Лекции по математическому анализу
(I-й курс, весенний семестр)
лектор С.В. Шапошников

«Так оставьте ненужные споры —
Я себе уже все доказал:
Лучше гор могут быть только горы,
На которых еще не бывал.»
В.С.Высоцкий, «Прощание с горами»

1. $\mathbb{R}^n$ – евклидово, нормированное и метрическое пространство.

Множество $\mathbb{R}^n$ является декартовым произведением n множеств $\mathbb{R}$, т. е. состоит из упорядоченных наборов $(x_1, x_2, \ldots, x_n)$ вещественных чисел x_i. На $\mathbb{R}^n$ можно ввести операцию сложения и операцию умножения на число:

- $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$,
- $\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$.

Легко проверить, что $\mathbb{R}^n$ является линейным пространством над полем $\mathbb{R}$. Вектора

$$e_1 = (1, 0, \ldots, 0), \quad e_2 = (0, 1, 0, \ldots, 0), \quad \ldots \quad e_n = (0, 0, \ldots, 1)$$

образуют базис, т. е. всякий вектор x единственным образом раскладывается по e_k:

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n.$$

Рассмотрим функцию $\langle \cdot\,, \cdot \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, заданную формулой:

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + \ldots + x_ny_n.$$

Эта функция удовлетворяет следующим свойствам:

(i) $\langle x, x\rangle \geqslant 0$ и $\langle x, x\rangle = 0 \Leftrightarrow x = 0$;

(ii) $\langle x, y\rangle = \langle y, x\rangle$;

(iii) $\langle \alpha_1x_1 + \alpha_2x_2, y\rangle = \alpha_1\langle x_1, y\rangle + \alpha_2\langle x_2, y\rangle$.

Всякая функция $\langle \cdot\,, \cdot \rangle$ на линейном пространстве, удовлетворяющая этим трем свойствам, называется **скалярным произведением**. Линейное пространство со скалярным произведением называется **евклидовым** пространством.

Теорема 1.1. (Неравенство Коши–Буняковского–Шварца) *Для всех $x, y \in \mathbb{R}^n$ выполняется неравенство*

$$|\langle x, y\rangle| \leqslant \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle},$$

причем равенство равносильно линейной зависимости x и y.

Доказательство. Заметим, что при $y = 0$ утверждение выполнено. Пусть теперь $y \neq 0$. Рассмотрим квадратичную функцию

$$f(t) = \langle x - ty, x - ty\rangle = \langle x, x\rangle - 2t\langle x, y\rangle + t^2\langle y, y\rangle.$$

Так как $f(t) \geqslant 0$ для всех t, то

$$D = 4|\langle x, y\rangle|^2 - 4\langle x, x\rangle\langle y, y\rangle \leqslant 0.$$

Из последнего неравенства немедленно следует требуемое утверждение. Остается заметить, что $D = 0$ тогда и только тогда, когда $x - ty = 0$ для некоторого t. $\square$

Величина $\|x\| = \sqrt{\langle x, x\rangle}$ называется нормой или длиной вектора x.

функция $\|\cdot\|\colon \mathbb{R}^n \to [0, +\infty)$ удовлетворяет следующим свойствам:

(i) $\|x\| = 0 \Leftrightarrow x = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$;

(iii) $\|x + y\| \leqslant \|x\| + \|y\|$.

Неравенство из последнего свойства называют **неравенством треугольника.** Проверим, что оно действительно выполняется:

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

Всякая неотрицательная функция на линейном пространстве, удовлетворяющая свойствам (i), (ii) и (iii), называется **нормой**. Линейное пространство с нормой называется **нормированным** пространством.

Наконец рассмотрим функцию $\varrho(\,\cdot\,,\,\cdot\,)\colon \mathbb{R}^n \times \mathbb{R}^n \to [0, +\infty)$, заданную формулой

$$\varrho(x, y) = \|x - y\|.$$

Для этой функции выполняются следующие свойства:

(i) $\varrho(x, y) = 0 \Leftrightarrow x = y$;

(ii) $\varrho(x, y) = \varrho(y, x)$;

(iii) $\varrho(x, y) \leqslant \varrho(x, z) + \varrho(z, y)$.

Последнее свойство называют **неравенством треугольника** и оно немедленно следует из соответствующего свойства нормы.

Неотрицательная функция $\varrho(\,\cdot\,,\,\cdot\,)$ на непустом множестве, удовлетворяющая (i), (ii) и (iii), называется **метрикой**, а множество с заданной на нем метрикой называется **метрическим пространством.**

Таким образом, $\mathbb{R}^n$ является евклидовым, нормированным и метрическим пространством.

2. Метрическое пространство и нормированное пространство.

Пусть X – произвольное непустое множество. Функция $\varrho(\,\cdot\,,\,\cdot\,)\colon X \times X \to [0, +\infty)$, которая обладает следующими свойствами:

(i) $\varrho(x, y) = 0 \Leftrightarrow x = y$;

(ii) $\varrho(x, y) = \varrho(y, x)$ (симметричность);

(iii) $\varrho(x, y) \leqslant \varrho(x, z) + \varrho(z, y)$ (неравенство треугольника),

называется **метрикой**, а пара (X, ϱ) называется **метрическим пространством**.

Примеры:

0. X – произвольное непустое множество и $\varrho(x, y) = 0$ при $x = y$ и $\varrho(x, y) = 1$ при $x \neq y$.

1. $X = \mathbb{R}$ и $\varrho(x, y) = |x - y|$.

2. $X = \mathbb{R}$ и $\varrho(x, y) = |\mathrm{arctg} x - \mathrm{arctg} y|$.

3. $X = \mathbb{R}^n$ и $\varrho(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$ (евклидова метрика).

4. $X = \mathbb{R}^n$ и $\varrho(x, y) = |x_1 - y_1| + \ldots + |x_n - y_n|$.

5. $X = \mathbb{R}^n$ и $\varrho(x, y = \max_{1 \leqslant k \leqslant n} |x_k - y_k|$.

Множество $B(a, r) = \{x \in X \colon \varrho(a, x) < r\}$ называется **открытым шаром**, а множество $\overline{B}(a, r) = \{x \in X \colon \varrho(a, x) \leqslant r\}$ называется **замкнутым шаром**.

Отметим, что в метрическом пространстве шар большего радиуса может содержаться строго внутри шара меньшего радиуса.

Говорят, что последовательность x_n точек метрического пространства (X, ϱ) сходится к точке x, если $\varrho(x_n, x) \to 0$ при $n \to \infty$.

Предложение 2.1. *Если* $x_n \to x$ *и* $x_n \to y$*, то* $x = y$*.*

Доказательство. По неравенству треугольника

$$\varrho(x,y) \leqslant \varrho(x,x_n) + \varrho(x_n,y) \to 0$$

и, следовательно, $\varrho(x,y) = 0$ и $x = y$. □

Предложение 2.2. *Если последовательность сходится, то она ограничена, т. е. лежит целиком в некотором шаре.*

Доказательство. Пусть $x_n \to x$. Тогда найдется N такое, что для всех $n > N$ выполняется $\varrho(x,x_n) < 1$. Пусть $R > \max\{\varrho(x_1,x),\dots,\varrho(x_N,x),1\}$. Тогда $x_n \in B(x,R)$ для всякого n. □

Лемма 2.1. *Для всякого* $x \in \mathbb{R}^n$ *имеют место неравенства*

$$\max_{1\leqslant m\leqslant n}|x_m| \leqslant \sqrt{\sum_{m=1}^{n} x_m^2} \leqslant \sqrt{n}\max_{1\leqslant m\leqslant n}|x_m|$$

Теорема 2.1. *Последовательность* $x^k \to x$ *в* $\mathbb{R}^n$ *с евклидовой метрикой тогда и только тогда, когда* $x_m^k \to x_m$ *для каждого* $m = 1,2,\dots,n$*.*

Доказательство. Утверждение следует из Леммы 2.1. □

Предложение 2.3. (Больцано) *Если последовательность* $x^k \in \mathbb{R}^n$ *ограниченная, то из нее можно выбрать сходящуюся подпоследовательность.*

Доказательство. Каждая числовая последовательность x_m^k является ограниченной и из нее можно выбрать сходящуюся подпоследовательность. Пусть $n = 2$, т. е. $k = 1$ или $k = 2$. Выбираем сходящуюся подпоследовательность $x_1^{k_m}$, а затем из последовательности $x_2^{k_m}$ выбираем сходящуюся подпоследовательность $x_2^{k_{m_j}}$. Подпоследовательность $x^{k_{m_j}}$ сходится. □

Последовательность $\{x_n\}$ элементов метрического пространства (X,ϱ) называется **фундаментальной**, если для всякого $\varepsilon > 0$ найдется N такое, что для всех $n,m > N$ выполняется неравенство $\varrho(x_n,x_m) < \varepsilon$.

Метрическое пространство (X,ϱ) называется **полным**, если всякая фундаментальная последовательность сходится.

Теорема 2.2. $\mathbb{R}^n$ *является полным пространством.*

Доказательство. Если последовательность x^k фундаментальна, то по Лемме 2.1 каждая из числовых последовательностей x_m^k фундаментальна. По критерию Коши для числовых последовательностей последовательность x_m^k сходится для каждого m. Следовательно, по доказанному выше, x^k сходится. □

Пример неполного пространства: числовая прямая с выколотой точкой.

Пусть X – линейное пространство над $\mathbb{R}$. Функция $\|\cdot\|\colon X \to [0,+\infty)$, которая обладает свойствами:

(i) $\|x\| = 0 \Leftrightarrow x = 0$;

(ii) $\|\alpha x\| = |\alpha|\|x\|$;

(iii) $\|x+y\| \leqslant \|x\| + \|y\|$ (неравенство треугольника),

называется **нормой**, а пара $(X,\|\cdot\|)$ называется **нормированным пространством**.

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

Предложение 2.4. *Пусть* $(X, \|\cdot\|)$ *– нормированное пространство. Пусть* $x_n, y_n \in X$ *сходятся соответственно к* x *и* y*, числовая последовательность* α_n *сходится к* α*. Тогда*

$$x_n + y_n \to x + y, \quad \alpha_n x_n \to \alpha x.$$

Предложение 2.5. *Пусть* $(X, \|\cdot\|)$ *– нормированное пространство. Если* $x_n \to x$*, то* $\|x_n\| \to \|x\|$*.*

Доказательство. Действительно, $|\|x_n\| - \|x\|| \leqslant \|x_n - x\| \to 0$. □

На $\mathbb{R}^n$ кроме евклидовой нормы $\|x\| = \sqrt{\sum_{k=1}^n x_k^2}$ можно ввести и другие нормы, например $\|x\| = \max_{1\leqslant k\leqslant n} |x_k|$ или $\|x\| = \sum_{1\leqslant k\leqslant n} |x_k|$. Однако, оказывается, что на $\mathbb{R}^n$ все нормы эквивалентны.

Теорема 2.3. *На* $\mathbb{R}^n$ *все нормы эквивалентны, т. е. для любых двух норм* $\|\,\|_1$ *и* $\|\,\|_2$ *существуют такие числа* γ_1, γ_2*, что*

$$\gamma_1\|x\|_1 \leqslant \|x\|_2 \leqslant \gamma_2\|x\|_1$$

для всех x*.*

Доказательство. Докажем, что все нормы эквивалентны евклидовой норме $\|\,\|_2$. Пусть задана произвольная норма $\|\,\|_1$ на $\mathbb{R}^n$. Заметим, что

$$\|x\|_1 = \|x_1e_1 + \ldots + x_ne_n\| \leqslant (|x_1| + \ldots + |x_n|)\max_k \|e_k\|_1 \leqslant n\|x\|_2 \max_k \|e_k\|_1.$$

Таким образом, остается лишь установить неравенство $\|x\|_2 \leqslant C\|x\|_1$ для всех x и некоторого $C > 0$. Предположим, что это неравенство не выполняется. Тогда для всякого $N \in \mathbb{N}$ найдется отличный от нуля вектор x^N такой, что $\|x^N\|_2 \geqslant N\|x^N\|_1$. Делим это неравенство на $\|x^N\|_2$ и на N. Получаем

$$1/N > \|y^N\|_1, \quad y^N = x^N/\|x^N\|_2, \quad \|y^N\|_2 = 1.$$

Итак, по норме $\|\,\|_1$ последовательность y^N сходится к нулю, а по евклидовой норме все векторы y^N имеют единичную длину. По теореме Больцано, можно выбрать сходящуюся по евклидовой норме подпоследовательность y^{N_j}. Пусть $y^{N_j} \to y$. Тогда $\|y\|_2 = 1$ и, следовательно, $y \neq 0$. Сходимость по евклидовой норме влечет сходимость по норме $\|\,\|_1$, по которой данная последовательность сходится к нулю. Противоречие. □

Пусть $(X, \|\,\|)$ – нормированное пространство. Формальная сумма $\sum_{n=1}^\infty x_n$ элементов $x_n \in X$ называется рядом. Ряд сходится, если сходится последовательность его частичных сумм $S_N = \sum_{n=1}^N x_n$.

Предложение 2.6. *Пусть* $(X, \|\,\|)$ *– полное нормированное пространство. Если сходится числовой ряд* $\sum_{n=1}^\infty \|x_n\|$*, то сходится ряд* $\sum_{n=1}^\infty x_n$*.*

Доказательство. Достаточно показать, что последовательность частичных сумм фундаментальна. По неравенству треугольника

$$\|S_N - S_M\| \leqslant \sum_{n=M+1}^N \|x_n\|.$$

Так как ряд из $\|x_n\|$ сходится, то для всякого $\varepsilon > 0$ найдется N_0 такое, что для всех $M, N > N_0$ выполняется $\sum_{n=M+1}^N \|x_n\| < \varepsilon$. □

Полное нормированное пространство называют **банаховым пространством.**

3. Равномерная сходимость и ее свойства.

Пусть X – непустое множество. Через $B(X)$ обозначаем линейное пространство ограниченных функций $f\colon X \to \mathbb{R}$. Пространство $B(X)$ является нормированным с нормой

$$\|f\| = \sup_{x\in X} |f(x)|.$$

Последовательность f_n сходится к f в $B(X)$, если $\|f_n - f\| \to 0$, т. е.

$$\sup_{x\in X} |f_n(x) - f(x)| \to 0.$$

Это уже знакомое нам определение равномерной сходимости.

Теорема 3.1. *Пространство $B(X)$ полное.*

Доказательство. Пусть f_n – фундаментальная последовательность. Тогда при каждом $x \in X$ числовая последовательность $f_n(x)$ является фундаментальной и, следовательно, сходится к некоторому числу $f(x)$. Покажем, что $f \in B(X)$ и $\|f_n - f\| \to 0$. Опять воспользуемся тем, что f_n фундаментальна. Для всякого $\varepsilon > 0$ найдется N такое, что для всех $n, m > N$ и всех $x \in X$ выполняется неравенство $|f_n(x) - f_m(x)| < \varepsilon$. Устремим $m \to \infty$. Получаем $|f_n(x) - f(x)| \leqslant \varepsilon$ для всех $x \in X$. $\square$

Таким образом, при исследовании равномерной сходимости можно пользоваться критерием Коши.

Следствие 3.1. (признак Вейерштрасса) *Если $f_n \in B(X)$, $\sup_{x\in X} |f_n(x)| \leqslant a_n$ и ряд $\sum_{n=1}^{\infty} a_n$ сходится, то ряд $\sum_{n=1}^{\infty} f_n(x)$ сходится в $B(X)$, т. е. сходится равномерно на множестве X.*

Доказательство. Так как $B(X)$ – полное пространство, то из сходимости ряда $\sum_n \|f_n\|$ следует сходимость $\sum_n f_n$. $\square$

Следствие 3.2. *Множество $C[a, b]$ непрерывных на отрезке $[a, b]$ функций является полным нормированным пространством с нормой $\|f\| = \max_{x\in[a,b]} |f(x)|$.*

Доказательство. Так как $C[a, b] \subset B[a, b]$, то фундаментальная последовательность f_n сходится по данной норме к некоторой функции $f \in B(X)$ и надо лишь проверить, что эта функция непрерывна. Это мы уже знаем: равномерный предел непрерывных функций является непрерывной функцией. $\square$

Кроме равномерной сходимости естественно рассматривать поточечную. Однако, поточечная сходимость в общем случае не задается метрикой.

Предложение 3.1. *На $C[a, b]$ не существует метрики такой, что последовательность f_n сходится к f по этой метрике тогда и только тогда, когда сходится поточечно.*

Доказательство. Предположим противное. Пусть ϱ – искомая метрика. Для всякого шара $B(0, r)$ и всякого интервала $(\alpha, \beta) \subset [a, b]$ найдется отрезок $[c, d] \subset (\alpha, \beta)$ такой, что все непрерывные функции, равные нулю вне этого отрезка, принадлежат шару $B(0, r)$. Действительно, иначе можно устроить последовательность попарно не пересекающихся отрезков $[c_n, d_n]$ и для каждого из них найти функцию f_n, которая равна нулю вне соответствующего отрезка и не принадлежит шару $B(0, r)$. Тогда последовательность f_n сходится поточечно к нулю, но не сходится по метрике. Строим систему вложенных отрезков I_n таких, что все функции, которые равны нулю вне I_n, принадлежат шару $B(0, 1/n)$. У отрезков есть общая точка c. Пусть f_n – непрерывная функция, которая равна нулю вне I_n и $f_n(c) = 1$. Тогда

$f_n \in B(0,1/n)$ и по метрике f_n сходятся к нулю. Поточечно эта последовательность к нулю не сходится так как в точке c все функции равны единице. $\square$

При некоторых дополнительных условиях из поточечной сходимости следует равномерная сходимость.

Предложение 3.2. *Пусть монотонные функции $f_n \in C[a,b]$ поточечно сходятся к некоторой непрерывной функции f. Тогда последовательность f_n сходится к f равномерно.*

Доказательство. Пусть $x_1 \leqslant x \leqslant x_2$. Тогда $f_n(x_1) \leqslant f_n(x) \leqslant f_n(x_2)$, $f(x_1) \leqslant f(x) \leqslant f(x_2)$ и

$(f_n(x_1)-f(x_1))+(f(x_1)-f(x_2)) \leqslant f_n(x)-f(x) \leqslant f_n(x_2)-f(x_1) = (f_n(x_2)-f(x_2))+(f(x_2)-f(x_1))$.

Следовательно, имеет место неравенство

$$|f_n(x)-f(x)| \leqslant \max\{|f_n(x_1)-f(x_1)|, |f_n(x_2)-f(x_2)|\} + |f(x_2)-f(x_1)|.$$

Пусть $\varepsilon > 0$. Разобьем отрезок $[a,b]$ точками $a = x_0 < x_1 < \ldots < x_K = b$ так, что

$$|f(x_j)-f(x_{j-1})| < \varepsilon$$

для всякого j. Выбираем N столь большим, что $|f_n(x_j)-f(x_j)| < \varepsilon$ для всякого $n > N$ и всякого j. Из полученной выше оценки следует неравенство $|f_n(x)-f(x)| < 2\varepsilon$. $\square$

Предложение 3.3. (признак Дини) *Пусть функции $f_n \in C[a,b]$ таковы, что в каждой точке $x \in [a,b]$ последовательность $f_n(x)$ монотонна и сходится к $f(x) \in C[a,b]$. Тогда последовательность f_n сходится к f равномерно на $[a,b]$.*

Доказательство. Вычитая f можно считать, что последовательность f_n сходится поточечно к нулю. Пусть $\varepsilon > 0$. Для каждой точки $x_0 \in [a,b]$ найдется номер N такой, что $|f_N(x_0)| < \varepsilon$. Из-за непрерывности существует такая окрестность $U(x_0)$, что $|f_N(x)| < 2\varepsilon$ для всякого $x \in U(x_0)$. Так как $f_n(x)$ – монотонная последовательность, то оценка $|f_n(x)| < 2\varepsilon$ выполняется для всех $n > N$. Остается выбрать конечное подпокрытие отрезка $[a,b]$ из окрестностей $U(x)$ и взять максимальный номер N. $\square$

Рассмотрим пространство $C^1[a,b]$ непрерывных на $[a,b]$ функций, имеющих непрерывную производную на $[a,b]$. Это пространство будет нормированным относительно нормы $\|f\| = \max_{x\in[a,b]} |f(x)|$, но не будет полным. Действительно, возьмем последовательность функций $\sqrt{x^2+n^{-2}}$ в пространстве $C^1[-1,1]$. Так как в $C[-1,1]$ эта последовательность сходится к $|x|$, то она является фундаментальной. Однако, данная последовательность не имеет предела в $C^1[a,b]$. Этот пример показывает, что равномерная сходимость не сохраняет дифференцируемость функций.

Теорема 3.2. *Пусть $f_n \in C^1[a,b]$. Если $f_n \rightrightarrows f$ и $f_n' \rightrightarrows g$, то $f' = g$.*

Доказательство. Фиксируем $x \in [a,b]$. Рассмотрим функции

$$g_n(y) = \frac{f_n(y)-f(x)}{y-x} \quad \text{при} \quad y \neq x, \quad g_n(x) = f_n'(x).$$

Заметим, что $g_n(y)$ – непрерывная функция. Кроме того, последовательность g_n фундаментальна в $C[a,b]$. Действительно, по теореме Лагранжа

$$g_n(y)-g_m(y) = \frac{(f_n(y)-f_m(y))-(f_m(x)-f_m(x)))}{y-x} = f_n'(c)-f_m'(c)$$

и

$$\max_{y\in[a,b]]} |g_n(y)-g_m(y)| \leqslant \max_{y\in[a,b]} |f_n'(y)-f_m'(y)|.$$

Значит, последовательность $g_n(y)$ сходится к непрерывной функции $h(y)$, причем

$$h(y) = \frac{f(y) - f(x)}{y - x} \quad \text{при} \quad y \neq x, \quad h(x) = g(x).$$

Непрерывность дает равенство

$$\lim_{y \to x} h(y) = h(x),$$

т. е. $f'(x) = g(x)$. □

Заметим, что условия этой теоремы можно ослабить: предполагать вместо равномерной сходимости f_n лишь сходимость в некоторой точке $x_0 \in [a, b]$. Действительно, по теореме Лагранжа

$$|f_n(x) - f_m(x)| \leqslant |f_n(x_0) - f_m(x_0)| + |f_n'(c) - f_m'(c)|$$

и из равномерной сходимости производных и сходимости $f_n(x_0)$ следует фундаментальность f_n в $C[a, b]$.

Следствие 3.3. *Пространство $C^k[a, b]$ с нормой*

$$\|f\| = \max_{x \in [a,b]} |f(x)| + \max_{x \in [a,b]} |f'(x)| + \ldots + \max_{x \in [a,b]} |f^{(k)}(x)|$$

является банаховым пространством.

4. Открытые и замкнутые множества. Компакты.

Множество U в метрическом пространстве называется **открытым**, если для всякого элемента $a \in U$ найдется шар $B(a, r)$ такой, что $B(a, r) \subset U$. Множество называется **замкнутым**, если его дополнение является открытым множеством.

Открытый шар является открытым множеством, а замкнутый шар является замкнутым множеством.

Предложение 4.1. *Объединение любого набора открытых множеств и пересечение конечного набора открытых множеств является открытым множеством. Пересечение любого набора замкнутых множеств и объединение конечного набора замкнутых множеств является замкнутым множеством.*

Отметим, что всякое открытое множество (просто в силу определения) является объединение шаров.

Точка a называется **граничной точкой** множества A в метрическом пространстве, если во всяком шаре $B(a, r)$ есть точки множества A и точки дополнения множества A. Множество граничных точек множества A называется **границей** A и обозначается через ∂A.

Точка a называется **предельной точкой** множества A, если во всяком шаре $B(a, r)$ бесконечно много точек множества A.

Предложение 4.2. *Следующие утверждения равносильны:*

(i) *A – замкнутое множество.*

(ii) *A содержит все свои граничные точки.*

(iii) *A содержит все свои предельные точки.*

(iv) *Если $a_n \in A$ и $a_n \to a$, то $a \in A$.*

Множество $\overline{A} = A \cup \partial A$ называется замыканием множества A.

Предложение 4.3. *$\overline{A}$ является замкнутым множеством. Более того, множество A замкнуто тогда и только тогда, когда $\overline{A} = A$.*

Доказательство. Предположим, что граничная точка a множества $\overline{A}$ не является граничной точкой множества A. Тогда существует шар $B(a, r)$, который не пересекается с A. Поскольку $B(a, r)$ – открытое множество, то никакая его точка не является граничной точкой множества A. Противоречие. □

Отметим, что в метрическом пространстве замкнутый шар не всегда совпадает с замыканием открытого шара. В нормированном пространстве замкнутый шар является замыканием открытого шара.

Теорема 4.1. *В полном метрическом пространстве всякая последовательность вложенных замкнутых шаров, радиусы которых стремятся к нулю, имеет общую точку.*

Доказательство. Достаточно заметить, что центры этих шаров образуют фундаментальную последовательность, которая в силу полноты сходится. Предел этой последовательности является общей точкой. □

Множество K в метрическом пространстве называется *компактом*, если из любого его покрытия открытыми множествами можно выбрать конечное подпокрытие.

Так как всякое открытое множество является объединением шаров, то в этом определении можно было бы рассматривать лишь покрытия шарами.

Пусть (X, ϱ) – метрическое пространство. Тогда всякое его непустое подмножество $A \subset X$ также является метрическим пространством с метрикой ϱ. Заметим, что шар $B_A(x, r)$ в A является пересечением шара $B(x, r)$ в X с множеством A.

Предложение 4.4. *Подмножество K метрического пространства (X, ϱ) является компактом тогда и только тогда, когда K является компактом в метрическом пространстве (K, ϱ).*

Доказательство. Как отмечалось выше при проверке компактности достаточно рассматривать покрытия шарами. Остается заметить, что шар $B_K(x, r)$ в K является пересечением шара $B(x, r)$ в X с множеством K. □

Следовательно, при доказательстве компактности множества K можно считать, что рассматривается метрическое пространство (K, ϱ).

Предложение 4.5. *Параллелепипед $[a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$ является компактом в пространстве $\mathbb{R}^n$.*

Доказательство. Предположим противное: существует бесконечное покрытие открытыми множествами, из которого нельзя выбрать конечное подпокрытие. Разделим каждое ребро $[a_k, b_k]$ пополам. Тогда весь параллелепипед разбивается на 2^n частей. Какой-то из параллелепипедов разбиения не имеет конечного подпокрытия. Продолжая построение получаем последовательность вложенных параллелепипедов, длины ребер которых стремятся к нулю, а сами ребра образуют систему вложенных отрезков. Следовательно, существует общая точка, которая покрывается некоторым открытым множеством, а значит этим множеством покрывается какой-то из построенных параллелепипедов. Противоречие. □

Теорема 4.2. *Пусть K – компакт в метрическом пространстве.*

(i) *K – ограниченное множество.*

(ii) *K – замкнутое множество.*

(iii) *всякое замкнутое подмножество K является компактом.*

(iv) *всякое бесконечное подмножество K имеет предельную точку в K.*

Следствие 4.1. (Критерий компактности в $\mathbb{R}^n$) *Множество $K \subset \mathbb{R}^n$ является компактом тогда и только тогда, когда множество K является ограниченным и замкнутым.*

Доказательство. Достаточно заметить, что ограниченное и замкнутое множество можно поместить внутрь параллелепипеда, который сам уже является компактом. □

Предложение 4.6. *Если множество K компактно, то всякая последовательность вложенных непустых замкнутых подмножеств K имеет общую точку.*

Доказательство. Пусть $F_n \subset K$, $F_n \subset F_{n+1}$ – последовательность замкнутых множеств. Тогда $U_n = X \setminus F_n$ – открытые множества. Если у множеств F_n нет общей точки, то U_n покрывают K, но выбрать конечное подпокрытие нельзя, что противоречит компактности K. Следовательно, общая точка существует. □

Предложение 4.7. *Если из всякой последовательности элементов некоторого множества K можно выбрать сходящуюся подпоследовательность, то для всякого $\varepsilon > 0$ множество K можно покрыть конечным набором шаров радиуса ε (говорят, что K имеет конечную ε-сеть).*

Доказательство. Предположим противное: найдется $\varepsilon > 0$ такое, что K нельзя покрыть конечным набором шаров радиуса ε. Пусть $x_1 \in K$. Тогда найдется элемент $x_2 \in K$, который не лежит в $B(x_1, \varepsilon)$. Если уже построили $x_1, x_2, \ldots, x_n$, то найдем $x_{n+1} \in K$ такой, что $x_{n+1} \notin \cup_{k=1}^{n} B(x_k, \varepsilon)$. Так как попарные расстояния между элементами построенной последовательности не меньше ε, то из нее нельзя выбрать сходящуюся подпоследовательность. □

Теорема 4.3. *Пусть (K, ϱ) – метрическое пространство. Множество K компактно тогда и только тогда, когда из всякой последовательности элементов K можно выбрать подпоследовательность, сходящуюся к элементу из K.*

Доказательство. Доказательства требует лишь достаточность. Покажем сначала, что всякая последовательность непустых замкнутых вложенных подмножеств $F_n \subset K$ имеет общую точку. Выбираем в каждом F_n точку x_n. По условию найдется подпоследовательность x_{n_k}, которая сходится к некоторому $x \in K$. В силу вложенности множеств F_n каждое из них содержит все элементы этой последовательности начиная с некоторого номера. Так как F_n замкнуты, то предел x лежит в каждом F_n.

Предположим теперь, что существует покрытие открытыми множествами, из которого нельзя выбрать конечное подпокрытие. Для всякого $\varepsilon > 0$ существует конечная ε-сеть. Пусть B_k^1 – конечный набор шаров, образующих конечную 1-сеть множества K. Хотя бы один из замкнутых шаров $\overline{B}_k^1$ не имеет конечного подпокрытия. Обозначим такой шар F_1. Пусть теперь $B_k^{1/2}$ – конечный набор шаров, образующих конечную $1/2$-сеть множества $\overline{B}_1$. Хотя бы одно из множеств $F_1 \cap \overline{B}_k^{1/2}$ не имеет конечного подпокрытия. Обозначаем это множество F_2. Строим $1/3$-сеть для множества F_2 и затем выбираем замкнутый шар $\overline{B}_k^{1/3}$ такой, что $F_2 \cap \overline{B}_k^{1/3}$ не имеет конечного подпокрытия. Обозначаем это множество F_3. Продолжаем построение. Получаем последовательность вложенных непустых замкнутых множеств. Следовательно, существует общая точка всех F_n. Эта точка содержится в некотором множестве покрытия. Так как радиусы шаров, пересечение с которыми и давали F_n, стремятся к нулю, то для достаточно большого номера n множество F_n содержится в этом множестве покрытия. Противоречие. □

Если даны метрические пространства (X, ϱ_x) и (Y, ϱ_y), то можно рассмотреть метрическое пространство $(X \times Y)$ с метрикой $\varrho((x_1, y_1), (x_2, y_2)) = \sqrt{\varrho_x(x_1, x_2)^2 + \varrho_y(y_1, y_2)^2}$.

Следствие 4.2. *Если (K_x, ϱ_x) и (K_y, ϱ_y) – компактные метрические пространства, тогда $K_x \times K_y$ является компактным метрическим пространством.*

В заключение отметим, что в бесконечномерном случае замкнутости и ограниченности не хватает для компактности. Например, единичный шар в $C[a,b]$ не является компактом. Пусть f_n – непрерывная функция, которая равна нулю вне отрезка $[a+\frac{b-a}{n+2}, a+\frac{b-a}{n+1}]$ и единице в середине этого отрезка. Легко видеть, что $\|f_n - f_m\| = 2$ при $n \neq m$. Следовательно, из этой последовательности нельзя выбрать сходящуюся подпоследовательность.

Теорема 4.4. *Если в нормированном пространстве замкнутый шар положительного радиуса является компактом, то пространство конечномерно.*

Доказательство. Пусть $\overline{B}(0,2)$ – компакт. Рассмотрим покрытие $\{B(a,1)\}_{a\in B(0,2)}$ компакта $\overline{B}(0,2)$. Выбираем конечное подпокрытие $B(a_k,1)$, $k = 1,2,\ldots,N$. Пусть L – линейное пространство, натянутое на вектора $a_1, a_2, \ldots a_N$.

Сначала покажем, что $X = L + B(0,1)$. По построению $L + B(0,1) \supset B(0,2)$. Следовательно, $L + B(0,1) \supset L + B(0,2)$. Однако, $L + B(0,2) = 2(L + B(0,1))$. Значит $L + B(0,2) \supset L + B(0,4)$. Продолжая эти рассуждения, заключаем, что $L + B(0,1) \supset L + B(0,2^n)$ для всякого n. Это немедленно дает равенство $X = L + B(0,1)$.

Теперь докажем, что $X = L$. Предположим противное. Найдется шар $B(q,r) \subset X \setminus L$. Заметим, что $B(\lambda q, \lambda r) \subset X \setminus L$ для всякого $\lambda > 0$. Следовательно, $\lambda q \notin L + B(0,\lambda r)$ для всякого $\lambda > 0$. Но при $\lambda = 1/r$ имеем $L + B(0,\lambda r) = X$. Противоречие. □

5. Предел функции. Непрерывные функции. Связные пространства.

Пусть (X, ϱ_X) и (Y, ϱ_Y) – метрические пространства и a – предельная точка X. Пусть задана функция $f\colon X \setminus \{a\} \to Y$.

(Определение Гейне) Элемент $b \in Y$ называется пределом функции f при $x \to a$, если для всякой последовательности $x_n \in X$ такой, что $x_n \to a$ и $x_n \neq a$, выполняется $f(x_n) \to b$.

(Определение Коши) Элемент $b \in Y$ называется пределом функции f при $x \to a$, если для всякого $\varepsilon > 0$ найдется $\delta > 0$ такое, что из неравенств $0 < \varrho_X(x,a) < \delta$ следует неравенство $\varrho_Y(f(x), b) < \varepsilon$.

Предложение 5.1. *Определения Коши и Гейне равносильны.*

Рассмотрим важную частную ситуацию. Пусть $f\colon \mathbb{R}^2 \to \mathbb{R}$. Можно ли предел $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ заменить повторными пределами

$$\lim_{x\to a}(\lim_{y\to b} f(x,y)) \quad \text{или} \quad \lim_{y\to a}(\lim_{x\to b} f(x,y))?$$

Несложно привести примеры, показывающие что без дополнительных условий ответ отрицательный.

Предложение 5.2. *Пусть* $\lim\limits_{(x,y)\to(a,b)} f(x,y) = A$ *и для всякого* $x \neq a$ *существует предел* $\lim_{y\to b} f(x,y) = \varphi(x)$. *Тогда существует предел* $\lim_{x\to a}\varphi(x) = A$.

Доказательство. Пусть $x_n \to a$ и $x_n \neq a$. Для каждого номера n найдем такое y_n, что $0 < |y_n - b| < 1/n$ и $|f(x_n,y_n) - \varphi(x_n)| < 1/n$. Тогда $f(x_n,y_n) \to A$ и

$$|\varphi(x_n) - A| \leqslant 1/n + |f(x_n,y_n) - A| \to 0.$$

□

Функция $f\colon X \to Y$ **непрерывна** в точке $a \in X$, если для всякого $\varepsilon > 0$ найдется $\delta > 0$ такое, что из неравенства $\varrho_X(x,a) < \delta$ следует неравенство $\varrho_Y(f(x), f(a)) < \varepsilon$.

Предложение 5.3. *Следующие утверждения равносильны.*

(i) f *– непрерывна в точке* a.

(ii) $x_n \to a \Rightarrow f(x_n) \to f(a)$.

(iii) a *– изолированная точка или* a *– предельная и* $\lim_{x\to a} f(x) = f(a)$.

Предложение 5.4. *Пусть* Y *– линейное пространство и метрика* ϱ_Y *задана нормой, т. е.* $\varrho_Y(y_1, y_2) = \|y_1 - y_2\|$. *Если* $f, g\colon X \to Y$ *– непрерывны в точке* a *и* $\alpha\colon X \to \mathbb{R}$ *– непрерывна в точке* a, *то* $f + g$ *и* αf *являются непрерывными в точке* a *функциями.*

Предложение 5.5. *Пусть* X, Y, Z *– метрические пространства. Если функция* $f\colon X \to Y$ *непрерывна в точке* $a \in X$ *и функция* $g\colon Y \to Z$ *непрерывна в точке* $f(a)$, *то функция* $g(f(x))$ *непрерывна в точке* a.

Будем говорить, что $f\colon X \to Y$ непрерывное отображение из X в Y или непрерывная функция на X, если f непрерывна в каждой точке множества X.

Теорема 5.1. $f\colon X \to Y$ *непрерывное отображение тогда и только тогда, когда для всякого открытого множества* $V \subset Y$ *множество* $f^{-1}(V)$ *открыто.*

Доказательство. Докажем необходимость. Пусть $a \in f^{-1}(V)$. Найдется $\varepsilon > 0$ такое, что $B(f(a), \varepsilon) \subset V$. Существует число $\delta > 0$ такое, что $B(a, \delta) \subset f^{-1}(B(f(a), \varepsilon))$. Следовательно, $B(a, \delta) \subset f^{-1}(V)$. Для доказательства достаточности заметим, что для всякого шара $B(f(a), \varepsilon)$ множество $f^{-1}(B(f(a), \varepsilon))$ открыто, а значит точка a содержится в этом множестве с некоторой своей окрестностью. □

Следствие 5.1. *При непрерывном отображении образ компакта является компактом.*

Доказательство. Достаточно заметить, что для всякого покрытия $\{U_\alpha\}$ открытыми множествами образа компакта открытые множества $f^{-1}(U_\alpha)$ образуют покрытие самого компакта. □

Следствие 5.2. (Теорема Вейерштрасса) *Пусть* $f\colon K \to \mathbb{R}$ *непрерывная функция и* K *– компакт. Тогда* f *ограничена и достигает наибольшего и наименьшего значений.*

Доказательство. Образ $f(K)$ является компактом в $\mathbb{R}$. Следовательно, $f(K)$ – ограниченное и замкнутое множество, которое содержит граничные значения $\inf f(K)$ и $\sup f(K)$. □

Следствие 5.3. (Теорема Кантора) *Пусть* $f\colon K \to Y$ *непрерывная функция и* K *– компакт. Тогда* f *равномерно непрерывна на* K, *т. е. для всякого* $\varepsilon > 0$ *существует* $\delta > 0$ *такое, что из неравенства* $\varrho_K(x_1, x_2) < \delta$ *следует неравенство* $\varrho_Y(f(x_1), f(x_2)) < \varepsilon$.

Доказательство. Так как K – компакт, то существует конечное покрытие шарами $B(a_k, \delta_k)$ такими, что $\varrho(f(x_1), f(x_2)) < \varepsilon$ для всех $x_1, x_2 \in B(a_k, 3\delta_k)$. Число $\delta = \min_k \delta_k$ искомое. □

Рассмотрим несколько примеров непрерывных функций на метрическом пространстве. Пусть (X, ϱ) – метрическое пространство

1. Функция $\varrho(x, y)$ является непрерывной функцией на $X \times X$. Действительно, из неравенства треугольника следует, что

$$\varrho(x,y) - \varrho(u,v) = \varrho(x,y) - \varrho(u,y) + \varrho(u,y) - \varrho(u,v) \leqslant \varrho(x,u) + \varrho(y,v) \leqslant 2\sqrt{\varrho(x,u)^2 + \varrho(y,v)^2}.$$

Если $K \subset X$ – компакт, то $K \times K$ – компакт в $X \times X$. Следовательно, непрерывная функция $\varrho(x, y)$ является ограниченной на $K \times K$ и достигает точной верхней грани своих значений. Число $d(K) = \max_{K\times K} \varrho(x, y)$ называется **диаметром** множества K.

2. Пусть F – замкнутое подмножество X. Положим $d(x,F) = \inf_{y\in F} \varrho(x,y)$. Величину $d(x,F)$ называют расстоянием от x до F. Отметим, что такое расстояние не всегда достигается на каком-то элементе множества F. Проверим, что $d(x,F)$ – непрерывная функция. Для всякого z и всякого $\varepsilon > 0$ найдется $y \in F$ такой, что

$$d(x,F) - d(z,F) \leqslant \varrho(x,y) - \varrho(z,y) + \varepsilon.$$

По неравенству треугольника

$$d(x,F) - d(z,F) \leqslant \varrho(x,z) + \varepsilon.$$

Устремляя ε к нулю, получаем неравенство $d(x,F) - d(z,F) \leqslant \varrho(x,z)$.

Рассмотрим множество $K(X)$ всех непустых компактных подмножеств X. На этом множестве можно ввести метрику Хаусдорфа

$$\varrho(K_1,K_2) = \max_{x\in K_1} d(x,K_2) + \max_{x\in K_2} d(x,K_1).$$

Интересно, что если X является полным пространством, то $K(X)$ также полно. Пусть $f_1,\ldots,f_N$ – сжимающие отображения $X \to X$, т. е. $\varrho(f_k(x),f_k(y)) \leqslant q_k\varrho(x,y)$ для всех x,y и некоторого $q_k \in (0,1)$. Можно доказать, что отображение

$$F\colon K \mapsto f_1(K) \cup f_2(K) \cup \ldots \cup f_N(K)$$

является сжимающим. Известно (и даже будет нами доказано позднее), что сжимающее отображение полного пространства имеет неподвижную точку, т. е. найдется $K \in K(X)$ такое, что $F(K) = K$. Это множество называют **неоднородным самоподобным фракталом**.

Пусть $X = [0,1]$, $\varrho(x,y) = |x-y|$, $f_1(x) = x/3$ и $f_2(x) = (x+2)/3$. Тогда неподвижное множество соответствующего отображения F называется **множеством Кантора** и играет важнейшую роль в анализе.

Пусть $f\colon X \to Y$ – непрерывная биекция. Обратное отображение может и не быть непрерывным. Действительно, пусть $X = [0,1) \cup [2,3]$ и $f(x) = x$ при $x \in [0,1)$, $f(x) = x-1$ при $x \in [2,3]$. Ясно, что это биективное отображение X на $[0,2]$, но обратное отображение разрывно в точке $y = 1$.

Если же обратное отображение непрерывно, то отображение f называют **гомеоморфизмом**.

Теорема 5.2. *Если $f\colon X \to Y$ – непрерывная биекция и X – компакт, то f – гомеоморфизм.*

Доказательство. Пусть $y_n \to y$. Из последовательности $x_n = f^{-1}(y_n)$ можно выбрать сходящуюся подпоследовательность $x_{n_k} \to x$. В силу непрерывности f имеем $f(x) = y$ и значит всякая сходящаяся подпоследовательность сходится к одному и тому же. Это влечет сходимость $f^{-1}(y_n) \to f^{-1}(y)$. ☐

Метрическое пространство X называется **несвязным**, если его можно представить в виде объединения двух непустых непересекающихся открытых (или замкнутых) множеств. Метрическое пространство, которое не является несвязным называют **связным**.

Пример: отрезок $[0,1]$ с метрикой $\varrho(x,y) = |x-y|$ является связным пространством.

Предложение 5.6. *Если X – связное пространство и функция $f\colon X \to \mathbb{R}$ непрерывна, то $f(X)$ является промежутком.*

Доказательство. Пусть $A,B \in f(X)$ и $A < C < B$. Рассмотрим открытые непересекающиеся множества $U^- = f^{-1}((-\infty,C))$ и $U^+ = f^{-1}((C,+\infty))$. Если значение C не принимается, то $X = U^- \cup U^+$ и это противоречит связности X. ☐

Если пространство несвязно, то на нем легко предъявить непрерывную функцию, принимающую всего два различных значения.

Предложение 5.7. *Если $f\colon X \to Y$ – непрерывная сюрьекция и X связно, то Y связно.*

Метрическое пространство X называется **линейно связным**, если для любых двух точек $x_0, x_1 \in X$ существует непрерывное отображение $\gamma\colon [0,1] \to X$ (непрерывный путь) такое, что $\gamma(0) = x_0$ и $\gamma(1) = x_1$.

Предложение 5.8. *Если X линейно связно, то X связно. Обратное неверно.*

Доказательство. Предположим, что X несвязно. Тогда существуют два непустых открытых непересекающихся множества U_1, U_2 такие, что $X = U_1 \cup U_2$. Возьмем в каждом из этих множеств по точке и соединим их непрерывным путем γ, тогда $\gamma^{-1}(U_1)$ и $\gamma^{-1}(U_2)$ – непересекающиеся непустые открытые множества, которые в объединении дают отрезок $[0,1]$, но отрезок является связным множеством. Противоречие. □

Используя соображения связности, можно показать, что прямая не гомеоморфна плоскости. Действительно, выкидывая точку из прямой, получаем несвязное пространство, а выкидывая образ этой точки из плоскости, получаем связное пространство. Однако, такие соображения не работают в большей размерности.

6. Гомотопия. Фундаментальная группа окружности.

Пусть X и Y – метрические пространства. Два непрерывных отображения $f, g\colon X \to Y$ называются гомотопными, если существует непрерывное отображение $F\colon [0,1] \times X \to Y$ такое, что $f(x) = F(0,x)$ и $g(x) = F(1,x)$. Например, если Y – нормированное пространство, то любые два непрерывных отображения гомотопны: $F(t,x) = tg(x) + (1-t)f(x)$. На множестве непрерывных отображений из X в Y задано отношение эквивалентности: два отображения эквивалентны тогда и только тогда, когда гомотопны. Множество классов эквивалентности обозначается через $\pi(X,Y)$. Фактически $\pi(X,Y)$ – множество компонент связности пространства непрерывных отображений $C(X,Y)$. Рассмотрим еще пример. Пусть X состоит из одной точки a. Тогда отображение f из X в Y отождествляется с точкой $f(a) \in Y$, а соответствующая гомотопия F – непрерывный путь, связывающий две точки $f(a)$ и $g(a)$. Следовательно, множество $\pi(X,Y)$ – множество компонент линейной связности в пространстве Y.

Пусть $X = S_1$ – единичная окружность $x^2 + y^2 = 1$ с метрикой из $\mathbb{R}^2$ и $x_0 = (1,0)$. Зафиксируем точку $y_0 \in Y$. Рассмотрим непрерывные отображения $f\colon X \to Y$, удовлетворяющие условию $f(x_0) = y_0$. Такие непрерывные отображения называем **петлями** в Y.

Две петли f, g гомотопны, если существует непрерывное отображение $F\colon [0,1] \times X \to Y$ такое, что $F(0,x) = f(x)$, $F(1,x) = g(x)$ и $F(t,x_0) = y_0$ для всех t. Множество петель опять разбивается на классы эквивалентности. Множество классов эквивалентности обозначаем через $\pi_1(Y, y_0)$.

Параметризуем окружность: $(\cos(2\pi t), \sin(2\pi t))$, $t \in [0,1]$.

Далее считаем, что непрерывное отображение $f\colon S_1 \to Y$, $f(x_0) = y_0$, это отображение f отрезка $[0,1]$ в Y такое, что $f(0) = f(1) = y_0$.

Если даны две петли f и g, то можно построить новую петлю $f * g$ по следующему правилу: $f * g(t) = f(2t)$ при $t \in [0,1/2)$ и $f * g(t) = g(1-2t)$ при $t \in [1/2,1]$.

Если f_1 и f_2 гомотопны и g_1 и g_2 – гомотопны, то $f_1 * g_1$ и $f_2 * g_2$ гомотопны. Пусть F_1 – гомотопия f_1, f_2 и F_2 – гомотопия g_1, g_2. Положим $F(s,t) = F_1(s,2t)$ при $(s,t) \in [0,1] \times [0,1/2)$ и $F(s,t) = F_2(s,1-2t))$ при $(s,t) \in [0,1] \times [1/2,1]$. Функция F задает гомотопию петель $f_1 * g_1$ и $f_2 * g_2$. Таким образом, данная операция на петлях переносится на классы эквивалентности $\pi_1(Y, y_0)$.

Предложение 6.1. *$\pi_1(Y, y_0)$ является группой.*

Доказательство. Единица группы – класс эквивалентности петли $f \equiv y_0$. Пусть f – петля. Положим $f^{-1}(t) = f(1-t)$. Покажем, что класс эквивалентности петли $f * f^{-1}$ совпадает с классом эквивалентности петли $f \equiv y_0$. Положим $F(s,t) = f(2st)$ при $t \in [0, 1/2)$ и $F(s,t) = F(s, 1-2t)$ при $t \in [1/2, 1]$. Отображение F – искомая гомотопия. □

Эту группу называют **фундаментальной группой** пространства Y.

Предложение 6.2. *Если Y – линейно связное пространство, то для любых двух точек y_0 и y_1 группы $\pi_1(Y, y_0)$ и $\pi_1(Y, y_1)$ изоморфны.*

Далее в случае линейно связных пространств фундаментальную группу обозначаем $\pi_1(Y)$.

Предложение 6.3. *Если линейно связные пространства Y_1 и Y_2 гомеоморфны, то группы $\pi(Y_1)$ и $\pi(Y_2)$ изоморфны.*

Обратное неверно. Например, фундаментальные группы прямой и плоскости состоят только из единицы, но эти пространства не гомеоморфны.

Пусть $A \subset Y$. Непрерывное отображение $h\colon Y \to A$ такое, что $h(a) = a$ для всех $a \in A$, называется **ретракцией**.

Предложение 6.4. *Если существует ретракция $h\colon Y \to A$ и $\pi_1(Y)$ – состоит только из единицы, то $\pi_1(A)$ также состоит только из единицы.*

Доказательство. Петли f, g в A можно считать петлями в Y. Пусть F – гомотопия f, g в Y. Тогда $h(F)$ – гомотопия f, g в A. □

Теорема 6.1. *$\pi_1(S_1) \simeq \mathbb{Z}$.*

Для доказательства нам потребуется следующая лемма.

Лемма 6.1. *Для всякой петли $f\colon S_1 \to S_1$ существует единственная непрерывная функция $\varphi\colon [0,1] \to \mathbb{R}$, что $f(t) = (\cos(2\pi\varphi(t)), \sin(2\pi\varphi(t)))$, $\varphi(0) = 0$ и $\varphi(1) \in \mathbb{Z}$. Более того, если две петли гомотопны, то и соответствующие функции φ также гомотопны, в частности величина $\varphi(1)$ совпадает у гомотопных петель. Число $\varphi(1)$ называют индексом петли.*

Доказательство. Пусть $0 = t_0 < t_1 < \ldots < t_N = 1$ такое разбиение $[0,1]$, что $f([t_{k-1}, t_k])$ целиком лежит внутри одной из четырех дуг окружности: $(-\pi/2, \pi/2)$, $(0, \pi)$, $(\pi/2, 2\pi/3)$ и $(\pi, 2\pi)$. Используя $\mathrm{arctg}x$ или $\mathrm{arcctg}x$, легко найти непрерывную функцию $\varphi_k(t)$ на отрезке $[t_{k-1}, t_k]$ такую, что $f(t) = (\cos(2\pi\varphi_k(t)), \sin(2\pi\varphi_k(t)))$ на этом отрезке. Ясно, что такая функция φ_k определена с точностью до добавления целого числа. Так как в t_k имеют место равенства $\cos(2\pi\varphi_k(t_k)) = \cos(2\pi\varphi_{k-1}(t_k))$ и $\sin(2\pi\varphi_k(t_k)) = \sin(2\pi\varphi_{k-1}(t_k))$, то разность $\varphi_k(t_k) - \varphi_{k-1}(t_k)$ является целым числом. Следовательно, сдвигая каждую φ_k на подходящее целое число, можно сшить из этих функций искомую непрерывную функцию φ. □

Доказательство теоремы: сопоставляя каждому классу эквивалентности индекс какой-нибудь петли из этого класса, получаем искомый изоморфизм групп.

Следствие 6.1. (Теорема Брауэра) *Пусть $f\colon \bar{B} \to \bar{B}$ – непрерывное отображение замкнутого круга $\bar{B} \subset \mathbb{R}^2$. Тогда существует точка $x \in \bar{B}$ такая, что $f(x) = x$.*

Доказательство. Предположим, что $f(x) \neq x$. Тогда сопоставим каждой точке $x \in B$ точку на окружности $S_1 = \partial B$, которая является пересечением с окружностью луча, выходящего из $f(x)$ и проходящего через x. Это отображение является непрерывной ретракцией $\bar{B}$ на S_1, но тогда фундаментальная группа окружности тривиальна, а это не так. Противоречие. $\square$

7. Лемма Шпернера и теоремы Брауэра.

Мы уже говорили о том, что не существует гомеоморфизма между прямой и плоскостью. Это следует из соображений связности. В общем случае так просто доказать отсутствие гомеоморфизма между $\mathbb{R}^n$ и $\mathbb{R}^m$ при $n \neq m$ не получается. Здесь мы приведем рассуждения, опирающиеся на лемму Шпернера.

Пусть в $\mathbb{R}^n$ заданы $(n+1)$ точка $a_0, a_1, \ldots, a_n$ так, что вектора $a_1 - a_0, a_2 - a_0, \ldots, a_n - a_0$ линейно независимы. Выпуклая оболочка этих точек, т. е. множество всех возможных точек

$$t_0 a_0 + t_1 a_1 + \ldots + t_n a_n,$$

где $t_i \geqslant 0$ и $t_0 + t_1 + \ldots + t_n = 1$, называется n-мерным **симплексом.** На числовой прямой – это отрезок $[a_0, a_1]$; на плоскости – это треугольник с вершинами в точках a_0, a_1, a_2. В пространстве это тетраэдр. Если рассматривается выпуклая оболочка всех точек, кроме какой-то одной a_i, то такое подмножество называют $(n-1)$-мерной гранью.

Рассмотрим следующее разбиение симплекса на симплексы. Если симплекс одномерный, т. е. отрезок $[a_0, a_1]$, то разбиение состоит из двух отрезков $[a_0, M]$ и $[M, a_1]$, где M – середина отрезка $[a_0, a_1]$. Если симплекс двумерный, т. е. треугольник с вершинами в точках a_0, a_1, a_2, то разбиение состоит из треугольников, которые получились после соединения центра масс (точка пересечения медиан) с вершинами и серединами сторон треугольника. Пусть уже умеем разбивать $(n-1)$-мерные симплексы. Рассмотрим n-мерный симплекс. Можно считать, что его грани уже разбиты. Пусть M – центр масс этого симплекса и M_i – центры масс его граней. Разбиение состоит из всех возможных симплексов с вершинами в точках M, M_i и каких-то $(n-1)$ вершинах разбиения i-й грани исходного симплекса. Такое разбиение можно повторить для каждого из вновь образовавшихся симплексов. Такое подразбиение симплекса называют **барицентрическими**.

Лемма 7.1. *При итерировании барицентрического подразбиения длины ребер получающихся симплексов стремятся к нулю.*

Лемма 7.2. (Шпернер) *Пусть дано барицентричекое подразбиение симплекса и у каждой вершины стоит номер от* 1 *до* $n+1$, *причем выполняется условие, что среди вершин исходного симплекса встречаются все номера и если какая-то вершина симплекса подразбиения попала на грань исходного симплекса, то ее номер совпадает с номером какой-нибудь из вершин данной грани. Тогда найдется симплекс разбиения, среди вершин которого встречаются все номера от* 1 *до* $n+1$.

Доказательство. Если симплекс подразбиения имеет грань с полным набором номеров от 1 до n, то либо такой симплекс имеет полную нумерацию вершин, либо в нем две грани с полным набором номеров от 1 до n. Следовательно, четность числа полных симплексов подразбиения совпадает с четностью числа граней (грани считаются по каждому симплексу), имеющих полную нумерацию от 1 до n. Если такая полная грань находится внутри симплекса, то при подсчете она учитывается дважды. Значит четность числа полных граней совпадает с четностью числа полных граней, лежащих на гранях большого симплекса, который мы собственно подразбиваем. Если полная грань симплекса подразбиения попала на грань большого симплекса, то согласно условию в вершинах этой грани встречаются все числа от 1 до n. Такая грань только одна в большом симплексе. Таким образом, четность числа симплексов подразделения, имеющих полный набор номеров, совпадает с четностью числа $(n-1)$-мерных симплексов на полной грани большого симплекса, имеющих полную нумерацию. Теперь утверждение легко доказывается по индукции. □

Теорема 7.1. (Брауэр) *Непрерывное отображение симплекса в себя имеет неподвижную точку.*

Доказательство. Пусть M – центр масс симплекса. Для всякой точки X симплекса существует единственный набор чисел $x_0, \ldots, x_n$ такой, что

$$XM = x_0\overrightarrow{Xa_0} + \ldots + x_n\overrightarrow{Xa_n}, \quad x_i \geqslant 0, \quad x_0 + \ldots + x_n = 1.$$

Этот набор чисел называется барицентрическими координатами. Заметим, что равенство $x_i = 0$ равносильно тому, что X лежит на грани с вершинами $a_0, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$. Пусть $x = (x_1, \ldots, x_n)$ – некоторая точка симплекса и $f(x) = (y_1, \ldots, y_n)$ – ее образ. Положим j – минимальный номер такой, что $y_j \leqslant x_j \neq 0$. Такой номер обязательно есть, иначе приходим к противоречию с тем, что сумма всех координат равна единице. Кроме того, если точка лежит в грани $a_{i_1}, \ldots, a_{i_s}$, то только координаты $x_{i_1}, \ldots, x_{i_s}$ отличны от нуля. Следовательно, этой точке будет присвоен номер из набора $\{i_1, \ldots, i_s\}$. Следовательно, для разбиения на симплексы будет выполняться условие леммы Шпернера. Значит для всякого разбиения найдется симплекс разбиения с полной нумерацией. Измельчаем разбиение и в симплексах полной нумерации выбираем по точке. Из полученной последовательности точек выбираем сходящуюся подпоследовательность T_m к точке T. Так как длины ребер выбранных симплексов стремятся к нулю, то их вершины приближаются к T. Следовательно, точка T обладает свойством $x_i \leqslant y_i$ для всех i. Учитывая, что сумма координат равна единице, получаем $x_i = y_i$ для всех i. Неподвижная точка построена. □

Теорема 7.2. (Брауэр) *Если $\mathbb{R}^n$ и $\mathbb{R}^m$ гомеоморфны, то $n = m$.*

Для доказательства нам потребуется вспомогательное утверждение, которое еще называют леммой Лебега о покрытии.

Лемма 7.3. (Лебег) *Для всякого n-мерного симплекса можно указать такое число $\delta > 0$, что для всякого конечного набора компактных множеств F_k диаметра меньше δ, покрывающих этот симплекс, найдется точка этого симплекса, которая принадлежит не менее чем $n+1$ множеству покрытия.*

Доказательство. Пусть диаметр множеств покрытия столь мал, что каждое из этих множеств не пересекает хотя бы одну из граней симплекса. Значит для каждого F_k существует хотя-бы одна грань, которую это множество не пересекает. Занумеруем все грани симплекса. Поставим каждому F_k в соответствие номер $n(k)$ какой-нибудь грани, которую это множество не пересекает. Заметим, что если множество покрытия содержит вершину симплекса, которая лежит напротив j-й грани, то этому множеству сопоставляется номер j. Пусть S_m – объединение всех F_k, которым сопоставлен номер m. Заметим, что S_m замкнуты и $\cup_m S_m$ содержит симплекс. Если мы докажем, что у новых множеств S_m есть общая точка, то тем самым это точка будет общей точкой не менее чем $n+1$ множества F_k. Действительно, каждое F_k входит лишь в одно из S_m. Сопоставим каждой точке симплекса наименьший из номеров m таких, что $x \in S_m$. Для всякого барицентрического разбиения симплекса расстановка номеров у вершин разбиения будет удовлетворять лемме Шпернера. Следовательно, найдется симплекс с полной нумерацией. Измельчая разбиение и выбирая в симплексах полной нумерации по точке, получаем последовательность точек. Из этой последовательности выбираем сходящуюся подпоследовательность к некоторой точке T. Точка T является предельной точкой для каждого множества S_m. Ввиду замкнутости S_m эта точка принадлежит всем множествам S_m. □

Отметим, что барицентрическое разбиение дает покрытие симплекса замкнутыми множествами сколь угодно малого диаметра, причем всякая точка исходного симплекса содержится в не более чем $(n+1)$ множестве покрытия.

Вернемся к доказательству теоремы Брауэра.

Доказательство. Предположим, что существует гомеоморфизм f из $\mathbb{R}^n \to \mathbb{R}^m$. Рассмотрим симплекс Δ_n в $\mathbb{R}^n$. Тогда его образ $f(\Delta_m)$ является компактом в $\mathbb{R}^m$ и, следовательно, содержится в некотором симплексе Δ_m. Возьмем покрытие Δ_m замкнутыми множествами F_k. Тогда $f^{-1}(F_k)$ является покрытием Δ_n. В силу равномерной непрерывности f^{-1} можно уменьшая диаметр F_k сделать сколь угодно малым диаметр $f^{-1}(F_k)$. Из леммы о покрытии сразу заключаем, что $m \geqslant n$. Аналогичные рассуждения дают неравенство $n \geqslant m$. □

8. Линейные отображения

Пусть X, Y – нормированные пространства. Отображение $L\colon X \to Y$ называется **линейным оператором**, если

$$L(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 L(x_1) + \alpha_2 L(x_2)$$

для всех $\alpha_1, \alpha_2 \in \mathbb{R}$, $x_1, x_2 \in X$. Если $Y = \mathbb{R}$, то линейные отображения называют **линейными функционалами**.

Предложение 8.1. *Всякое линейное отображение $L\colon \mathbb{R}^n \to \mathbb{R}^m$ имеет вид $L(x) = Ax$, где A – матрица $n \times m$, в частности, в случае $m = 1$ имеем $L(x) = a_1x_1 + a_2x_2 + \ldots + a_nx_n$.*

Из этого утверждения немедленно следует, что всякое линейное отображение конечномерных пространств является непрерывным. В случае бесконечномерного пространства X всегда можно построить разрывный линейный функционал.

Предложение 8.2. *Следующие утверждения равносильны:*

(i) *Линейный оператор $L\colon X \to Y$ непрерывен.*

(ii) *Линейный оператор L непрерывен в нуле.*

(iii) *Существует такое число $C > 0$, что $\|L(x)\| \leqslant C\|x\|$ для всех x.*

Наименьшее из чисел C, для которых выполняется неравенство пункта (iii) называют нормой линейного оператора L, т. е.

$$\|L\| = \sup_{x \neq 0} \frac{\|L(x)\|}{\|x\|}.$$

Предложение 8.3. *Пространство линейных операторов является нормированным пространством.*

Пусть $H\colon Z \to X$ и $L\colon X \to Y$ – линейные непрерывные операторы. Композиция линейных операторов $L \circ H$ является линейным непрерывным оператором и выполняется неравенство $\|L \circ H\| \leqslant \|L\|\|H\|$.

9. Дифференцируемость функции многих переменных. Градиент и матрица Якоби.

Пусть X, Y – нормированные пространства и отображение $f\colon X \to Y$ определено в окрестности $a \in X$. Будем говорить, что f **дифференцируемо** в a, если существует такой линейный непрерывный оператор $L\colon X \to Y$, что

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

где $\alpha\colon X \to Y$ удовлетворяет условию $\lim_{h\to 0} \alpha(h) = 0$. Это определение можно переформулировать следующим образом:

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} = 0.$$

Линейный оператор L называется **дифференциалом** отображения f и обозначается через $df(h)$ или $df(a,h)$.

Теорема 9.1. *Если отображение f дифференцируемо в точке a, то f непрерывно в точке a.*

Если функция $f\colon \mathbb{R}^n \to \mathbb{R}$ дифференцируема в точке $a \in \mathbb{R}^n$, то ее дифференциал $df(a,h)$ является линейным функционалом $\mathbb{R}^n \to \mathbb{R}$ и имеет вид

$$df(h) = c_1h_1 + c_2h_1 + \ldots + c_nh_n.$$

Вектор $(c_1, c_2, \ldots, c_n)$ называется **градиентом** функции f в точке a и обозначается $\nabla f(a)$.

Предложение 9.1. *Если функция $f\colon \mathbb{R}^n \to \mathbb{R}$ дифференцируема в точке $a \in \mathbb{R}^n$ и вектор $(c_1, c_2, \ldots, c_n)$ является градиентом f, то*

$$c_k = \lim_{t\to 0} \frac{f(a+te_k) - f(a)}{t}.$$

Доказательство. Имеем

$$f(a+te_k) - f(a) = tdf(e_k) + \alpha(te_k)t.$$

Следовательно,

$$\frac{f(a+te_k) - f(a)}{t} = c_k + \alpha(te_k).$$

□

Предел

$$\lim_{t\to 0} \frac{f(a+te_k) - f(a)}{t}$$

может существовать и для не дифференцируемой функции f. Этот предел называют **частной производной** в точке a по переменной x_k и обозначают через

$$\frac{\partial f}{\partial x_k}(a).$$

Теорема 9.2. *Если все частные производные функции f существуют в окрестности точки a и непрерывны в точке a, то функция f дифференцируема в точке a.*

Доказательство. Рассмотрим случай функции двух переменных. Заметим, что

$$f(a_1+h_1, a_2+h_2) - f(a_1,a_2) = f(a_1+h_1, a_2+h_2) - f(a_1, a_2+h_2) + f(a_1, a_2+h_2) - f(a_1,a_2).$$

Применяем теорему Лагранжа к первым двум и ко вторым двум слагаемым:

$$f(a_1+h_1, a_2+h_2) - f(a_1,a_2) = \frac{\partial f}{\partial x_1}(c_1, a_2+h_2)h_1 + \frac{\partial f}{\partial x_2}(a_1, c_2)h_1,$$

где c_1 лежит между a_1 и a_1+h_1, а c_2 лежит между a_2 и a_2+h_2. Представим правую часть в следующем виде:

$$\frac{\partial f}{\partial x_1}(a_1,a_2)h_1+\frac{\partial f}{\partial x_2}(a_1,a_2)h_1+\alpha(h)\|h\|,$$

где

$$\alpha(h)=\Big(\frac{\partial f}{\partial x_1}(c_1,a_2+h_2)-\frac{\partial f}{\partial x_1}(a_1,a_2)\Big)\frac{h_1}{\|h\|}+\Big(\frac{\partial f}{\partial x_2}(a_1,c_2)-\frac{\partial f}{\partial x_2}(a_1,a_2)\Big)\frac{h_2}{\|h\|}.$$

Ясно, что $\lim_{h\to 0}\alpha(h)=0$. □

Так как $dx_k(h)=h_k$, то выражение дифференциала функции f часто записывают в таком виде:

$$df=\frac{\partial f}{\partial x_1}dx_1+\frac{\partial f}{\partial x_2}dx_2+\ldots+\frac{\partial f}{\partial x_n}dx_n.$$

Далее мы увидим, что данный вид является инвариантным при замене переменных.

Пусть отображение f из $\mathbb{R}^n$ в $\mathbb{R}^m$ дифференцируемо в точке a. Тогда дифференциал является линейным оператором $\mathbb{R}^n\to\mathbb{R}^m$ и имеет вид $df(h)=Ah$, где A – матрица $n\times m$. Матрица A называется **матрицей Якоби** отображения f в точке a и обозначается через $f'(a)$ или $Df(a)$.

Всякое отображение $f\colon\mathbb{R}^n\to\mathbb{R}^m$ имеет вид

$$f(x)=\begin{pmatrix} f_1(x)\\ f_2(x)\\ \ldots\\ f_m(x)\end{pmatrix},$$

где $f_i\colon\mathbb{R}^n\to\mathbb{R}$.

Предложение 9.2. *Отображение f дифференцируемо в точке a тогда и только тогда, когда в a дифференцируема каждая функция f_i. Если f дифференцируемо в точке a, то элементы матрицы Якоби имеют вид*

$$\frac{\partial f_i}{\partial x_j}(a).$$

Доказательство. Равенство

$$\lim_{h\to 0}\frac{\|f(a+h)-f(a)-L(h)\|}{\|h\|}=0$$

равносильно системе равенств

$$\lim_{h\to 0}\frac{|f_i(a+h)-f_i(a)-L_i(h)|}{\|h\|}=0,\quad i=1,2,\ldots,m.$$

Здесь $L_i(h)$ – i-я координата вектора $L(h)$, т. е. $a_{i1}h_1+a_{i2}h_2+\ldots+a_{in}h_n$, где a_{ij} – элементы матрицы Якоби отображения f. Остается заметить, что

$$L_i(h)=a_{i1}h_1+a_{i2}h_2+\ldots+a_{in}h_n,$$

а выше было показано, что

$$a_{ij}=\frac{\partial f_i}{\partial x_j},\quad j=1,2,\ldots,n.$$

□

Таким образом, матрица Якоби отображения f – матрица, у которой строки являются градиентами функций f_i.

10. Правила дифференцирования. Дифференцирование сложной функции. Производная вдоль векторного поля.

Теорема 10.1. *Пусть $f, g\colon X \to Y$ и $\alpha\colon X \to \mathbb{R}$ – дифференцируемые отображения в точке a, тогда*

(i) *$f+g$ дифференцируемо в точке a и $d(f+g) = df + dg$;*

(ii) *αf дифференцируемо в точке a и $d(\alpha f) = \alpha\, df + f\, d\alpha$;*

(iii) *если $\alpha(a) \neq 0$, то $\frac{f}{\alpha}$ дифференцируемо в точке a и*

$$d\Big(\frac{f}{\alpha}\Big) = \frac{\alpha df + f\, d\alpha}{\alpha^2}.$$

Доказательство. Докажем только второе утверждение. По условию

$$\alpha(a+h) - \alpha(a) = d\alpha(h) + \gamma_1(h)\|h\|, \quad f(a+h) - f(a) = df(h) + \gamma_2(h)\|h\|,$$

где $\lim_{h\to 0}\gamma_1(h) = 0$ и $\lim_{h\to 0}\gamma_2(h) = 0$. Получаем

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

где

$$r(h) = (\alpha(a+h) - \alpha(a))df(h)\|h\|^{-1} + \alpha(a+h)\gamma_2(h) + \gamma_1(h)f(a).$$

Заметим, что в силу непрерывности линейного оператора $df(h)$ существует число $C > 0$ такое, что $\|df(h)\| \leqslant C\|h\|$. Следовательно, $\lim_{h\to 0} r(h) = 0$. □

Теорема 10.2. *Пусть $g\colon Z \to X$ – дифференцируемо в точке a и $f\colon X \to Y$ – дифференцируемо в точке $g(a)$. Тогда $f\circ g$ дифференцируема в точке a и $d(f\circ g) = df \circ dg$.*

Доказательство. По условию

$$g(a+h) - g(a) = dg(h) + \gamma_1(h)\|h\|, \quad f(g(a)+p) - f(g(a)) = df(p) + \gamma_2(p)\|p\|,$$

где $\lim_{h\to 0}\gamma_1(h) = 0$ и $\lim_{h\to 0}\gamma_2(h) = 0$. Определим отображение γ_2 в нуле нулем. Теперь γ_2 непрерывно в нуле. Имеем

$$f(g(a+h)) - f(g(a)) = df(g(a+h)-g(a)) + \gamma_2(g(a+h)-g(a))\|g(a+h)-g(a)\| == df(dg(h)) + r(h)\|h\|,$$

где

$$r(h) = df(\gamma_1(h)) + \gamma_2(g(a+h) - g(a))\|g(a+h) - g(a)\|\|h\|^{-1}.$$

Существуют такие числа C_1 и C_2, что $\|dg(h)\| \leqslant C_1\|h\|$ и $\|df(p)\| \leqslant C_2\|p\|$. Тогда

$$\|df(\gamma_1(h))\| \leqslant C_2\|\gamma_1(h)\| \to 0,$$

$$\|\gamma_2(g(a+h) - g(a))\|\|g(a+h) - g(a)\|\|h\|^{-1} \leqslant (C_1 + \|\gamma_1(h)\|)\|\gamma_2(g(a+h) - g(a))\| \to 0$$

при $h \to 0$. Получаем равенство $\lim_{h\to 0} r(h) = 0$. □

Пусть $g\colon \mathbb{R}^n \to \mathbb{R}^m$ и $f\colon \mathbb{R}^m \to \mathbb{R}^k$ – дифференцируемые отображения. Тогда их композиция дифференцируема и дифференциал композиции является композицией дифференциалов. Следовательно, матрица Якоби композиции является произведением матриц Якоби.

Пусть $n = 1$ и $k = 1$, т. е. композиция является функцией вида $f(g_1(t), g_2(t), \ldots, g_m(t))$.

Матрица Якоби отображения g:

$$\begin{pmatrix} \frac{d}{dt}g_1 \\ \frac{d}{dt}g_2 \\ \ldots \\ \frac{d}{dt}g_m \end{pmatrix},$$

матрица Якоби отображения f:

$$\Big(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m}\Big).$$

Следовательно, имеем

$$\frac{d}{dt} f\big(g_1(t), g_2(t), \ldots, g_m(t)\big) =$$

$$= \Big(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_m}\Big) \begin{pmatrix} \frac{d}{dt} g_1 \\ \frac{d}{dt} g_2 \\ \ldots \\ \frac{d}{dt} g_m \end{pmatrix} =$$

$$= \frac{\partial f}{\partial x_1}\frac{d}{dt}g_1 + \frac{\partial f}{\partial x_2}\frac{d}{dt}g_2 + \ldots + \frac{\partial f}{\partial x_m}\frac{d}{dt}g_m.$$

Предел

$$\lim_{t \to 0} \frac{f(a+vt) - f(a)}{t}$$

называется **производной вдоль вектора** v и обозначается через $\frac{\partial f}{\partial v}$.

Предложение 10.1. *Пусть функция* $f\colon \mathbb{R}^n \to \mathbb{R}$ *дифференцируема в точке* a. *Тогда производная вдоль всякого вектора* v *существует и равна*

$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x_1}v_1 + \frac{\partial f}{\partial x_2}v_2 + \ldots + \frac{\partial f}{\partial x_n}v_n = \langle \nabla f, v\rangle.$$

Более того,

$$\max_{\|v\|=1} \frac{\partial f}{\partial v} \quad \text{и} \quad \min_{\|v\|=1} \frac{\partial f}{\partial v}$$

достигаются соответственно на векторах $\frac{\nabla f}{\|\nabla f\|}$ *и* $-\frac{\nabla f}{\|\nabla f\|}$.

Доказательство. Первое утверждение немедленно следует из правила дифференцирования сложной функции, а второе утверждение из свойств скалярного произведения. □

Теперь рассмотрим более общую ситуацию, когда только $k = 1$. Матрица Якоби композиции получается умножением строки

$$\Big(\frac{\partial f}{\partial y_1}, \frac{\partial f}{\partial y_2}, \ldots, \frac{\partial f}{\partial y_m}\Big)$$

на матрицу

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \ldots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \ldots & \frac{\partial g_2}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \ldots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}.$$

Выпишем элемент матрицы Якоби композиции (в данном случае элементы вектора градиента):

$$\frac{\partial f(g(x))}{\partial x_i} = \frac{\partial f}{\partial y_1}\frac{\partial g_1}{\partial x_i} + \frac{\partial f}{\partial y_2}\frac{\partial g_2}{\partial x_i} + \ldots + \frac{\partial f}{\partial y_m}\frac{\partial g_m}{\partial x_i}.$$

Заметим, что этот результат можно формально получить из выражения

$$df = \frac{\partial f}{\partial y_1}dy_1 + \frac{\partial f}{\partial y_2}dy_2 + \ldots + \frac{\partial f}{\partial y_m}dy_m,$$

подставляя вместо y_k функцию g_k и вычисляя дифференциалы dg_k. Это наблюдение называют **инвариантностью первого дифференциала**.

Теорема 10.3. *Пусть $f\colon X \to Y$, U – окрестность точки a в X, V – окрестность точки $f(a)$ в Y. Предположим, что $f\colon U \to V$ гомеоморфизм, f дифференцируемо в точке a и df имеет непрерывный обратный оператор. Тогда $f^{-1}\colon V \to U$ дифференцируема в точке $f(a)$ и $df^{-1} = (df)^{-1}$.*

Доказательство. Найдем предел
$$\lim_{p\to 0} \frac{\|f^{-1}(f(a)+p) - f^{-1}(f(a)) - (df)^{-1}(p)\|}{\|p\|}.$$
Сделаем замену переменных $h = f^{-1}(f(a)+p) - a$. Ясно, что $h \to 0$ при $p \to 0$. Имеем
$$p = f(a+h) - f(a) = df(h) + \gamma(h)\|h\|, \quad \lim_{h\to 0}\gamma(h) = 0.$$
Получаем
$$\frac{\|f^{-1}(f(a)+p) - f^{-1}(f(a)) - (df)^{-1}(p)\|}{\|p\|} == \frac{\|h\|\|(df)^{-1}(\gamma(h))\|}{\|df(h)+\gamma(h)\|h\|\|} \leqslant \frac{\|h\|\|(df)^{-1}(\gamma(h))\|}{\|df(h)\| - \|\gamma(h)\|\|h\|\|}.$$
Существует такое число $C > 0$, что $\|df(h)\| \geqslant C\|h\|$. Следовательно,
$$\frac{\|f^{-1}(f(a)+p) - f^{-1}(f(a)) - (df)^{-1}(p)\|}{\|p\|} \leqslant \frac{\|(df)^{-1}(\gamma(h))\|}{C - \|\gamma(h)\|} \to 0, \quad h \to 0.$$
□

Отметим, что матрица Якоби обратного отображения является обратной матрицей к матрице Якоби прямого отображения.

В заключение приведем еще одно полезное утверждение.

Предложение 10.2. *Пусть отображение $f\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируемо в шаре $B(a,r)$ и все частные производные $\frac{\partial f_i}{\partial x_j}$ ограниченны на этом шаре. Положим*
$$C = \sup_{x\in B(a,r)} \|df(x)\|.$$
Тогда для всяких $x, y \in B(a,r)$ выполняется неравенство:
$$\|f(x) - f(y)\| \leqslant C|x-y|.$$

Доказательство. Пусть $m = 1$. Рассмотрим функцию $\varphi(t) = f(y + t(x-y)$. По теореме Лагранжа
$$f(x) - f(y) = \varphi(1) - \varphi(0) = \varphi'(c) = \langle \nabla f(y + c(x-y)), x-y\rangle.$$
□

11. Теорема об обратной функции и теорема о неявных функциях.

Рассмотрим уравнение $x = f(y)$. Здесь $x, y \in \mathbb{R}^1$. Пусть существует такая точка (x_0, y_0), что $x_0 = f(y_0)$. Возникает естественный вопрос: можно ли в некоторой окрестности этой точки выразить y через x, а точнее существуют ли такие отрезки $I_x = [x_0 - \alpha, x_0 + \alpha]$ и $I_y = [y_0 - \beta, y_0 + \beta]$, что функция $f\colon I_y \to I_x$ обратима, т. е. определена обратная функция $f^{-1}\colon I_x \to I_y$? В одномерном случае мы уже исследовали этот вопрос и знаем, что для существования непрерывной обратной достаточно непрерывности и строгой монотонности функции f. Если $f'(y_0) > 0$, то в силу непрерывности f' найдется такое $\beta > 0$, что на отрезке $I_y = [y_0 - \beta, y_0 + \beta]$ функция f строго монотонна и значит на отрезке $I_x = f(I_y)$ определена непрерывно дифференцируемая обратная функция f^{-1}.

Соображения монотонности хорошо работают в размерности один, но плохо переносятся на многомерную ситуацию. Поэтому, оставаясь пока в размерности один, рассмотрим другой метод построения обратной функции.

Первое, что мы хотим сделать – доказать, что для x достаточно близких к x_0 уравнение $x = f(y)$ имеет решение. Перепишем это уравнение в новом виде:

$$y = y + q(x - f(y)),$$

где q некоторое отличное от нуля число, которое мы укажем позднее. Ясно, что это уравнение имеет решение в точности тогда, когда исходное уравнение имело решение. Положим $G_x(y) = y + q(x - f(y))$. При каждом фиксированном x это функция от y и мы ищем такое y, что $y = G_x(y)$, т. е. ищем неподвижную точку. Нам потребуется следующее важное утверждение.

Теорема 11.1. (Банах) *Пусть (X, ϱ) – полное метрическое пространство и $f\colon X \to X$ является сжимающим отображением, т. е. выполняется неравенство*

$$\varrho(f(x), f(y)) \leqslant \lambda\varrho(x, y), \quad 0 < \lambda < 1.$$

Тогда существует единственная точка x такая, что $f(x) = x$.

Доказательство. Пусть x_0 – произвольная точка X. Положим $x_{n+1} = f(x_n)$. Так как

$$\varrho(x_{n+1}, x_n) \leqslant \lambda\varrho(x_n, x_{n-1}) \leqslant \ldots \leqslant \lambda^n\varrho(x_1, x_0),$$

то для $n > m$

$$\varrho(x_n, x_m) \leqslant \frac{\lambda^m}{1-\lambda}\varrho(x_1, x_0).$$

Следовательно, x_n является фундаментальной последовательностью и имеет предел. Пусть $x_n \to x$. В силу непрерывности f заключаем, что $f(x) = x$. Предположим, что кроме x есть другая неподвижная точка y. Тогда

$$\varrho(x, y) = \varrho(f(x), f(y)) \leqslant \lambda\varrho(x, y)$$

и, следовательно, $x = y$. □

Попробуем так выбрать отрезок $I_y = [y_0 - \beta, y_0 + \beta]$ и отрезок $I_x = [x_0 - \alpha, x_0 + \alpha]$, что для всякого $x \in I_x$ отображение $G_x\colon I_y \to I_y$ является сжимающим. По теореме Лагранжа

$$|G_x(y) - G_x(z)| = |G'_x(c)||y - z|,$$

где

$$G'_x(c) = 1 - qf'(c).$$

Возьмем $q = 1/f'(y_0)$. Тогда $G'_x(y_0) = 0$. По условию f' непрерывная функция. Выбирая β достаточно малым, можно считать, что на $I_x \times I_y$ выполняется неравенство $|G'_x| \leqslant 1/2$. Теперь осталось сделать так, что значения G_x лежат в I_y. Заметим, что $G_{x_0}(y_0) = y_0$. Имеем

$$|G_x(y) - y_0| \leqslant |G_x(y) - G_x(y_0)| + |G_x(y_0) - G_{x_0}(y_0)| \leqslant \frac{1}{2}|y - y_0| + |q||x - x_0| \leqslant \frac{1}{2}\beta + |q|\alpha.$$

Выбираем α столь малым, что $|q|\alpha < \beta/2$. Итак, мы выбрали I_y и I_x так, что $G_x\colon I_y \to I_y$ является сжимающим отображением для всякого $x \in I_x$. По теореме Банаха существует единственное $y \in I_y$ такое, что $G_x(y) = y$. Следовательно, построена обратная функция $y = f^{-1}(x)$. Однако, нам еще надо обосновать непрерывность и дифференцируемость этой функции. В этом нам поможет следующее утверждение.

Предложение 11.1. *Пусть (X, ϱ) – полное метрическое пространство и $f, g\colon X \to X$ – сжимающие отображения с коэффициентом сжатия $\lambda \in (0, 1)$. По теореме Банаха существуют единственные неподвижные точки $x = f(x)$ и $y = g(y)$. Тогда имеет место оценка*

$$\varrho(x, y) \leqslant (1 - \lambda)^{-1}\varrho(f(x), g(x)).$$

Доказательство. Имеем

$$\varrho(x,y) = \varrho(f(x), g(y)) \leqslant \varrho(f(x), g(x)) + \lambda\varrho(x,y).$$

Следовательно,

$$\varrho(x,y) \leqslant (1-\lambda)^{-1}\varrho(f(x), g(x)).$$

□

Заметим, что

$$|G_x(y) - G_z(y)| = |q||x-z|.$$

Следовательно для соответствующих неподвижных точек $y(x)$ и $y(z)$ выполняется оценка

$$|y(x) - y(z)| \leqslant (1-\lambda)^{-1}|q||x-z|.$$

Таким образом, обратная функция является не просто непрерывной, а даже липшицевой.

Осталось показать, что обратная функция дифференцируема, но это следует из доказанной выше теоремы о дифференцируемости обратной функции.

Теперь заметим, что эти рассуждения практически без изменений переносятся на многомерный случай.

Теорема 11.2. (об обратной функции) *Пусть $f\colon \mathbb{R}^n \to \mathbb{R}^n$ – непрерывно дифференцируема в окрестности точки y_0 и $f(y_0) = x_0$. Предположим, что матрица Якоби $f'(y_0)$ обратима. Тогда существуют такие окрестности $U(x_0)$ и $V(y_0)$, что $f\colon V(y_0) \to U(x_0)$ обратима и обратное отображение f^{-1} дифференцируемо в каждой точке $U(x_0)$.*

Доказательство. Действуем также как и выше. Рассмотрим отображение

$$G_x(y) = y + Q(x - f(y))$$

, где $Q = f'(y_0)^{-1}$. Построим такие шары $B_x = \{x\colon \|x - x_0\| \leqslant \alpha\}$ и $B_y = \{y\colon \|y - y_0\| \leqslant \beta\}$, что для всякого $x \in B_x$ отображение $G_x\colon B_y \to B_y$ является сжимающим. Выбираем такое $\beta > 0$, что на $B(y_0, \beta)$ выполняется $\|G'_x(y)\| \leqslant 1/2$. Следовательно, выполняется неравенство

$$\|G_x(y) - G_x(z)\| \leqslant \frac{1}{2}\|y - z\|.$$

Теперь сделаем так, чтобы G_x принимало значения в B_y. Имеем

$$\|G_x(y) - y_0\| \leqslant \frac{1}{2}\|y - y_0\| + \|Q\|\|x - x_0\| \leqslant \frac{1}{2}\beta + \|Q\|\alpha.$$

Выбираем α так, что $\|Q\|\alpha < \beta/2$. Применяем теорему Банаха и следующее за ней предложение. Получаем искомое обратное отображение. □

Взаимно однозначное дифференцируемое отображение, которое имеет дифференцируемое обратное, называется **диффеоморфизмом**.

Следствие 11.1. (Теорема о неявных функциях) *Пусть отображение $F\colon \mathbb{R}^n_x \times \mathbb{R}^m_y \to \mathbb{R}^m$ дифференцируемо в окрестности точки (x_0, y_0), матрица Якоби $D_yF(x_0, y_0)$ обратима и $F(x_0, y_0) = 0$. Тогда существуют такие окрестности $U(x_0)$ и $V(y_0)$, непрерывно дифференцируемое отображение $f\colon U(x_0) \to V(y_0)$ такие, что на $U(x_0) \times V(y_0)$*

$$F(x,y) = 0 \Leftrightarrow y = f(x).$$

Доказательство. Рассмотрим отображение $H\colon \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ заданное следующим образом: $H^i(x,y) = x_i$ при $i = 1, 2, \ldots, n$ и $H^i(x,y) = F^{i-n}(x,y)$ при $i = n+1, \ldots, n+m$. Заметим, что $H(x_0, y_0) = (x_0, 0)$ и матрица Якоби отображения H имеет вид:

$$\begin{pmatrix} E & 0 \\ * & D_yF \end{pmatrix}$$

Так как матрица $D_yF(x_0, y_0)$ обратима, то и матрица Якоби $H'(x_0, y_0)$ обратима. По теореме об обратной функции существует обратная функция $(x, y) = H^{-1}(z)$. При $z = (x, 0)$ получаем функции $y_i = (H^{-1})^i(x, 0)$ такие, что $F(x, y(x)) = 0$. $\square$

12. Гладкая поверхность в $\mathbb{R}^n$ и касательное пространство.

Множество M_k в $\mathbb{R}^n$ называется **гладкой k-мерной поверхностью**, если для всякой точки $p \in M_k$ существует окрестность $U(p)$ точки p в $\mathbb{R}^n$, открытое непустое множество $V \subset \mathbb{R}^k$ и непрерывно дифференцируемое **инъективное** отображение $f\colon V \to \mathbb{R}^n$ такие, что $f(V) = M_k \bigcap U(p)$ и ранг матрицы Якоби $f'(u)$ в каждой точке $u \in V$ равен k.

Предложение 12.1. *Следующие утверждения равносильны:*

(i) *M_k – гладкая k-мерная поверхность в $\mathbb{R}^n$;*

(ii) *для всякой точки $p \in M_k$ существует окрестность $U(p)$ и непрерывно дифференцируемое отображение $F\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ такое, что ранг матрицы Якоби $F'(p)$ равен $(n-k)$ и M_k совпадает с множеством $F(x) = 0$ в $U(p)$.*

Доказательство. Немедленно следует из теоремы о неявных функциях. $\square$

Пусть M_k – гладкая k-мерная поверхность в $\mathbb{R}^n$. Пусть $\gamma\colon (-1, 1) \to \mathbb{R}^n$ – гладкая кривая такая, что $\gamma(t) \in M_k$ и $\gamma(0) = p$. Вектор скорости $\dot\gamma(0)$ называется касательным вектором к поверхности M_k в точке p.

Предложение 12.2. *Множество касательных векторов является k-мерным линейным пространством.*

Доказательство. Пусть поверхность M_k локально задана отображениями $x_i = f_i(u_1, \ldots, u_k)$. Тогда всякая кривая $\gamma(t)$ на M_k задается так: $x_i(t) = f_i(u_1(t), \ldots, u_k(t))$. Найдем $\dot\gamma(0)$:

$$\dot\gamma(0) = \dot u_1(0)\frac{\partial f}{\partial u_1}(0) + \ldots + \dot u_k(0)\frac{\partial f}{\partial u_k}(0).$$

Таким образом, всякий касательный вектор является линейной комбинацией линейно независимых касательных векторов

$$\frac{\partial f}{\partial u_1}, \quad \frac{\partial f}{\partial u_2}, \quad \ldots \quad \frac{\partial f}{\partial u_k}.$$

Легко проверить, что всякая линейная комбинация этих векторов является касательным вектором. $\square$

Линейное пространство касательных векторов в точке p называется касательным пространством к M_k в точке p и обозначается через T_pM_k.

Предложение 12.3. *Пусть для точки $p \in M_k$ существует окрестность $U(p)$ и непрерывно дифференцируемое отображение $F\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ такое, что ранг матрицы Якоби $F'(p)$ равен $(n-k)$ и M_k совпадает с множеством $F(x) = 0$ в $U(p)$. Тогда*

$$\xi \in T_pM_k \Leftrightarrow \begin{cases} \langle \nabla F_1, \xi \rangle = 0 \\ \langle \nabla F_2, \xi \rangle = 0 \\ \ldots \\ \langle \nabla F_{n-k}, \xi \rangle = 0 \end{cases}$$

Доказательство. Необходимость этих условий немедленно вытекает из дифференцирования равенств $F_i(\gamma(t)) = 0$. Следовательно, касательное пространство является подпространством решений этой однородной системы. Размерность касательного пространства равна k и размерность решений этой системы равна k. Значит эти пространства совпадают. $\square$

В частности, если поверхность является множеством уровня $F(x) = const$, то ее касательное пространство ортогонально вектору градиента ∇F.

13. Производные высокого порядка. Равенство повторных производных.

Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$. Предположим, что у f в каждой точке некоторой окрестности $U(a)$ есть частная производная $\frac{\partial f}{\partial x_i}$ и функция $x \mapsto \frac{\partial f}{\partial x_i}(x)$ в точке a имеет частную производную по x_j. Тогда выражение
$$\frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}(x)\Big)\Big|_{x=a}$$
называется **частной производной второго порядка** по переменным x_i и x_j и обозначается через
$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$
Аналогично определяются **частные производные высокого порядка**:
$$\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_m}}(x) := \frac{\partial}{\partial x_{i_1}}\Big(\frac{\partial}{\partial x_{i_2}}\Big(\dots\Big(\frac{\partial f}{\partial x_{i_m}}(x)\Big)\Big)\Big).$$

Предположим теперь, что f дифференцируема в каждой точке некоторой окрестности $U(a)$ точки a. Тогда в этой окрестности определены новые функции (частные производные функции f)
$$x \mapsto \frac{\partial f}{\partial x_i}(x), \quad i = 1, 2, \dots n.$$
Если все частные производные дифференцируемы в точке a, то говорят, что функция f **дважды дифференцируема** в точке a.

Рассмотрим другой способ определить, что значит функция дважды дифференцируема в точке a. Пусть функция f дифференцируема в каждой точке некоторой окрестности $U(a)$ точки a, тогда задано отображение $x \mapsto df(x, \cdot)$ из $U(a)$ в нормированное пространство линейных функционалов. Будем говорить, что функция f дважды дифференцируема в точке a, если отображение $x \mapsto df(x, \cdot)$ дифференцируемо в точке a.

Предложение 13.1. *Оба определения дважды дифференцируемой функции равносильны.*

Доказательство. Предположим, что выполнено второе определение. Рассмотрим отображение L_i, которое сопоставляет каждому линейному функционалу l на $\mathbb{R}^n$ значение $l(e_i)$. Поскольку L_i – линейное отображение, то оно дифференцируемо. Заметим, что отображение $x \mapsto \frac{\partial}{\partial x_i}(x)$ является композицией отображения $x \mapsto df(x, \cdot)$ и отображения L_i, а значит является дифференцируемым отображением.

Теперь предположим, что выполняется первое определение. Отображение $x \to \frac{\partial f}{\partial x_i}(x)h_i$ дифференцируемо как произведение дифференцируемой функции $\frac{\partial f}{\partial x_i}(x)$ на дифференцируемую (даже постоянную) функцию $g(x) \equiv l_i(h)$. Остается заметить, что $df(x, h)$ является суммой дифференцируемых отображений. $\square$

Найдем дифференциал отображения $x \mapsto df(x, \cdot)$. Имеем
$$df(a+h, v) - df(a, v) = \sum_{i=1}^{n}\Big(\frac{\partial f}{\partial x_i}(a+h) - \frac{\partial f}{\partial x_i}(a)\Big)v_i = \sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_j \partial x_i}(a)v_i h_j + o_{h\to 0}(\|h\|).$$
Таким образом, дифференциал данного отображения – линейны оператор, сопоставляющий каждому вектору h линейный функционал
$$l_h(v) = \sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_j \partial x_i}(a)v_i h_j.$$

Этот линейный оператор полностью определяется билинейной формой:

$$B(h,v)=\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_j\partial x_i}v_ih_j.$$

Именно эту билинейную форму часто называют вторым дифференциалом функции f. Далее мы увидим, что если функция дважды дифференцируема в точке a, то порядок вычисления вторых производных не имеет значения. Следовательно, данная билинейная форма симметрична и полностью определяется своей квадратичной формой:

$$B(h,h)=\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_j\partial x_i}h_ih_j$$

и именно это выражение мы далее называем вторым дифференциалом функции f и обозначаем $d^2f(a,h)$ или $d^2f(h)$.

Далее при определении дифференцируемости более высокого порядка мы будем следовать первому способу определения через частные производные.

Пусть уже определено, что такое m-раз дифференцируемая функция. Предположим, что функция f m раз дифференцируема в каждой точке некоторой окрестности $U(a)$. Если все ее частные производные m-го порядка

$$\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_m}}$$

являются дифференцируемыми функциями в точке a, то функция f называется $(m+1)$ раз дифференцируемой в точке a.

Дифференциал m-го порядка:

$$d^mf(a,h)=\sum_{1\leqslant i_1,i_2,\dots i_m\leqslant n}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_m}}(a)h_{i_1}h_{i_2}\dots h_{i_m}.$$

Вообще говоря значение частной производной зависит от порядка дифференцирования, но если функция f дифференцируема соответствующее число раз, то порядок не важен.

Теорема 13.1. (Юнг) *Если $f\colon\mathbb{R}^2\to\mathbb{R}$ дважды дифференцируема в точке a, то*

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

Доказательство. Будем считать, что $a=(0,0)$. Составим разностное соотношение соответствующее частной производной второго порядка по x и y:

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

Заметим, что по теореме Лагранжа, примененной к функции $g(u)=f(t,u)-f(0,u)$, имеем

$$F(t,t)=g(t)-g(0)=g'(c)t=\Big(\frac{\partial f}{\partial y}(t,c)-\frac{\partial f}{\partial y}(0,c)\Big)t,$$

где c лежит между 0 и t. Так как f дважды дифференцируема в точке $(0,0)$, то

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

Следовательно,

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

Таким образом, имеем

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

Аналогично можно получить равенство:

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

Приравниваем правые части двух последних равенств и сокращаем на t^2. Устремляя $t \to 0$, получаем равенство

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

□

Похожими рассуждениями можно получить аналогичный результат без предположения о том, что функция дважды дифференцируема.

Теорема 13.2. (Шварц) *Если у функции $f\colon \mathbb{R}^2 \to \mathbb{R}$ в некоторой окрестности $U(a)$ существуют смешанные частные производные второго порядка по переменным x и y, которые непрерывны в точке a, то*

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

14. Формула Тейлора. Локальный экстремум.

Пусть в окрестности точки a задана $(n+1)$-раз дифференцируемая функция $f\colon \mathbb{R}^n \to \mathbb{R}$. Рассмотрим функцию $g(t) = f(a+th)$. По индукции легко показать, что

$$g^{(k)}(t) = \sum_{i_1,\ldots,i_k} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(a+th)h_{i_1} \cdots h_{i_k} = d^k f(a+th,h).$$

Выписывая для функции g формулу Тейлора в точке $t = 0$ с остаточным членом в форме Лагранжа получаем следующее утверждение.

Теорема 14.1. *Для функции f выполняется равенство:*

$$f(a+h) = f(a) + df(a,h) + \frac{d^2 f(a,h)}{2!} + \ldots + \frac{d^n f(a,h)}{n!} + \frac{d^{n+1} f(a+ch,h)}{(n+1)!}.$$

Следствие 14.1. *Пусть частные производные $(n+1)$-го порядка непрерывны в точке a. Тогда*

$$f(a+h) = f(a) + df(a,h) + \frac{d^2 f(a,h)}{2!} + \ldots + \frac{d^n f(a,h)}{n!} + \frac{d^{n+1} f(a,h)}{(n+1)!} + o(\|h\|^{n+1}).$$

Доказательство. Заметим, что $\frac{h_i}{\|h\|} \leqslant 1$ и

$$\frac{\partial^{n+1} f}{\partial x_{i_1} \ldots \partial x_{i_{n+1}}}(a) - \frac{\partial^{n+1} f}{\partial x_{i_1} \ldots \partial x_{i_{n+1}}}(a+ch) \to 0$$

при $h \to 0$. Следовательно,

$$d^{n+1} f(a+ch,h) - d^{n+1} f(a,h) = \Big(d^{n+1} f(a+ch, \frac{h}{\|h\|}) - d^{n+1} f(a, \frac{h}{\|h\|})\Big)\|h\|^{n+1} \to 0.$$

□

Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$ и существует такая окрестность $U(a)$ точки a, что f определена в этой окрестности и $f(x) \leqslant f(a)$ $(f(x) < f(a))$ для всякого $x \in U'(a)$, то точка a называется **точкой локального максимума (строгого локального максимума)** функции f. Аналогично определяется точка локального минимума (строгого локального минимума).

Теорема 14.2. *Если функция f дифференцируема в точке своего локального экстремума, то в этой точке $df = 0$.*

Теорема 14.3. *Пусть f дважды дифференцируема в окрестности точки a и частные производные второго порядка непрерывны в точке a. Предположим, что $df(a,h) \equiv 0$. Тогда*

(i) *если $d^2f(a,h) > 0$ для всякого $h \neq 0$, то a – точка строгого локального минимума;*

(ii) *если $d^2f(a,h) < 0$ для всякого $h \neq 0$, то a – точка строгого локального максимума;*

(iii) *если существуют такие h и v, что $d^2f(a,h) > 0$ и $d^2f(a,v) < 0$, то точка a не является точкой локального экстремума.*

Доказательство. По формуле Тейлора

$$f(a+h) - f(a) = (2^{-1}d^2f(a, \frac{h}{\|h\|}) - \alpha(h))\|h\|^2, \quad \lim_{h\to 0}\alpha(h) = 0.$$

Так как функция $v \mapsto d^2f(a,v)$ непрерывна, а единичная сфера в $\mathbb{R}^n$ компактна, то по теореме Вейерштрасса эта функция достигает своего наибольшего и наименьшего значения. Пусть $m = \min_{\|v\|=1} d^2f(a,v) > 0$. Найдем такую окрестность $U(0)$, что $|\alpha(h)| < m/4$ для всех $h \in U'(0)$. Тогда для всех $h \in U'(0)$ верно неравенство $f(a+h) - f(a) > 0$. Аналогично разбираем случай, когда $M = \max_{\|v\|=1} d^2f(a,v) < 0$. Пусть теперь существуют такие h и v, что $d^2f(a,h) > 0$ и $d^2f(a,v) < 0$. Функция $\varphi(t) = f(a+th)$ удовлетворяет следующим условиям: $\varphi'(0) = 0$ и $\varphi''(0) = d^2f(a,h) > 0$. Следовательно, у этой функции точка 0 является точкой строгого локального минимума. В тоже самое время функция $\psi(t) = f(a+tv)$ имеет в точке 0 строгий локальный максимум. Значит в любой окрестности точки a встречаются точки (на прямой $a+th$), в которых значение строго больше $f(a)$, и встречаются точки (на прямой $a+tv$), в которых значение строго меньше $f(a)$. Значит точка a не является точкой локального экстремума. □

15. Условный экстремум.

Пусть M_k – гладкая k-мерная поверхность в $\mathbb{R}^n$ и $f\colon \mathbb{R}^n \to \mathbb{R}$ – непрерывно дифференцируемая функция. Точка $p \in M_k$ является **точкой локального условного максимума (строгого локального условного максимума)** функции f, если существует такая окрестность $U(p)$, что для всех $x \in U'(p)\bigcap M_k$ верно неравенство $f(x) \leqslant f(p)$ ($f(x) < f(p)$). Аналогично определяется точка локального условного минимума (строгого локального условного минимума).

Теорема 15.1. Необходимое условие *Если p является точкой локального условного экстремума, то $\nabla f(p) \perp T_pM_k$. Более того, если $\nabla f \neq 0$, то в точке p локального условного экстремума f касательное пространство T_pM_k является подпространством касательного пространства к поверхности $\{x\colon\ f(x) = f(p)\}$ в точке p.*

Доказательство. Пусть $\xi \in T_pM_k$. Тогда существует такая кривая γ на M_k, что $\gamma(0) = p$ и $\dot\gamma(0) = \xi$. функция $f(\gamma(t))$ имеет в точке $t = 0$ локальный экстремум. Следовательно,

$$0 = \frac{d}{dt}f(\gamma(t))\Big|_{t=0} = \langle \nabla f(p), \dot\gamma(0)\rangle = \langle \nabla f(p), \xi\rangle.$$

□

Пусть M_k задано уравнениями $F_1 = 0, \ldots, F_{n-k} = 0$. Так как $\nabla F_1(p), \nabla F_2(p), \ldots, \nabla F_{n-k}(p)$ составляют базис ортогонального дополнения к пространству T_pM_k, то

$$\nabla f(p) = \lambda_1\nabla F_1(p) + \lambda_2\nabla F_2(p) + \ldots + \lambda\nabla F_{n-k}(p),$$

причем λ_i определены единственным образом.

Это наблюдение называют **правилом множителей Лагранжа**: для поиска точек экстремума функции f при условии $F_1 = 0, \dots, F_{n-k} = 0$ надо составить функцию Лагранжа

$$L(x, \lambda) = f(x) - \lambda_1 F_1(x) - \ldots - \lambda_{n-k} F_{n-k}(x)$$

и приравнять к нулю ее частные производные по x_i и λ_j.

Теперь обсудим достаточное условие.

Теорема 15.2. *Пусть в точке p выполняются необходимые условия и $L(x) = L(x, \lambda)$. Тогда*

(i) *если $d_x^2 L(p, h) > 0$ для всякого $h \neq 0$ и $h \in T_p M_k$, то p – точка строгого локального условного минимума;*

(ii) *если $d_x^2 L(p, h) < 0$ для всякого $h \neq 0$ и $h \in T_p M_k$, то p – точка строгого локального условного максимума;*

(iii) *если существуют такие $h, v \in T_p M_k$, что $d_x^2 L(p, h) > 0$ и $d_x^2 L(p, v) < 0$, то точка p не является точкой локального условного экстремума.*

Доказательство. Заметим, что $f(x) = L(x)$ при $x \in M_k$. Пусть $x_i = g_i(t_1, \dots, t_k)$ задают M_k в окрестности точки p. Вычислим второй дифференциал функции $H(t) = L(g(t))$:

$$d^2 H(0, u) = \frac{\partial^2 L}{\partial x_i \partial x_j} \frac{\partial g_j}{\partial t_k} \frac{\partial g_i}{\partial t_m} u_k u_m.$$

Заметим, что $d^2 H(0, u) = d_x^2 L(p, h)$, где вектор

$$h = \frac{\partial g}{\partial t_1} u_1 + \ldots + \frac{\partial g}{\partial t_k} u_k$$

является касательным к M_k в точке p. Теперь достаточно заметить, что условный экстремум для f является обычным локальным экстремумом для $H(t)$ и можно применить достаточные условия локального экстремума. □

16. Неопределенный интеграл.

Пусть $f\colon (\alpha, \beta) \to \mathbb{R}$. Функция $F\colon (\alpha, \beta) \to \mathbb{R}$ называется первообразной или неопределенным интегралом функции f, если F дифференцируема на (α, β) и $F' = f$.

Предложение 16.1. *Любые две первообразные F_1 и F_2 функции f отличаются на константу.*

Далее через

$$\int f(x)\, dx$$

обозначаем произвольную первообразную функции f. Соответственно, если F – первообразная функции f, то

$$\int f(x)\, dx = F + C.$$

Из определения немедленно следуют равенства:

$$d\Big(\int f(x)\, dx\Big) = f(x)\, dx,$$

$$\int F'(x)\, dx = \int dF = F + C.$$

Найти первообразные некоторых элементарных функций можно просто прочитав справа налево таблицу производных.

Свойства первообразных являются переформулировкой на языке интегралов свойств дифференцирования.

1. (**Линейность**) Пусть на интервале (a, b) заданы дифференцируемые функции F и G, причем $F' = f$ и $G' = g$. Свойство линейности для производных имеет вид

$$(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g.$$

Записывая это равенство с помощью интеграла получаем

$$\int (\alpha f(x) + \beta g(x))\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx + C.$$

2. (**Формула интегрирования по частям**) Пусть на интервале (a, b) заданы дифференцируемые функции f и g. Правило Лейбница имеет вид

$$(fg)' = f'g + fg'.$$

Запишем это равенство с помощью интегралов:

$$\int (fg)'\,dx = \int (f'g + fg')\,dx + C.$$

Так как левая часть равна $fg + C$, то приходим к равенству

$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx + C,$$

которое называют формулой интегрирования по частям.

3. (**Формула замены переменных**) Пусть функция F дифференцируема на интервале (a, b) и $F' = f$. Пусть также $\varphi\colon (\alpha, \beta) \to (a, b)$ – дифференцируемая функция. По теореме о дифференцировании сложной функции композиция $F(\varphi(t))$ дифференцируема и

$$F(\varphi(t))' = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

Записывая это равенство с помощью интегралов получаем формулу замены переменных:

$$\int f(\varphi(t))\varphi'(t)\,dt = \int f(x)\,dx\Big|_{x=\varphi(t)}.$$

Основная идея вычисления первообразных состоит в применении указанных правил интегрирования и сведении исходного интеграла к табличному. В некоторых случаях отыскание первообразных алгоритмизированно. Например, для отыскания первообразной рациональной функции надо разложить эту функцию в сумму простейших дробей и затем вычислить интеграл отдельно от каждого слагаемого, а эти интегралы мало отличаются от табличных. В некоторых других ситуациях подходящей заменой переменной подынтегральное выражение можно преобразовать к рациональной функции. Однако в общем случае вычисление первообразных является сложной задачей. Более того, во многих случаях первообразная не выражается в элементарных функциях, например такова первообразная функции e^{-x^2}.

17. Интеграл Римана.

Дадим следующие определения:

- **Разбиением** $\mathbb{T}$ отрезка $[a, b]$ называется набор точек

$$a = x_0 < x_1 < x_2 < \ldots < x_n = b.$$

- Отрезки $\Delta_i = [x_{i-1}, x_i]$, где $i = 1, 2, \ldots, n$, называются **отрезками разбиения**.
- Число $\lambda(\mathbb{T}) = \max_{1\leqslant i\leqslant n} |x_i - x_{i-1}|$ называется **параметром разбиения** или **масштабом**.
- **Отмеченным разбиением** $(\mathbb{T}, \xi)$ отрезка $[a, b]$ называется разбиение $\mathbb{T}$ отрезка $[a, b]$ с отмеченными точками $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$, где $\xi_k \in [x_{k-1}, x_k]$.

- **Римановой интегральной суммой** функции f соответствующей отмеченному разбиению $(\mathbb{T}, \xi)$ отрезка $[a, b]$ называется выражение
$$\sigma(f, \mathbb{T}, \xi) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}).$$

Пусть функция f определена на отрезке $[a, b]$. Говорят, что f интегрируема по Риману на отрезке $[a, b]$ и число I является ее интегралом, если для всякого $\varepsilon > 0$ найдется такое $\delta > 0$, что для всякого отмеченного разбиения $(\mathbb{T}, \xi)$ с параметром разбиения $\lambda(\mathbb{T}) < \delta$ выполняется
$$|\sigma(f, \mathbb{T}, \xi) - I| < \varepsilon.$$
Для числа I обычно используют следующее обозначение:
$$\int_a^b f(x)\, dx.$$

Примеры.

1. Пусть $f \equiv 1$ на $[a, b]$. Тогда для всякого отмеченного разбиения риманова сумма равна $(b - a)$. Следовательно, имеем
$$\int_a^b 1\, dx = b - a.$$

2. Пусть $a \leqslant c \leqslant b$ и $f(x) = 0$ при $x < c$, $f(x) = 1$ при $x > c$, а $f(c)$ – некоторое число. Тогда для всякого отмеченного разбиения $(\mathbb{T}, \xi)$ отрезка $[a, b]$ точка c или совпадает с какой-нибудь отмеченной точкой или лежит внутри одного отрезка разбиения. Рассмотрим сначала случай, когда c лежит строго внутри отрезка разбиения Δ_k. Распишем интегральную сумму
$$\sigma(f, \mathbb{T}, \xi) = \sum_{i \leqslant k-1} f(\xi_i)|\Delta_i| + f(\xi_k)|\Delta_k| + \sum_{k+1 \leqslant i} f(\xi_i)|\Delta_i| = f(\xi_k)|\Delta_k| + (b - c) + (c - x_k).$$
Следовательно, при стремлении масштаба к нулю интегральная сумма стремится к $(b - c)$. Теперь рассмотрим случай, когда $c = x_k$. Тогда
$$\sigma(f, \mathbb{T}, \xi) = f(\xi_k)|\Delta_k| + f(\xi_{k+1})|\Delta_{k+1}| + (b - c) + (x_k - x_{k+1})$$
В этом случае как и выше при стремлении масштаба к нулю интегральная сумма стремится к $(b - c)$. Таким образом, f интегрируема и
$$\int_a^b f\, dx = b - c.$$

3. Функция Дирихле $D(x) = 1$ при $x \in \mathbb{Q}$ и $D(x) = 0$ при $x \notin \mathbb{Q}$ не интегрируема по Риману ни на каком отрезке $[a, b]$. Действительно, выбором отмеченных точек риманову сумму можно сделать равной нулю или равной $b - a$ при любом масштабе разбиения.

Предложение 17.1. *Если функция интегрируема по Риману на отрезке, то эта функция ограниченна на этом отрезке.*

Доказательство. Если функция не является ограниченной, то для всякого разбиения найдется отрезок разбиения, на котором эта функция принимает сколь угодно большие значения. Следовательно, для всякого разбиения выбором отмеченных точек интегральную сумму можно сделать сколь угодно большой, что противоречит определению интеграла Римана. $\square$

Из определения немедленно следуют два важнейших свойства интеграла: линейность и монотонность.

Предложение 17.2. (Линейность интеграла) *Если f и g интегрируемы на $[a,b]$, то функция $\alpha f+\beta g$ интегрируема на $[a,b]$ и верно равенство:*

$$\int_a^b(\alpha f+\beta g)\,dx=\alpha\int_a^b f\,dx+\beta\int_a^b g\,dx.$$

Доказательство. Это утверждение немедленно следует из определения и равенства

$$\sigma(\alpha f+\beta g,\mathbb{T},\xi)=\alpha\sigma(f,\mathbb{T},\xi)+\beta\sigma(g,\mathbb{T},\xi).$$

□

Пусть $a\leqslant\alpha<\beta\leqslant b$. Функция $f(x)=0$ при $x\notin[\alpha,\beta]$ и $f(x)=C$ при (α,β), а вточках α и β принимает произвольные значения. Положим $g(x)=0$ при $x\leqslant 0$, $g(x)=1$ при $x>0$ и $h(x)=0$ при $x<0$, $h(x)=1$ при $x\geqslant 0$. Функцию f можно представить в виде

$$C(g(x-\alpha)-h(x-\beta))+f(\alpha)(h(x-\alpha)-g(x-\alpha))+f(\beta)(h(x-\beta)-g(x-\beta)).$$

Следовательно по свойству линейности функция f интегрируема и

$$\int_a^b f\,dx=C(\beta-\alpha).$$

Пусть задано разбиение $a<x_0<x_1<\ldots<x_n=b$ отрезка $[a,b]$. Предположим, что функция f определена на $[a,b]$ и $f\equiv c_k$ на (x_{k-1},x_k). Тогда f интегрируема и

$$\int_a^b f\,dx=\sum_k c_k|\Delta_k|.$$

Такие функции называются кусочно постоянными.

Предложение 17.3. (Монотонность интеграла) *Если f и g интегрируемы на $[a,b]$ и $f\leqslant g$, то*

$$\int_a^b f\,dx\leqslant\int_a^b g\,dx.$$

Доказательство. Из-за линейности интеграла это свойство достаточно доказать в случае $f=0$, т. е. из неотрицательности g вывести неотрицательность интеграла от g. Это немедленно следует из неотрицательности римановой суммы для g. □

Следствие 17.1. (Теорема о среднем) *Пусть f интегрируема на $[a,b]$ и $m=\inf_{[a,b]}f$, $M=\sup_{[a,b]}f$. Тогда найдется число $\mu\in[m,M]$ такое, что*

$$\int_a^b f\,dx=\mu(b-a).$$

Более того, если f непрерывна, то $\mu=f(c)$ для некоторого $c\in[a,b]$.

Доказательство. Достаточно заметить, что из-за монотонности интеграла верно следствие:

$$m\leqslant f(x)\leqslant M\Rightarrow m\leqslant\frac{1}{b-a}\int_a^b f\,dx\leqslant M.$$

□

В качестве следствия получим следующий важный результат.

Теорема 17.1. *Если последовательность интегрируемых на* $[a,b]$ *функций* f_n *равномерно на* $[a,b]$ *сходится к* f*, то* f *интегрируема на* $[a,b]$ *и верно равенство:*

$$\lim_{n\to\infty}\int_a^b f_n\,dx = \int_a^b f\,dx.$$

Доказательство. По теореме о среднем

$$\Big|\int_a^b f_n\,dx - \int_a^b f_m\,dx\Big| = |\mu_{n,m}||b-a|,$$

где

$$\inf_{[a,b]}(f_n(x)-f_m(x)) \leqslant \mu_{n,m} \leqslant \sup_{[a,b]}(f_n(x)-f_m(x)).$$

Следовательно, $|\mu_{n,m}| \leqslant \sup_{[a,b]}|f_n - f_m|$ и $\mu_{m,n}\to 0$ при $n,m\to\infty$. Таким образом числовая последовательность

$$\int_a^b f_n\,dx$$

сходится к некоторому числу I. Покажем, что функция f интегрируема и ее интеграл равен I. Пусть $(\mathbb{T},\xi)$ – отмеченное разбиение отрезка $[a,b]$. Имеет место неравенство:

$$\Big|\sigma(f,\mathbb{T},\xi)-I\Big| \leqslant \sigma(|f-f_n|,\mathbb{T},\xi) + \Big|\sigma(f_n,\mathbb{T},\xi) - \int_a^b f_n\,dx\Big| + \Big|\int_a^b f_n\,dx - I\Big|.$$

Пусть $\varepsilon>0$. Выбираем номер n столь большим, что первое и третье слагаемые меньше ε. Фиксируем такое n. Теперь выбираем масштаб разбиения так, что второе слагаемое меньше ε для всех разбиений с таким или меньшим масштабом. Получаем оценку:

$$\Big|\sigma(f,\mathbb{T},\xi)-I\Big| < 3\varepsilon.$$

□

Следствие 17.2. *Если функция* f *непрерывна на отрезке* $[a,b]$*, то функция* f *интегрируема на* $[a,b]$*.*

Доказательство. Рассмотрим разбиение отрезка $[a,b]$, заданное точками $x_k = a + \frac{k(b-a)}{n}$. Определим функцию f_n следующим образом: $f_n(x) = f(x_k)$ при $x\in[x_{k-1},x_k)$ и $f_n(b) = f(b)$. Функция f_n интегрируема на $[a,b]$. Так как f равномерно непрерывна на $[a,b]$, то для всякого $\varepsilon>0$ найдется N такое, что для всех $n>N$ выполняется $|f(x)-f(x_k)|<\varepsilon$ для всех $x\in[x_{k-1},x_k]$. Следовательно, для всех $n>N$ верна оценка $|f(x)-f_n(x)|<\varepsilon$ для всех $x\in[a,b]$. Таким образом, f_n равномерно сходятся к f, а мы уже знаем, что это влечет интегрируемость f. □

Несложно модифицировать это рассуждение и доказать, что всякая кусочно непрерывная функция интегрируема.

Функция f называется **кусочно непрерывной** на отрезке $[a,b]$, если есть такое разбиение $a = x_0 < x_1 < \ldots < x_n = b$, что функция f непрерывна на каждом интервале (x_{i-1},x_i) и существуют конечные пределы $\lim_{x\to x_i-0} f(x)$ и $\lim_{x\to x_i+0} f(x)$ в каждой точке x_i.

Однако, для получения критерия интегрируемости нам потребуются дополнительные построения.

18. Критерий Дарбу. Множества меры ноль. Критерий Лебега.

Пусть f – ограниченная на отрезке $[a,b]$ функция и $\mathbb{T}$ – разбиение $[a,b]$. Выражения

$$s(f,\mathbb{T})=\sum_k \inf_{\Delta_k} f(x)|\Delta_k|,\quad S(f,\mathbb{T})=\sum_k \sup_{\Delta_k} f(x)|\Delta_k|$$

называются *нижней и верхней суммой Дарбу* соответственно. Здесь $\Delta_k=[x_{k-1},x_k]$.

Лемма 18.1. *Имеют место следующие утверждения:*

(i) *для всякого разбиения* $\mathbb{T}$

$$s(f,\mathbb{T})\leqslant \inf_\xi \sigma(f,\mathbb{T},\xi)\leqslant \sigma(f,\mathbb{T},\xi)\leqslant \sup_\xi \sigma(f,\mathbb{T},\xi)\leqslant S(f,\mathbb{T}),$$

(i) *если разбиение* $\mathbb{T}\subset\mathbb{T}'$ *(говорят, что* $\mathbb{T}'$ *является измельчением* $\mathbb{T}$*), то*

$$s(f,\mathbb{T})\leqslant s(f,\mathbb{T}'),\quad S(f,\mathbb{T}')\leqslant S(f,\mathbb{T}).$$

(iii) *если* $\mathbb{T}_1$ *и* $\mathbb{T}_2$ *– разбиения отрезка* $[a,b]$*, то* $s(f,\mathbb{T}_1)\leqslant S(f,\mathbb{T}_2)$*.*

Нижним интегралом Дарбу функции f называется величина

$$\underline{I}=\sup_{\mathbb{T}} s(f,\mathbb{T}),$$

а *верхним интегралом Дарбу* функции f называется величина

$$\overline{I}=\inf_{\mathbb{T}} S(f,\mathbb{T}).$$

Лемма 18.2. *Для всякого* $\varepsilon>0$ *найдется* $\delta>0$ *такое, что для всякого разбиения* $\mathbb{T}$ *с* $\lambda(\mathbb{T})<\delta$

$$\underline{I}-s(f,\mathbb{T})<\varepsilon.$$

Аналогичное утверждение выполнено для $S(f,\mathbb{T})$*.*

Доказательство. Пусть $\varepsilon>0$. Найдется $\mathbb{T}_\varepsilon$ такое, что

$$\underline{I}-s(f,\mathbb{T}_\varepsilon)<\varepsilon.$$

Пусть $\mathbb{T}$ – произвольное разбиение. Тогда

$$\underline{I}-s(f,\mathbb{T})=\underline{I}-s(f,\mathbb{T}_\varepsilon)+s(f,\mathbb{T}_\varepsilon)-s(f,\mathbb{T})<\varepsilon+s(f,\mathbb{T}_\varepsilon\cup\mathbb{T})-s(f,\mathbb{T}).$$

Рассмотрим разность $s(f,\mathbb{T}_\varepsilon\cup\mathbb{T})-s(f,\mathbb{T})$. Пусть $a=x_0<x_1<\ldots<x_N$ – точки разбиения $\mathbb{T}_\varepsilon$, Δ'_k – отрезки разбиения $\mathbb{T}_\varepsilon\cup\mathbb{T}$, Δ_k – отрезки разбиения $\mathbb{T}$. Если $x_j\notin\Delta'_k$, то Δ'_k совпадает с одним из отрезков разбиения $\mathbb{T}$. Следовательно, имеем

$$s(f,\mathbb{T}_\varepsilon\cup\mathbb{T})-s(f,\mathbb{T})=\sum_{k:\,x_j\in\Delta'_k}\inf_{\Delta'_k} f|\Delta'_k|-\sum_{k:\,x_j\in\Delta_k}\inf_{\Delta_k} f|\Delta_k|\leqslant 4N\sup_{[a,b]}|f|\lambda(\mathbb{T}).$$

Пусть $\delta>0$ столь мало, что $4N\sup\limits_{[a,b]}|f|<\varepsilon$. Тогда для всякого $\mathbb{T}$ с $\lambda(\mathbb{T})<\delta$ выполняется неравенство

$$\underline{I}-s(f,\mathbb{T})<\varepsilon.$$

□

Из последних двух лемм следует критерий интегрируемости.

Теорема 18.1. (Критерий Дарбу) *Ограниченная функция* f *интегрируема по Риману на отрезке* $[a,b]$ *тогда и только тогда, когда ее нижний и верхний интегралы Дарбу совпадают.*

Следствие 18.1. *Ограниченная функция f интегрируема по Риману на отрезке $[a,b]$ тогда и только тогда, когда для всякого $\varepsilon > 0$ найдется разбиение $\mathbb{T}$ отрезка $[a,b]$ такое, что*

$$S(f,\mathbb{T}) - s(f,\mathbb{T}) < \varepsilon.$$

Заметим, что

$$S(f,\mathbb{T}) - s(f,\mathbb{T}) = \sum_k (\sup_{\Delta_k} f - \inf_{\Delta_k} f)|\Delta_k|.$$

Разность $\sup\limits_{\Delta_k} f - \inf\limits_{\Delta_k} f$ обозначают через $\omega(f,\Delta_k)$ и называют колебанием функции f на Δ_k. Таким образом, интегрируемость ограниченной функции f равносильна тому, что для всякого $\varepsilon > 0$ найдется разбиение $\mathbb{T}$ отрезка $[a,b]$ такое, что

$$\sum_k \omega(f,\Delta_k)|\Delta_k| < \varepsilon.$$

Следствие 18.2. *Если f ограниченная и монотонная функция на $[a,b]$, то f интегрируема на $[a,b]$.*

Доказательство. Пусть f не убывает. Тогда $\omega(f,\Delta_k) = f(x_k) - f(x_{k-1})$. Возьмем разбиение $[a,b]$ на отрезки равной длины $|\Delta_k| = (b-a)/N$. Получаем

$$\sum_k \omega(f,\Delta_k)|\Delta_k| = \frac{b-a}{N}\sum_k (f(x_k) - f(x_{k-1})) = \frac{(b-a)(f(b)-f(a))}{N}.$$

Остается заметить, что правая часть стремится к нулю при $N \to \infty$. □

Следствие 18.3. *Если функция f интегрируема на $[a,b]$, то функция $|f|$ интегрируема на $[a,b]$ и имеет место оценка*

$$\Big|\int_a^b f\,dx\Big| \leqslant \int_a^b |f|\,dx.$$

Доказательство. Достаточно заметить, что $\omega(|f|,\Delta_k) \leqslant \omega(f,\Delta_k)$. □

Теорема 18.2. (Аддитивность интеграла)

(i) *Если $[c,d] \subset [a,b]$ и f интегрируема на $[a,b]$, то f интегрируема на $[c,d]$.*

(ii) *Если $a < c < b$ и функция f интегрируема на $[a,c]$ и на $[c,b]$, то f интегрируема на $[a,b]$.*

(iii) *Если f интегрируема на $[a,b]$ и $a < c < b$, то*

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx.$$

Доказательство. (i) Так как f интегрируема на $[a,b]$, то для всякого $\varepsilon > 0$ найдется разбиение $\mathbb{T}$ отрезка $[a,b]$ такое, что $S(f,\mathbb{T}) - s(f,\mathbb{T}) < \varepsilon$. Добавим к $\mathbb{T}$ точку c. Новое разбиение разбивается на два разбиения: одно $\mathbb{T}_{[a,c]}$) для $[a,c]$, а другое $\mathbb{T}_{[c,b]}$) для $[c,b]$. Имея ввиду явный вид $S(f,\mathbb{T}) - s(f,\mathbb{T})$, получаем

$$S(f,\mathbb{T}_{[a,c]}) - s(f,\mathbb{T}_{[a,c]}) \leqslant S(f,\mathbb{T}\cup\{c\}) - s(f,\mathbb{T}\cup\{c\}) \leqslant S(f,\mathbb{T}) - s(f,\mathbb{T}) < \varepsilon.$$

(ii) Требуемое в критерии Дарбу разбиение для $[a,b]$ получаем с помощью объединения соответствующих разбиений для отрезков $[a,c]$ и $[c,b]$.

(iii) Поскольку функция f интегрируема на $[a,b]$, то по любой последовательности разбиений, у которых параметр разбиения стремится к нулю, сходится к интегралу от f. Остается заметить, что если отмеченное разбиение $(\mathbb{T}, \xi)$ является объединением разбиений $(\mathbb{T}_{[a,c]}, \xi)$ и $(\mathbb{T}_{[c,b]}, \xi)$, то

$$\sigma(f, \mathbb{T}, \xi) = \sigma(f, \mathbb{T}_{[a,c]}, \xi) + \sigma(f, \mathbb{T}_{[c,b]}, \xi).$$

□

При внимательном взгляде на выражение $\sum_k \omega(f, \Delta_k)|\Delta_k|$ становится ясно, что малость этой суммы зависит от малости $\omega(f, \Delta_k)$ и от малости $|\Delta_k|$. Малость колебания функции на малых отрезках связана с непрерывностью функции. Справиться с точками разрыва функции можно с помощью малости $|\Delta_k|$, но таких точек не должно быть слишком много, иначе малые длины отрезков разбиения в сумме могут дать уже не очень малую величину. Точное описание непрерывности интегрируемой функции дает критерий Лебега.

Множество $E \subset \mathbb{R}$ называется *множеством меры ноль по Лебегу*, если для всякого $\varepsilon > 0$ существует не более чем счетный набор интервалов I_k такой, что $E \subset \bigcup\limits_k I_k$ и $\sum_k |I_k| < \varepsilon$.

Предложение 18.1. (i) *Точка является множеством меры ноль.*

(ii) *Отрезок не является множеством меры ноль.*

(iii) *Не более чем счетное объединение множеств меры ноль является множеством меры ноль.*

Отметим, что множество меры ноль может быть континуальным.

Теорема 18.3. (Критерий Лебега) *Функция интегрируема тогда и только тогда, когда она ограничена и ее множество точек разрыва имеет меру ноль по Лебегу.*

Следствие 18.4. *Если функция f интегрируема на $[a, b]$ и ψ непрерывна на $[-\inf f, \sup f]$, то $\psi(f)$ интегрируема на $[a, b]$.*

19. Формула Ньютона–Лейбница.

Пусть f интегрируема на отрезке $[a, b]$. Функция F, заданная равенством

$$F(x) = \int_a^x f(t)\, dt,$$

называется *интегралом с переменным верхним пределом.*

Предложение 19.1. *Имеет место оценка*

$$|F(x) - F(y)| \leqslant M|x - y|, \quad M = \sup |f|,$$

в частности, функция F непрерывна.

Доказательство. Пусть $x > y$. Тогда по аддитивности интеграла

$$F(x) - F(y) = \int_y^x f(t)\, dt.$$

Следовательно, приходим к оценке

$$|F(x) - F(y)| \leqslant \int_y^x |f(t)|\, dt \leqslant M|x - y|.$$

□

Теорема 19.1. *Если функция f непрерывна в точке $c \in [a, b]$, то F дифференцируема в точке c и $F'(c) = f(c)$.*

Доказательство. Используя определение F, получаем
$$\frac{F(c+h)-F(c)}{h}-f(c)=h^{-1}\int_c^{c+h}(f(t)-f(c))\,dt.$$
Следовательно, приходим к оценке
$$\Big|\frac{F(c+h)-F(c)}{h}-f(c)\Big|\leqslant \sup_{[c,c+h]}|f(t)-f(c)|,$$
где правая часть из-за непрерывности f в точке c стремится к нулю. □

Следствие 19.1. (Формула Ньютона–Лейбница) *Если f непрерывна на $[a,b]$, то существует первообразная функции f и для всякой первообразной $\mathcal{F}$ выполняется равенство:*
$$\mathcal{F}(b)-\mathcal{F}(a)=\int_a^b f(t)\,dt.$$

Следствие 19.2. (Формула интегрирования по частям) *Пусть f,g – непрерывно дифференцируемые на отрезке $[a,b]$ функции. Тогда*
$$\int_a^b f(x)g'(x)\,dx=f(b)g(b)-f(a)g(a)-\int_a^b f'(x)g(x)\,dx.$$

Доказательство. Достаточно заметить, что по формуле Ньютона–Лейбница
$$\int_a^b (fg)'\,dx=f(b)g(b)-f(a)g(a)$$
и вспомнить правило Лейбница $(fg)'=f'g+fg'$. □

Используя последнее следствие можно показать, что если функция f на отрезке с концами x_0 и x имеет n непрерывных производных, то остаточный член в формуле Тейлора представляется в следующем виде
$$r_{n-1}(x_0,x)=\frac{1}{(n-1)!}\int_{x_0}^x f^{(n)}(t)(x-t)^{n-1}\,dt.$$

Следствие 19.3. (Формула замены переменных) *Пусть f – непрерывная на отрезке $[a,b]$ функция и $\varphi:[\alpha,\beta]\to[a,b]$ – непрерывно дифференцируемая функция, причем, $\varphi(\alpha)=a$ и $\varphi(\beta)=b$. Тогда*
$$\int_a^b f(x)\,dx=\int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt.$$

Доказательство. Пусть F – первообразная f. Тогда $F(\varphi(t))$ – первообразная $f(\varphi(t))\varphi'(t)$. Применяя формулу Ньютона–Лейбница получаем
$$\int_a^b f\,dx=F(b)-F(a)=F(\varphi(\beta))-F(\varphi(\alpha))=\int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt.$$
□

20. Интеграл с параметром: непрерывность и дифференцируемость.

В различных приложениях появляется необходимость работать с выражениями вида
$$I(y)=\int_{g(y)}^{h(y)} f(x,y)\,dx,$$
которые принято называть интегралом с параметром.

Теорема 20.1. *Пусть f непрерывна на $[a,b]\times[c,d]$ и $h,g\colon[c,d]\to[a,b]$ непрерывные функции. Тогда $I(y)$ непрерывная на $[c,d]$ функция.*

Доказательство. Так как

$$I(y) = \Phi(h(y), y) - \Phi(g(y), y), \quad \Phi(u, y) = \int_a^u f(x, y)\, dx,$$

то достаточно доказать, что функция $\Phi(u, y)$ непрерывна. Имеем

$$|\Phi(u_0, y_0) - \Phi(u, y)| \leqslant |\Phi(u_0, y) - \Phi(u, y)| + |\Phi(u_0, y_0) - \Phi(u_0, y)|.$$

Первое слагаемое в правой части оценивается через $\max|f||u_0 - u|$. Так как f равномерно непрерывна на $[a, b]\times[c, d]$, то для всякого $\varepsilon > 0$ найдется $\delta > 0$ такое, что $(x-s)^2+(y-t)^2 < \delta$ влечет $|f(x, y) - f(s, t)| < \varepsilon$, в частности, для всех $x \in [a, b]$ имеем $|f(x, y) - f(x, y_0)| < \varepsilon$ при $|y - y_0| < \delta$. Следовательно, выражение $|\Phi(u_0, y_0) - \Phi(u_0, y)|$ меньше $\varepsilon(b-a)$ при $|y - y_0| < \delta$. Итак, $\Phi(u, y) \to \Phi(u_0, y_0)$ при $(u, y) \to (u_0, y_0)$. $\square$

Теорема 20.2. *Пусть f и $\frac{\partial}{\partial y} f$ являются непрерывными функциями на $[a, b] \times [c, d]$, функции $g, h\colon [c, d] \to [a, b]$ дифференцируемы. Тогда $I(y)$ дифференцируемая функция и*

$$I'(y) = f(h(y), y)h'(y) - f(g(y), y)g'(y) + \int_{g(y)}^{h(y)} \frac{\partial f}{\partial y}(x, y)\, dx.$$

Доказательство. Достаточно доказать дифференцируемость $\Phi(u, y)$ по y при фиксированном u. Полное утверждение будет немедленно следовать из теоремы о дифференцируемости сложной функции.

$$\frac{\Phi(u, y + p) - \Phi(u, y)}{p} - \int_a^u \frac{\partial f}{\partial y}(x, y)\, dx = \int_a^b \Big[\frac{f(x, y + p) - f(x, y)}{p} - \frac{\partial f}{\partial y}(x, y)\Big]\, dx.$$

По теореме Лагранжа

$$\frac{f(x, y + p) - f(x, y)}{p} = \frac{\partial f}{\partial y}(x, \xi),$$

где ξ лежит между y и $y + p$. Следовательно,

$$\Big|\frac{f(x, y + p) - f(x, y)}{p} - \frac{\partial f}{\partial y}(x, y)\Big| \leqslant \max_x \max_{|z-y|\leqslant p} \Big|\frac{\partial f}{\partial y}(x, z) - \frac{\partial f}{\partial y}(x, y)\Big| = r(p).$$

В силу непрерывности частной производной f по y заключаем, что $r(p) \to 0$ при $p \to 0$. Остается только выписать оценку:

$$\Big|\frac{\Phi(u, y + p) - \Phi(u, y)}{p} - \int_a^u \frac{\partial f}{\partial y}(x, y)\, dx\Big| \leqslant (b - a)r(p).$$

$\square$

Литература

1. Зорич В.А. Математический анализ. М.: МЦНМО, 2007.
2. Шварц Л. Анализ I и II. М.: Мир, 1972.
3. Львовский С.М. Лекции по математическому анализу. М.: МЦНМО, 2008.
4. Гелбаум Б., Олмстед Дж. Контрпримеры в анализе. М.: Мир, 1967.
5. Прасолов В.В. Элементы комбинаторной и дифференциальной топологии. М.: МЦНМО, 2014.