{\partial y_1}(a) & \dots & \frac{\partial(g \circ f)}{\partial y_k}(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial g}{\partial x_1}(f(a)) & \dots & \frac{\partial g}{\partial x_m}(f(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a) & \dots & \frac{\partial f_1}{\partial y_k}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(a) & \dots & \frac{\partial f_m}{\partial y_k}(a) \end{pmatrix}$$

Отсюда, во-первых получаем правило вычисления частной производной сложной функции:  $\frac{\partial(g \circ f)}{\partial y_j}(a) = \frac{\partial g}{\partial x_1}(f(a)) \frac{\partial f_1}{\partial y_j}(a) + \dots + \frac{\partial g}{\partial x_m}(f(a)) \frac{\partial f_m}{\partial y_j}(a)$ . Во-вторых, получаем следующее свойство инвариантности первого дифференциала: для дифференциала выполнено равенство  $dg = \frac{\partial g}{\partial x_1} dx_1 + \dots + \frac{\partial g}{\partial x_m} dx_m$ , где нам не важно, являются ли  $dx_1, \dots, dx_m$  — дифференциалами независимых переменных или же являются дифференциалами некоторых функций  $x_j = f_j(y_1, \dots, y_k)$ .

**Пример.** Пусть  $f(x, y) = \varphi(u, v, w)$ , где  $u = xy, v = x + y, w = x - y$ . Тогда

$$\begin{aligned} df &= \frac{\partial \varphi}{\partial u} du + \frac{\partial \varphi}{\partial v} dv + \frac{\partial \varphi}{\partial w} dw = \frac{\partial \varphi}{\partial u} d(xy) + \frac{\partial \varphi}{\partial v} d(x + y) + \frac{\partial \varphi}{\partial w} d(x - y) \\ &= \frac{\partial \varphi}{\partial u} (x dy + y dx) + \frac{\partial \varphi}{\partial v} (dx + dy) + \frac{\partial \varphi}{\partial w} (dx - dy). \end{aligned}$$

В частности,  $\frac{\partial f}{\partial x} = y \frac{\partial \varphi}{\partial u}(xy, x + y, x - y) + \frac{\partial \varphi}{\partial v}(xy, x + y, x - y) + \frac{\partial \varphi}{\partial w}(xy, x + y, x - y)$  и  $\frac{\partial f}{\partial y} = x \frac{\partial \varphi}{\partial u}(xy, x + y, x - y) + \frac{\partial \varphi}{\partial v}(xy, x + y, x - y) - \frac{\partial \varphi}{\partial w}(xy, x + y, x - y)$ .

## 9 Дифференциал обратного отображения

**Теорема.** Пусть  $f: \mathbb{R}^k \rightarrow \mathbb{R}^k$  — есть непрерывная биекция между окрестностями  $U(a)$  и  $V(f(a))$ , причем обратное отображение  $f^{-1}: V(f(a)) \rightarrow U(a)$  также непрерывно (т.е.  $f$  — гомеоморфизм между  $U(a)$  и  $V(f(a))$ ).

Предположим, что  $f$  — дифференцируемо в точке  $a$  и  $df$  — обратимое линейное отображение. Тогда  $f^{-1}$  — дифференцируемо в точке  $f(a)$  и  $df^{-1}|_{f(a)} = (df|_a)^{-1}$ .

*Доказательство.* Нам нужно проверить, что

$$\lim_{\|q\| \rightarrow 0} \frac{\|f^{-1}(f(a) + q) - f^{-1}(f(a)) - (df)^{-1}(q)\|}{\|q\|} = 0.$$

Пусть  $h = f^{-1}(f(a) + q) - f^{-1}(f(a)) = f^{-1}(f(a) + q) - a$ , тогда  $q = f(a + h) - f(a)$  и  $\|q\| \rightarrow 0$  тогда и только тогда, когда  $\|h\| \rightarrow 0$ .

Так как  $f$  — дифференцируемо в точке  $a$ , то

$$f(a + h) - f(a) = df(h) + \alpha(h)\|h\|,$$

где  $\lim_{\|h\| \rightarrow 0} \|\alpha(h)\| = 0$ .

Таким образом,

$$\lim_{\|q\| \rightarrow 0} \frac{\|f^{-1}(f(a) + q) - f^{-1}(f(a)) - (df)^{-1}(q)\|}{\|q\|} = \lim_{\|h\| \rightarrow 0} \frac{\|h - (df)^{-1}(df(h) + \alpha(h)\|h\|)\|}{\|df(h) + \alpha(h)\|h\|}.$$

Числитель в последнем выражении равен  $\|h\| \|(df)^{-1}(\alpha(h))\|$ . Для линейного отображения  $(df)^{-1}$  найдется число  $C > 0$ , для которого  $\|(df)^{-1}(p)\| \leq C\|p\|, \forall p \in \mathbb{R}^k$ .

Отсюда, подставив  $p = df(h)$ , получаем  $C^{-1}\|h\| \leq \|df(h)\|$ . Тем самым

$$\|df(h) + \alpha(h)\|h\| \geq \|df(h)\| - \|h\|\|\alpha(h)\| \geq \|h\|(C^{-1} - \|\alpha(h)\|).$$

Таким образом,

$$\frac{\|h - (df)^{-1}(df(h) + \alpha(h)\|h\|)\|}{\|df(h) + \alpha(h)\|h\|} \leq \frac{C\|h\|\|\alpha(h)\|}{\|h\|(C^{-1} - \|\alpha(h)\|)} = \frac{C\|\alpha(h)\|}{(C^{-1} - \|\alpha(h)\|)} \rightarrow 0$$

при  $\|h\| \rightarrow 0$ . ■

**Замечание.** Отметим, что матрица обратного линейного отображения есть обратная матрица к матрице исходного линейного отображения. Тем самым, матрица Якоби обратного отображения  $J_{f^{-1}}(y)$  является обратной к матрице Якоби исходного отображения, т.е. равна  $(J_f(f^{-1}(y)))^{-1}$ .

## Теорема о неявной функции.

Пусть в  $\mathbb{R}^2$  у нас имеется соотношение  $F(x, y) = 0$ . Нам бы хотелось понять при каких условиях данное уравнение возможно разрешить относительно  $y$  в виде явной зависимости  $y = f(x)$ .

Рассмотрим например  $F(x, y) = x^2 + y^2 - 1$ . Тогда уравнение  $F(x, y) = 0$  задает обычную окружность и все решения данного уравнения относительно  $y$  имеют вид  $y = \pm\sqrt{1 - x^2}$ . Ясно, что произвольный выбор знаков в разных точках  $x$  будет давать бесконечно много решений данного уравнения.

В тоже время в малой окрестности произвольной точки  $(x_0, y_0)$  на окружности (кроме  $x_0 = \pm 1$ ) кривая  $F(x, y) = 0$  единственным образом представима в виде графика непрерывной функции  $y = f(x)$ . В окрестности же точек  $(\pm 1, 0)$  никакая дуга окружности не может быть представлена в виде графика функции  $y = f(x)$ . Зато эти дуги в окрестности точек  $(\pm 1, 0)$  хорошо расположены относительно оси  $y$  и могут быть представлены в виде графика  $x = g(y)$ .

Чем же обусловлена такая особенность точек  $(\pm 1, 0)$  в случае окружности? Заметим, что локально функция  $F(x, y)$  представима в виде  $F(x, y) = \frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0)(y - y_0) + \bar{o}(\sqrt{|x - x_0|^2 + |y - y_0|^2})$ .

Таким образом, пренебрегая малыми более высокого порядка, наше уравнение  $F(x, y) = 0$  в окрестности точки  $(x_0, y_0)$  похоже на линейное уравнение  $\frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial y}(x_0, y_0)(y - y_0) = 0$ , которое в свою очередь разрешимо относительно  $y$  только в случае  $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$ .

В частности, в случае окружности как раз  $\frac{\partial F}{\partial y}(\pm 1, 0) = 0$ . Из данного эвристического рассуждения возникает гипотеза, что уравнение  $F(x, y) = 0$  разрешимо относительно переменной  $y$  в некоторой окрестности данной точки  $(x_0, y_0)$ , если производная  $\frac{\partial F}{\partial y}(x_0, y_0)$  отлична от нуля. Именно это мы и докажем в следующей теореме уже в строго сформулированном виде.

Для сокращения всех записей будем использовать обозначение  $F'_y(x, y) := \frac{\partial F}{\partial y}(x, y)$ .

**Теорема.** Пусть  $F: \mathbb{R}^2 \rightarrow \mathbb{R}$  — определена и непрерывно дифференцируема (т.е. частные производные непрерывно зависят от точки) в некоторой окрестности  $U$  точки  $(a, b) \in \mathbb{R}^2$ . Пусть 1)  $F(a, b) = 0$  и 2)  $F'_y(a, b) \neq 0$ .

Тогда найдутся промежутки  $I_x = (a - \alpha, a + \alpha)$  и  $I_y = (b - \beta, b + \beta)$  и непрерывно дифференцируемая функция  $f: I_x \rightarrow I_y$ , для которых  $I_x \times I_y \subset U$  и для каждой точки  $(x, y) \in I_x \times I_y$  выполнено  $F(x, y) = 0 \Leftrightarrow y = f(x)$ .

Кроме того,  $f'(x) = -\frac{F'_x(x, f(x))}{F'_y(x, f(x))}$ .

*Доказательство.* 1. Для определенности считаем, что  $F'_y(a, b) > 0$ . Так как производные функции  $F$  непрерывны в  $U$ , то в малой окрестности  $\{(x, y): \sqrt{|x - a|^2 + |y - b|^2} < 2\beta\}$  точки  $(a, b)$  также выполнено  $F'_y(x, y) > 0$ .

Так как  $F'_y(a, y) > 0$  на отрезке  $[b - \beta, b + \beta]$ , то функция  $y \mapsto F(a, y)$  монотонно возрастает на этом отрезке, откуда

$$F(a, b - \beta) < F(a, b) = 0 < F(a, b + \beta).$$

Так как  $F$  непрерывна в  $U$ , то найдется такое число  $\alpha < \beta$ , что  $F(x, b - \beta) < 0 < F(x, b + \beta)$  при  $x \in (a - \alpha, a + \alpha)$ .

При каждом  $x \in (a - \alpha, a + \alpha)$  рассмотрим функцию  $y \mapsto F(x, y)$ , заданную на отрезке  $[b - \beta, b + \beta]$ . Рассматриваемая функция есть непрерывная строго возрастающая функция на отрезке, причем

на концах отрезка данная функция принимает значения разных знаков. Поэтому при каждом  $x \in (a - \alpha, a + \alpha)$  существует единственная точка  $y = f(x)$ , для которой  $F(x, f(x)) = 0$ . Тем самым построена окрестность точки  $(a, b)$  вида  $I_x \times I_y$  в которой построено единственное решение уравнения  $F(x, y) = 0$  относительно переменной  $y$ .

2. Проверим непрерывность построенного решения в точке  $a$ . Ясно, что  $f(a) = b$  в силу единственности нуля у функции  $y \mapsto F(a, y)$  на  $I_y$ . Пусть теперь фиксировано некоторое  $\varepsilon \in (0, \beta)$ . Повторяя рассуждения первой части для интервала  $(b - \varepsilon, b + \varepsilon)$  найдем интервал  $(a - \delta, a + \delta)$  с  $\delta < \alpha$  и функцию  $\tilde{f}: (a - \delta, a + \delta) \rightarrow (b - \varepsilon, b + \varepsilon)$  для которых  $F(x, y) = 0$  при  $(x, y) \in (a - \delta, a + \delta) \times (b - \varepsilon, b + \varepsilon) \Leftrightarrow y = \tilde{f}(x), x \in (a - \delta, a + \delta)$ .

Так как  $(a - \delta, a + \delta) \subset I_x$  и  $(b - \varepsilon, b + \varepsilon) \subset I_y$ , то в силу единственности решения  $f$  в  $I_x \times I_y$  получаем, что  $f(x) = \tilde{f}(x)$  при  $x \in (a - \delta, a + \delta)$ . Это означает, что  $|f(x) - b| < \varepsilon$  при  $|x - a| < \delta$ .

Теперь проверим непрерывность  $f$  в произвольной точке  $x \in I_x$ . Для этого просто примем за начальную точку построения произвольную точку  $(x, y)$  с  $x \in I_x, y \in I_y$  и повторим рассуждение выше.

3. Докажем непрерывную дифференцируемость  $f$  на  $I_x$  и докажем формулу для вычисления производной. Пусть  $x \in I_x$  и рассмотрим достаточно малое  $\Delta x$ , для которого  $x + \Delta x \in I_x$ . Пусть  $y = f(x) \in I_y$  и  $\Delta y = f(x + \Delta x) - f(x)$ .

Применим теорему Лагранжа к функции  $t \mapsto F(x + t\Delta x, y + t\Delta y)$ :

$$0 = F(x + \Delta x, y + \Delta y) - F(x, y) = F'_x(x + \xi\Delta x, y + \xi\Delta y)\Delta x + F'_y(x + \xi\Delta x, y + \xi\Delta y)\Delta y,$$

где  $\xi \in (0, 1)$ . Т.к.  $F'_y \neq 0$  в  $I_x \times I_y$ , то

$$\frac{\Delta y}{\Delta x} = -\frac{F'_x(x + \xi\Delta x, y + \xi\Delta y)}{F'_y(x + \xi\Delta x, y + \xi\Delta y)}.$$

В силу непрерывности  $f$  при  $\Delta x \rightarrow 0$  выполнено, что и  $\Delta y \rightarrow 0$ , поэтому, в силу непрерывности производных функции  $F$  в  $I_x \times I_y$ , получается, что  $f'(x) = -\frac{F'_x(x, y)}{F'_y(x, y)}$ , где  $y = f(x)$ . Из формулы следует и непрерывность производной. ■

Аналогично доказывается следующий многомерный аналог предыдущей теоремы.

**Теорема.** Пусть  $F: \mathbb{R}^{k+1} \rightarrow \mathbb{R}$  — определена и непрерывно дифференцируема (то есть частные производные непрерывно зависят от точки) в некоторой окрестности  $U$  точки  $(a, b) = (a_1, \dots, a_k, b) \in \mathbb{R}^{k+1}$ .

Пусть 1)  $F(a, b) = 0$  и 2)  $F'_y(a, b) \neq 0$ . Найдутся  $I_x = (a_1 - \alpha_1, a_1 + \alpha_1) \times \dots \times (a_k - \alpha_k, a_k + \alpha_k)$  и  $I_y = (b - \beta, b + \beta)$  и непрерывно дифференцируемая функция  $f: I_x \rightarrow I_y$ , для которых  $I_x \times I_y \subset U$  и для каждой точки  $(x, y) \in I_x \times I_y$  выполнено  $F(x, y) = 0 \Leftrightarrow y = f(x)$ .

Кроме того, 
$$\frac{\partial f}{\partial x_j}(x) = -\frac{F'_{x_j}(x, f(x))}{F'_y(x, f(x))}.$$

*Доказательство.* Существование  $I_x \times I_y$  и функции  $f$ , а также ее непрерывность, дословно повторяют рассуждение из предыдущей теоремы.

Если теперь фиксировать все переменные, кроме  $x_j$  и  $y$ , мы попадем в ситуацию предыдущей теоремы, откуда следует формула для вычисления частной производной. Из формулы следует непрерывность этой частной производной, а значит и непрерывная дифференцируемость  $f$ . ■

**Замечание.** Отметим, что формула для подсчет производной берется из дифференцирования тождества  $F(x, f(x)) = 0$ . Действительно,  $\frac{\partial F}{\partial x_1}dx_1 + \dots + \frac{\partial F}{\partial x_k} + \frac{\partial F}{\partial y}df = 0$ , откуда выражая  $df$  и получаем нужную нам формулу.