Математический анализ, II семестр

Весна 2015, лектор: Храбров Александр Игоревич

Авторы: Анастасия Гайдашенко, Юрий Кравченко, Надежда Бугакова, Анастасия Старкова, Никита Подгузов, Дмитрий Лапшин, Егор Суворов

Собрано: 6 сентября 2015 г. 04:52

Оглавление

Глава 1

Интегральное исчисление (продолжение)

1.1. Определенный интеграл

𝔇𝔢𝔣 1.1.1. $\mathcal{F}$ — множество ограниченных подмножеств $\mathbb{R}^2$

𝔇𝔢𝔣 1.1.2. Площадь:

$$S\colon \mathcal{F} \to [0;+\infty)$$

Должны выполняться следующие свойства:

1. $E_1 \cap E_2 = \emptyset \Rightarrow S(E_1 \cup E_2) = S(E_1) + S(E_2)$

   Следствие 1. $E_1 \subset E_2 \Rightarrow S(E_1) \leqslant S(E_2)$

   ▶ $E_2 = E_1 \cup (E_2 \setminus E_1); S(E_2) = S(E_1) + S(E_2 \setminus E_1); S(E_2 \setminus E_1) \geqslant 0$ ◀

2. Если E_1 переводится движением в E_2, то $S(E_1) = S(E_2)$.

3. Площадь прямоугольника есть произведение длин сторон.

Теорема 1.1.1 (Теорема Банаха)**.** Объект, обладающий такими свойствами существует, но не единственнен.

Теорема 1.1.2 (Хаусдорфа)**.** Для любого $n \geqslant 3$ в $\mathbb{R}^n$ аналогичным образом определенного объекта не существует.

Мы будем пользоваться упрощением понятия площадь:

𝔇𝔢𝔣 1.1.3. $\sigma\colon \mathcal{F} \to [0;+\infty)$ со свойствами:

1. Площадь прямоугольника со сторонами, *параллельными осям координат*, есть произведение его сторон.

2. $E_1 \subset E_2 \Rightarrow \sigma(E_1) \leqslant \sigma(E_2)$

3. Разрешаем разбивать множество на части только горизонтальными/вертикальными прямыми. Обозначаем: E_-, E_+

   (a) $E_- \cap E_+ = \emptyset$

   (b) $E_- \cup E_+ = E$

𝔇𝔢𝔣 1.1.4. P — прямоугольник.

$$|P| = \sigma(\langle a, b\rangle \times \langle c, d\rangle) = (b-a)(d-c)$$

𝔇𝔢𝔣 1.1.5.

$$\sigma(E) = \inf \sum_{k=1}^{n} |P_k| : E \subset \bigcup_{k=1}^{n} P_k$$

Теорема 1.1.3 (σ есть площадь). 1. σ — площадь

2. σ не меняется при параллельном переносе

▶ 1. Площадь прямоугольника:

$$\sum_{k=1}^{n} |P_k| \geqslant P_0 \Rightarrow \inf \sum |P_k| \geqslant (b-a)(d-c)$$

Мы можем взять покрытие из одного прямоугольника:

$$\inf \leqslant (b-a)(d-c)$$

2. $E_1 \subset E_2 \Rightarrow \sigma(E_1) \leqslant \sigma(E_2)$:
$$E_2 \subset \bigcup P_k \Rightarrow E_1 \subset \bigcup P_k$$
Таким образом, класс покрытий E_1 шире класса покрытий E_2, и его инфимум не больше, то есть $\sigma(E_1) \leqslant \sigma(E_2)$.

3. $E = E_- \cup E_+$. Докажем, что : $\sigma(E) = \sigma(E_-) + \sigma(E_+)$. Рассмотрим $\bigcup P_k^{\pm}$:

$$\sum |P_k^{\pm}| \leqslant \sigma(E_{\pm}) + \varepsilon \Rightarrow \sigma(E) \leqslant \sigma(E_-) + \sigma(E_+) + 2\varepsilon$$

Обратим ε в ноль, получим
$$\sigma(E) \leqslant \sigma(E_-) + \sigma(E_+)$$
Рассмотрим $\bigcup P_k$: $E \subset \bigcup P_k$: Разделим прямой: прямоугольники, которые пересеклись с ней, образуют два новых прямоугольника. Мы получили покрытия $E_{\pm}$. Инфинум суммы не меньше суммы инфинумов, поэтому

$$\sigma(E) \geqslant \sigma(E_-) + \sigma(E_+)$$

◀

𝔇𝔢𝔣 1.1.6. $f : E \to \mathbb{R}$.

$$f_{\pm} \colon E \to [0; +\infty) \quad f_+ = \max\{f, 0\} \quad f_- = \max\{-f; 0\}$$

Свойства:

1. $f = f_+ - f_-$
2. $|f| = f_+ + f_-$
3. $2f_{\pm} = |f| \pm f$
4. f — непрерывна $\Rightarrow f_{\pm}$ непрерывна

𝔇𝔢𝔣 1.1.7. Подграфик функции $f \geqslant 0$:

$$P_f = \{(x, y) \mid x \in E, y \in [0; f(x)]\}$$

Замечание 1.1.1.
$$f \in C[a, b] \Rightarrow P_f \text{ — ограниченное множество}$$

𝔇𝔢𝔣 1.1.8. $f \in C[a,b]$. $\int_a^b f$ — определенный интеграл функции f на $[a;b]$:

$$\int\limits_a^b f = \sigma(P_{f_+}) - \sigma(P_{f_-})$$

Свойства:

1. $$f \geqslant 0 \Rightarrow \int\limits_a^b f = \sigma(P_f) \geqslant 0$$

2. $$\int\limits_a^b f = \int\limits_a^b f_+ - \int\limits_a^b f_-$$

3. $f \geqslant 0$:
$$\int\limits_a^b f = 0 \Rightarrow f \equiv 0$$

▶ От противного: пусть есть точка, значение в которой не ноль, тогда и в её окрестности не ноль (по теореме о стабилизации знака). Выбираем в этой окрестности минимум по значениям и получаем прямоугольник с площадью, отличной от нуля. Площадь графика больше этой площади, но равна нулю — противоречие. ◀

4. $$\int\limits_a^b -f = -\int\limits_a^b f$$

▶
$$(-f)_+ = f_-; (-f)_- = f_+$$
◀

5. $$\int\limits_a^b C = C(b-a)$$

1.2. Свойства интеграла

Замечание 1.2.1. Свойство $\sigma(E) = \sigma(E_-) + \sigma(E_+)$ не испортится, если убрать требование про непересечение, то есть могут быть общие точки на разделяющей прямой.

▶ Прямая — вырожденный прямоугольник со стороной, равной нулю, поэтому его σ равна нулю, значит он не влияет на сумму. ◀

Теорема 1.2.1 (Аддитивность интеграла)**.** $f \in C[a,b]$, $c \in [a,b]$. Тогда

$$\int\limits_a^b f = \int\limits_a^c f + \int\limits_c^b f$$

▶

$$\int\limits_a^b f = \sigma\left(P_{f_+}\right) - \sigma\left(P_{f_-}\right) =$$

$$= \sigma\left(P_{f_+}([a,c])\right) - \sigma\left(P_{f_-}([a,c])\right) + \sigma\left(P_{f_+}([c,b])\right) - \sigma\left(P_{f_-}([c,b])\right) = \int\limits_a^c f + \int\limits_c^b f$$

◀

Теорема 1.2.2 (Монотонность интеграла)**.** $f, g \in C[a,b]$, $f \geqslant g$. Тогда

$$\int\limits_a^b f \geqslant \int\limits_a^b g$$

▶

$$\int\limits_a^b f = \sigma\left(P_{f_+}\right) - \sigma\left(P_{f_-}\right) \quad \int\limits_a^b g = \sigma\left(P_{g_+}\right) - \sigma\left(P_{g_-}\right)$$

$$\begin{array}{llllllll} f & \geqslant g & \Rightarrow & f_+ & \geqslant g_+ & \Rightarrow & P_{g_+} & \subset P_{f_+} & \Rightarrow & \sigma\left(P_{f_+}\right) & \geqslant \sigma\left(P_{g_+}\right) \\ -g & \geqslant -f & \Rightarrow & g_- & \geqslant f_- & \Rightarrow & P_{f_-} & \subset P_{g_-} & \Rightarrow & \sigma\left(P_{f_-}\right) & \leqslant \sigma\left(P_{g_-}\right) \end{array}$$

◀

Следствие 1.2.2.1.

$$(b-a)\min f \leqslant \int\limits_a^b f \leqslant (b-a)\max f$$

▶ Проинтегрируем неравенства $\min f \leqslant f \leqslant \max f$ и получим то, что надо. ◀

Следствие 1.2.2.2.

$$\left|\int\limits_a^b f\right| \leqslant \int\limits_a^b |f|$$

▶

$$-|f| \leqslant f \leqslant |f|$$

$$\int\limits_a^b -|f| \leqslant \int\limits_a^b f \leqslant \int\limits_a^b |f|$$

$$-\int\limits_a^b |f| \leqslant \int\limits_a^b f \leqslant \int\limits_a^b |f|$$

$$\left|\int\limits_a^b f\right| \leqslant \int\limits_a^b |f|$$

◀

Следствие 1.2.2.3 (Интегральная теорема о среднем).

$$\exists c\colon \int\limits_a^b f = (b-a)f(c)$$

▶

$$\min f \leqslant \frac{1}{b-a}\int\limits_a^b f \leqslant \max f$$

Заметим, что $\min f$ и $\max f$ достигаются, а значит и $\frac{1}{b-a}\int_a^b f$ достигается в некоторой точке. Возьмём её за c. ◀

𝔇𝔢𝔣 1.2.1. ${}_a^b\, f$ — среднее значение функции f на отрезке $[a,b]$

$${}_a^b\, f = \frac{1}{b-a}\int\limits_a^b f$$

𝔇𝔢𝔣 1.2.2. $f \in C[a,b]$, $\varPhi\colon [a,b] \to \mathbb{R}$. $\varPhi(x) = \int_a^x f$ — интеграл с переменным верхним пределом.

Теорема 1.2.3 (Барроу)**.** $f \in C[a,b]$

$$\varPhi'(x) = f(x)$$

▶ Пусть $x, y \in [a,b]$, $x < y$. Тогда

$$\varPhi'(x) = \lim_{y\to x}\frac{\varPhi(y)-\varPhi(x)}{y-x}$$

$$\varPhi(y)-\varPhi(x) = \int\limits_a^y f - \int\limits_a^x f = \int\limits_x^y f$$

$$\varPhi'(x) = \lim_{y\to x}\frac{1}{y-x}\int\limits_x^y f$$

$$\exists c_y \in [x,y]\colon \lim_{y\to x}\frac{1}{y-x}\int\limits_x^y f = \lim_{y\to x} f(c_y)$$

Неформальное объяснение: пусть $y \to x$, тогда $c_y \to x \Rightarrow \lim_{y\to x} f(c_y) = f(x) \Rightarrow \varPhi' = f(x)$. Формальное объяснение: докажем, что

$$\forall \varepsilon > 0, \exists \delta\colon \forall y \in (x-\delta, x+\delta), |f(c_y) - f(x)| < \varepsilon$$

По определению непрерывности в точке x:

$$\forall \varepsilon > 0, \exists \delta\colon \forall y \in (x-\delta, x+\delta), |f(y) - f(x)| < \varepsilon$$
$$y \in (x-\delta, x+\delta) \Rightarrow c_y \in (x-\delta, x+\delta) \Rightarrow |f(c_y) - f(x)| < \varepsilon$$

Аналогично можно доказать для $y < x$. ◀

𝔇𝔢𝔣 1.2.3. $f \in C[a,b]$, $\varPsi\colon [a,b] \to \mathbb{R}$. $\varPsi(x) = \int_x^b f$ — интеграл с переменным нижним пределом.

Следствие 1.2.3.1.
$$\Psi'(x) = -f(x)$$

▶

$$\Phi(x) + \Psi(x) = \int\limits_a^x f + \int\limits_x^b f = \int\limits_a^b f$$
$$\Psi(x) = \int\limits_a^b f - \Phi(x)$$
$$\Psi'(x) = 0 - \Phi'(x) = -f(x)$$

◀

Следствие 1.2.3.2.
$$f \in C\langle a, b\rangle \Rightarrow \exists F\colon F \text{ — первообразная } f$$

▶ Возьмём $c \in (a, b)$:

$$\Phi\colon [c, b\rangle \to \mathbb{R}; \Phi(x) = \int\limits_c^x f; \Phi'(x) = f(x)$$
$$\Psi\colon \langle a, c] \to \mathbb{R}; \Psi(x) = \int\limits_x^c f; \Psi'(x) = -f(x)$$
$$F(x) = \begin{cases} \Phi(x) & x \in [c, b\rangle \\ -\Psi(x) & x \in \langle a, c] \end{cases}$$

◀

Теорема 1.2.4 (Формула Ньютона — Лейбница)**.** $f \in C[a, b]$, F — первообразная f. Тогда
$$\int\limits_a^b f = F(b) - F(a)$$

▶

$$\Phi(x) = \int\limits_a^x f \text{ — первообразная } f$$
$$F(x) = \Phi(x) + C$$
$$F(b) - F(a) = \Phi(b) + C - \Phi(a) - C = \Phi(b) = \int\limits_a^b f$$

◀

Следствие 1.2.4.1. Определённый интеграл не зависит от выбора площади σ, используемой в определении.

Теорема 1.2.5 (Линейность интеграла)**.** $f, g \in C[a, b]$, $\alpha, \beta \in \mathbb{R}$. Тогда

$$\int\limits_a^b (\alpha f + \beta g) = \alpha \int\limits_a^b f + \beta \int\limits_a^b g$$

▶ Пусть F — первообразная f, G — первообразная g. Тогда $\alpha F + \beta G$ — первообразная $\alpha f + \beta g$:

$$\int\limits_a^b (\alpha f + \beta g) = \alpha F(b) + \beta G(b) - (\alpha F(a) + \beta G(a)) = \alpha(F(a) - F(b)) + \beta(G(a) - G(b)) = \alpha \int\limits_a^b f + \beta \int\limits_a^b g$$

◀

$\mathfrak{Def}$ **1.2.4.** $F\Big|_a^b$ — подстановка:

$$F\Big|_a^b = F(b) - F(a)$$

Теорема 1.2.6 (Формула интегрирования по частям)**.** $u, v \in C^1[a, b]$. Тогда

$$\int\limits_a^b uv' = uv\Big|_a^b - \int\limits_a^b u'v$$

▶

$$(uv)' = uv' + u'v$$
$$uv = \int uv' + \int u'v$$
$$\int uv' = uv - \int u'v$$
$$F(x) = u(x)v(x) - G(x)$$

$$\int\limits_a^b uv' = F(b) - F(a) = (u(b)v(b) - G(b)) - (v(a)u(a) - G(a)) =$$

$$= (u(b)v(b) - v(a)u(a)) - (G(b) - G(a)) = uv|_a^b - \int\limits_a^b u'v$$

◀

Пример 1.2.1. Интеграл логарифма:

$$\int\limits_1^n \ln x = x \ln x\Big|_1^n - \int\limits_1^n x(\ln x)' = n \ln n - 1 \ln 1 - \int\limits_1^n \frac{x}{x} = n \ln n - n + 1$$

$\mathfrak{Def}$ **1.2.5.** $\mathrm{d}x$ — параметр интегрирования:

$$\int\limits_a^b f(x)\mathrm{d}x$$

Пока не очень будем задумываться над глубоким смыслом этой записи, воспринимаем как «скобки».

$\mathfrak{Def}$ **1.2.6.** $a > b$:

$$\int\limits_a^b f = -\int\limits_b^a f$$

Теорема 1.2.7 (Замена переменных в определённом интеграле)**.** $f \in C\langle a, b\rangle$, $\varphi \in C^1[\alpha, \beta]$, $\varphi([\alpha, \beta]) \subset \langle a, b\rangle$. Тогда

$$\int\limits_\alpha^\beta f(\varphi(t))\varphi'(t)\mathrm{d}t = \int\limits_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\mathrm{d}x$$

▶ F — первообразная f, $F(\varphi(t))$ — первообразная $f(\varphi(t)) \cdot \varphi'(t)$.

$$\int\limits_\alpha^\beta f(\varphi(t))\varphi'(t)\mathrm{d}t = F(\varphi(\beta)) - F(\varphi(\alpha)) = \int\limits_{\varphi(\alpha)}^{\varphi(\beta)} f(x)dx$$

◀

Пример 1.2.2.

$$\int\limits_\alpha^\beta \frac{t}{1+t^4}\mathrm{d}t =$$

$$f(x) = \frac{1}{x^2} \quad \varphi(t) \leftrightharpoons t^2$$

$$= \int\limits_\alpha^\beta \frac{1}{2} f(\varphi(t))\varphi'(t) = \int\limits_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\mathrm{d}x = \frac{1}{2}\int_{\alpha^2}^{\beta^2} \frac{1}{1+x^2}\mathrm{d}x = \frac{1}{2}(\operatorname{arctg} t)\Big|_{\alpha^2}^{\beta^2} = \frac{\operatorname{arctg}\beta^2 - \operatorname{arctg}\alpha^2}{2}$$

Пример 1.2.3.

$$W_n = \int\limits_0^{\pi/2} \sin^n x\mathrm{d}x = \int\limits_0^{\pi/2} \cos^n \mathrm{d}x$$

$$\int\limits_0^{\pi/2} \cos^n x\mathrm{d}x = \int\limits_0^{\pi/2} \cos^{n-1} x(\sin x)'\mathrm{d}x = \cos^{n-1} x \sin x\Big|_0^{\pi/2} - \int\limits_0^{\pi/2} (\cos^{n-1} x)' \sin x\mathrm{d}x =$$

$$= -\int\limits_0^{\pi/2} (\cos^{n-1} x)' \sin x\mathrm{d}x = (n-1)\int\limits_0^{\pi/2} \cos^{n-2} x \sin^2 x\mathrm{d}x =$$

$$= (n-1)\int\limits_0^{\pi/2} \cos^{n-2} x(1-\cos^2 x)\mathrm{d}x = (n-1)(W_{n-2} - W_n)W_n = \frac{n-1}{n}W_{n-2}$$

$$W_0 = \frac{\pi}{2}$$

$$W_1 = 1$$

$$W_{2n} = \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdot \frac{2n-5}{2n-4} \cdot \ldots \cdot \frac{1}{2} \cdot \frac{\pi}{2} = \frac{(2n-1)!!}{(2n)!!} \cdot \frac{\pi}{2}$$

$$W_{2n+1} = \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdot \ldots \cdot \frac{2}{3} \cdot 1 = \frac{(2n)!!}{(2n+1)!!}$$

Утверждение 1.2.1 (Формула Валлиса).

$$\lim_{n\to\infty} \frac{(2n)!!}{(2n-1)!!} \cdot \frac{1}{\sqrt{2n+1}} = \sqrt{\frac{\pi}{2}}$$

▶

$$\cos^{2n} x \geqslant \cos^{2n+1} x \geqslant \cos^{2n+2} x \quad \text{при } x \in [0, \pi/2]$$
$$W_{2n} \geqslant W_{2n+1} \geqslant W_{2n+2}$$
$$\frac{(2n-1)!!}{(2n)!!} \cdot \frac{\pi}{2} \geqslant \frac{(2n)!!}{(2n+1)(2n-1)!!} \geqslant \frac{(2n+1)!!}{(2n+2)!!}\frac{\pi}{2}$$
$$\frac{\pi}{2} \geqslant \frac{((2n)!!)^2}{(2n+1)((2n-1)!!)^2} \geqslant \frac{2n+1}{2n+2} \cdot \frac{\pi}{2} \to \frac{\pi}{2}$$
$$\lim_{n\to\infty} \frac{((2n)!!)^2}{(2n+1)((2n-1)!!)^2} = \frac{\pi}{2}$$

◀

Следствие 1.2.7.1. $C_{2n}^n \sim \frac{4^n}{\sqrt{\pi n}}$

▶

$$(2n)!! = 2^n n! \quad (2n-1)!!\,(2n)!! = (2n)!$$
$$\frac{(2n)!!}{(2n-1)!!} = \frac{((2n)!!)^2}{(2n)!} = \frac{(2^n n!)^2}{(2n)!} = 4^n \frac{n!\,n!}{(2n)!} = \frac{4^n}{C_{2n}^n}$$
$$\frac{4^n}{C_{2n}^n} \cdot \frac{1}{\sqrt{2n+1}} \sim \sqrt{\frac{\pi}{2}}$$
$$C_{2n}^n \sim \frac{4^n}{\sqrt{\pi/2}} \cdot \underbrace{\frac{1}{\sqrt{2n+1}}}_{\sim \frac{1}{\sqrt{2n}}} \sim \frac{4^n}{\sqrt{\pi n}} \Rightarrow C_{2n}^n \sim \frac{4^n}{\sqrt{\pi n}}$$

◀

Пример 1.2.4.

$$H_n = \frac{1}{n!} \int\limits_0^{\pi/2} \left(\left(\frac{\pi}{2}\right)^2 - x^2 \right)^n \cos x \mathrm{d}x$$

1. $H_n \to 0$

$$0 < H_n \leqslant \frac{1}{n!}\frac{\pi}{2} \max_{x\in[0,\pi/2]} \left(\left(\frac{\pi}{2}\right)^2 - x^2 \right)^n \cos x \leqslant \frac{1}{n!}\frac{\pi}{2}\left(\frac{\pi}{2}\right)^{2n} \to 0$$

   Так как показательная функция растёт медленнее факториала.

2. $H_n = (4n-2)H_{n-1} - \pi^2 H_{n-2}$, $H_0 = 1$, $H_1 = 2$.

<u>Пояснение:</u>

$$n!\,H_n = \int\limits_0^{\pi/2} \left((\pi/2)^2 - x^2\right)^n (\sin x)' \mathrm{d}x =$$

$$= (\text{Подстановка, которая равна } 0) - \int\limits_0^{\pi/2} (\left((\pi/2)^2 - x^2\right)^n)' \sin x \mathrm{d}x =$$

$$= -2n \int\limits_0^{\pi/2} \left((\pi/2)^2 - x^2\right)^{n-1} x (\cos x)' \mathrm{d}x$$

Затем ещё раз интегрируем по частям, используя равенство $x^2 = (\pi/2)^2 - ((\pi/2)^2 - x^2)$

3. $H_n = P_n(\pi^2)$, где P_n - многочлен n-ой степени с целыми коэффициентами.

Теорема 1.2.8. π и π^2 иррациональны.

▶ От противного. Пусть $\pi = \frac{a}{b}$.

$$H_n = P_n\left(\frac{a^2}{b^2}\right) \Rightarrow b^{2n} H_n = b^{2n} P_n \left(\frac{a^2}{b^2}\right) - \text{целое число.}$$

$$H_n > 0 \Rightarrow b^{2n} H_n > 0$$

$$1 \leqslant b^{2n} H_n \leqslant b^{2n} \frac{\pi}{2} \left(\frac{\pi}{2}\right)^{2n} \frac{1}{n!} = \frac{\pi}{2} \frac{(b \cdot \pi/2)^{2n}}{n!} \to 0$$

◀

Теорема 1.2.9 (Формула Тейлора с остатком в интегральной форме)**.** $f \in C^{n+1} \langle a, b \rangle$, $x_0 \in \langle a, b \rangle$. Тогда

$$f(x) = T_{n,x_0} f(x) + \frac{1}{n!} \int\limits_{x_0}^{x} (x-t)^n f^{(n+1)}(t) \mathrm{d}t$$

▶ Докажем по индукции.

База $n = 0$:

$$f(x) = f(x_0) + \int\limits_{x_0}^{x} f'(t) \mathrm{d}t = f(x_0) + f(x) - f(x_0)$$

Переход:

$$f(x) = T_{n,x_0} f(x) + \frac{1}{n!} \int\limits_{x_0}^{x} (x-t)^n f^{(n+1)}(t) \mathrm{d}t = T_{n,x_0} f(x) - \frac{1}{(n+1)!} \int\limits_{x_0}^{x} f^{(n+1)}(t) ((x-t)^{n+1})' \mathrm{d}t =$$

$$= T_{n,x_0} f(x) - \left(\frac{f^{(n+1)}(t)}{(n+1)!} (x-t)^{(n+1)} \Big|_{x_0}^{x} - \frac{1}{(n+1)!} \int\limits_{x_0}^{x} (x-t)^{n+1} f^{n+2}(t) \mathrm{d}t \right) =$$

$$= T_{n,x_0} f(x) + \frac{f^{(n+1)} x_0}{(n+1)!} (x - x_0)^{(n+1)} + \frac{1}{(n+1)!} \int\limits_{x_0}^{x} (x-t)^{n+1} f^{(n+2)}(t) \mathrm{d}t$$

◀

1.3. Интегральные суммы.

𝔇𝔢𝔣 1.3.1. $f\colon X \to Y$, X, Y — метрические пространства.

$$\omega_f(\delta) = \sup\{\rho_y(f(x_1), f(x_2)) \mid \rho_x(x_1, x_2) < \delta\}$$

$\omega_f(\delta)$ — модуль непрерывности.

Свойства:

1. $\omega_f(\delta) \geqslant 0$
2. $\omega_f(0) = 0$
3. $\omega_f(\delta) \nearrow$
4. f равномерно непрерывно на X тогда и только тогда, когда w_f непрерывна в 0.

   ▶ **⇒:** f равномерно непрерывно на X:

   $$\forall \varepsilon > 0, \exists \delta > 0\colon \forall x_1, x_2 \in X\colon \rho_x(x_1, x_2) \leqslant \delta, \rho_y(f(x_1), f(x_2)) < \varepsilon \Rightarrow$$
   $$\Rightarrow \underbrace{\sup\{\rho_y(f(x_1, f(x_2)) \mid \rho_x(x_1, x_2) \leqslant \delta\}}_{=w_f(\delta)} \leqslant \varepsilon$$
   $$\forall \varepsilon > 0, \exists \delta > 0\colon w_f(\delta) \leqslant \varepsilon \Rightarrow 0 \leqslant w_f(\gamma) \leqslant w_f(\delta) \leqslant \varepsilon \quad \text{при } \gamma \leqslant \delta$$

   Таким образом w_f непрерывна в 0.

   ⇐:

   $$\forall \varepsilon < 0, \exists \delta > 0\colon w_f(\delta) < \varepsilon \Rightarrow$$
   $$\Rightarrow ((\rho_x(x_1, x_2) \leqslant \delta) \Rightarrow (\rho_y(f(x_1), f(x_2) < \sup\{\rho_y(f(x_1), f(x_2)) \mid \rho_x(x_1, x_2) \leqslant \delta\} < \varepsilon))$$

   ◀

5. Если X — компакт, то f непрерывна на X тогда и только тогда, когда w_f непрерывна в 0.

𝔇𝔢𝔣 1.3.2. $[a, b]$. Дробление отрезка (иногда называют пунктиром):

$$\tau\colon x_0 = a < x_1 < x_2 < \dots < x_n = b$$

Мелкость дробления (ранг дробления):

$$|\tau| = \max_{k=1..n} (x_k - x_{k-1})$$

Оснащение дробления:

$$\xi_1, \xi_2, \dots, \xi_n \quad \xi_k \in [x_{k-1}, x_k]$$

𝔇𝔢𝔣 1.3.3. Сумма Римана: $f\colon [a, b] \to \mathbb{R}$

$$\sigma(f, \tau, \xi) = \sum_{k=1}^{n} f(\xi_k)\,(x_k - x_{k-1})$$

Теорема 1.3.1. $f \in C[a,b]$.

$$\forall \varepsilon > 0, \exists \delta > 0 \colon \forall (\tau, \xi), |\tau| < \delta \Rightarrow \left| \sigma(f, \tau, \xi) - \int\limits_a^b f(x)\mathrm{d}x \right| < \varepsilon$$

В частности, если (τ_n, ξ_n) — последовательность дроблений такая, что $|\tau_n| \to 0$, то

$$\sigma(f, \tau_n, \xi_n) \to \int\limits_a^b f(x)\mathrm{d}x$$

▶

$$\varDelta = \sigma(f, \tau, \xi) - \int\limits_a^b f(x)\mathrm{d}x = \sum_{k=1}^n f(\xi_k)(x_k - x_{k-1}) - \sum_{k=1}^n \int\limits_{x_{k-1}}^{x_k} f(x)\mathrm{d}x =$$

$$= \sum_{k=1}^n \left(\int\limits_{x_{k-1}}^{x_k} f(\xi_k)\mathrm{d}x - \int\limits_{x_{k-1}}^{x_k} f(x)\mathrm{d}x \right) = \sum_{k=1}^n \int\limits_{x_{k-1}}^{x_k} (f(\xi_k) - f(x))\mathrm{d}x$$

$$|\varDelta| \leqslant \sum_{k=1}^n \left| \int\limits_{x_{k-1}}^{x_k} (f(\xi_k) - f(x))\mathrm{d}x \right| \leqslant \sum_{x=1}^n \int\limits_{x_{k-1}}^{x_k} |f(\xi_k) - f(x)|\mathrm{d}x \leqslant$$

$$\overset{|\xi_k - x| < \tau}{\leqslant} \sum_{k=1}^n \int\limits_{x_{k-1}}^{x_k} w_f(|\tau|)\mathrm{d}x = \sum_{k=1}^n w_f(|\tau|)(x_k - x_{k-1})$$

$$|\varDelta| \leqslant w_f(|\tau|)(b-a) < \varepsilon(b-a)$$

◀

𝔇𝔢𝔣 1.3.4. $f \colon [a,b] \to \mathbb{R}$. f интегрируема по Риману, если

$$\exists I \colon \forall \varepsilon, \exists \delta > 0 \colon \forall |\tau| < \delta |\sigma(f, \tau, \xi) - I| < \varepsilon$$

и I называется интегралом Римана.

$$\int\limits_a^b f(x)\mathrm{d}x$$

Замечание 1.3.1. В этом определении не обязательна непрерывность функции, но для не непрерывных будет непонятно, существует ли интеграл.

Пример 1.3.1. Выясним, с какой скоростью возрастают суммы.

$$S_p(n) = 1^p + 2^p + \dots + n^p$$

$$S_p(n) < n \cdot n^p = n^{p+1}$$

$$S_p(n) > \left(\frac{n}{2}\right)^p + \dots + n^p > \frac{n}{2} \cdot \left(\frac{n}{2}\right)^p = \frac{n^{p+1}}{2^{p+1}}$$

$$\lim_{n \to \infty} \frac{S_p(n)}{n^{p+1}} = ?$$

$$\frac{S_p(n)}{n^{p+1}} = \frac{1}{n} \cdot \left(\left(\frac{1}{n}\right)^p + \left(\frac{2}{n}\right)^p + \left(\frac{3}{n}\right)^p + \dots + \left(\frac{n}{n}\right)^p \right)$$

$f(x) = x^p$ на $[0,1]$, $p \geqslant 0$.

$$x_0 = 0, x_1 = \frac{1}{n}(\xi_1), x_2 = \frac{2}{n}(\xi_2), \dots, x_n = 1(\xi_n)$$

Мелкость равна $\frac{1}{n}$

$$\sigma(f,\tau,\xi) \to \int\limits_0^1 x^p \mathrm{d}x = (\frac{x^{p+1}}{p+1})\Big|_0^1 = \frac{1}{p+1}$$

$$\sigma(f,\tau,\xi) = \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}) = \frac{1}{n}\sum_{k=1}^{n}\left(\frac{k}{n}\right)^p$$

$$\lim_{n\to\infty}\frac{S_p(n)}{n^{p+1}} = \frac{1}{p+1} \Rightarrow S_p(n) \backsim \frac{n^{p+1}}{p+1}$$

Оценим модуль непрерывности дифференцируемой функции:

$$f(x_1) - f(x_2) = f'(c)(x_1 - x_2)$$
$$|f(x_1) - f(x_2) \leqslant \max f'|x_1 - x_2||$$
$$\omega_f(\delta) \leqslant \max f' \cdot \delta$$

Теорема 1.3.2. Оценка погрешности в формуле трапеций $f \in C^2[a,b]$.

$$\left|\int\limits_a^b f(x)\mathrm{d}x - \sum_{k=1}^{n}\frac{f(x_{k-1}) - f(x_k)}{2}(x_k - x_{k-1})\right| \leqslant \frac{|\tau|^2}{8}\int\limits_a^b |f''(x)|\,\mathrm{d}x$$

▶ $y_k = \frac{x_{k-1}+x_k}{2}$

$$\Delta \leftrightharpoons \left(\int\limits_a^b f(x)\mathrm{d}x - \sum_{k=1}^{n}\frac{f(x_{k-1}) + f(x_k)}{2}(x_k - x_{k-1})\right) =$$

$$= \sum_{k=1}^{n}\left(\int\limits_{x_{k-1}}^{x_k} f(x)\mathrm{d}x - \frac{f(x_{k-1}) + f(x_k)}{2}(x_k - x_{k-1})\right) = \sum_{k=1}^{n}\int\limits_{x_{k-1}}^{x_k}\left(f(x) - \frac{f(x_{k-1}) - f(x_k)}{2}\right)\mathrm{d}x$$

$$\int\limits_{x_{k-1}}^{x_k} f(x)\mathrm{d}x = \int\limits_{x_{k-1}}^{x_k} f(x)(x-y_k)'\mathrm{d}x = f(x)(x-y_k)\Big|_{x=x_{k-1}}^{x_k} - \int\limits_{x_{k-1}}^{x_k} f'(x)(x-y_k)\mathrm{d}x =$$

$$= f(x_k)\frac{x_k - x_{k-1}}{2} - f(x_{k-1})\frac{x_{k-1} - x_k}{2} - \int\limits_{x_{k-1}}^{x_k} f'(x)(x-y_k)\mathrm{d}x =$$

$$= \frac{f(x_k) + f(x_{k-1})}{2}(x_k - x_{k-1}) - \int\limits_{x_{k-1}}^{x_k} f'(x)(x-y_k)\mathrm{d}x$$

$$\varDelta = -\sum_{k=1}^{n}\int\limits_{x_{k-1}}^{x_k} f'(x)(x-y_k)\mathrm{d}x$$

$$((x-x_{k-1})(x_k-x))' = (x_k-x)(1)+(x-x_{k-1})(-1) = x_{k-1}+x_k-2x = 2(y_k-x)$$

$$-\int\limits_{x_{k-1}}^{x_k} f'(x)(x-y_k)\mathrm{d}x = \int\limits_{x_{k-1}}^{x_k} f'(x)\left(\frac{(x-x_{k-1})(x_k-x)}{2}\right)'\mathrm{d}x =$$

$$= \underbrace{\left(f'(x)\left(\frac{(x-x_{k-1})(x_k-x)}{2}\right)\right)\Bigg|_{x=x_{k-1}}^{x_k}}_{=0} - \int\limits_{x_{k-1}}^{x_k} f''(x)\frac{(x-x_{k-1})(x_k-x)}{2}\mathrm{d}x$$

$$\varDelta = -\sum_{k=1}^{n}\int\limits_{x_{k-1}}^{x_k} f''(x)\frac{(x-x_{k-1})(x_k-x)}{2}\mathrm{d}x$$

$$\frac{(x-x_{k-1})(x_k-x)}{2} \leqslant \left(\frac{x_k-x_{k-1}}{2}\right)^2 \leqslant \frac{|\tau^2|}{4}$$

$$|\varDelta| \leqslant \sum_{k=1}^{n}\int\limits_{x_{k-1}}^{x_k} |f''(x)|\,\mathrm{d}x\cdot\frac{|\tau|^2}{8} = \frac{|\tau|^2}{8}\int\limits_a^b |f''(x)|\,\mathrm{d}x$$

◀

𝔇𝔢𝔣 1.3.5.

$$\sum_{i=1}^{n}{}' a_i = \frac{a_1+a_2}{2}+\frac{a_2+a_3}{2}+\dots+\frac{a_{n-1}+a_n}{2} = \frac{a_1}{2}+a_2+\dots+a_{n-1}+\frac{a_n}{2}$$

Следствие 1.3.2.1. $f\in C^2[a,b]$.

$$\left|\int\limits_a^b f(x)\mathrm{d}x - \frac{b-a}{n}\sum_{k=0}^{n}{}' f\left(a+\frac{n}{k}(b-a)\right)\right| \leqslant \frac{(b-a)^2}{8n^2}\int\limits_a^b |f''(x)|\,\mathrm{d}x$$

▶ В качестве x_k возьмём точки на равных расстояниях

$$x_k = a+\frac{k}{n}(b-a)$$

Тогда

$$|\tau| = \frac{b-a}{n}$$

$$\sum_{k=1}^{n}\frac{f(x_{k-1})-f(x_k)}{2}(x_k-x_{k-1}) = \frac{b-a}{n}\sum_{k=0}^{n}\frac{f(x_{k-1})+f(x_k)}{2} = \frac{b-a}{n}\sum_{k=0}^{n}{}' f\left(a+\frac{k}{n}(b-a)\right)$$

◀

Теорема 1.3.3 (Формула Эйлера-Маклорена)**.**

$$\sum_{k=m}^{n}{}' f(k) = \int\limits_m^n f(x)\mathrm{d}x + \frac{1}{2}\int\limits_n^m f''(x)\{x\}(1-\{x\})\mathrm{d}x$$

▶

$$\int\limits_m^n f(x)\mathrm{d}x - \sum_{k=m}^{n}{}' f(k) = \int\limits_m^n f(x)\mathrm{d}x - \sum_{k=m+1}^{n} \frac{f(k-1)+f(k)}{2} =$$

$$= -\sum_{k=m+1}^{n} \int\limits_{k-1}^{k} f''(x)\frac{(x-(k-1))(k-x)}{2}\mathrm{d}x = -\sum_{k=m+1}^{n} \int\limits_{k-1}^{k} f''(x)\frac{\{x\}(1-\{x\})}{2}\mathrm{d}x =$$

$$= -\int\limits_m^n f''(x)\frac{\{x\}(1-\{x\})}{2}\mathrm{d}x$$

◀

Пример 1.3.2. 1.

$$S_p(n) = 1^p + 2^p + ... + n^p$$

$$S_p(n) = \sum_{k=1}^{n}{}' k^p + \frac{1^p+n^p}{2} = \int_1^n x^p\mathrm{d}x + \frac{1}{2}\int_1^n p(p-1)x^{p-2}\{x\}(1-\{x\}) + \frac{1+n^p}{2}$$

$$= \frac{1+n^p}{2} + \frac{n^{p+1}-1}{p+1} + \frac{p(p-1)}{2}\int_1^n x^{p-2}\{x\}(1-\{x\})$$

Пусть $p \in (-1, 1)$, тогда

$$0 \leqslant \int_1^n x^{p-2}\{x\}(1-\{x\})\mathrm{d}x \leqslant \int_1^n x^{p-2} = \frac{n^{p-1}-1}{p-1} = O(1)$$

$$S_p(n) = \frac{n^{p+1}}{p+1} + \frac{n^p}{2} + O(1)$$

Если $p = 1$, тогда

$$S_p(n) = \frac{n^{p+1}}{p+1} + \frac{n^p}{2} + O(\log n)$$

Если $p > 1$, тогда

$$S_p(n) = \frac{n^{p+1}}{p+1} + \frac{n^p}{2} + O(n^{p-1})$$

2. Рассмотрим гармонические числа.

$$H_n = \sum_{k=1}^{n}{}' \frac{1}{k} + \frac{1+\frac{1}{n}}{2} = \frac{1+\frac{1}{n}}{2} + \int_1^n \frac{\mathrm{d}x}{x} + \frac{1}{2}\int_1^n \frac{2\{x\}(1-\{x\})}{x^3}\mathrm{d}x =$$

$$= \frac{1}{2} + \frac{1}{2n} + \ln n + \int_1^n \frac{\{x\}(1-\{x\})}{x^3}\mathrm{d}x$$

$$H_n = \ln n + O(1)$$

$$H_n - \ln n = \frac{1}{2} + \frac{1}{2n} + a_n$$

$$a_n = \int_1^n \frac{\{x\}(1-\{x\})}{x^3}\mathrm{d}x$$

$$a_n \leqslant a_{n+1}$$

Поймем, что a_n ограничено

$$a_n = \int_1^n \frac{\{x\}(1-\{x\})}{x^3}\mathrm{d}x \leqslant \int_1^n \frac{\mathrm{d}x}{x^3} = -\frac{1}{2x^2}\Big|_{x=1}^{x=n} = \frac{1}{2} - \frac{1}{2n} \leqslant \frac{1}{2}$$

$$\Rightarrow \exists \lim_{n\to\infty} a_n \Rightarrow \exists \lim_{n\to\infty}(H_n - \ln n) = \gamma$$

$\gamma = 0,57721566\ldots$ — постоянная Эйлера.

Итог: $H_n = \ln n + \gamma + o(1)$

3. Формула Стирлинга.

$$\ln n! = \ln 1 + \ln 2 + \ldots + \ln n$$

$$\ln n! = \sum_{k=1}^{n}{}' \ln k + \frac{\ln n}{2} = \int_1^n \ln x\mathrm{d}x - \frac{1}{2}\int_1^n \frac{\{x\}(1-\{x\})}{x^2}\mathrm{d}x + \frac{\ln n}{2} =$$

$$= n\ln n - n + \frac{\ln n}{2} + 1 - \frac{1}{2}\int_1^n \frac{\{x\}(1-\{x\})}{x^2}\mathrm{d}x$$

$$\int_1^n \frac{\{x\}(1-\{x\})}{x^2}\mathrm{d}x = b_n$$

$$b_n \leqslant b_{n+1}$$

$$b_n = \int_1^n \frac{\{x\}(1-\{x\})}{x^2}\mathrm{d}x \leqslant \int_1^n \frac{\mathrm{d}x}{x^2} = -\frac{1}{x}\Big|_{x=1}^{x=n} = 1 - \frac{1}{n} \leqslant 1$$

$$\Rightarrow \exists \lim_{n\to\infty} b_n$$

$$\ln n! = n\ln n + \frac{\ln n}{2} - n + C + o(1)$$

$$n! = n^n e^{-n}\sqrt{n}e^c e^{o(1)} = n^n e^{-n}\sqrt{n}e^c(1+o(1)) \sim n^n e^{-n}\sqrt{n}e^c$$

$$\frac{4^n}{\sqrt{\pi n}} = C_{2n}^n = \frac{(2n)!}{(n!)^2} \sim \frac{(2n)^{2n}e^{-2n}\sqrt{2n}e^c}{(n^n e^{-n}\sqrt{n}e^c)^2} = \frac{4^n n^{2n}e^{-2n}\sqrt{2n}e^c}{n^{2n}e^{-2n}ne^{2c}} = \frac{4^n\sqrt{2n}}{ne^c}$$

$$\frac{4^n}{\sqrt{\pi n}} \sim \frac{4^n\sqrt{2n}}{ne^c}$$

$$e^c \sim \sqrt{2\pi} \Rightarrow e^c = \sqrt{2\pi}$$

Формула Стирлинга:

$$n! \sim n^n e^{-n}\sqrt{2\pi n}$$

Замечание 1.3.2. Можно использовать для биномиальных коэффициентов. Записать их через факториал и расписать.

1.4. Длина кривой

𝔇𝔢𝔣 1.4.1. Путь — непрерывное отображение $\gamma\colon [a,b] \to \mathbb{R}^d$.

𝔇𝔢𝔣 1.4.2. Начало пути — $\gamma(a)$.

𝔇𝔢𝔣 1.4.3. Конец пути — $\gamma(b)$.

𝔇𝔢𝔣 1.4.4. Носитель пути — $\gamma([a,b])$.

𝔇𝔢𝔣 1.4.5. Замкнутый путь — путь, в котором $\gamma(a) = \gamma(b)$.

𝔇𝔢𝔣 1.4.6. Несамопересекающийся (простой) путь — путь, в котором

$$\gamma(t_1) = \gamma(t_2) \Rightarrow \left[\begin{array}{c} t_1 = t_2 \\ t_1 = a, t_2 = b \\ t_1 = b, t_2 = a \end{array}\right.$$

𝔇𝔢𝔣 1.4.7. C^r-гладкий путь — путь

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_d \end{pmatrix}$$

в котором

$$\forall i = 1..d, \gamma_i \in C^r[a, b]$$

𝔇𝔢𝔣 1.4.8. Гладкий путь — C^1-гладкий путь.

𝔇𝔢𝔣 1.4.9. Кусочно-гладкий путь — путь $\gamma\colon [a, b] \to \mathbb{R}^d$ такой, что

$$\exists c_0 = a \leqslant c_1 \leqslant c_2 \leqslant ... \leqslant c_{n-1} \leqslant c_n = b\colon \gamma\Big|_{[c_i, c_{i+1}]} \text{— гладкий путь}$$

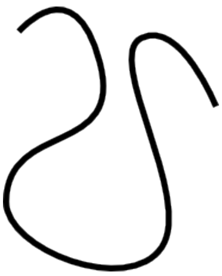


Рис. 1.1: Простой путь

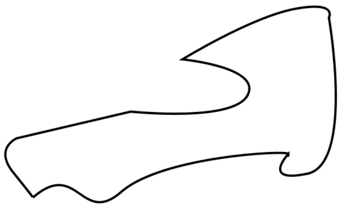


Рис. 1.2: Замкнутый путь

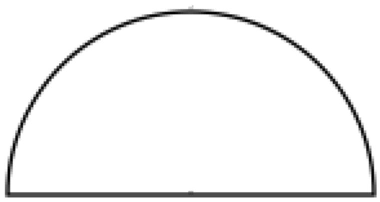


Рис. 1.3: Кусочно-гладкий путь

𝔇𝔢𝔣 1.4.10. $\gamma\colon [a, b] \to \mathbb{R}^d$, $\tilde{\gamma}\colon [c, d] \to \mathbb{R}^d$. Если

$$\exists \tau\colon [a, b] \to [c, d]\colon \tau(a) = c \land \tau(b) = d \land \tau \uparrow \land \gamma = \tilde{\gamma} \circ \tau$$

то γ и $\tilde{\gamma}$ — эквивалентные пути.

Пример 1.4.1.

$$\gamma(t) = (\cos t, \sin t), 0 \leqslant t \leqslant \pi$$
$$\tilde{\gamma}(x) = (-x, \sqrt{1 - x^2}), -1 \leqslant x \leqslant 1$$
$$\tau(t) = -\cos t$$

Свойства:

1. Носители эквивалентных путей одинаковы
2. Это отношение эквивалентности

3. Одинаковый порядок точек

𝔇𝔢𝔣 1.4.11. Кривая — класс эквивалентности путей.

𝔇𝔢𝔣 1.4.12. Параметризация кривой — конкретный представитель класса.

𝔇𝔢𝔣 1.4.13. Носитель кривой — носитель пути из класса эквивалентности.

𝔇𝔢𝔣 1.4.14. Гладкая (C^r-гладкая) кривая — кривая, у которой существует гладкая (C^r-гладкая) параметризация.

𝔇𝔢𝔣 1.4.15. Кусочно-гладкая кривая — кривая, у которой существует кусочно-гладкая параметризация.

Пример 1.4.2. $\gamma(t) = (\cos t, \sin t)$, $0 \leqslant t \leqslant 2\pi$ — гладкая параметризация. $\tilde{\gamma}(x) = (\cos\sqrt{t}, \sin\sqrt{t})$, $0 \leqslant t \leqslant (2\pi)^2$ — негладкая параметризация.

𝔇𝔢𝔣 1.4.16. $\gamma\colon [a,b] \to \mathbb{R}^d$ — параметризация. Противоположная кривая — кривая, ориентированная в другую сторону:

$$\tilde{\gamma}\colon [a,b] \to \mathbb{R}^d$$
$$\tilde{\gamma}(t) = \gamma(b+a-t)$$
$$\tilde{\gamma}(a) = \gamma(b) \quad \tilde{\gamma}(b) = \gamma(a)$$

Замечание 1.4.1. У разных кривых может быть один и тот же носитель.

$$\gamma(t) = (\cos t, \sin t), 0 \leqslant t \leqslant 2\pi$$
$$\tilde{\gamma}(t) = (\cos t, \sin t), 0 \leqslant t \leqslant 3\pi$$

𝔇𝔢𝔣 1.4.17. $\gamma\colon [a,b] \to \mathbb{R}^d$. Разобьем отрезок $[a,b]$ на n кусков.

$$t_0 = a \leqslant t_1 \leqslant t_2 \leqslant \ldots \leqslant t_{n-1} \leqslant t_n = b$$

Проведем ломаную через точки

$$\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_{n-1}), \gamma(t_n)$$

Длина кривой $l(\gamma)$ — супремум длин таких вписанных ломаных.

Свойства:

1. Длина не зависит от параметризации.
2. Длины противоположных путей (кривых) равны.
3. $l(\gamma) \geqslant \|\gamma(b) - \gamma(a)\|$ — длина отрезка, соединяющего концы.
4. $l(\gamma) \geqslant$ длины любой вписанной в γ ломаной.

Теорема 1.4.1. $\gamma\colon [a,b] \to \mathbb{R}^d$, $c \in (a,b)$, $\gamma_1 = \gamma\big|_{[a,c]}$, $\gamma_2 = \gamma\big|_{[c,b]}$. Тогда $l(\gamma) = l(\gamma_1) + l(\gamma_2)$.

▶ $l(\gamma) \geqslant l(\gamma_1) + l(\gamma_2)$**:** $l(\gamma_1)$ — супремум вписанных в γ_1 ломаных. $l(\gamma_2)$ — супремум вписанных в γ_2 ломаных. Объединение этих ломаных $l_1 \cup l_2$ — ломаная, вписанная в $\gamma_1 \cup \gamma_2 = \gamma$.

$$|l_1 \cup l_2| \leqslant l(\gamma)$$

Переходим к sup по l_1, а затем к sup по l_2.

$l(\gamma) \leqslant l(\gamma_1) + l(\gamma_2)$**:** Берем ломаную l, вписанную в γ. Делаем из нее ломаную $\tilde{l} = l_1 \cup l_2$, добавляя точку разбиения c.

$$|\tilde{l}| \geqslant |l|$$
$$l(\gamma_1) + l(\gamma_2) \geqslant |l_1||l_2| = |\tilde{l}| \geqslant |l|$$

Переходим к sup по l.

◀

𝔇𝔢𝔣 **1.4.18.** Кривая спрямляемая, если её длина конечна.

Теорема 1.4.2. $\gamma\colon [a,b] \to \mathbb{R}^d$ — гладкая кривая, $\|x\| = \sqrt{x_1^2 + \dots + x_d^2}$,

$$\gamma' = \begin{pmatrix} \gamma_1' \\ \gamma_2' \\ \vdots \\ \gamma_d' \end{pmatrix}$$

Тогда

$$l(\gamma) = \int\limits_a^b \|\gamma'(t)\| \mathrm{d}t = \int\limits_a^b \sqrt{\gamma_1'(t)^2 + \gamma_2'(t)^2 + \dots + \gamma_d'(t)^2} \mathrm{d}t$$

▶ Рассмотрим какой-то подотрезок Δ, лежащий в отрезке $[a,b]$, и сужение $\gamma\big|_\Delta\colon \Delta \to \mathbb{R}^d$.

$$m_\Delta^{(i)} = \min_{t\in\Delta} |\gamma_i'(t)|\,,\ \text{достигается в точке}\ \zeta_\Delta^{(i)}, m_\Delta^2 = \sum_{i=1}^d \left(m_\Delta^{(i)}\right)^2$$
$$M_\Delta^{(i)} = \max_{t\in\Delta} |\gamma_i'(t)|\,,\ \text{достигается в точке}\ \eta_\Delta^{(i)}, M_\Delta^2 = \sum_{i=1}^d \left(M_\Delta^{(i)}\right)^2$$

Лемма 1.4.1.

$$m_\Delta|\Delta| \leqslant l\left(\gamma\big|_\Delta\right) \leqslant M_\Delta|\Delta|$$

▶ Обозначим отрезок Δ за $[\alpha, \beta]$. Рассмотрим разбиение Δ:

$$t_0 = \alpha, t_1, t_2, \dots, t_{n-1}, t_n = \beta$$

и ломаную, построенную на точках $\gamma(t_0), \gamma(t_1), \dots, \gamma(t_n)$.
Пусть длина i-го звена равна a_i. Длина ломаной тогда будет $\sum_{i=1}^n a_i$.

$$a_i^2 = (\gamma_1(t_i) - \gamma_1(t_{i-1}))^2 + \dots + (\gamma_d(t_i) - \gamma_d(t_{i-1}))^2$$
$$\gamma_1(t_i) - \gamma_1(t_{i-1}) = \gamma_1'(\mu_{i1})(t_i - t_{i-1}) \leqslant M_\Delta^{(1)}(t_i - t_{i-1})$$
$$\vdots$$
$$\gamma_d(t_i) - \gamma_d(t_{i-1}) = \gamma_d'(\mu_{id})(t_i - t_{i-1}) \leqslant M_\Delta^{(d)}(t_i - t_{i-1})$$
$$a_i^2 \leqslant (M_\Delta^{(1)}(t_i - t_{i-1}))^2 + \dots + (M_\Delta^{(d)}(t_i - t_{i-1}))^2 = M_\Delta^2(t_i - t_{i-1})^2$$

Аналогично, $a_i^2 \geqslant m_\Delta^2(t_i - t_{i-1})^2$.

$$m_\Delta(t_i - t_{i-1}) \leqslant a_i \leqslant M_\Delta(t_i - t_{i-1})$$

Просуммируем по всем i:

$$m_\Delta \sum_{i=1}^{n}(t_i - t_{i-1}) \leqslant \sum_{i=1}^{n} a_i \leqslant M_\Delta \sum_{i=1}^{n}(t_i - t_{i-1})$$
$$m_\Delta(\beta - \alpha) \leqslant \sum_{i=1}^{n} a_i \leqslant M_\Delta(\beta - \alpha)$$

Переходим к супремумам.

$$m_\Delta|\Delta| \leqslant l(\gamma|_\Delta) \leqslant M_\Delta|\Delta|$$

◀

Разобьем отрезок $[a, b]$ на n кусков.

$$t_0 = a, t_1, t_2, \dots, t_{n-1}, t_n = b$$

Обозначим $m_i = m_{[t_{i-1}, t_i]}$, $M_i = M_{[t_{i-1}, t_i]}$.

$$l(\gamma) = l\left(\gamma\big|_{[t_0,t_1]}\right) + l\left(\gamma\big|_{[t_1,t_2]}\right) + \dots + l\left(\gamma\big|_{[t_{n-1},t_n]}\right)$$
$$m_i(t_i - t_{i-1}) \leqslant l(\gamma|_{[t_{i-1},t_i]}) \leqslant M_i(t_i - t_{i-1})$$

Просуммируем:

$$\sum_{i=1}^{n} m_i(t_i - t_{i-1}) \leqslant l(\gamma) \leqslant \sum_{i=1}^{n} M_i(t_i - t_{i-1})$$

Посмотрим на сумму Римана для интеграла $\int\limits_a^b \|\gamma'(t)\| \mathrm{d}t$ в точках $\xi_i \in [t_{i-1}, t_i]$:

$$\sum_{i=1}^{n} \|\gamma'(\xi_i)\|(t_i - t_{i-1})$$
$$\|\gamma'(\xi_i)\|^2 = |\gamma_1'(\xi_i)|^2 + |\gamma_2'(\xi_i)|^2 + \dots + |\gamma_d'(\xi_i)|^2$$
$$|\gamma_j'(\xi_i)|^2 \leqslant (M^{(j)}_{[t_{i-1},t_i]})^2$$

Просуммируем по $j = 1..d$:

$$\|\gamma'(\xi_i)\|^2 \leqslant M^2_{[t_{i-1},t_i]} = M_i^2$$
$$m_i \leqslant \|\gamma'(\xi_i)\| \leqslant M_i$$

Просуммируем по i:

$$\sum_{i=1}^{n} m_i(t_i - t_{i-1}) \leqslant \sum_{i=1}^{n} \|\gamma'(\xi_i)\|(t_i - t_{i-1}) \leqslant \sum_{i=1}^{n} M_i(t_i - t_{i-1})$$

Хотим доказать, что $\sum\limits_{i=1}^{n}(M_i - m_i)(t_i - t_{i-1}) \to 0$.

$$M_i - m_i = \frac{M_i^2 - m_i^2}{M_i + m_i} = \frac{\sum\limits_{j=1}^{d}((M^{(j)}_{[t_{i-1},t_i]})^2 - (m^{(j)}_{[t_{i-1},t_i]})^2)}{M_i + m_i} = \sum_{j=1}^{d}(M^{(j)}_{[t_{i-1},t_i]} - m^{(j)}_{[t_{i-1},t_i]}) \underbrace{\frac{M^{(j)}_{[t_{i-1},t_i]} + m^{(j)}_{[t_{i-1},t_i]}}{M_i + m_i}}_{\leqslant 1}$$

$$M_i - m_i \leqslant \sum_{j=1}^{d}(M^{(j)}_{[t_{i-1},t_i]} - m^{(j)}_{[t_{i-1},t_i]})$$

Осталось доказать последнее неравенство: $\sum\limits_{j=1}^{d}(M^{(j)}_{[t_{i-1},t_i]} - m^{(j)}_{[t_{i-1},t_i]}) \leqslant \varepsilon d$.

$$\left|M^{(j)}_{[t_{i-1},t_i]} - m^{(j)}_{[t_{i-1},t_i]}\right| = |\gamma_j'(\eta_i) - \gamma_j'(\zeta_i)|$$

Знаем, что γ_j' непрерывно на $[a,b]$, а значит равномерно непрерывно (по теореме Кантора):

$$\forall\varepsilon > 0, \exists\delta > 0 : \forall\eta,\zeta : |\eta - \zeta| < \delta \Rightarrow |\gamma_j'(\eta) - \gamma_j'(\zeta)| < \varepsilon$$

Если наше разбиение t_i имеет мелкость менее δ, то верно, что $|\gamma_j'(\eta_i) - \gamma_j'(\zeta_i)| < \varepsilon$.

$$|M_i - m_i| = |\gamma_j'(\eta_i) - \gamma_j'(\zeta_i)| < \varepsilon d$$

$$\sum_{i=1}^{n}(M_i - m_i)(t_i - t_{i-1}) < \varepsilon d\sum_{i=1}^{n}(t_i - t_{i-1}) = \varepsilon d(b-a) \to 0$$

◀

Следствия:

1. $f : [a,b] \to \mathbb{R}$, $f \in C^1[a,b]$. Длина графика равна

$$\int\limits_a^b \sqrt{1 + f'(x)^2}\mathrm{d}x$$

▶ Кривая с параметризацией $(x, f(x)), a \leqslant x \leqslant b$. ◀

2. Длина кривой в полярных координатах есть

$$\int\limits_\alpha^\beta \sqrt{r(\varphi)^2 + r'(\varphi)^2}\mathrm{d}\varphi$$

▶

$$\gamma(\varphi) = (r(\varphi)\cos\varphi, r(\varphi)\sin\varphi)$$
$$(r(\varphi)\cos\varphi)' = r'\cos\varphi - r\sin\varphi$$
$$(r(\varphi)\sin\varphi)' = r'\sin\varphi + r\cos\varphi$$
$$((r(\varphi)\cos\varphi)')^2 + ((r(\varphi)\sin\varphi)')^2 = r(\varphi)^2 + r'(\varphi)^2$$

◀

3. $l(\gamma) \leqslant (b-a)\max\limits_{t\in[a,b]}\|\gamma'(t)\|$

Пример 1.4.3. Посчитаем длину эллипса:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, a < b$$
$$x(t) = a\cos t, x'(t) = -a\sin t$$
$$y(t) = b\sin t, y'(t) = b\cos t$$
$$0 \leqslant t < 2\pi$$
$$x'(t)^2 + y'(t)^2 = a^2\sin^2(t) + b^2\cos^2(t) = b^2 - (b^2 - a^2)\sin^2(t)$$
$$l = \int\limits_0^{2\pi}\sqrt{b^2 - (b^2 - a^2)\sin^2 t}\mathrm{d}t = b\int\limits_0^{2\pi}\sqrt{1 - \varepsilon^2\sin^2 t}\mathrm{d}t$$

где $\varepsilon = \frac{\sqrt{b^2 - a^2}}{b}$ и называется эксцентриситетом эллипса.

$\mathfrak{Def}$ 1.4.19. $\gamma\colon [0,S] \to \mathbb{R}^d$ — спрямляемая кривая. Параметризация называется натуральной, если

$$\forall s \in [0,S], l\left(\gamma\big|_{[0,s]}\right) = s$$

Теорема 1.4.3. У всякой гладкой спрямляемой кривой существует натуральная параметризация.

▶ $\tilde{\gamma}[a,b] \to \mathbb{R}^d$ — произвольная параметризация. Введём отображение $\tau\colon [a,b] \to [0, l(\tilde{\gamma})]$, $\tau(t) = l(\tilde{\gamma}|_{[a,t]})$. Оно строго монотонно. τ непрерывна, так как

$$|l(\tilde{\gamma}(s_2)) - l(\tilde{\gamma}(s_1))| \leqslant \left|\max_{t\in[s_1,s_2]} \|\tilde{\gamma}'(t)\|\right| (s_2 - s_1) \to 0$$

$\tau^{-1}\colon [0, l(\tilde{\gamma})] \to [a,b]$ непрерывна и строго монотонна. $\gamma = \tilde{\gamma} \circ \tau^{-1}$ — натуральная параметризация. ◀

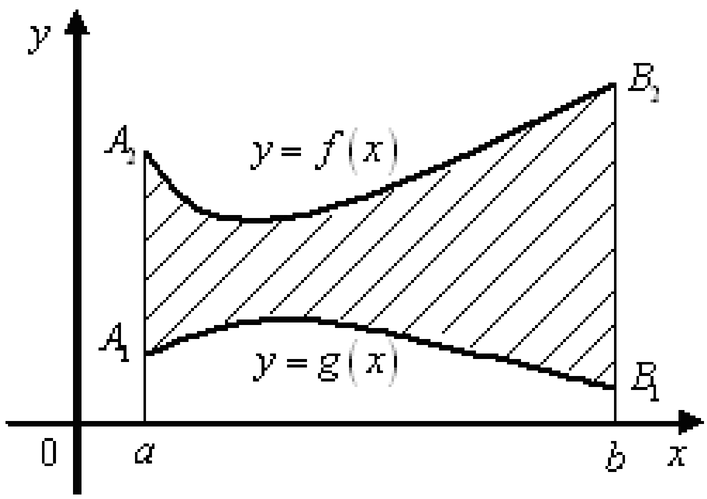
Замечание 1.4.2. Для кусочно-гладких кривых теоремы и следствия тоже верны.

1.5. Формулы для вычисления площади

$\mathfrak{Def}$ 1.5.1. Ограниченное множество $E \subset \mathbb{R}^2$ квадрируемо, если площадь E определена однозначно (то есть не зависит от σ из определения площади).

Замечание 1.5.1. Подграфики непрерывных функций квадрируемы.

$$S = \int\limits_a^b f(x)\mathrm{d}x$$



Введем два дополнительных обозначения: точки $A_0 = (a,0)$ и $B_0 = (b,0)$.

Утверждение 1.5.1. $f, g \in C[a,b]$, $f \geqslant g$ Рассмотрим криволинейную трапецию — множество

$$(x,y) \mid a \leqslant x \leqslant b, g(x) \leqslant y \leqslant f(x)$$

Криволинейная трапеция квадрируема и

$$S = \int\limits_a^b (f(x) - g(x))\mathrm{d}x$$

▶

$$S = S_f - S_g = \int\limits_a^b f(x)\mathrm{d}x - \int\limits_a^b g(x)\mathrm{d}x = \int\limits_a^b (f(x) - g(x))\mathrm{d}x$$

где S_f — площадь фигуры $A_0A_2B_2B_0$, а S_g — площадь фигуры $A_0A_1B_1B_0$. ◀

Теорема 1.5.1. E — множество, ограниченное замкнутой простой кусочно-гладкой кривой $\gamma\colon [a, b] \to \mathbb{R}^2$, $\gamma(t) = (x(t), y(t))$.

$$S_E = \int\limits_a^b x(t)y'(t)\mathrm{d}t = -\int\limits_a^b y(t)x'(t)\mathrm{d}t = \frac{1}{2}\int\limits_a^b (x(t)y'(t) - y(t)x'(t))\mathrm{d}t$$

▶ 1. Все формулы дают один и тот же результат.

$$\int\limits_a^b x(t)y'(t)\mathrm{d}t + \int\limits_a^b y(t)x'(t)\mathrm{d}t = x(t)y(t)\Big|_a^b = x(b)y(b) - x(a)y(a) = 0$$

так как $(x(a), y(a)) = (x(b), y(b))$.

2. Формула верна для криволинейной трапеции. Обратимся к картинке (переобозначив $\alpha = a, \beta = b$).
   A_1B_1**:** $x(t) = t, y(t) = g(t)$
   B_1B_2**:** $x(t) = \beta$
   B_2A_2**:** $x(t) = -t + c, y(t) = f(-t + c)$
   A_2A_1**:** $x(t) = \alpha$

$$-\int\limits_a^b y(t)x'(t)\mathrm{d}t = -\int\limits_a^b g(t)\mathrm{d}t + \int\limits_a^b f(t)\mathrm{d}t$$

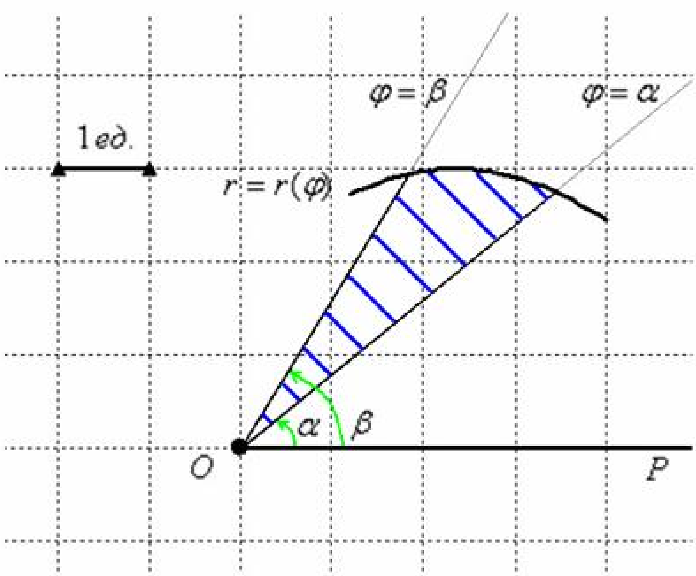
3. Теперь мы умеем делать вертикальные и горизонтальные разрезы, покажем это. Пусть было какое-то множество, мы его разделили вертикальным разрезом. Но тогда $\int yx'\mathrm{d}t$ даст 0 на этом вертикальном отрезке (две части сократятся). С горизонтальным аналогично, только рассматривать надо интеграл $\int xy'\mathrm{d}t$.
   На самом деле, это нам дает еще и способность объединять множества по общему вертикальному/горизонтальному отрезку.

4. Имеется кусочно-гладкая кривая. Проведем через все точки, где меняется монотонность x, все точки, где меняется монотонность y, и все негладкие точки горизонтальные и вертикальные разрезы. Получили много криволинейных трапеций, на каждой из которых формула работает. После чего склеиваем все обратно.

◀

Следствие 1.5.1.1. Площадь в полярных координатах.

$$S = \frac{1}{2}\int\limits_\alpha^\beta r^2(\varphi)d\varphi$$



▶ $\gamma(\varphi) = (r(\varphi)\cos(\varphi), r(\varphi)\sin(\varphi))$

$$\frac{1}{2}\int\limits_{\alpha}^{\beta}(y'(\varphi)x(\varphi) - x'(\varphi)y(\varphi))d\varphi = \frac{1}{2}\int\limits_{\alpha}^{\beta} r^2(\varphi)d\varphi$$

◀

Глава 2

Функции нескольких переменных

2.1. Линейные операторы

$\mathfrak{Def}$ **2.1.1.** X — линейное (векторное) пространство над $\mathbb{R}$, $x_1, x_2, \dots, x_n \in X$, $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$. $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$ — линейная комбинация векторов $x_1, x_2, \dots, x_n$.

$\mathfrak{Def}$ **2.1.2.** $x_1, x_2, ..., x_n \in X$ — линейно независимы, если

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = 0$$

только при

$$\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$$

$\mathfrak{Def}$ **2.1.3.** Размерность пространства — максимальное число линейно независимых векторов

$\mathfrak{Def}$ **2.1.4.** $x_1, x_2, ..., x_n \in X$ — базис, если

1. $x_1, x_2, ..., x_n$ линейно независимы
2. $$\forall x \in X, \exists \lambda_1, \lambda_2, \dots, \lambda_n \colon x = \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n$$

Пример 2.1.1. $X = \mathbb{R}^n$.
$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$ — базис, потому что любой вектор $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ можно представить как

$$a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + a_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

$\mathfrak{Def}$ **2.1.5.** X, Y — линейные пространства. $A \colon X \to Y$ — линейный оператор, если

1. $\forall x, y \in X, A(x + y) = A(x) + A(y)$
2. $\forall x \in X, \forall \lambda \in \mathbb{R}, A(\lambda x) = \lambda A(x)$

Свойства:

1. $\forall A, A(0_x) = 0_y$

   ▶ $\lambda = 0, A(0x) = 0A(x)$ ◀

2. $$A\left(\sum_{k=1}^{n} \lambda_k x_k\right) = \sum_{k=1}^{n} \lambda_k A\left(x_k\right)$$

   ▶
   $$A\left(\sum_{k=1}^{n} \lambda_k x_k\right) = \sum_{k=1}^{n} A\left(\lambda_k x_k\right) = \sum_{k=1}^{n} \lambda_k A\left(x_k\right)$$
   ◀

3. Множество линейных операторов — линейное пространство

   ▶ (a) $\forall A, B, (A+B)(x) \equiv (B+A)(x)$
   (b) $\forall A, B, C, ((A+B)+C)(x) \equiv (A+(B+C))(x)$
   (c) $\exists O\colon \forall A, (A+O)(x) \equiv A(x) + O(x)$
   (d) $\forall A, \exists -A\colon (A+-A)(x) \equiv O(x)$
   (e) $\forall \alpha, \beta, A, \alpha(\beta A)(x) \equiv (\alpha\beta)A(x)$
   (f) $\forall A, (1 \cdot A)(x) \equiv A(x)$
   (g) $\forall \alpha, \beta, A, (\alpha+\beta)A(x) \equiv \alpha A(x) + \beta A(x)$
   (h) $\forall \alpha, A, B, \alpha(A+B)(x) \equiv \alpha A(x) + \alpha B(x)$
   Проверьте сами ◀

𝔇𝔢𝔣 2.1.6. Композиция линейных операторов $A\colon X \to Y$ и $B\colon Y \to Z$:

$$B \circ A\colon X \to Z \quad (B \circ A)(x) = B(A(x))$$

Теорема 2.1.1. Композиция линейных операторов — линейный оператор.

▶ 1. $(B \circ A)(x+y) = B(A(x+y)) = B(A(x)+A(y)) = B(A(x)) + B(A(y)) = (B \circ A)(x) + (B \circ A)(y)$
2. $(B \circ A)(\lambda x) = B(A(\lambda x)) = B(\lambda A(x)) = \lambda B(A(x)) = \lambda(A \circ B)(x)$

◀

𝔇𝔢𝔣 2.1.7. $A\colon X \to Y$, $B\colon Y \to X$. B — обратный оператор к A, если:

1. $\forall x, (B \circ A)(x) = x$

2. $\forall y, (A \circ B)(y) = y$

$$B = A^{-1}$$

<u>Свойства:</u>

1. Если обратный оператор существует, то он единственный

2. $(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}$

   ▶
   $$\left(\frac{1}{\lambda} A^{-1}\right) \circ (\lambda A)(x) = \frac{1}{\lambda}\lambda\left(A^{-1} \circ A\right)(x) = x$$

   ◀

Теорема 2.1.2. Теперь рассмотрим случай $Y = X$. $A\colon X \to X$. Множество обратимых операторов образует группу относительно композиции.

▶ 1. $\exists E\colon E(x) = x$

2. Проверим замкнутость операций. Для этого докажем, что

$$(A \circ B)^{-1}(x) = (B^{-1} \circ A^{-1})(x)$$

▶

$$B^{-1} \circ A^{-1} \circ (A \circ B)(x) = B^{-1}(A^{-1}(A(B(x))) = B^{-1}(B(x)) = x$$

◀

◀

Частный случай $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $A\colon \mathbb{R}^n \to \mathbb{R}^m$.

$$\bar{x} = x_1\bar{e}_1 + \dots + x_n\bar{e}_n$$
$$A(\bar{x}) = A(x_1\bar{e}_1 + \dots + x_n\bar{e}_n) = x_1A\bar{e}_1 + \dots + x_nA\bar{e}_n$$

Напишем это же равенство через матрицы:

$$A(\bar{x}) = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 + \dots + a_{1,n}x_n \\ \vdots \\ a_{m,1}x_1 + \dots + a_{m,n}x_n \end{pmatrix} = x_1 \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix}$$

𝔇𝔢𝔣 2.1.8. Норма оператора: X и Y — нормированные пространства (заданы $\|\|_x$ и $\|\|_y$). $A\colon X \to Y$ — линейный оператор.

$$\|A\| = \sup_{\|x\|_x \leqslant 1} \|A(x)\|_y$$

или то же самое

$$\|A\| = \sup\{\|A(x)\|_y \mid \|x\|_x \leqslant 1\}$$

𝔇𝔢𝔣 2.1.9. Если $\|A\| < +\infty$, то A — ограниченный оператор.

Свойства:

1. $\|A + B\| \leqslant \|A\| + \|B\|$

   ▶ $\|A+B\| = \sup \|(A+B)(x)\|_y = \sup \|A(x) + B(x)\|_y \leqslant \sup \|A(x)\|_y + \sup \|B(x)\|_y = \|A\| + \|B\|$ ◀

2. $\|\lambda A\| = |\lambda| \cdot \|A\|$

   ▶ $\|\lambda A\| = \sup \|(\lambda A)(x)\|_y = \sup \|\lambda A(x)\|_y = \sup(|\lambda| \cdot \|A(x)\|_y) = |\lambda| \sup \|A(x)\|_y = |\lambda| \cdot \|A\|$ ◀

3. $\|A\| = 0 \Rightarrow A$ — нулевой оператор

   ▶ $\|A\| = 0 \Rightarrow \sup \|A(x)\|_y = 0 \Rightarrow \|A(x)\|_y = 0, \forall x\colon \|x\|_x \leqslant 1$
   Пусть $\|\tilde{x}\|_x > 1$, тогда возьмём $x = \frac{\tilde{x}}{\|\tilde{x}\|_x}$ и $\|x\|_x = 1$
   $A(\tilde{x}) = A(\|\tilde{x}\|_x \cdot x) = \|\tilde{x}\|_x \cdot A(x) = 0$ ◀

4. Норма для операторов, которую мы задали, действительно является нормой в линейном пространстве операторов.

Замечание 2.1.1. Ограниченный оператор не означает ограниченное отображение. Если $A\colon X \to Y$ — ограниченное отображение, то $A = O$.

▶ Пусть $A \neq O$

$$\exists x_0 \in X\colon A(x_0) \neq 0_y \Rightarrow \|A(x_0)\|_y \neq 0$$
$$\|A(\lambda x_0)\|_y = |\lambda| \cdot \|Ax_0\|_y$$

Если A — ограниченное отображение, то $A(\lambda x_0)$ — ограниченное множество, поэтому

$$\forall \lambda \in \mathbb{R}, \|A(\lambda x_0)\| \leqslant M$$

но

$$\|A(\lambda x_0)\| = |\lambda| \cdot \|Ax_0\|$$

— противоречие ◀

Теорема 2.1.3. $A\colon X \to Y$ — линейный оператор, X, Y — нормированные пространства. Тогда

$$\sup_{\|x\|_x \leqslant 1} \|A(x)\|_y = \sup_{\|x\|_x < 1} \|A(x)\|_y = \sup_{\|x\|_x = 1} \|A(x)\|_y = \sup_{x \neq 0} \frac{\|A(x)\|_y}{\|x\|_x} = \inf\{c \mid \forall x \in X, \|A(x)\|_y \leqslant c\|x\|_x\}$$

▶ Для удобства обозначим части равенств N_1, N_2, N_3, N_4, N_5 соответственно.

$N_1 \geqslant N_2$ **и** $N_1 \geqslant N_3$**:** Очевидно.

$N_3 = N_4$**:**

$$\frac{\|A_x\|_y}{\|x\|_x} = \frac{1}{\|x\|_x} \cdot \|A(x)\|_y = \|A(\frac{x}{\|x\|_x})\|_y$$

$\frac{x}{\|x\|_x}$ — единичный вектор.

$N_4 = N_5$**:**

$$\inf\{c \mid \forall x \in X, \|A(x)\|_y \leqslant c\|x\|_x\} = \inf\{c \mid \forall x \in X\colon x \neq 0, \|A(x)\|_y \leqslant c\|x\|_x\} =$$
$$= \inf\{c \mid \forall x \in X\colon x \neq 0, \frac{\|A(x)\|_y}{\|x\|_x} \leqslant c\}$$

— наименьшая из верхних границ для $\frac{\|A(x)\|_y}{\|x\|_x}$ при $x \neq 0$, что является супремумом по определению.

$N_2 \geqslant N_1$**:** Докажем, что

$$(1+\varepsilon)N_2 \geqslant N_1, \forall \varepsilon > 0$$
$$(1+\varepsilon)N_2 = \sup_{\|x\|_x < 1} \|A(x)\|_y \cdot (1+\varepsilon)$$

Зафиксируем $x \in X, \|x\|_x \leqslant 1$:

$$\tilde{x} = \frac{x}{1+\varepsilon}$$
$$\|\tilde{x}\|_x = \|\frac{x}{1+\varepsilon}\|_x = \frac{1}{1+\varepsilon}\|x\|_x \leqslant \frac{1}{1+\varepsilon} < 1$$
$$\|A(x)\|_y = \|A((1+\varepsilon)\tilde{x})\|_y = (1+\varepsilon)\|A(\tilde{x})\|_y \leqslant (1+\varepsilon)N_2$$

$N_3 \geqslant N_1$**:**

$$\|A(x)\|_y = \|A(\|x\|_x \cdot \frac{x}{\|x\|_x})\|_y = \|x\|_x \cdot \|A(\frac{x}{\|x\|_x})\|_y \leqslant \|A(\frac{x}{\|x\|_x})\|_y \leqslant N_3$$

◀

Следствие 2.1.3.1. 1. $\forall x \in X, \|A(x)\|_y \leqslant \|A\| \cdot \|x\|_x$

▶ $\|A\| = \sup_{x\neq 0} \frac{\|A(x)\|_y}{\|x\|_x}$ ◀

2. $\|B \circ A\| \leqslant \|B\| \cdot \|A\|$

▶

$$\|(B\circ A)(x)\|_z = \|B(A(x))\|_z \leqslant \|B\| \cdot \|A(x)\|_y \leqslant \|B\| \cdot \|A\| \cdot \|x\|_x$$
$$\|B\circ A\| = \sup_{\|x\|_x \leqslant 1} \|(B\circ A)(x)\|_z \leqslant \|B\| \cdot \|A\|$$

◀

Теорема 2.1.4. $A\colon X \to Y$ — линейный оператор. Тогда следующие условия равносильны:

1. A ограниченный оператор
2. A непрерывен в 0
3. A непрерывен
4. A равномерно непрерывен

▶ **$4 \Rightarrow 3 \Rightarrow 2$:** Очевидно.

$1 \Rightarrow 4$:

$$\|A(x_1) - A(x_2)\|_y = \|A(x_1 - x_2)\|_y \leqslant \|A\| \cdot \|x_1 - x_2\|_x$$
$$\forall \varepsilon > 0, \exists \delta > 0\colon \forall x_1, x_2, \|x_1 - x_2\|_x < \delta \Rightarrow \|A(x) - A(y)\|_y < \varepsilon$$

Возьмём $\delta = \frac{\varepsilon}{\|A\|}$

$$\|x_1 - x_2\|_x < \frac{\varepsilon}{\|A\|} \Rightarrow \|A(x_1) - A(x_2)\|_y \leqslant \|A\| \cdot \|x_1 - x_2\|_x < \|A\| \cdot \frac{\varepsilon}{\|A\|} = \varepsilon$$

$2 \Rightarrow 1$:

$$\forall \varepsilon > 0, \exists \delta > 0\colon \|x - 0\|_x < \delta \Rightarrow \|A(x) - A(0)\|_y < \varepsilon$$

Возьмём $\varepsilon = 1$ и соответствующий ему δ:

$$\|x - 0\|_x < \delta \Rightarrow \|A(x) - A(0)\|_y < 1$$

Возьмём $x \in X\colon \|x\|_x < 1$, тогда $\tilde{x} = \delta x$

$$\|\tilde{x}\|_x < \delta \Rightarrow 1 > \|A\tilde{x}\|_y = \|A(\delta x)\|_y = \delta \|A(x)\|_y$$
$$\|A(x)\|_y < \frac{1}{\delta}, \forall x \in X\colon \|x\|_x < 1$$
$$\sup_{\|x\|_x < 1} \|A(x)\|_y \leqslant \frac{1}{\delta}$$

◀

Теорема 2.1.5. $A\colon \mathbb{R}^n \to \mathbb{R}^m$

$$\|A\| \leqslant \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}^2}$$

▶

$$x = x_1e_1 + \dots + x_ne_n$$
$$\|x\|^2 = x_1^2 + \dots + x_n^2$$
$$\|A(x)\| = \|A(x_1e_1 + \dots + x_ne_n)\| = \|x_1A(e_1) + \dots + x_nA(e_n)\| \leqslant |x_1| \cdot \|A(e_1)\| + \dots + |x_n| \cdot \|A(e_n)\|$$
$$\|A(x)\|^2 \leqslant (|x_1| \cdot \|A(e_1)\| + \dots + |x_n| \cdot \|A(e_n)\|)^2 \leqslant (|x_1|^2 + \dots + |x_n|^2) \cdot (\|A(e_1)\|^2 + \dots + \|A(e_n)\|^2) =$$
$$= \|x^2\|(\|A(e_1)\|^2 + \dots + \|A(e_n)\|^2) = \|x\|^2 \cdot \sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}^2$$

$$A(e_k) = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{pmatrix}$$

$$\|A(e_k)\|^2 = a_{1,k}^2 + \dots + a_{m,k}^2$$
$$\|A(x)\| \leqslant \|x\|\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}^2} \leqslant \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}^2} \quad \text{при } \|x\| \leqslant 1$$
$$\sup_{\|x\|\leqslant 1} \|A(x)\| \leqslant \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}^2}$$

◀

𝔇𝔢𝔣 2.1.10. X — линейное пространство. Две нормы называются эквивалентными, если

$$\exists c_1, c_2 > 0\colon \forall x \in X, c_1\|x\|_1 \leqslant \|x\|_2 \leqslant c_2\|x\|_1$$

Замечание 2.1.2.

1. Если нормы эквивалентны, то сходимости по этим нормам означают одно и то же. Пусть $x_n \to x_0$ по норме 1, тогда $\|x_n - x_0\|_2 \leqslant c_2\|x_n - x_0\|_1$, но $x_n - x_0 \to 0 \Rightarrow x_n \to x_0$ по норме 2.
2. Непрерывности относительно эквивалентных норм совпадают

Теорема 2.1.6. В $\mathbb{R}^d$ все нормы эквивалентны

▶

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$$

Докажем, что все нормы эквивалентны данной Пусть p — другая норма.

$$p(x) = p\left(\sum_{k=1}^{d} x_ke_k\right) \leqslant \sum_{k=1}^{d} |x_k|p(e_k) \leqslant \left(\sum_{k=1}^{d} |x_k|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{d} p(e_k)^2\right)^{\frac{1}{2}} =$$
$$= \|x\| \cdot \left(\sum_{k=1}^{d} p(e_k)^2\right)^{\frac{1}{2}}$$

Возьмём $c_2 = \sqrt{\sum_{k=1}^{d} p(e_k)^2}$:

$$p(x) \leqslant c_2\|x\|$$
$$|p(x) - p(y)| \leqslant p(x - y) \leqslant c_2\|x - y\| \Rightarrow p \text{ — непрерывная функция}$$

$S^{d-1} = \{x \in \mathbb{R}^d \mid \|x\| = 1\}$ — сфера в n-мерном пространстве. Она компактна (ограничена и замкнута), поэтому по теореме Вейерштрасса p достигает наименьшее значение $p(y)$. $p(y) \neq 0$, потому что иначе $y \notin S^{d-1}$.

Пусть $c_1 = p(y) > 0$, тогда

$$c_1 \leqslant p\left(\frac{x}{\|x\|}\right) = p\left(\frac{1}{\|x\|} \cdot x\right) = \frac{1}{\|x\|} \cdot p(x)$$
$$c_1\|x\| \leqslant p(x)$$

◀

2.2. Дифференцируемость в $\mathbb{R}^n$

Def 2.2.1. $f\colon D \subset \mathbb{R}^n \to \mathbb{R}^m$, $a \in \operatorname{int} D$. f дифференцируема в точке a, если существует линейный оператор $T\colon \mathbb{R}^n \to \mathbb{R}^m$ такой, что

$$f(a+h)_{h\to 0} = f(a) + T(h) + o(\|h\|)$$

Замечание 2.2.1. T определен однозначно: зафиксируем h, $t > 0$:

$$f(a+th)_{t\to 0} = f(a) + T(th) + o(\|th\|) = f(a) + T(th) + o(t)$$
$$T(h) + o(1) = \frac{f(a+th) - f(a)}{t}$$
$$T(h) = \lim_{t\to 0} \frac{f(a+th) - f(a)}{t}$$

Def 2.2.2. Матрица, соответствующая оператору T — матрица Якоби.

Частный случай $m = 1$:

$$f\colon D \subset R^n \to R$$
$$f(a+h) = f(a) + T(h) + o(\|h\|)$$
$$T\colon \mathbb{R}^n \to \mathbb{R}$$
$$T(h) = \begin{pmatrix} t_1 & \dots & t_n \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = t_1 h_1 + \dots + t_n h_n = \langle V, h \rangle$$
$$V = \begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}$$
$$\exists V \in \mathbb{R}^n\colon f(a+h) = f(a) + \langle V, h \rangle + o(\|h\|)$$

Def 2.2.3. V — градиент функции f:

$$V = \operatorname{grad} f = \nabla f$$

Теорема 2.2.1. $f\colon D \subset \mathbb{R}^n \to \mathbb{R}^m$, $a \in \operatorname{int} D$,

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$$

Тогда f дифференцируема в точке a тогда и только тогда, когда все f_k дифференцируемы в точке a.

▶

$$f(a+h) = f(a) + T(h) + o(\|h\|)$$

$\Rightarrow$:

$$f_k(a+h) = f_k(a) + (T(h))_k + o(\|h\|)$$

$\Leftarrow$:

$$\alpha_k(h) = f_k(a+h) - f_k(a) - \langle V_k, h\rangle = o(\|h\|)$$
$$\frac{\alpha_k(h)}{\|h\|} \to 0$$
$$\alpha(h) = f(a+h) - f(a) - T(h) = o(\|h\|)$$
$$\alpha(h) = \begin{pmatrix} \alpha_1(h) \\ \vdots \\ \alpha_n(h) \end{pmatrix}$$
$$\|\alpha(h)\| = \sqrt{\alpha_1^2(h) + \dots + \alpha_n^2(h)}$$
$$\frac{\|\alpha(h)\|}{\|h\|} = \sqrt{(\frac{\|\alpha_1(h)\|}{\|h\|})^2 + \dots + (\frac{\|\alpha_n(h)\|}{\|h\|})^2} \to 0$$

потому что

$$\frac{\alpha_k(h)}{\|h\|} \to 0$$

◀

Ещё раз вспомним:

$$f(a+h) = f(a) + Th + o(\|g\|) \quad \text{при } h \to \infty$$
$$f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m$$
$$T\colon \mathbb{R}^n \to \mathbb{R}^m\text{— дифференциал в точке } a$$
$$\mathrm{d}_a f \quad Th \quad \mathrm{d}_a f(h)$$

Пример 2.2.1. 1. f — константа: $f(x) = v \in \mathbb{R}^m$. Она дифференцируема во всех точках:

$$f(a+h) = f(a) + \underbrace{\mathrm{d}_a f(h) + o(\|h\|)}_{=0}$$
$$\mathrm{d}_a f = 0$$

Как и раньше, производная константы — ноль.

2. f — линейное отображение: $f(x) = Tx$. Оно также дифференцируемо везде:

$$f(a+h) = T(a+h) = Ta + Th = f(a) + Th + 0$$
$$\mathrm{d}_a f = T$$

Теорема 2.2.2 (Линейность дифференциала)**.** $f, g\colon E \subset \mathbb{R}^n \to \mathbb{R}^m$, $a \in \operatorname{int} E$, f и g дифференцируемы в a, $\lambda \in \mathbb{R}$. Тогда $f+g$ и λf дифференцируемы в a и

$$\mathrm{d}_a(f+g) = \mathrm{d}_a f + \mathrm{d}_a g$$
$$\mathrm{d}_a(\lambda f) = \lambda \mathrm{d}_a f$$

▶

$$f(a+h) = f(a) + \mathrm{d}_a f(h) + o(\|h\|)$$
$$g(a+h) = g(a) + \mathrm{d}_a g(h) + o(\|h\|)$$
$$(f+g)(a+h) = f(a+h) + g(a+h) = \underbrace{f(a)+g(a)}_{(f+g)(a)} + \underbrace{\mathrm{d}_a f(h) + \mathrm{d}_a g(h)}_{\text{линейно по } h} + o(\|h\|)$$

Получили, что

$$\mathrm{d}_a(f+g) = \mathrm{d}_a f + \mathrm{d}_a g$$

Собственно, в этом рассуждении мы размерностью нигде не пользуемся. Эти рассуждения действуют и на одномерный случай без изменений.

$$(\lambda f)(a+h) = \lambda f(a+h) = \lambda f(a) + \lambda \mathrm{d}_a f(h) + o(\|h\|)\mathrm{d}_a(\lambda f) = \lambda \mathrm{d}_a f$$

◀

Теорема 2.2.3 (Дифференцируемость композиции)**.** $f\colon E \subset \mathbb{R}^n \to \mathbb{R}^m$, $g\colon D \subset \mathbb{R}^m \to \mathbb{R}^l$, $f(E) \subset D$, $a \in \operatorname{int} E$, f дифференцируема в a, g дифференцируема в $f(a)$. Тогда $g \circ f$ тоже дифференцируема и

$$\mathrm{d}_a(g \circ f) = \mathrm{d}_{f(a)} g \circ \mathrm{d}_a f$$

Замечание 2.2.2.

$$\mathrm{d}_a f\colon \mathbb{R}^n \to \mathbb{R}^m$$
$$\mathrm{d}_{f(a)} g\colon \mathbb{R}^m \to \mathbb{R}^l$$
$$\mathrm{d}_a(g \circ f)\colon \mathbb{R}^n \to \mathbb{R}^l$$

Все размерности сходятся.

▶ $b \leftrightharpoons f(a) \in D$

$$f(a+h) = f(a) + \mathrm{d}_a f(h) + \alpha(h)\|h\| \quad \alpha\colon E \to \mathbb{R}^m \quad \lim_{h\to\vec{0}} \alpha(h) = \vec{0}$$
$$g(b+k) = g(b) + \mathrm{d}_b g(k) + \beta(k)\|k\| \quad \beta\colon D \to \mathbb{R}^l \quad \lim_{k\to\vec{0}} \beta(k) = \vec{0}$$
$$k = \mathrm{d}_a f(h) + \alpha(h)\|h\|$$
$$(g\circ f)(a) = (g\circ f)(a) + \mathrm{d}_{f(a)} g(k) + \beta(k)\|k\| = (g\circ f)(a) + \mathrm{d}_{f(a)} g(\mathrm{d}_a f(h) + \alpha(h)\|h\|) + \beta(k)\|k\| =$$
$$= (g\circ f)(a) + \underbrace{\mathrm{d}_{f(a)} g(d_a f(h))}_{\mathrm{d}_a(g\circ f)(h)} + \mathrm{d}_{f(a)} g(\alpha(h))\|h\| + \beta(k)\|k\|$$

Проверим, что «хвост» стремится к нулю при h, стремящемуся к нулю:

$$\mathrm{d}_{f(a)} g(\alpha(h)) + \frac{\beta(k)\|k\|}{\|h\|}$$
$$\left\|\mathrm{d}_{f(a)} g(\alpha(h))\right\| \leqslant \underbrace{\left\|\mathrm{d}_{f(a)} g\right\|}_{\text{конечно}} \circ \underbrace{\|\alpha(h)\|}_{\to\vec{0}} \to \vec{0}$$

Поймём, что $k \to \vec{0}$ при $h \to \vec{0}$

$$k = \mathrm{d}_a f(h) + \alpha(h)\|h\|$$
$$\|k\| \leqslant \|\mathrm{d}_a f(h)\| + \|\alpha(h)\|\|h\| \leqslant (\|\mathrm{d}_a f\| + \|\alpha(h)\|)\,\|h\| \to \vec{0}$$

Значит при $h \to \vec{0}$ и $\beta(k) \to \vec{0}$. Осталось понять, что $\frac{\|k\|}{\|h\|}$ ограничено.

$$\frac{\|k\|}{\|h\|} \leqslant \frac{(\|\mathrm{d}_a f\| + \|\alpha(h)\|)\,\|h\|}{\|h\|} = \|\mathrm{d}_a f\| + \|\alpha(h)\| \text{ — ограниченно}$$

◀

Давайте задумаемся, что такое произведение векторных функций. Для начала обратимся к умножению на скаляр:

Теорема 2.2.4 (Дифференциал произведение векторной и скалярной функций)**.** $f\colon E\subset\mathbb{R}^n\to\mathbb{R}^m$, $\lambda\colon E\to\mathbb{R}$, $a\in\operatorname{int}E$, f и λ дифференцируемы в a. Тогда $\lambda f\colon E\to\mathbb{R}^m$ дифференцируемо в a и

$$\mathrm{d}_a(\lambda f)(h)=\mathrm{d}_a\lambda(h)f(a)+\lambda(a)\mathrm{d}_af(h)$$

Замечание 2.2.3. Размерности сходятся:

$$\mathrm{d}_af\colon\mathbb{R}^n\to\mathbb{R}^m$$
$$\mathrm{d}_a\lambda\colon\mathbb{R}^n\to\mathbb{R}$$
$$\mathrm{d}_a(\lambda f)\colon\mathbb{R}^n\to\mathbb{R}^m$$

▶ Перейдём к случаю $m=1$, а потом воспользуемся дифференцируемостью координатных функций.

$$(\lambda f)(a+h)=\lambda(a+h)f(a+h)=\left(\lambda(a)+\mathrm{d}_a\lambda(h)+o(\|h\|)\right)\left(f(a)+\mathrm{d}_af(h)+o\|h\|\right)=$$
$$\lambda(a)f(a)+\mathrm{d}_a\lambda(h)f(a)+\lambda(a)\mathrm{d}_af(h)+$$
$$+\lambda(a)o(\|h\|)+f(a)o(\|h\|)+\mathrm{d}_a\lambda(h)\mathrm{d}_af(h)+$$
$$+\mathrm{d}_a\lambda(h)o(\|h\|)+\mathrm{d}_af(h)o(\|h\|)+o(\|h\|)$$

Осталось увидеть, что последние шесть слагаемых есть $o(\|h\|)$

$$|\mathrm{d}_a\lambda(h)|\leqslant\|\mathrm{d}_a\lambda\|\|h\| \text{ — ограниченно при } h\to\vec{0}$$
$$\|\mathrm{d}_af(h)\|\leqslant\|\mathrm{d}_af\|\|h\| \text{ — ограниченно при } h\to\vec{0}$$
$$\|\mathrm{d}_a\lambda(h)\mathrm{d}_af(h)\|\leqslant\|\mathrm{d}_a\lambda\|\|\mathrm{d}_af\|\|h\|^2=o(\|h\|)$$

◀

Теорема 2.2.5 (Дифференциал скалярного произведения)**.** $f,g\colon E\subset\mathbb{R}^n\to\mathbb{R}^m$, $a\in\operatorname{int}E$, f и g дифференцируемы в a. Тогда $\langle f,g\rangle$ дифференцируемо в a и

$$\mathrm{d}_a\left\langle f,g\right\rangle(h)=\langle\mathrm{d}_af(h),g(a)\rangle+\langle f(a),\mathrm{d}_ag(h)\rangle$$

▶ Разложим покоординатно:

$$\langle f,g\rangle=f_1g_1+f_2g_2+\dots f_mg_m$$

Достаточно понять, что происходит с каждым слагаемым.

$$f_k,g_k\colon E\to\mathbb{R}$$
$$\mathrm{d}_a(f_kg_k)(h)=\mathrm{d}_af_k(h)g_k(a)+f_k(a)\mathrm{d}_ag_k(h)$$
$$\mathrm{d}_a\left\langle f,g\right\rangle(h)=\sum_{k=1}^m\mathrm{d}_a(f_kg_k)(h)=\sum_{k=1}^m\left(\mathrm{d}_af_k(h)g_k(a)+f_k(a)\mathrm{d}_ag_k(h)\right)=\langle\mathrm{d}_af(h)g(a),f(a)\mathrm{d}_ag(h)\rangle$$

◀

Теорема 2.2.6 (Теорема Лагранжа векторнозначных функций)**.** $f\colon[a,b]\to\mathbb{R}^m$, f непрерывна на $[a,b]$, дифференцируема на (a,b). Тогда

$$\exists c\in(a,b)\colon\|f(b)-f(a)\|\leqslant\|f'(c)\|(b-a)$$

▶ $\varphi(t) \leftrightharpoons \langle f(t), f(b) - f(a)\rangle$, $\varphi\colon [a,b] \to \mathbb{R}$.

$$\varphi(b) - \varphi(a) = \langle f(b), f(b) - f(a)\rangle - \langle f(a), f(b) - f(a)\rangle = \|f(b) - f(a)\|^2$$

Применяем линейную теорему Лагранжа для φ:

$$\exists c \in (a,b)\colon \varphi(b) - \varphi(a) = \underbrace{\varphi'(c)}_{\mathbb{R}\to\mathbb{R}}(b-a)$$

$$\varphi'(c)h = \mathrm{d}_c\varphi(h) = \left\langle \underbrace{\mathrm{d}_c f(h)}_{=f'(c)h}, f(b) - f(a) \right\rangle$$

$$\varphi(b) - \varphi(a) = \varphi'(c)(b-a) = \langle f'(c), f(b) - f(a)\rangle\,(b-a) \leqslant$$

Неравенство Коши-Буняковского:

$$\leqslant \|f'(c)\|\|f(b) - f(a)\|(b-a)$$
$$\|f(b) - f(a)\| \leqslant \|f'(c)\|(b-a)$$

◀

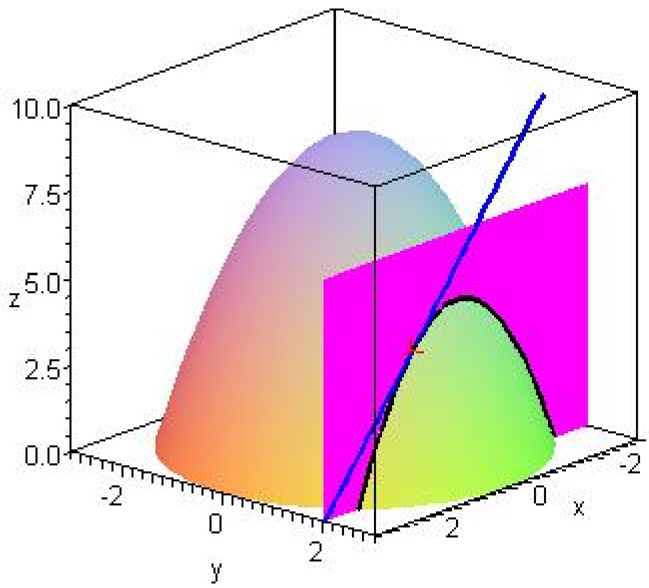
Пример 2.2.2. $f(t) = (\cos t, \sin t)$, $b = 2\pi$, $a = 0$:

$$f(b) - f(a) = \vec{0}$$
$$\|f(b) - f(a)\| = 0$$
$$f'(t) = (-\sin t, \cos t)$$
$$\|f'(t)\| = \sqrt{\sin^2 t + \cos^2 t} = 1$$

Очевидно, равенство здесь недостижимо.

2.3. Частные производные

Если есть функия от многих переменных, мы можем взять сечение её графика плоскостью, и рассматривать производную оставшейся функции.



http://moodle.capilanou.ca/mod/book/view.php?id=328667&chapterid=1398

$\mathfrak{Def}$ 2.3.1. Пусть есть функция $f\colon D \to \mathbb{R}$, точка $a \in \operatorname{int} D$ и вектор направления $\vec{h}\colon \|h\| = 1$. Рассмотрим функцию

$$F(t) = f(a + th), t \in \mathbb{R}$$

Тогда производной f по направлению h в точке a называют

$$\frac{\partial f}{\partial h}(a) = F'(0)$$

Замечание 2.3.1.

$$F'(0) = \lim_{t\to 0} \frac{F(t) - F(0)}{t} = \lim_{t\to 0} \frac{f(a + th) - f(a)}{t}$$

Очень похоже на определение производной функции одной переменной.

Утверждение 2.3.1. $f\colon D \to \mathbb{R}$, $a \in \operatorname{int} D$, f дифференцируема в a, h — ненулевой вектор.

$$\frac{\partial f}{\partial h}(a) = \mathrm{d}_a f(h)$$

▶

$$\frac{\partial f}{\partial h}(a) = F'(0)$$

$$\mathrm{d}_a f\colon t \mapsto Tt$$

$$\mathrm{d}_0 F = \underbrace{\mathrm{d}_a f}_{t\mapsto Tt} \circ \underbrace{\mathrm{d}_0(a + th)}_{t\mapsto ht} = (t \mapsto Tht)$$

$$\mathrm{d}_0 F = (t \mapsto F'(0) \cdot t)$$

$$F'(0) = Th = (t \mapsto Tt)(h) = \mathrm{d}_a f(h)$$

◀

Теорема 2.3.1 (Экстремальное свойство градиента)**.** $f\colon D \to \mathbb{R}$, $a \in \operatorname{int} D$, f дифференцируема в a, h — напралвение, $\nabla f(a) \neq 0$. Тогда

$$\left|\frac{\partial f}{\partial h}(a)\right| \leqslant \|\nabla f(a)\|$$

и равенство достигается только в случае $h \parallel \nabla f(a)$.

▶

$$\left|\frac{\partial f}{\partial h}(a)\right| = |\langle \nabla f(a), h\rangle| \leqslant \|\nabla f(a)\|\|h\| = \|\nabla f(a)\|$$

В неревенстве Коши— Буняковского равенство было как раз в случае параллельности векторов $\nabla f(a)$ и h. ◀

То есть градиент указывает, где производная максимальна. Физический смысл градиента такой: пусть есть поверхность. Положим на неё шарик. Он покатится в направлении, противоположном градиенту.

$\mathfrak{Def}$ 2.3.2. Частная производная — производная по направлению базисного вектора

$$e_k = (0, \ldots, \underbrace{1}_{k}, \ldots, 0)$$

Записывается:

$$\frac{\partial f}{\partial e_k}(a)$$

Замечание 2.3.2.

$$\frac{\partial f}{\partial e_k}(a) = \lim_{t\to 0}\frac{f(a_1, \dots, a_k + t, \dots, a_n) - f(a_1, \dots, a_k, \dots, a_n)}{t} =$$
$$= \lim_{t\to 0}\frac{g(a_k+t) - g(a_k)}{t} = \frac{\mathrm{d}g}{\mathrm{d}x}(a_k) \quad g(x) \leftrightharpoons f(a_1, \dots, x, \dots, a_n)$$

Таким образом, частная производная — производная всей функции по одной координате.

Обозначения:

$$\tfrac{\partial f}{\partial x_k} \quad D_{x_k} f \quad f'_{x_k}$$

Пример 2.3.1. $f(x,y) = x^y$:

$$\frac{\partial f}{\partial x}(x,y) = y\cdot x^{y-1}$$
$$\frac{\partial f}{\partial y}(x,y) = \ln x\cdot x^y$$

Следствие 2.3.1.1. $f\colon D\subset R^n \to \mathbb{R}$, $a\in \operatorname{int} D$, f дифференцируема в a.

$$\frac{\partial f}{\partial e_k}(a) = \langle \nabla f(a), e_k\rangle = (\nabla f(a))_k$$

Таким образом,

$$\nabla f(a) = \left(\frac{\partial f}{\partial e_1}(a), \dots, \frac{\partial f}{\partial e_n}(a)\right)$$

Следствие 2.3.1.2. $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $a\in \operatorname{int} D$, f дифференцируема в a. Тогда матрица Якоби (матрица дифференциала)

$$\begin{pmatrix} \frac{\partial f_1}{\partial e_1}(a) & \frac{\partial f_1}{\partial e_2}(a) & \cdots & \frac{\partial f_1}{\partial e_n}(a) \\ \frac{\partial f_2}{\partial e_1}(a) & \frac{\partial f_2}{\partial e_2}(a) & \cdots & \frac{\partial f_2}{\partial e_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial e_1}(a) & \frac{\partial f_m}{\partial e_2}(a) & \cdots & \frac{\partial f_m}{\partial e_n}(a) \end{pmatrix}$$

▶

$$f(a+h) = f(a) + \mathrm{d}_a f(h) + o(h)$$
$$f_i(a+h) = f_i(a) + (\mathrm{d}_a f(h))_i + o(h) = f_i(a) + \mathrm{d}_a f_i(h) + o(h) = f_i(a) + \langle \nabla f_i(a), h\rangle + o(h)$$

Отсюда видим, что i-я строка матрицы есть вектор частных производных:

$$\begin{pmatrix} \nabla f_1(a) \\ \nabla f_2(a) \\ \vdots \\ \nabla f_m(a) \end{pmatrix}$$

◀

Следствие 2.3.1.3. Формула для дифференциала композиции: $f\colon D\subset R^n \to R^m$, $g\colon E\subset R^m \to R^l$, $f(D)\subset E$, $a\in\operatorname{int} D$, $f(a)\in \operatorname{int} E$, f дифференцируема в a, g — в $f(a)$. Тогда

$$\mathrm{d}_a(g\circ f) = \mathrm{d}_{f(a)} g\circ \mathrm{d}_a f$$

то есть матрица $\mathrm{d}_a(g \circ f)$ есть

$$\begin{pmatrix} \frac{\partial g_1}{\partial e_1}(a) & \frac{\partial g_1}{\partial e_2}(a) & \cdots & \frac{\partial g_1}{\partial e_n}(a) \\ \frac{\partial g_2}{\partial e_1}(a) & \frac{\partial g_2}{\partial e_2}(a) & \cdots & \frac{\partial g_2}{\partial e_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_l}{\partial e_1}(a) & \frac{\partial g_l}{\partial e_2}(a) & \cdots & \frac{\partial g_l}{\partial e_m}(a) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial e_1}(a) & \frac{\partial f_1}{\partial e_2}(a) & \cdots & \frac{\partial f_1}{\partial e_n}(a) \\ \frac{\partial f_2}{\partial e_1}(a) & \frac{\partial f_2}{\partial e_2}(a) & \cdots & \frac{\partial f_2}{\partial e_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial e_1}(a) & \frac{\partial f_m}{\partial e_2}(a) & \cdots & \frac{\partial f_m}{\partial e_n}(a) \end{pmatrix}$$

и

$$\frac{\partial (g \circ f)_i}{\partial e_k} = \sum_{j=1}^{m} \frac{\partial g_i}{\partial e_j}(f(a)) \frac{\partial f_j}{\partial e_k}(a)$$

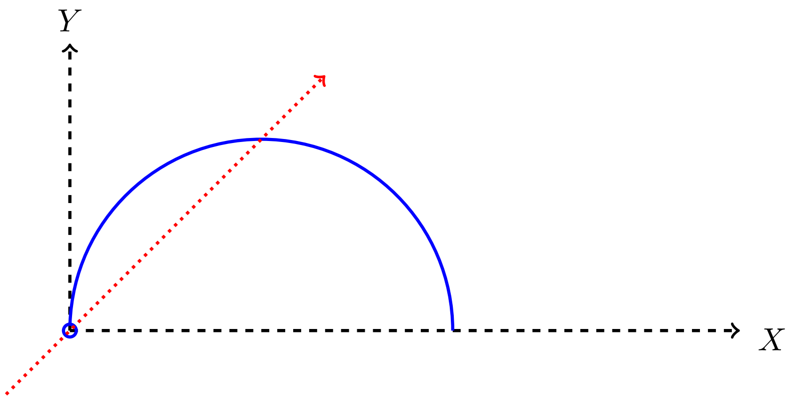
Замечание 2.3.3. Наличие частных производных ничего не говорит о дифференцируемости:

$$f(x,y) = \begin{cases} 0 & xy = 0 \\ 1 & xy \neq 0 \end{cases}$$

$$f'_x(0,0) = 0 \quad f'_y(0,0) = 0$$

Сама функция даже не непрерывна!

Даже если потребовать про все направления, ситуацию это не исправит. Рассмотрим функцию, равную единице только на дуге



Там все производные по направлению равны нулю, так как суженная на направление функция равна нулю везде, кроме одно фиксированной точки, отделённой от нуля. То есть сужение в некоторой окрестности нуля равно нулю, соответственно, производная в нуле тоже ноль.

Теорема 2.3.2. $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$, $a \in \operatorname{int} D$, все частные производные в окрестности a существуют и непрерывны. Тогда f дифференцируема в a.

▶ Хотим доказать (по определению дифференцируемости), что

$$f(a+h) = f(a) + \langle \nabla f(a), h \rangle + o(h) \qquad \text{здесь } \langle a, b \rangle \text{ — скалярное произведение}$$

При этом ∇f мы уже знаем, поэтому остаётся:

$$R(h) = f(a+h) - f(a) - \langle \nabla f(a), h \rangle \stackrel{?}{=} o(h)$$

Распишем скалярное произведение:

$$\langle \nabla f(a), h \rangle = \sum_{k=1}^{n} \frac{\partial f}{\partial e_k}(a) h_k$$

Теперь покоординатно будем смотреть, что происходит. Положим $\vec{b_k} = (a_1+h_1, a_2+h_2, \dots, a_k+h_k, a_{k+1}, a_{k+2}, \dots, a_n)$ (это мы сдвинулись на h только по первым k координатам). Очевидно, что:

$$b_n = a + h$$
$$b_0 = a$$
$$f(a+h) - f(a) = f(b_n) - f(b_{n-1}) + f(b_{n-1}) - f(b_{n-2}) + \dots + f(b_1) - f(b_0) = \sum_{k=1}^{n} f(b_k) - f(b_{k-1})$$

$$\begin{aligned} f(b_k) - f(b_{k-1}) = {} & f(a_1+h_1, a_2+h_2, \dots, a_{k-1}+h_{k-1}, a_k+h_k, a_{k+1}, \dots, a_n) \\ & - f(a_1+h_1, a_2+h_2, \dots, a_{k-1}+h_{k-1}, a_k \qquad\quad, a_{k+1}, \dots, a_n) \end{aligned}$$

Введём функцию одной переменной (заменили в b_k координату k на переменную):

$$F_k(t) = f(a_1+h_1, \dots, a_{k-1}+h_{k-1}, t, a_{k+1}, \dots, a_n)$$

Заметим, что производная F_k в окрестности a_k — это в точности частная производная f в окрестности a (мы вольны выбирать очень маленькое h так, чтобы f было дифференцируемо во всех b_k по всем координатам). Значит можем написать формулу Лагранжа:

$$\exists \theta_k \in [0,1] \colon F_k(a_k+h_k) - F_k(a_k) = F_k'(a_k + \theta_k h_k) h_k = \frac{\partial f}{\partial e_k}(b_{k-1} + \theta_k h_k e_k) h_k$$

Собираем обратно ($F_k(a_k+h_k) - F_k(a_k) = f(b_k) - f(b_{k-1})$):

$$f(a+h) - f(a) = \sum_{k=1}^{n} \frac{\partial f}{\partial e_k}(b_{k-1} + \theta_k h_k e_k) h_k$$
$$R(h) = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial e_k}(b_{k-1} + \theta_k h_k e_k) h_k - \frac{\partial f}{\partial e_k}(a) h_k \right)$$

Теперь оцениваем $R(h)$.

$$\begin{aligned} R(h) &= \sum_{k=1}^{n} \left(\frac{\partial f}{\partial e_k}(b_{k-1} + \theta_k h_k e_k) - \frac{\partial f}{\partial e_k}(a) \right) h_k = \\ &= \left\langle \left(\frac{\partial f}{\partial e_1}(b_0 + \theta_1 h_1 e_1) - \frac{\partial f}{\partial e_1}(a), \dots, \frac{\partial f}{\partial e_n}(b_{n-1} + \theta_n h_n e_n) - \frac{\partial f}{\partial e_n}(a) \right), h \right\rangle \end{aligned}$$

По Коши-Буняковскому:

$$|R(h)| \leqslant \left(\sum_{k=1}^{n} \left(\frac{\partial f}{\partial e_k}(b_{k-1} + \theta_k h_k e_k) - \frac{\partial f}{\partial e_k}(a) \right)^2 \right)^{1/2} \|h\|$$
$$R(h) \stackrel{?}{=} o(\|h\|) \iff \frac{|R(h)|}{\|h\|} \xrightarrow{?} 0$$
$$\sum_{k=1}^{n} \left(\frac{\partial f}{\partial e_k}(b_{k-1} + \theta_k h_k e_k) - \frac{\partial f}{\partial e_k}(a) \right)^2 \xrightarrow{?} 0$$

Осталось понять, что каждое слагамое стремится к нулю:

$$\frac{\partial f}{\partial e_k}(b_{k-1} + \theta_k h_k e_k) - \frac{\partial f}{\partial e_k}(a) \xrightarrow{?} 0$$

То есть надо понять, что разность $\frac{\partial f}{\partial e_k}$ в точках $b_{k-1} + \theta_k h_k e_k$ и a стремится к нулю. Так как частная производная непрерывна в точке a, нам достаточно понять, что $b_{k-1} + \theta_k h_k e_k$ стремится к a:

$$\begin{aligned} b_{k-1} + \theta_k h_k e_k - a &= (a_1 + h_1, \dots, a_{k-1} + h_{k-1}, a_k + \theta_k h_k, a_{k+1}, \dots, a_n) - (a_1, \dots, a_n) = \\ &= (h_1, \dots, h_{k-1}, \theta_k h_k, 0, \dots, 0) \end{aligned}$$

Тогда длина этого вектора не больше, чем длина h, так как модуль каждой координаты не больше: первые $k-1$ совпадают, k-я не больше, следующие равны нулю.

◀

Замечание 2.3.4. Эта теорема очень удобна для проверки дифференцируемости: иногда по определению муторно и долго, а частные производные посчитать и посмотреть просто.

Пример 2.3.2. Пусть $f(x, y) = \sqrt[3]{xy}$. Хотим понять, что с дифференцируемостью. Найдём частные производные:

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{\sqrt[3]{y}}{3\sqrt[3]{x^2}} \text{ при } x \neq 0 \\ \frac{\partial f}{\partial y}(x, y) &= \frac{\sqrt[3]{x}}{3\sqrt[3]{y^2}} \text{ при } y \neq 0 \end{aligned}$$

Несложно понять, что если $x, y \neq 0$, то частные производные еще и непрерывны. То есть f точно дифференцируема при $x, y \neq 0$.

Теперь посмотрим на частную производную по x при $x = 0$:

$$\frac{\partial f}{\partial x}(0, y) = \lim_{h \to 0} \frac{f(h, y) - f(0, y)}{h} = \lim_{h \to 0} \frac{\sqrt[3]{hy}}{h} = \lim_{h \to 0} \frac{\sqrt[3]{y}}{h^{2/3}}$$

То есть при $y = 0$ получаем ноль, а при $y \neq 0$ частная производная не существует. Аналогично с производной по y:

$$\frac{\partial f}{\partial y}(x, y) = \begin{cases} 0, & x = 0 \\ \text{не существует}, & x \neq 0 \end{cases}$$

То есть в точках $(0, y)$ $(y \neq 0)$ и $(x, 0)$ $(x \neq 0)$ дифференцируемости нет, так как нет даже частных производных. А вот про $(0, 0)$ мы ничего сказать не можем: частные производные там есть, но не непрерывны (даже не заданы в окрестности). Поэтому условие теоремы не выполнено, но она на этот счёт ничего не говорит, надо проверять руками.

Разбираемся по определению.

$$f(h, k) = f(0, 0) + \underbrace{A}_{\text{матрица}} \cdot \begin{pmatrix} h \\ k \end{pmatrix} + o\left(\left|\begin{matrix} h \\ k \end{matrix}\right|\right)$$

Так как A — матрица из частных производных, то она тождественный ноль (так как частные производные нули).

$$\begin{aligned} f(h, k) &= f(0, 0) + o(\sqrt{h^2 + k^2}) \\ \sqrt[3]{hk} &\stackrel{?}{=} o(\sqrt{h^2 + k^2}) \end{aligned}$$

При $h = k$ имеем $h^{2/3} \neq o(h)$, то есть это неверно и дифференцируемости в $(0, 0)$ нет.

Пример 2.3.3. Мы уже знаем, что из непрерывности частных производных следует дифференцируемость. И знаем, что из дифференцируемости следует существование частных производных. Следует ли их непрерывность? Оказывается, что нет.

Возьмём функцию:

$$f(x,y) = \begin{cases} x^2+y^2, & \text{если ровно одна из } x,\ y \text{ рациональна} \\ 0, & \text{если обе рациональны или обе иррациональны} \end{cases}$$

Покажем, что она дифференцируема в $(0,0)$. Легко заметить, что $f(x,y) = f(0,0) + o(\sqrt{x^2+y^2})$ (при малых x, y), так как $f(x,y) \leqslant (x^2+y^2) = o(\sqrt{x^2+y^2})$.

Покажем, что с частными производыми всё не очень хорошо Возьмём $y_0 \in \mathbb{Q}$:

$$f(x,y_0) = \begin{cases} x^2+y^2, & x \notin \mathbb{Q}, \\ 0, & x \in \mathbb{Q} \end{cases}$$

Эта функция даже не непрерывна, и уж тем более не дифференцируема, то есть частные производные отсутствуют. Таким образом частных производных нет нигде, кроме $(0,0)$ и совсем странно будет рассуждать про их непрерывность.

Замечание 2.3.5. Вообще говоря, можно еще определить непрерывную дифференцируемость и вот там как раз получится равносильное условие.

𝔇𝔢𝔣 2.3.3. $f\colon D \to \mathbb{R}$ $(D \subset \mathbb{R}^n)$, $a \in \operatorname{int} D$ и f дифференцируема в окрестности точки a. Если $\|\mathrm{d}_x f - \mathrm{d}_a f\| \to 0$ при $x \to a$, то f — непрерывно дифференцируема в точке a. Перефразируя: есть дифференциал и он непрерывен.

Теорема 2.3.3. Пусть $f\colon D \to \mathbb{R}$, D открыто. Тогда: f непрерывно дифференцируема во всех точках $\iff$ все частные производные f существуют и непрерывны во всех точках.

▶ $\Rightarrow$ Пусть f непрерывно дифференцируема в точке a. Это значит, что f дифференцируема там же. То есть в окрестности имеются все частные производные. Осталось доказать их непрерывность

$$\frac{\partial f}{\partial e_k}(x) = \mathrm{d}_x f(e_k)$$

$$\left|\frac{\partial f}{\partial e_k}(x) - \frac{\partial f}{\partial e_k}(a)\right| = |\mathrm{d}_x f(e_k) - \mathrm{d}_a f(e_k)| =$$

$$= |(\mathrm{d}_x f - \mathrm{d}_a f)(e_k)| \leqslant \underbrace{\|\mathrm{d}_x f - \mathrm{d}_a f\|}_{\to 0,\text{ при } x \to a} \cdot \underbrace{\|e_k\|}_{=1}$$

$\Leftarrow$ Возьмём точку a. Так как частные производные существуют и непрерывны везде (и в окрестности точки a), то f дифференцируема в точке a.

Было когда-то утверждение: норма отображения не превосходит корня из суммы квадратов всех коэффициентов.

Распишем в лоб:

$$\|\mathrm{d}_x f - \mathrm{d}_a f\|^2 \leqslant \sum_{k=1}^{n} \left(\frac{\partial f}{\partial e_k}(x) - \frac{\partial f}{\partial e_k}(a)\right)^2 \to 0$$

Стремится к нулю, так как частные производные непрерывны $\Rightarrow$ каждое слагаемое непрерывно.

◀

Замечание 2.3.6. Если бы R была векторозначной, то теорема была бы верна, только доказательство несколько усложнится и надо будет говорить про вообще все частные производные.

2.4. Производные высших порядков

𝔇𝔢𝔣 2.4.1. Пусть $f : D \to \mathbb{R}$ ($D \subset \mathbb{R}^n$), $a \in \operatorname{int} D$. Пусть $g = \frac{\partial f}{\partial e_i}$ существует в окрестности точки a и у g: окрестность $a \to \mathbb{R}$ существует частная производная $\frac{\partial g}{\partial e_j}(a)$. Тогда её называют второй частной производной, обозначения:

- $\frac{\partial^2 f}{\partial e_j \partial e_i}$. Порядок можно запомнить так: $\frac{\partial}{\partial e_i}$ — это операция над функцией. Если применим сначала одну, потом другую, то получим $\frac{\partial}{\partial e_j}\frac{\partial}{\partial e_i}$, а выше — укороченная запись.
- $f''_{x_i x_j}$
- $D_{x_j x_i}$

Замечание 2.4.1. К сожалению, каноничного порядка индексов нет. Связано с тем, что обычно от порядка не зависит, за исключением специальных случаев.

Замечание 2.4.2. Укороченная запись: $\frac{\partial^2 f}{\partial e_i \partial e_i} = \frac{\partial^2 f}{\partial e_i^2}$

В том же духе можно ввести производные третьего, четвертого и больших порядков.

Пример 2.4.1. Пусть $f(x, y) = x^y$ и $x > 0$. Хотим посчитать вторые частные производные.

$$\frac{\partial f}{\partial x}(x, y) = yx^{y-1}$$
$$\frac{\partial f}{\partial y}(x, y) = \ln x x^y$$
$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial}{\partial y}(yx^{y-1}) = x^{y-1} + y\frac{\partial}{\partial y}(x^{y-1}) = x^{y-1} + y \ln x \cdot x^{y-1}$$
$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial}{\partial x}(\ln x \cdot x^y) = \frac{1}{x} \cdot x^y + \ln x \cdot \frac{\partial}{\partial x}(x^y) = x^{y-1} + \ln x \cdot yx^{y-1}$$

Вне зависимости от порядка значение второй частной производной совпало и это не случайность. Но так бывает не всегда.

Пример 2.4.2.

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

Посчитаем первые производные в точках, отличных от $(0, 0)$:

$$\frac{\partial f}{\partial x}(x, y) = y \cdot \frac{x^2 - y^2}{x^2 + y^2} + xy \cdot \frac{2x(x^2 + y^2) - 2x(x^2 - y^2)}{(x^2 + y^2)^2} = y \cdot \frac{x^2 - y^2}{x^2 + y^2} + xy \cdot \frac{4xy^2}{(x^2 + y^2)^2}$$
$$\frac{\partial f}{\partial y}(x, y) = -x \cdot \frac{y^2 - x^2}{y^2 + x^2} - xy \cdot \frac{4x^2 y}{(y^2 + x^2)^2}$$

А в $(0, 0)$ производные ноль, так как $xy = o(\sqrt{x^2 + y^2})$, а $\frac{x^2 - y^2}{x^2 + y^2}$ ограничено.

Теперь считаем вторые производные в $(0, 0)$:

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \lim_{y \to 0} \frac{\frac{\partial f}{\partial x}(0, y) - \frac{\partial f}{\partial x}(0, 0)}{y} = \lim_{y \to 0} \frac{-y - 0}{y} = -1$$
$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \lim_{x \to 0} \frac{\frac{\partial f}{\partial y}(x, 0) - \frac{\partial f}{\partial y}(0, 0)}{x} = \lim_{x \to 0} \frac{x - 0}{x} = 1$$

Видим, что не совпало.

Замечание 2.4.3. В следующий раз будет теорема, рассказывающая о том, когда неважен порядок. Если есть функция двух переменных, то порядок неважен, если частные производные непрерывны. В последнем примере проблема была именно в этом.

Теорема 2.4.1. $f\colon D \subset \mathbb{R}^2 \to \mathbb{R}$, $(x_0, y_0) \in D$, D — открыто, f''_{xy} и f''_{yx} существуют в окрестности (x_0, y_0) и непрерывны в (x_0, y_0). Тогда

$$f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0)$$

▶

$$\Delta = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0)$$

Два раза применяем теорему Лагранжа

$$\varphi(s) = f(s, y_0 + k) - f(s, y_0) \quad \varphi \text{ дифференцируема на } [x_0, x_0 + h]$$
$$\Delta = \varphi(x_0 + h) - \varphi(x_0) = \varphi'(x_0 + \theta_1 h)h \quad \theta_1 \in (0, 1)$$
$$\Delta = \varphi'(x_0 + \theta_1 h)h = (f'_x(x_0 + \theta_1 h, y_0 + k) - f'_x(x_0 + \theta_1 h, y_0))h$$
$$\tilde{\varphi}(t) = f'_x(x_0 + \theta_1 h, t) \quad \tilde{\varphi} \text{ дифференцируема на } [y_0, y_0 + k]$$
$$\Delta = (\tilde{\varphi}(y_0 + k) - \tilde{\varphi}(y_0))h = \tilde{\varphi}'(y_0 + \theta_2 k)kh \quad \theta_2 \in (0, 1)$$
$$\Delta = f''_{xy}(x_0 + \theta_1 h, y_0 + \theta_2 k)hk$$

Теперь всё в другом порядке

$$\psi(t) = f(x_0 + h, t) - f(x_0, t) \quad \psi \text{ дифференцируема на } [y_0, y_0 + k]$$
$$\Delta = \psi(y_0 + k) - \psi(y_0) = \psi'(y_0 + \theta_3 k)k \quad \theta_3 \in (0, 1)$$
$$\Delta = \psi'(y_0 + \theta_3 k)k = (f'_y(x_0 + h, y_0 + \theta_3 k) - f'_x(x_0, y_0 + \theta_3 k))k$$
$$\tilde{\psi}(s) = f'_y(s, y_0 + \theta_3 k) \quad \tilde{\psi} \text{ дифференцируема на } [x_0, x_0 + h]$$
$$\Delta = (\tilde{\psi}(x_0 + h) - \tilde{\psi}(x_0))k = \tilde{\psi}'(x_0 + \theta_4 k)kh \quad \theta_4 \in (0, 1)$$
$$\Delta = f''_{yx}(x_0 + \theta_4 h, y_0 + \theta_3 k)hk$$
$$f''_{xy}(x_0 + \theta_1 h, y_0 + \theta_2 k)hk = f''_{yx}(x_0 + \theta_4 h, y_0 + \theta_3 k)hk$$

Устремим h и k к нулю

$$f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0)$$

◀

𝔇𝔢𝔣 2.4.2. $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$, D открыто. Говорим, что f непрерывно дифференцируема r раз (r— гладкая), если у f существуют все частные производные порядка не более r и они все непрерывны.

Или, что равносильно:
Говорим, что f непрерывно дифференцируема r раз (r— гладкая), если у f координатные функции непрерывно дифференцируемы r раз.

$$f \in C^r(D)$$

Замечание 2.4.4. Свойства r— гладкости сохраняются при арифметических операциях и композиции.

Теорема 2.4.2. $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$, D открыто, $f \in C^r(D)$, $(i_1, i_2, \dots, i_r)$ и $(j_1, j_2, \dots, j_r)$ — перестановки. Тогда

$$\frac{\partial^r f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_r}} = \frac{\partial^r f}{\partial x_{j_1} \partial x_{j_2} \dots \partial x_{j_r}}$$

▶ $(i_1, i_2, \dots, i_r)$ и $(j_1, j_2, \dots, j_r)$. Тогда перестановка из одного в другое получается через несколько транспозиций (на алгебре было). Достаточно доказать теорему для транспозиции:

$$\frac{\partial^r f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k} \partial x_{i_{k+1}} \dots \partial x_{i_r}} \overset{(!)}{=} \frac{\partial^r f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_{k+1}} \partial x_{i_k} \dots \partial x_{i_r}}$$

Возьмём функцию

$$g(s,t) = \frac{\partial^{k-1} f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_{k-1}}}(\dots, \underbrace{s}_{k}, \underbrace{t}_{k+1}, \dots)$$

По предыдущей теореме знаем, что

$$g''_{st} = g''_{ts}$$
$$\frac{\partial^{k+1} f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k} \partial x_{i_{k+1}}} = \frac{\partial^{k+1} f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_{k+1}} \partial x_{i_k}}$$

Додифференцируем это дальше по всем оставшимся переменным. ◀

Замечание 2.4.5. Если очень надо, можно строго записать порядок производных:

$$\frac{\partial}{\partial x} \frac{\partial}{\partial y} f$$

2.5. Формула Тейлора

Много обозначений:

𝔇𝔢𝔣 2.5.1. Мультииндекс — вектор из целых неотрицательных чисел

$$k = (k_1, k_2, \dots, k_r)$$

𝔇𝔢𝔣 2.5.2. Высота мультииндекса $|k|$ — сумма его координат.

𝔇𝔢𝔣 2.5.3. Факториал мультииндекса:

$$k! = k_1!\, k_2! \dots k_n!$$

𝔇𝔢𝔣 2.5.4. $h \in \mathbb{R}^n$, k — мультииндекс.

$$h^k = h_1^{k_1} h_2^{k_2} \dots h_n^{k_n}$$

$0^0 = 1$

𝔇𝔢𝔣 2.5.5. Производная порядка мультииндекса:

$$f^{(k)}(x) = \frac{\partial^{|k|} f}{\partial x_1^{k_1} \partial x_2^{k_2} \dots \partial x_r^{k_r}}$$

Лемма 2.5.1. $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$, D открыто, $f \in C^r(D)$, $[x_0, x_0 + h] \subset D$ (отрезок). $F(t) \leftrightharpoons f(x_0 + th)$, $t \in [0,1]$.

$$F^{(r)}(t) = r! \sum_{\substack{|k|=r \\ k\text{ — мультииндекс}}} \frac{f^{(k)}(x_0 + th)}{k!} h^k$$

Замечание 2.5.1. Рассмотрим для прояснения двумерный случай:

$$f = (i, r-i)$$
$$f^{(k)} = \frac{\partial^r f}{\partial x^i \partial y^{r-i}}$$
$$h = (u, v)$$
$$h^k = u^i v^{r-i}$$
$$\frac{r!}{k!} = \frac{r!}{i!\,(r-i)!} = \binom{r}{i}$$
$$f(x,y) \quad F(t) = f(x_0 + tu, y_0 + tv)$$
$$F^{(r)}(t) = \sum_{i=0}^{r} \binom{r}{i} \frac{\partial^r f}{\partial x^i \partial y^{r-i}} (x_0 + tu, y_0 + tv) u^i v^{r-i}$$

▶ Докажем индукцией по r:

База $n = 0$: Написано, что $F(t) = f(x_0 + th)$.

Переход: Возьмём и продифференцируем ещё раз:

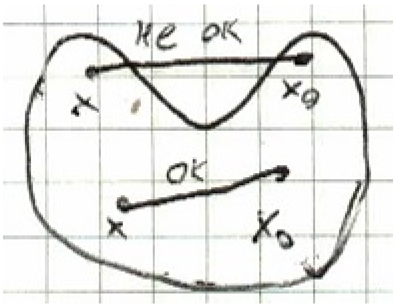
$$\left(F^{(r)}(t)\right)' = \left(r! \sum_{|k|=r} \frac{f^{(k)}(x_0 + th)}{k!} h^k\right)'_t = r! \sum_{|k|=r} \frac{\left(f^{(k)}(x_0 + th)\right)'_t}{k!} h^k$$
$$g \leftrightharpoons f^{(k)}$$
$$g(x_0 + ht)'_t = \frac{\partial g}{\partial x_1}(x_0 + th)h_1 + \frac{\partial g}{\partial x_2}(x_0 + th)h_2 + \dots + \frac{\partial g}{\partial x_n}(x_0 + th)h_n = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(x_0 + th)h_i$$
$$k + e_i = (k_1, k_2, \dots, k_i + 1, \dots, k_n)$$
$$\left(F^{(r)}(t)\right)' = r! \sum_{|k|=r} \sum_{i=1}^{n} \frac{f^{(k+e_i)}(x_0 + th)}{k!} \underbrace{h^k h_i}_{=h^{k+e_i}} = r! \sum_{|k|=r} \sum_{i=1}^{n} \frac{f^{(k+e_i)}(x_0 + th)}{(k+e_i)!} h^{k+e_i} \frac{(k+e_i)!}{k!} =$$
$$\frac{(k+e_i)!}{k!} = \frac{k_1!\, k_2! \dots (k_i + 1)! \dots k_n!}{k_1!\, k_2! \dots k_i! \dots k_n!} = k_i + 1$$
$$= (r+1)! \sum_{i=1}^{n} \sum_{|k|=r} \frac{f^{(k+e_i)}(x_0 + th)}{(k+e_i)!} h^{k+e_i} \cdot \frac{k_i + 1}{r+1} =$$
$$\frac{k_i + 1}{r+1} = \frac{l_i}{r+1}$$

Заметим, что под двумя суммами мы считаем производные по всем мультииндексам вида $k + e_i$, где $|k| = r$. Все они — мультииндексы высоты $r + 1$. Каждый мультииндекс мы могли получить несколько раз. Рассмотрим мультииндекс l высоты $r + 1$. Мы его могли получить его, прибавив к некоторому k e_i. Это могло произойти если только $l_i \neq 0$. Как теперь видно, последний множитель ($\frac{k_i+1}{r+1}$) или равен нулю для заданного l, которое таким прибавлением всё равно не удалость получить, или как раз характеризует, сколько раз могли получить l из x.
Осталось заметить, что все l_i просуммируются в $|l| = r + 1$. Таким образом,

$$= (r+1)! \sum_{|l|=r+1} \frac{f^l(x_0 + th)}{l!} h^l$$

◀

Теорема 2.5.1. Пусть есть многомерная функция $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$. f $(r+1)$ раз непрерывно дифференцируема в D. Также есть отрезок $[x, x_0] \subseteq D$ (отрезок по прямой).



Тогда можно написать формулу (здесь k — мультииндексы):

$$f(x) = \sum_{|k| \leqslant r} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \sum_{|k|=r+1} \frac{f^{(k)}(x_0 + \theta(x-x_0))}{k!}(x-x_0)^k$$

▶ Возьмём $F\colon [0,1] \to \mathbb{R}$ такое: $F(t) = f(x_0 + t\underbrace{(x-x_0)}_{h})$. Это обычная функция из $[0,1]$ в $\mathbb{R}$, она дифференцируема $r+1$ раз (так как все точки отрезка $[x, x_0]$ лежат в D). Напишем формулу Тейлора для неё:

$$F(1) = \sum_{l=0}^{r} \frac{F^{(l)}(0)}{l!} \cdot 1^l + \frac{F^{(r+1)}(\theta)}{(r+1)!} \cdot 1^{r+1}$$

Вспомним нашу последнюю лемму, которая выражает l-ю производную F:

$$F^{(l)}(t) = \sum_{|k|=l} \frac{l!}{k!} f^{(k)}(x_0 + th)h^k$$

Теперь подставим это дело в формулу Тейлора для F:

$$\begin{aligned}
F^{(l)}(0) &= \sum_{|k|=l} \frac{l!}{k!} f^{(k)}(x_0)h^k \\
F^{(l)}(\theta) &= \sum_{|k|=l} \frac{l!}{k!} f^{(k)}(x_0 + \theta h)h^k \\
F(1) &= \sum_{l=0}^{r} \sum_{|k|=l} \frac{1}{k!} f^{(k)}(x_0)h^k + \sum_{|k|=r+1} \frac{1}{k!} f^{(k)}(x_0 + \theta h)h^k \\
&= \sum_{|k| \leqslant r} \frac{f^{(k)}(x_0)}{k!}h^k + \sum_{|k|=r+1} \frac{f^{(k)}(x_0+\theta h)}{k!}h^k
\end{aligned}$$

◀

Замечание 2.5.2. $\sum_{|k|\leqslant r} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$ — многочлен Тейлора степени r.

Замечание 2.5.3. Если $r = 0$, то получим многомерную формулу Лагранжа:

$$f(x) = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0 + \theta h) \cdot h_i = f(x_0) + \langle \nabla f(x_0 + \theta h), h \rangle$$

Следствие 2.5.1.1. Формула Тейлора с остатком в форме Лагранжа:

$$f(x_0 + h) = f(x_0) + \langle \nabla f(x_0 + \theta h), h \rangle$$

Замечание 2.5.4. Формула Тейлора при $n = 2$ и произвольном r (двумерный случай) в точке (x_0, y_0) с сдвигом (h, k):

$$\begin{aligned} f(x_0 + h, y_0 + k) = f(x_0, y_0) &+ \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k + \\ &+ \frac{1}{2}\left(\frac{\partial^2 f}{\partial x \partial x}(x_0, y_0)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)hk + \frac{\partial^2 f}{\partial y \partial y}(x_0, y_0)k^2\right) + \\ &+ \dots + \frac{1}{l!}\sum_{i=0}^{l}\binom{l}{i}\frac{\partial^l f}{\partial x^i \partial y^{l-i}}(x_0, y_0)h^i k^{l-i} + \dots + \\ &+ \frac{1}{(r+1)!}\sum_{i=0}^{r+1}\binom{r+1}{i}\frac{\partial^{r+1} f}{\partial x^i \partial y^{r+1-i}}(x_0 + \theta h, y_0 + \theta k)h^i k^{r+1-i} \end{aligned}$$

Теорема 2.5.2. Формула Тейлора с остатком в форме Пеано:

$$f(x_0 + h) = \sum_{|k| \leqslant r}\frac{f^{(k)}(x_0)}{k!}h^k + o(\|h\|^r) \text{ при } \|h\| \to 0$$

▶ Выпишем формулу из последней теоремы (только уменьшим r на единицу для удобства):

$$\begin{aligned} f(x + h) &= \sum_{|k| \leqslant r-1}\frac{f^{(k)}(x_0)}{k!}h^k + \sum_{|k| = r}\frac{f^{(k)}(x_0 + \theta h)}{k!}h^k = \\ &= \sum_{|k| \leqslant r}\frac{f^{(k)}(x_0)}{k!}h^k + \sum_{|k| = r}\frac{f^{(k)}(x_0 + \theta h) - f^{(k)}(x_0)}{k!}h^k \end{aligned}$$

Надо оценить вторую сумму. Так как там слагаемых конечное количество, достаточно оценить каждое слагаемое. Пусть у нас имеется мультииндекс $(k_1, \dots, k_n)$, причём $\sum k_i = r$. На $k!$ тоже забьём, это константа:

$$\left(f^{(k)}(x_0 + \theta h) - f^{(k)}(x_0)\right)\frac{h^k}{\|h\|^r} \overset{?}{\to} 0 \text{ при } h \to 0$$

$$\left(f^{(k)}(x_0 + \theta h) - f^{(k)}(x_0)\right) \to 0$$

$$\frac{h^k}{\|h\|^r} = \frac{h_1^{k_1}h_2^{k_2}\dots h_n^{k_n}}{\|h\|^r} = \frac{h_1^{k_1}h_2^{k_2}\dots h_n^{k_n}}{\|h\|^{k_1}\|h\|^{k_2}\dots\|h\|^{k_n}} = \prod_{i=1}^{n}\frac{h_i^{k_i}}{\|h\|^{k_i}}$$

$$\frac{|h^k|}{\|h\|^r} = \prod_{i=1}^{n}\frac{|h_i|^{k_i}}{\|h\|^{k_i}} \leqslant 1$$

$$(\to 0)\cdot(\leqslant 1) \to 0$$

◀

Теорема 2.5.3 (Полиномиальная формула)**.**

$$\left(\sum_{i=1}^{n}x_i\right)^r = \sum_{|k|=r}\frac{r!}{k!}x^k = \sum_{k_1+k_2+\dots+k_n=r}\binom{r}{k_1, \dots, k_n}x_1^{k_1}x_2^{k_2}\dots x_n^{k_n}$$

где

$$\frac{r!}{k!} = \frac{r!}{k_1!\,k_2!\dots k_n!} = \binom{r}{k_1, k_2, \dots, k_n}$$

(мультиномиальный коэффициент).

▶ Можно по индукции, это несложно. Но давайте лучше жахнем из формулы Тейлора. Пусть

$$f(x) = \left(\sum_{i=1}^{n} x_i\right)^r$$

Найдём его частную производную по x_i:

$$\frac{\partial f}{\partial x_i}(x) = r\left(\sum_{i=1}^{n} x_i\right)^{r-1}$$

$$\frac{\partial^l f}{\partial x_{i_1} \dots \partial x_{i_l}}(x) = r(r-1)(r-2)\dots(r-l+1)\left(\sum_{i=1}^{n} x_i\right)^{r-l}$$

Посмотрим на неё в точке $x = 0$. При $l > r$ у нас остаток точно занулится (потому что встретится множитель ноль). При $l < r$ у нас остаток тоже занулится (так как последний множитель есть ноль в ненулевой степени). А при $l = r$ остаток получится равен $r!$.

Подставляем в формулу Тейлора с остатком в форме Лагранжа ($x_0 = 0$):

$$f(h) = \underbrace{\sum_{|k|\leqslant r} \frac{f^{(k)}(0)}{k!}h^k}_{=0 \text{ при } |k|<r} + \underbrace{\sum_{|k|=r+1} \frac{f^{(k)}(\theta h)}{k!}h^k}_{=0} = \sum_{|k|=r} \frac{r!}{k!}h^k$$

◀

2.6. Экстремумы функций

𝔇𝔢𝔣 2.6.1. $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ и $a \in D$. Говорим, что a — точка минимума, если $\exists U$ — открытая окрестность a такая, что

$$\forall x \in U \cap D\colon f(x) \geqslant f(a)$$

Замечание 2.6.1. Можно аналогично определить a — точка максимума.

𝔇𝔢𝔣 2.6.2. Говорим, что a — точка строгого минимума, если $\exists \mathring{U}$ — проколотая открытая окрестность a такая, что

$$\forall x \in \mathring{U} \cap D\colon f(x) > f(a)$$

Дальше будем учиться по производным делать какие-то выводы о наличии экстремумов в точках. По тому, что хочется — полная аналогия с одномерным случаем. По тому, что получится — намного больше наворотов.

Теорема 2.6.1 (Необходимое условие экстремума)**.** $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ и $a \in \operatorname{int} D$. Если a — точка экстремума (максимума/минимума) и $\frac{\partial f}{\partial x_i}$ существует в точке a, то $\frac{\partial f}{\partial x_i}(a) = 0$.

Замечание 2.6.2. В частности, если f дифференцируема в точке a, то $\nabla f(a) = 0$.

Замечание 2.6.3. На самом деле, если f дифференцируема, то у нас производная по любому направлению обращается в ноль. Но они все выражаются через $\frac{\partial f}{\partial x_i}$, поэтому про них говорить излишне.

▶ Пусть a — максимум. Взяли какую-нибудь окрестность из определения максимума, пересечём её с множеством D (в определении не требовалось $U \subseteq D$). Рассмотрим одну координату x_k. Положим $\varphi(t) = f(a_1, \dots, a_{k-1}, t, a_{k+1}, \dots a_n)$. Тогда, очевидно, a_k — точка максимума для $\varphi(t)$, так как φ — это просто сужение f на прямую. А если есть частная производная по x_k, то φ дифференцируема. По одномерной теореме $\varphi'(a_k) = 0$, то есть $\frac{\partial f}{\partial x_k}(a) = 0$. ◀

𝔇𝔢𝔣 2.6.3. a — стационарная точка, если $\nabla f(a) = 0$.

Теперь хотим узнавать, какие из стационарных точек являются экстремумами. В одномерном случае была такая же проблема: там были «сёдла». Мы решали эту проблему, написав многочлен Тейлора до следующего слагаемого. Попробуем порассуждать здесь так же.

Давайте возьмём стационарную точку и напишем многочлен Тейлора до второго слагаемого. Выглядит она вот так:

$$f(a+h) = f(a) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o(\|h\|^2)$$

▶

$$f(a+h) = f(a) + \sum_{|k|=2} \frac{f^{(k)}(a)h^k}{k!} + o(\|h\|^2)$$

У нас бывают мультииндексы двух типов:

1. $k = (0, \dots, 0, 2, 0, \dots, 0)$. Тогда у нас это слагаемое встретится в исходной сумме один раз и в знаменателе будет $k! = 2! = 2$. То есть выносится коэффициент $\frac{1}{2}$.
2. $k = (0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots, 0)$. Это слагаемое встретится в исходной сумме два раза (для (i, j) и для (j, i)) вместо одного, но в знаменателе будет $k! = 1! \cdot 1! = 1$. То есть тоже выносится коэффициент $\frac{1}{2}$.

◀

Теперь надо убедиться, что остаток везде имеет одинаковый знак.

2.6.1. Квадратичные формы

𝔇𝔢𝔣 2.6.4. Пусть C — симметричная матрица $n \times n$. Тогда квадратичная форма $Q(h) \colon \mathbb{R}^n \to \mathbb{R}$ по матрице C есть:

$$Q(h) = \sum_{i,j=1}^{n} c_{i,j} h_i h_j$$

Замечание 2.6.4. Если h — вектор-столбец, то $Q(h) = h^\top C h$. Но мы этим пользоваться не будем.

𝔇𝔢𝔣 2.6.5. Q — строго положительно определённа, если

$$\forall h \neq 0 \colon Q(h) > 0$$

Замечание 2.6.5. Аналогично можно определить нестрого положительно определённую. И отрицательно определённые.

Пример 2.6.1. Если $C = E$, то у нас получается строго положительно определённая форма, потому что $Q(h) = \sum h_i^2$.

Теорема 2.6.2 (Критерий Сильвестра)**.** Пусть есть матрица

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & \dots \\ c_{21} & c_{22} & c_{23} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Рассмотрим определители:

$$\det\left(c_{11}\right)$$
$$\det\begin{pmatrix}c_{11} & c_{12}\\ c_{21} & c_{22}\end{pmatrix}$$
$$\det\begin{pmatrix}c_{11} & c_{12} & c_{13}\\ c_{21} & c_{22} & c_{23}\\ c_{31} & c_{32} & c_{33}\end{pmatrix}$$
$$\vdots$$
$$\det C$$

Строгая положительная определённость эквивалентна тому, что все эти определители строго больше нуля. Для нестрогого — нестрого.

Замечание 2.6.6. Будет когда-нибудь доказано на алгебре.

Замечание 2.6.7. Чтобы сказать что-то для отрицательной определённости, надо рассмотреть матрицу $-C$. Определители поменяют знаки в зависимости от чётности размера аргумента.

2.6.2. Связь экстремумов и квадратичных форм

Теорема 2.6.3. $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ и $a \in \operatorname{int} D$. Пусть f дважды непрерывно дифференцируема в окрестности точки a, и a — стационарная точка. Пишем квадратичную форму

$$Q(h) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j$$

- Если Q — строго положительно определённая, то a — строгий минимум
- Если Q — строго отрицательно определённая, то a — строгий максимум
- Если a — нестрогий минимум, то Q нестрого положительно определена
- Если a — нестрогий максимум, то Q нестрого отрицательно определена

Замечание 2.6.8. Есть некоторая свобода. Может получиться так, что форма нестрого определена, а экстремума нет. Такая же проблема бывала и в одномерном случае: когда у нас вторая производная обращалась в ноль, приходилось работать головой и руками в каждом отдельном случае.

Лемма 2.6.1. Если $Q(h)$ — строго положительно определена, то существует $c > 0$, что

$$\forall h, Q(h) \geqslant c\|h\|^2$$

▶ Рассмотрим Q на единичной сфере. Получим непрерывную на компакте функцию, поэтому

$$\exists h_0\colon \underbrace{Q(h_0)}_{>0} = \min\{Q(h)\} \eqcolon c$$

$$Q(h) = Q\left(\|h\| \cdot \frac{h}{\|h\|}\right) = \|h\|^2 Q\left(\frac{h}{\|h\|}\right) \geqslant \|h\|^2 c$$

◀

Докажем теперь теорему:

▶

$$f(a+h)-f(a)=\underbrace{\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(a)h_ih_j}_{=Q(h)}+o(\|h\|^2)$$

1. $Q(h)$ — строго положительно определена, откуда $Q(h)\geqslant c\|h\|^2$ для некоторого $c>0$ по лемме:

$$f(a+h)-f(a)=\frac{1}{2}Q(h)+o(\|h\|^2)\geqslant\frac{c\|h\|^2}{2}+o(\|h\|^2)=\|h\|^2\underbrace{\left(\frac{c}{2}+o(1)\right)}_{>0\text{ при малых }h}$$

2. $f\to -f$
3. a нестрогий минимум, поэтому $f(a+h)\geqslant f(a)$ при достаточно малых h. Зафиксируем h и будем рассматривать $f(a+th)\geqslant f(a)$ при достаточно малых t.

$$f(a+th)-f(a)=\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(a)t^2h_ih_j+\underbrace{o(\|th\|^2)}_{=t^2o(1)}=\frac{t^2}{2}Q(h)+t^2o(1)$$

$$0\leqslant\frac{f(a+th)-f(a)}{t^2}=\frac{1}{2}Q(h)+o(1)$$

Переходим к пределу по t (h и t независимы):

$$0\leqslant\frac{1}{2}Q(h)$$

4. $f\to -f$

◀

Дальше Тейлора выписывать не рекомендуют: вместо матриц выйдет табличка четырёхмерных коэффициентов.

𝔇𝔢𝔣 2.6.6. Гессиан или матрица Гессе — та самая матрица, квадратичную форму которой мы и смотрели:

$$\left(\frac{\partial^2 f}{\partial x_i\partial x_j}\right)_{i,j=1}^{n}$$

2.7. Обратные отображения

Ещё одна теорема больше из алгебры:

Теорема 2.7.1. $A\colon\mathbb{R}^n\to\mathbb{R}^n$ — линейный оператор. Тогда следующее равносильно:

1. A обратим
2. $A(\mathbb{R}^n)=\mathbb{R}^m$
3. Уравнение $Ax=0$ имеет нулевое решение
4. $\det A\neq 0$

▶ Докажем не очень строго: перезапишем всё выше в терминах систем уравнений:

1. Для всех y уравнение $Ax=y$ имеет единственное решение
2. Для всех y уравнение $Ax=y$ имеет решение

3. Уравнение $Ax = 0$ имеет нулевое решение
4. $\det A \neq 0$

Получилось, что система линейных уравнений имеет единственное решение для любых свободных коэффициентов, в том числе для нулевых свободных коэффициентов — тривиальное, тогда и только тогда, когда определитель матрицы системы не ноль. ◀

Теорема 2.7.2. $A\colon \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор,

$$\exists m > 0\colon \forall x \in \mathbb{R}^n, \|Ax\| \geqslant m\|x\|$$

Тогда A обратим и $\|A^{-1}\| \leqslant \frac{1}{m}$.

▶

$$Ax = 0 \Rightarrow 0 = \|Ax\| \geqslant m\|x\| \Rightarrow x = 0$$

Смотрим теорему выше.

$$\|A^{-1}\| = \inf\{c > 0 \mid \|A^{-1}x\| \leqslant c\|x\|\}$$

Нам достаточно показать, что

$$\|A^{-1}x\| \leqslant \frac{1}{m}\|x\|$$

$y \leftrightharpoons A^{-1}x$, $x = Ay$

$$\|y\| \leqslant \frac{1}{m}\|Ay\|$$

Это мы уже знаем ◀

Замечание 2.7.1.

$$\forall x, \|Ax\| \geqslant \frac{1}{\|A^{-1}\|}\|x\|$$

▶

$$\|A^{-1}\|\,\|Ax\| \geqslant \|x\|$$

$$\|x\| = \|A^{-1}(Ax)\| \leqslant \|A^{-1}\|\,\|Ax\|$$

◀

Теорема 2.7.3 (Обратнимость линейных операторов, близких к обратимым)**.** $A\colon \mathbb{R}^n \to \mathbb{R}^m$ — обратимый линейный оператор, $\|A^{-1}\| = \frac{1}{\alpha}$.

1. Если $B\colon \mathbb{R}^n \to \mathbb{R}^m$ — линейный оператор, $\|B - A\| = \beta < \alpha$, то B обратим и

$$\|B^{-1} - A^{-1}\| < \frac{\beta}{\alpha(\alpha - \beta)}$$

2. Пусть Ω — множество обратимых операторов. Тогда $f\colon B \to B^{-1}$ непрерывна на Ω.

▶ 1.

$$\|Ax\| = \|Bx + (A - B)x\| \leqslant \|Bx\| + \|(A - B)x\| \leqslant \|Bx\| + \|(A - B)\|\|x\| = \|Bx\| + \beta\|x\|$$

$$\|Ax\| \geqslant \frac{1}{\|A^{-1}\|}\|x\| = \alpha\|x\|$$

$$(\alpha - \beta)\|x\| \leqslant \|Bx\|$$

По предыдущей теореме уже получили обратимость.

$$B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1}$$

$$\|B^{-1} - A^{-1}\| = \|B^{-1}(A - B)A^{-1}\| \leqslant \|B^{-1}\|\,\|A - B\|\,\|A^{-1}\| = \frac{1}{\alpha - \beta}\beta\frac{1}{\alpha} = \frac{\beta}{\alpha(\alpha - \beta)}$$

2.
$$B_k \to A \Rightarrow \|B_k - A\| \to 0 \overset{?}{\Rightarrow} \|B_k^{-1} - A^{-1}\| \to 0 \Rightarrow B_k^{-1} \to A^{-1}$$

$\|A^{-1}\| = \frac{1}{\alpha}$

$$\|B_k - A\| \leqslant \frac{\alpha}{2}$$
$$\|B_k^{-1} - A^{-1}\| \leqslant \frac{\|B_k - A\|}{\alpha(\alpha - \|B_k - A\|)} \leqslant \frac{2}{\alpha^2}\|B_k - A\| \to 0$$

◀

Теорема 2.7.4 (Об обратной функции)**.** $f\colon D \subset \mathbb{R}^n \to \mathbb{R}^n$, $a \in \operatorname{int} D$, f непрерывно дифференцируема в окрестности a, $\mathrm{d}_a f$ — обратимое отображение, $b = f(a)$. Тогда

1. Существует окрестности U точки a и V точки b, что $f\colon V \to U$ — биекция.
2. Пусть $g\colon V \to U$ — обратная к f. Тогда g — непрерывно дифференцируема.

Замечание 2.7.2. В случае верности теоремы

$$g(f(x)) = x$$

$$\mathrm{d}_{f(x)} g \circ \mathrm{d}_x f = id$$

$$\mathrm{d}_{f(x)} g = (\mathrm{d}_x f)^{-1}$$

Замечание 2.7.3. Это самая адовая теорема курса! Доказателсьтво длинное, разбито на важные шаги.

▶ $A = \mathrm{d}_a f$. Выберем $\lambda\colon 4\lambda\|A^{-1}\| = 1$. В качестве U возьмём шар с центром в a, что

$$\forall x \in U, \|\mathrm{d}_x f - \mathrm{d}_a f\| \leqslant 2\lambda$$

Шаг 1: Докажем, что f инъективна на U. Возьмём две точки $x, x+h \in U$. Покажем, что значения $f(x)$ и $f(x+h)$ различны. В шаре содержится отрезок $[x, x+h]$.

$$F\colon [0,1] \to \mathbb{R}^n; F(t) = f(x+th) - tAh$$
$$F'(t) = \mathrm{d}_{x+th} f(h) - Ah$$
$$\|F(1) - F(0)\| \leqslant \|F'(t)\| = \|\mathrm{d}_{x+th} f(h) - Ah\| \leqslant \quad \text{при некотором } t \text{ по теореме Лагранжа}$$
$$\leqslant \underbrace{\|\mathrm{d}_{x+th} f - A\|}_{x+th \in U} \|h\| < 2\lambda\|h\|$$
$$F(0) = f(x)$$
$$F(1) = f(x+h) - Ah$$
$$\|f(x+h) - Ah - f(x)\| \leqslant 2\lambda\|h\| = \frac{1}{2\,\|A^{-1}\|}\|h\| = \frac{\|A^{-1}Ah\|}{2\|A^{-1}\|} \leqslant \frac{\|Ah\|}{2}$$
$$\|f(x+h) - f(x)\| = \|f(x+h) - f(x) - Ah + Ah\| \geqslant \|Ah\| - \|f(x+h) - Ah - f(x)\| \geqslant$$
$$\geqslant \|Ah\| - \frac{\|Ah\|}{2} = \frac{\|Ah\|}{2}$$
$$f(x+h) = f(x) \Rightarrow 0 \geqslant \frac{\|Ah\|}{2} \Rightarrow \|Ah\| = 0 \Rightarrow h = 0$$

Таким образом, $f(x+h) \neq f(x)$.

Шаг 2: Покажем, что $f(U)$ открыто. Для этого покажем, что для $x_0 \in U$ образ $f(x_0)$ есть окрестность в $f(U)$. Выберем $r > 0$ такой, что $\bar{B}_r(x_0) \subset U$. Докажем, что $B_{\lambda r}(f(x_0)) \subset f(U)$. Возьмём $y \in B_{\lambda r}(f(x_0))$, $\|y - f(x_0)\| < \lambda r$ и найдём $x \in U$, что $f(x) = y$. Рассмотрим

$$\varphi(x)\colon \bar{B}_r(x_0) \to R \quad \varphi(x) = \|y - f(x)\|$$
$$2\lambda\|x - x_0\| < \|f(x) - f(x_0)\| = \|y - f(x_0) - (y - f(x))\| \leqslant$$
$$\leqslant \|y - f(x_0)\| + \|y - f(x)\| = \varphi(x) + \varphi(x_0)$$

Возьмём x на границе $\bar{B}_r(x_0)$, то есть $\|x - x_0\| = r$.

$$2\lambda r < \varphi(x) + \varphi(x_0) = \varphi(x) + \|y - f(x_0)\| < \varphi(x) + \lambda r$$

Таким образом,

$$2\lambda r < \varphi(x) + \lambda r \Rightarrow \varphi(x) > \lambda r > \varphi(x_0)$$

Получили, что у φ и φ^2 на границе значения больше, чем центре.
φ^2 — непрерывно на компакте $\bar{B}_r(x)$, значит φ^2 достигает наименьшего значения в x_*, причём не на границе. φ^2 дифференцируема:

$$\varphi^2(x) = \sum_{i=1}^{n} (y_k - f_k(x))^2$$
$$\nabla\varphi^2(x_*) = 0$$
$$\frac{\partial\varphi^2}{\partial x_i}(x_*) = \sum_{i=1}^{n} 2\left(-\frac{\partial f_k}{\partial x_i}(x_*)\right)(y_k - f_k(x_*))$$
$$\underbrace{\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_*) & \cdots & \frac{\partial f_n}{\partial x_1}(x_*) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n}(x_*) & \cdots & \frac{\partial f_n}{\partial x_n}(x_*) \end{pmatrix}}_{=\left(\mathrm{d}_{x_*} f\right)^T} \begin{pmatrix} y_1 - f_1(x_*) \\ \vdots \\ y_n - f_n(x_*) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\mathrm{d}_{x_*} f$ обратимо, а значит $\left(\mathrm{d}_{x_*} f\right)^T$ обратима и невырожденна, откуда

$$\begin{pmatrix} y_1 - f_1(x_*) \\ \vdots \\ y_n - f_n(x_*) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$
$$y - f(x_*) = 0$$

Получили, что

$$\forall y \in B_{\lambda r} f(x_0), \exists x_* \in B_r(x_0) \subset U\colon f(x_*) = y$$

и $B_{\lambda r} f(x_0) \subset f(U)$.

Шаг 3: Покажем, что g дифференцируема. Возьмём $y, y + k \in V = f(U)$. Покажем, что

$$g(y + k) - g(y) = Bk + o(\|k\|)$$

Заметим:

$$y \in V \Rightarrow \exists x \in U\colon f(x) = y$$
$$y + k \in V \Rightarrow \exists x + h \in U\colon f(x + h) = y + k$$
$$g(y + k) - g(y) = x + h - x = h$$
$$k = f(x + h) - f(x) = \mathrm{d}_x f(h) + r(h) \quad \frac{r(h)}{\|h\|} \to 0$$

$B = (\mathrm{d}_x f)^{-1}$

$$Bk = \underbrace{B\left(\mathrm{d}_x f(h)\right)}_{=h} + Br(h)$$

$$g(y+k) - g(y) = h = Bk - Br(h)$$

Осталось показать, что $\frac{Br(h)}{\|k\|} \to 0$ при $k \to 0$:

$$\frac{\|Br(h)\|}{\|k\|} \leqslant \frac{\|B\|\|r(h)\|}{\|h\|}\frac{\|h\|}{\|k\|}$$

$$\|k\| = \|f(x+h) - f(x)\| > 2\lambda\|h\|$$

$$\frac{\|h\|}{\|k\|} < \frac{1}{2\lambda}$$

$$\frac{\|B\|\|r(h)\|}{\|h\|}\frac{\|h\|}{\|k\|} \to 0$$

Шаг 4: Докажем, что g непрерывно дифференцируема. Мы уже знаем производную g из замечения перед доказательтвом.

$$\mathrm{d}_{f(x)} g = (\mathrm{d}_x f)^{-1} \text{— непрерывна}$$

$$\mathrm{d}_y g = \left(\mathrm{d}_{g(y)} f\right)^{-1}$$

◀

Следствие 2.7.4.1. $f\colon D \subset \mathbb{R}^n \to \mathbb{R}^n$, D открыто, f непрерывно дифференцируемо на D. Тогда если $W \subset D$ — открыто, то открыто и $f(W)$

▶

$$W = \bigcup_{x \in W} B_{r(x)}(x)$$

Докажем для одного шарика $f\left(B_{r(x)}(x)\right)$. Берём для него U и V, подходящее под условие теоремы при $a = x$. $g = f^{-1}\colon V \to U$ — непрерывна, поэтому для g прообраз открытого открыт, откуда для f образ открытого открыт, что и требовалось. ◀

Рассмотрим $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, тогда $(x, y) \in \mathbb{R}^{n+m}$. Рассмотрим также линейный оператор $A\colon \mathbb{R}^{n+m} \to \mathbb{R}^n$, обладающий свойством

$$A(h, 0) = 0 \Rightarrow h = 0$$

Тогда при любых $k \in \mathbb{R}^m$, $b \in \mathbb{R}^n$ уравнение $A(x, k) = b$ имеет ровно одно решение:

$$(x, k) = (x, 0) + (0, k)$$

$$A(x, k) = A(x, 0) + A(0, k)$$

$$A(x, 0) = b - A(0, k)$$

Получили систему n линейных уравнений на n координат вектора x со столбцом свободных членов $b - A(0, k)$. Она совместна и имеет одно решение, так как соотвествующая ей однородная система $A(x, 0) = 0$ имеет одно решение.

Теорема 2.7.5 (О неявной функции)**.** $f\colon D \subset \mathbb{R}^{n+m} \to \mathbb{R}^n$, D открыто, $(\underbrace{a}_{\in\mathbb{R}^n}, \underbrace{b}_{\in\mathbb{R}^m}) \in \mathbb{R}^{n+m}$, f непрерывно дифференцируема, $A = \mathrm{d}_{(a,b)} f$, $f(a, b) = 0$. Пусть A удовлетворяет условию $A(h, 0) =$

$0 \Rightarrow h = 0$, тогда существует окрестность W точки b и непрерывно дифференцируемая функция $g\colon W \to \mathbb{R}^n$, что

$$\forall y \in W, f(g(y), y) = 0$$
$$g(b) = a$$

Пример 2.7.1. Простая окружность:

$$f(x, y) = x^2 + y^2 - 1$$
$$\mathrm{d}_{(a,b)} f = (2a, 2b)$$
$$\mathrm{d}_{(a,b)} f(h, 0) = 2a \cdot h + 2b \cdot 0 = 0 \xRightarrow{a \neq 0} h = 0$$

Как можно заметить, если $a \neq 0$, то можно выбрать дугу окружности вокруг точки (a, b), что условие теоремы на ней выполнено.

▶ Рассмотрим $F\colon D \to \mathbb{R}^{n+m}$: $F(x, y) = (f(x, y), y)$ и проверим условие об обратимой функции. Мы знаем, что f, то есть первые коорлинаты F, дифференцируема:

$$f(a + h, b + k) - f(a, b) = A(h, k) + r(h, k) \quad r(h, k) = o(\|(h, k)\|)$$

а на последних координатах F она совсем дифференцируема:

$$b + k - b = k$$

Тогда

$$F(a + h, b + k) - F(a, b) = B(h + k) + (r(h, k), 0_m)$$
$$B(h, k) = (A(h, k), k)$$

Даже видно, что F непрерывно дифференцируема.

Проверим, что B обратимо. Проверим, что однородное уравнение $B(h, k) = (0, 0)$ имеет только нулевое решение:

$$(A(h, k), k) = B(h, k) = (0, 0) \Rightarrow k = 0$$
$$A(h, 0) = 0 \Rightarrow h = 0$$

Тогда по теореме об обратной функции существют окрестность U точки (a, b), окрестность V точки $F(a, b) = (0, b)$ и $G\colon V \to U$, обратная к F. Поймём, как устроенна G:

$$(f(x, y), y) = F(x, y) = (z, w)$$
$$G(z, w) = (x, y) = (\varphi(z, w), w)$$

φ — некая непрерывно дифференцируемая функция.

$$g(w) := \varphi(0, w)$$

У нас есть V вокруг $(0, b)$, нам нужно выбрать из неё срез только по последним координатам (первые всегда 0). Обозначим её за W — область определения g.

Проверим свойтсва g:

$$F(a, b) = (0, b) \quad g(b) = \varphi(0, b) = a$$
$$f(g(y), y) = f(\varphi(0, y), y) = (f \circ G)(0, y) = \Bigg(\underbrace{(F \circ G)(0, y)}_{(0,y)}\Bigg)_{\text{только первые координаты}} = 0$$

◀

Пример 2.7.2. Есть уравнение $x^5 + ax + b = 0$ и знаем какой-то корень, $(a, b) \in D$. Хотим найти самое большое значение корня. Тогда по нашей теореме можно увидеть, что функция корня от (a, b) дифференцируема, причём производная задана неявно, саму функцию выписвать не нужно (тем более в данном случая общей формулы нет).

2.8. Условные экстремумы

$\mathfrak{Def}$ 2.8.1. $f\colon D \subset \mathbb{R}^{n+m} \to \mathbb{R}$, $\Phi\colon D \to \mathbb{R}^m$, $a \in D$, $\Phi(a) = 0_m$. a называется строгим локальном максимумом при $\Phi = 0$, если

$$\exists U_a\colon \forall z \in \mathring{U}\colon \Phi = 0, f(z) > f(a)$$

Аналогичные определения минимума и нестрогих максимума и минимума.

Теорема 2.8.1 (Метод множителей Лагранжа)**.** $f\colon D \subset \mathbb{R}^{n+m} \to \mathbb{R}$ — непрерывно дифференцируема, $\Phi\colon D \to \mathbb{R}^m$ — непрерывно дифференцируема, $a \in \operatorname{int} D\colon \Phi(a) = 0_m$ — точка условного экстремума. Тогда

$$\nabla f, \nabla\Phi_1, \nabla\Phi_2, \dots, \nabla\Phi_m$$

линейно зависимы.

$\mathfrak{Def}$ 2.8.2. Если $\nabla\Phi_i$ линейно **не**зависимы, то

$$\nabla f = \lambda_1 \nabla\Phi_1 + \lambda_2 \nabla\Phi_2 + \dots + \lambda_m \nabla\Phi_m$$

λ_i называются множителями Лагранжа.

Замечание 2.8.1. Выше написана система на $n+m$ координат x, m множителей Лагранжа, $n+m$ уравнений из линейной комбинации и ещё m из условия $\Phi(x) = 0$. Итого $n+2m$ переменных и $n+2m$ уравнений. К сожалению, это всё совсем не обязательно линейно.

Замечание 2.8.2. $z = (x, y)$. Выпишем матрицу Якоби:

$$\begin{pmatrix} \frac{\partial\Phi_1}{\partial z_1} & \frac{\partial\Phi_1}{\partial z_2} & \cdots & \frac{\partial\Phi_1}{\partial z_{n+m}} \\ \frac{\partial\Phi_2}{\partial z_1} & \frac{\partial\Phi_2}{\partial z_2} & \cdots & \frac{\partial\Phi_2}{\partial z_{n+m}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\Phi_m}{\partial z_1} & \frac{\partial\Phi_m}{\partial z_2} & \cdots & \frac{\partial\Phi_m}{\partial z_{n+m}} \end{pmatrix}$$

Необходимые градиенты $\nabla\Phi_i$ — строки этой матрицы. Заметим, что строк меньше столбцов. Тогда условие независимости $\nabla\Phi_i$ равносильно тому, что ранг матрицы Якоби максимально возможный.

▶ Доказывать имеет смысл в случае независимости $\nabla\Phi_i$.

Пусть не равный 0 определитель — определитель с последними m столбцами. $a = (b, c)$, $b \in \mathbb{R}^n$, $c \in \mathbb{R}^m$, $\operatorname{rang} \mathrm{d}_a\Phi = m$, $A \leftrightharpoons \mathrm{d}_a\Phi$,

$$\det \begin{pmatrix} \frac{\partial\Phi_1}{\partial y_1} & \frac{\partial\Phi_1}{\partial y_2} & \cdots & \frac{\partial\Phi_1}{\partial y_m} \\ \frac{\partial\Phi_2}{\partial y_1} & \frac{\partial\Phi_2}{\partial y_2} & \cdots & \frac{\partial\Phi_2}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\Phi_m}{\partial y_1} & \frac{\partial\Phi_m}{\partial y_2} & \cdots & \frac{\partial\Phi_m}{\partial y_m} \end{pmatrix} \neq 0$$

Проверим условие теоремы об обратной функции:

$$A(0_n, h) = 0_m \Rightarrow h = 0_m$$

Тогда из теоремы о неявной функции знаем, что существует W — окрестность b и $g\colon W \to \mathbb{R}^m$, что $\Phi(x, g(x)) = 0$ и $g(b) = c$. Если a — условный локальный минимум, то существует окрестность W, что

$$\forall x \in W, f(b, g(b)) = f(a) \leqslant f(x, g(x))$$

Таким образом b — локальный минимум функции $H(x) = f(x, g(x))$, $b \in \operatorname{int} W$. Значит

$$\nabla H(b) = 0$$
$$\frac{\partial H}{\partial x_k}(b) = \frac{\partial f}{\partial x_k}(a) + \sum_{i=1}^{m} \frac{\partial f}{\partial y_i}(a)\frac{\partial g_i}{\partial x_k}(b)$$
$$\Phi(x, g(x)) = 0$$
$$0 = \frac{\partial \Phi_j}{\partial x_k}(a) + \sum_{i=1}^{m} \frac{\partial \Phi_j}{\partial y_i}(a)\frac{\partial g_i}{\partial x_k}(b)$$

Анонс: домножим последнее равенство на λ_j и сложим их для разных j так, чтобы всё ненужное — g_i — сократилось. Это система уравнений, которая разрешима, потому что её определитель написан наверху и не ноль. Мы хотим

$$\nabla f(a) = \lambda_1 \nabla \Phi_1(a) + \lambda_2 \nabla \Phi_2(a) + \dots + \lambda_m \nabla \Phi_m(a)$$

То есть хотим

$$F \leftrightharpoons f - \lambda_1 \Phi_1 - \lambda_2 \Phi_2 - \dots - \lambda_m \Phi_m$$
$$\nabla F = 0$$
$$\frac{\partial F}{\partial y_k}(a) = \frac{\partial f}{\partial y_k}(a) - \sum_{j=1}^{m} \lambda_j \frac{\partial \Phi_j}{\partial y_k}(a)$$

Получили матрицу:

$$\underbrace{\begin{pmatrix} \frac{\partial \Phi_1}{\partial y_1} & \frac{\partial \Phi_1}{\partial y_2} & \cdots & \frac{\partial \Phi_1}{\partial y_m} \\ \frac{\partial \Phi_2}{\partial y_1} & \frac{\partial \Phi_2}{\partial y_2} & \cdots & \frac{\partial \Phi_2}{\partial z_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \Phi_m}{\partial y_1} & \frac{\partial \Phi_m}{\partial y_2} & \cdots & \frac{\partial \Phi_m}{\partial z_m} \end{pmatrix}}_{\det \neq 0} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial y_1} \\ \frac{\partial f}{\partial y_2} \\ \vdots \\ \frac{\partial f}{\partial y_m} \end{pmatrix}$$

Эта линейная система имеет решение для λ_i. Возьмём его. Так все частные производные по y_k уже нули, проверим по x_k:

$$\frac{\partial F}{\partial x_k} = \frac{\partial f}{\partial x_k} - \sum_{j=1}^{m} \lambda_j \frac{\partial \Phi_j}{\partial x_k} = -\sum_{i=1}^{m} \frac{\partial f}{\partial y_i}\frac{\partial g_i}{\partial x_k} - \sum_{j=1}^{m} \lambda_j \sum_{i=1}^{m} \frac{\partial \Phi_j}{\partial y_i}\frac{\partial g_i}{\partial x_k} =$$
$$= -\sum_{i=1}^{m} \left(\frac{\partial f}{\partial y_i}\frac{\partial g_i}{\partial x_k} - \sum_{j=1}^{m} \lambda_j \frac{\partial \Phi_j}{\partial y_i}\frac{\partial g_i}{\partial x_k} \right) = -\sum_{i=1}^{m} \underbrace{\left(\frac{\partial f}{\partial y_i} - \sum_{j=1}^{m} \lambda_j \frac{\partial \Phi_j}{\partial y_i} \right)}_{=0} \frac{\partial g_i}{\partial x_k} = 0$$

◀

Пример 2.8.1. $Q(c) = \langle Ax, x \rangle$ — квадратичная форма. Найдём экстремумы квадратичной формы на единичной сфере.

Теорема 2.8.2. Наибольшее и наименьшее значения квадратичной формы на единичной сфере — наибольшее и наименьшее её собственное число, и достигаются в соотвествующих собственных векторах.

▶

$$f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$
$$\Phi(x) = x_1^2 + x_2^2 + \dots + x_n^2 - 1$$
$$\operatorname{rang}\begin{pmatrix} \frac{\partial \Phi}{\partial x_1} & \frac{\partial \Phi}{\partial x_2} & \cdots & \frac{\partial \Phi}{\partial x_n} \end{pmatrix} = 1$$
$$F(x) = f(x) - \lambda \Phi(x)$$
$$a \text{ — точка условного экстремума} \Rightarrow \exists \lambda \in \mathbb{R} \colon \nabla F(a) = 0$$
$$\frac{\partial F}{\partial x_k} = a_{kk} 2x_k + \sum_{i \neq k} a_{ik} x_i + \sum_{j \neq k} a_{kj} x_j - \lambda 2x_k = 2\left(\sum_{i=1}^{n} a_{ik} x_i - \lambda x_k\right)$$
$$\sum_{i=1}^{n} a_{ik} x_i = \lambda x_i$$
$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Итого, λ — собственное число A, x — отнормированный собственный вектор.

$$Q(x) = \langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda \|x\|^2 = \lambda$$

◀

Следствие 2.8.2.1. Если есть оператор A, то можно посчитать его норму:

$$\|A\| = \max\left\{ \sqrt{\lambda} \mid \lambda \text{ — собственное число матрицы } A^T A \right\}$$

▶

Утверждение 2.8.1. *Из алгебры:*

$$\langle Ax, y \rangle = \langle x, A^T y \rangle$$

▶ Для доказательства заметим, что достаточно проверить для базиса, так как всё линейно.

$$\left\langle Ae_i, e_j \right\rangle = \left\langle \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}, e_j \right\rangle = a_{ji}$$
$$\left\langle e_i, A^T e_j \right\rangle = \left\langle e_i, \begin{pmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{pmatrix} \right\rangle = a_{ji}$$

◀

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$
$$\|A\|^2 = \max_{\|x\|=1} \|Ax\|^2$$
$$\|Ax\|^2 = \langle Ax, Ax \rangle = \left\langle x, A^T A x \right\rangle = x^T (A^T A) x$$

Получили квадратичную форму $A^T A$

$$\|A\|^2 = \max\{\lambda \mid \lambda \text{ — собственное число матрицы } A^T A\}$$
$$\|A\| = \max\left\{\sqrt{\lambda} \mid \lambda \text{ — собственное число матрицы } A^T A\right\}$$

◀

Следствие 2.8.2.2. Внезапно, все собственные числа матрицы $A^T A$ неотрицательны.

Глава 3

Кратные интегралы

3.1. Мера Жордана

$\mathfrak{Def}$ **3.1.1.** Живём в R^n. Ячейка (полуоткрытый параллелепипед):

$$(a_1, b_1] \times (a_2, b_2] \times ... \times (a_n, b_n]$$

Мера ячейки — её объём:

$$\mu P = (b_1 - a_1)(b_2 - a_2) \dots (b_n - a_n)$$
$$\mu \emptyset = 0$$

Свойства меры:

1. $\mu P \geqslant 0$

2. μ аддитивна для дизъюнктного объединения:

$$P = \bigcup_{k=1}^{m} P_k \quad P_i \cap P_j = \emptyset$$
$$P = \bigsqcup_{k=1}^{m}$$
$$\mu P = \sum_{i=1}^{m} \mu P_k$$

▶ Продлим все разделяющие координаты, разрежем по всем координатам. При каждом разрезе равенство сумм мер сохраняется. ◀

$\mathfrak{Def}$ **3.1.2.** Конечное объединение непересекающихся ячеек называется клеточным множеством.

Лемма 3.1.1. Объединение, пересечение и разность клеточных множеств — клеточное множество.

▶

$$A = \bigsqcup_{k=1}^{m} P_k \quad B = \bigsqcup_{j=1}^{l} Q_j$$

Далее везде всё дизъюнктное:

Пересечение:

$$A \cap B = \left(\bigcup_{k=1}^{m} P_k\right) \cap \left(\bigcup_{j=1}^{l} Q_j\right) = \bigcup_{k=1}^{m}\bigcup_{j=1}^{l} P_k \cap Q_j$$

$$P = (a_1, b_1] \times ... \times (a_n, b_n] \quad Q = (a'_1, b'_1] \times ... \times (a'_n, b'_n]$$

$$P \cap Q = (\max a_1, a'_1, \min b_1, b'_1] \times ... \times (\max a_n, a'_n, \min b_n, b'_n]$$

Разность:

$$A \setminus B = \left(\bigcup_{k=1}^{m} P_k\right) \setminus \left(\bigcup_{j=1}^{l} Q_j\right) = \bigcup_{k=1}^{m}\left(P_k \cap \bigcup_{j=1}^{l} Q_j\right) = \bigcup_{k=1}^{m}\left(\bigcap_{j=1}^{l} P_k \setminus Q_j\right)$$

Разность двух ячеек — клеточное множество (продлим границы, переразрежем, объединим).

Объединение:

$$A \cup B = (A \setminus B) \cup B$$

◀

$\mathfrak{Def}$ **3.1.3.** Мера клеточного множества — сумма мер его ячеек.

Утверждение 3.1.1.

1. На ячейках ничего нового.
2. Определение корректно.

▶

$$A = \bigsqcup_{k=1}^{m} P_k = \bigsqcup_{j=1}^{l} Q_j$$

Надо показать, что

$$\sum_{k=1}^{m} \mu P_k = \sum_{j=1}^{l} \mu Q_j$$

Покажем:

$$\mu P_k = \sum_{j=1}^{l} \mu(P_k \cap Q_j) \text{ — дизъюнктное}$$

$$\mu Q_j = \sum_{k=1}^{m} \mu(P_k \cap Q_j) \text{ — дизъюнктное}$$

$$\sum_{k=1}^{m} \mu P_k = \sum_{k=1}^{m}\sum_{j=1}^{l} \mu(P_k \cap Q_j) = \sum_{j=1}^{l}\sum_{k=1}^{m} \mu(P_k \cap Q_j) = \sum_{j=1}^{l} \mu Q_j$$

◀

Теорема 3.1.1. Пусть A и B — клеточные множества. Тогда

1. (монотонность меры) $A \subset B \Rightarrow 0 \leqslant \mu A \leqslant \mu B$
2. (полуаддитивность меры) $\mu(A \cup B) \leqslant \mu A + \mu B$
3. (аддитивность меры) $A \cap B = \emptyset \Rightarrow \mu(A + B) = \mu A + \mu B$
4. $A \subset B \Rightarrow \mu(B \setminus A) = \mu B - \mu A$

▶

$$A = \bigsqcup_{k=1}^{m} P_k \quad B = \bigsqcup_{j=1}^{l} Q_j$$

3.

$$A \cap B = \emptyset \Rightarrow P_k \cap Q_j = \emptyset$$
$$A \cup B = \bigsqcup_{k=1}^{m} P_k \sqcup \bigsqcup_{j=1}^{l} Q_j \text{ — дизъюнктное}$$
$$\mu A = \sum_{k=1}^{m} \mu P_k \quad \mu B = \sum_{j=1}^{l} \mu Q_l$$
$$\mu(A \cup B) = \sum_{k=1}^{m} \mu P_k + \sum_{j=1}^{l} \mu Q_l$$

4.

$$B = (B \setminus A) \sqcup A$$
$$\mu B = \mu(B \setminus A) + \mu A$$

1.

$$\mu B - \mu A = \mu(B \setminus A) \geqslant 0$$

2.

$$A \cup B = (B \setminus (A \cap B)) \sqcup A$$
$$\mu(A \cup B) = \mu(B \setminus (A \cap B)) + \mu A \leqslant \mu B + \mu A$$

2'.

$$\mu(A \cup B) + \mu(A \cap B) = \mu A + \mu B$$

◀

𝔇𝔢𝔣 3.1.4. Пусть E — ограниченное множество в $\mathbb{R}^n$. Тогда внешняя мера Жордана есть $\mu^* E = \inf\{\mu A \mid A \supset E\}$, где A — клеточное.

𝔇𝔢𝔣 3.1.5. Внутренняя мера Жордана: $\mu_* E = \sup\{\mu A \mid A \subset E\}$, где A — клеточное.

Теорема 3.1.2. Если E — ограниченное множество, то:

1. $0 \leqslant \mu_* E \leqslant \mu^* E < +\infty$
2. Монотонность: если $E \subset F$ (F ограничено), то $\mu_* E \leqslant \mu_* F$ и $\mu^* E \leqslant \mu^* F$
3. Полуаддитивность *внешней* меры: $\mu^*(E \cup F) \leqslant \mu^* E + \mu^* F$

▶ 1. • Так как $\emptyset \subset E$, то $\mu\emptyset \leqslant \sup \mu A = \mu_* E$
• Если $A \subset E \subset B$, то $\mu A \leqslant \mu B$. Зафиксируем B, напишем супремум по A:

$$\mu_* E = \sup_{A \subset E} \mu A \leqslant \mu B$$
$$\forall B\colon \mu_* E \leqslant \mu B \forall B \supset E$$
$$\mu_* E \leqslant \inf_{B \supset E} \mu B = \mu^* E$$

- Если E ограничено, то $\exists P \supset E$, причём P — ячейка. Тогда $\mu^* E \leqslant \mu P$.

2. - Если $A \subset E$, то $A \subset F$. Поэтому

$$\mu_* E = \sup_{A \subset E} \mu A \leqslant \sup_{A \subset F} \mu A$$

так как справа супремум идёт по большему множеству.

- Если $B \supset F$, то $B \supset E$. Поэтому

$$\mu^* F = \inf_{B \supset F} \mu B \geqslant \inf_{B \supset E} \mu B = \mu^* E$$

Монотонность: если $E \subset F$ (F ограничено), то $\mu_* E \leqslant \mu_* F$ и $\mu^* E \leqslant \mu^* F$

3. Пусть есть клеточные множества A и B: $E \subset A$, $F \subset B$. Тогда $E \cup F \subset A \cup B$ и $A \cup B$ — клеточное.

$$\mu^*(E \cup F) = \inf_{C \supset E \cup F} \mu C \leqslant \inf_{A \supset E B \supset F} \mu(A \cup B) \leqslant \inf_{A \supset E B \supset F} \mu A + \mu B = \inf_{A \supset E} \mu A + \inf_{B \supset F} \mu B = \mu^* A + \mu^* B$$

◀

Def 3.1.6. Пусть E — ограниченное множество. Если $\mu_* E = \mu^* E$, то E является *измеримым по Жордану* и $\mu E \overset{\text{Def}}{=} \mu_* E = \mu^* E$ — его мера Жордана.

Замечание 3.1.1. Так мы доопределили меру на каком-то дополнительном классе множеств, правда, не очень понятно, на каком. Зато есть шанс, что у нас реально возникающие множества станут измеримыми (хотя бы замкнутые/открытые ячейки, а не только полуоткрытые)

Замечание 3.1.2. Дальше вместо «измеримо по Жордану» будем говорить «измеримо» — так короче, а других измеримостей у нас пока не предвидится.

Упражнение 3.1.1. Доказать, что $\mu E > 0 \iff \operatorname{int} E \neq \emptyset$

Пример 3.1.1. Пусть есть два параллелепипеда, один открытый, другой замкнутый и они зажимают множество P:

$$(a_1, b_1) \times \ldots \times (a_n, b_n) \subset P \subset [a_1, b_1] \times \ldots \times [a_n, b_n]$$

Тогда P измеримо и $\mu P = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$.

▶ Возьмём ячейки, вложенные в P, отрезав немного с правых концов:

$$P_m = \left(a_1, b_1 - \frac{1}{m}\right] \times \ldots \times \left(a_n, b_n - \frac{1}{m}\right] \subset (a_1, b_1) \times \ldots \times (a_n, b_n) \subset P$$

Возьмём ячейки, покрывающие P, отрезав немного с левых концов:

$$Q_m = \left(a_1 - \frac{1}{m}, b_1\right] \times \ldots \times \left(a_n - \frac{1}{m}, b_n\right] \supset [a_1, b_1] \times \ldots \times [a_n, b_n] \supset P$$

Отсюда видно:

$$\mu_* P \geqslant \sup \mu P_m = \sup \left(b_1 - a_1 - \frac{1}{m}\right) \cdot \ldots \cdot \left(b_n - a_n - \frac{1}{m}\right) = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$$

Аналогично для внешней меры:

$$\mu^* Q \leqslant \inf \mu Q_m = \inf \left(b_1 - a_1 + \frac{1}{m}\right) \cdot \ldots \cdot \left(b_n - a_n + \frac{1}{m}\right) = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$$

Собираем вместе:

$$(b_1 - a_1) \cdot \ldots \cdot (b_n - a_n) \geqslant \mu^* P \geqslant \mu_* P \geqslant (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$$

Таким образом верхняя и нижняя меры существуют, равны и мы даже знаем, чему именно. ◀

Пример 3.1.2. Если E — множество из конечного числа точек, то оно измеримо и $\mu E = 0$.

▶ Накроем каждую из n точек ячейкой вида $(x_1 - \frac{1}{m}, x_1] \times ... \times (x_n - \frac{1}{m}, x_n] = P_m$. Видно, что внешняя мера E ограничивается суммарной мерой таких ячеек, устремляем m к бесконечности, получаем $\mu^* E \leqslant 0$. То есть $\mu^* E = 0$. ◀

Упражнение 3.1.2. Найти внешнюю и внутреннюю меру $E = \{\frac{1}{m} \mid m \in \mathbb{N}\} \subset \mathbb{R}$.

Пример 3.1.3. Пусть $E = [0,1] \cap \mathbb{Q}$. Тогда $\mu_* E = 0$, и $\mu^* E = 1$, то есть оно не измеримо по Жордану.

▶ $\mu_* E = 0$, т.к. E не содержит внутренних точек и, как следствие, не содержит ячеек.

Покажем, что $\mu^* E = 1$. Для этого возьмём какое-то покрытие E ячейками:

$$E \subset \bigcup_{k=1}^{n} P_k$$

Пририсуем слева и справа замыкания:

$$[0,1] = \operatorname{cl} E \subset \cup_{k=1}^{n} \operatorname{cl} P_k$$

У нас есть конечное число замкнутых отрезков, которые покрывают отрезок $[0,1]$, мы интересуемся суммой длин. Ясно, что она хотя бы единица (можно показать по индукции). Значит $\mu^* e \geqslant 1$. Вывод: E неизмеримо по Жордану. ◀

Пример 3.1.4. Если E измеримо, $\mu^* E = 0$ и $F \subset E$, то F измеримо и $\mu F = 0$.

▶

$$\mu E = 0 \Rightarrow \mu^* E = 0 \Rightarrow 0 \leqslant \mu^* F \leqslant \mu^* E = 0 \Rightarrow \mu^* F = 0$$

То есть $0 \leqslant \mu_* F \leqslant \mu^* F = 0$, что и требовалось доказать. ◀

Пример 3.1.5. Возьмём в $\mathbb{R}^2$ горизонтальный отрезок $[a,b] \times \{c\}$. Его мера равна нулю, так как он является замкнутым параллелепипедом.

Замечание 3.1.3. Это пример бесконечного множества, которое имеет меру ноль. Можно придумать такой пример даже в $\mathbb{R}$.

Пример 3.1.6. Пример связного открытого неизмеримого множества.

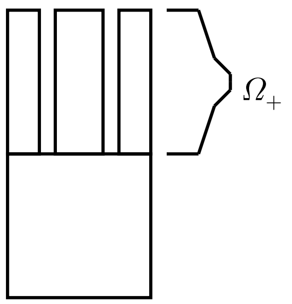
Пусть $\{r_k\}$ — это как-то занумерованные рациональные числа из $(0,1)$. Также возьмём достаточно маленькое $\varepsilon > 0$ (как минимум $\varepsilon < \frac{1}{2}$). Построим множество:

$$\mathcal{D} = \bigcup_{k=1}^{\infty} \left(r_k - \frac{\varepsilon}{2^k}, r_k + \frac{\varepsilon}{2^k}\right) \cap (0,1)$$

Оно является объединением открытых, то есть само оно тоже открыто. Заметим, что весь отрезок $[0,1]$ мы не накроем, так как сумма длин этих интервалов равна $\sum_{k=1}^{\infty} \frac{2\varepsilon}{2^k} = 2\varepsilon$ (на самом деле, это еще надо формально проверить, но для интуитивного понимания достаточно).

Давайте приклеим над множеством $\mathcal{D}$ цилиндры, получив $\mathcal{D} \times [0,1)$ (пока не слишком открытое). А потом снизу еще кое-что приклеим, получим:

$$\Omega = (\mathcal{D} \times [0,1]) \cup ((0,1) \times (-1,0))$$



Это открытое и связное множество. Покажем, что оно неизмеримо (на самом деле, $\mathcal{D}$ тоже неизмеримо, но связность — это круто).

▶ Обозначим за $\varOmega_+$ верхнюю половину множества (включая границу на нуле). Хотим посчитать $\mu^*\varOmega_+$, покроем ячейками и напишем замыкание:

$$\varOmega_+ \subset \bigcup_{k=1}^{m} P_k$$

$$[0,1]^2 = \operatorname{cl}\varOmega_+ \subset \bigcup_{k=1}^{m} \operatorname{cl} P_k$$

$$1 = \mu[0,1]^2 \leqslant \sum_{k=1}^{m} \mu(\operatorname{cl} P_k) = \sum_{k=1}^{m} \mu P_k$$

То есть $\mu^*\varOmega_+ \geqslant 1$. С другой стороны, всё $\varOmega_+$ покрывается при помощи $[0,1]^2$, то есть $\mu^*\varOmega_+ = 1$.

Теперь будем смотреть на внутреннюю меру. Хотим показать, что она не больше ε. Взяли какие-то ячейки:

$$\bigcup_{k=1}^{m} P_k \subset \varOmega_+$$

Отсюда ясно, что объединение проекций P_k на ось X содержится в $\mathcal{D}$. Возьмём клеточное множество вида «объединение проекций со столбиком $[0,1]$ наверх». Очевидно, оно накрывает объединение P_k и все его клетки не пересекаются, а его площадь не больше 2ε. Значит, площадь объединения P_k тоже не больше 2ε (так как они тоже не пересекаются). То есть $\mu_*\varOmega_+ \leqslant \varepsilon < \mu^*\varOmega+$.

Таким образом $\varOmega_+$ неизмеримо. Значит, $\varOmega$ тоже неизмеримо, так как нижний квадратик просто добавляет к μ_* и μ^* единичку. Были неравными — остались неравными. ◀

3.2. Свойства измеримых множеств

Теорема 3.2.1 (критерий измеримости)**.** Пусть E ограничено. Тогда E измеримо тогда и только тогда, когда $\forall\varepsilon > 0, \exists A_\varepsilon \subset E \subset B_\varepsilon$, причём $\mu B_\varepsilon - \mu A_\varepsilon < \varepsilon$.

▶ **⇒:** Так как E измеримо, то $\mu E = \mu^* E = \inf_{B\supset E} \mu B$. Значит, можем найти такое клеточное B, что $\mu B < \mu^* E + \frac{\varepsilon}{2}$. Аналогично для $\mu_* E$: можно найти такое A, что $\mu A > \mu_* E - \frac{\varepsilon}{2}$. Соберём вместе, получим, что $\mu B - \mu A < \varepsilon$.

⇐:

$$\mu^* E = \inf_{B\supset E} \mu B \leqslant \inf \mu B_\varepsilon$$

$$\mu_* E = \sup_{A\subset E} \mu A \geqslant \sup \mu A_\varepsilon$$

$$\mu A_\varepsilon \leqslant \mu_* E \leqslant m^* E \leqslant \mu B_\varepsilon$$

$$0 \leqslant \mu B_\varepsilon - \mu A_\varepsilon < \varepsilon$$

$$0 \leqslant \mu^* E - \mu_* E \leqslant \varepsilon$$

$$\mu^* E = \mu_* E$$

◀

𝔇𝔢𝔣 3.2.1. A — множество, x — точка.

$$\rho(x, A) = \inf\{\rho(x,y) \mid y \in A\}$$

$\mathfrak{Def}$ **3.2.2.** δ-окрестность множества $U_\delta(A)$ (иногда обозначается A_δ):

$$U_\delta(A) = \{x \mid \rho(x, A) < \delta\} = \{x \mid \exists y \in A\colon \rho(x, y) < \delta\}$$

Замечание 3.2.1. δ-окрестность является объединением открытых шаров радиуса δ с центрами из A:

$$U_\delta(A) = \bigcup_{x\in A} B_\delta(x)$$

Теорема 3.2.2. E — измерима, тогда $\mu^*(U_\delta(E)) \to \mu E$ при $\delta \to 0$.

▶ E — **ячейка:**

$$E = (a_1, b_1] \times \ldots \times (a_n, b_n]$$
$$E \subset U_\delta(A) \subset (a_1 - \delta, b_1 + \delta] \times (a_n - \delta, a_n + \delta]$$
$$\prod_{k=1}^{n}(b_k - a_k) = \mu E \leqslant \mu^* U_\delta(E) \leqslant \mu((a_1 - \delta, b_1 + \delta] \times (a_n - \delta, a_n + \delta]) = \prod_{k=1}^{n}(b_k - a_k + 2\delta)$$

E — **клеточное множество:**

$$E = \bigcup_{k=1}^{m} P_k$$
$$U_\delta(E) = \bigcup_{k=1}^{m} U_\delta(P_k)$$
$$\mu E = \mu^* U_\delta(E) \leqslant \sum_{k=1}^{m} \mu^*(U_\delta(P_k)) \to \sum_{k=1}^{m} \mu P_k$$

E — **измеримое:**

$$\forall \varepsilon > 0, \exists B_\varepsilon \supset E\colon \mu B_\varepsilon < \mu E + \varepsilon$$
$$U_\delta(B_\varepsilon) \supset U_\delta(E)$$
$$\mu^*(U_\delta(B_\varepsilon)) \to \mu B_\varepsilon$$
$$\exists \delta_0 > 0\colon \forall \delta < \delta_0, \mu E \leqslant \mu^* U_\delta(E) \leqslant \mu^* U_\delta(B_\varepsilon) < \mu B_\varepsilon + \varepsilon < \mu E + 2\varepsilon$$

◀

Следствие 3.2.2.1. E — множество нулевой меры. Тогда существует клеточное множество C_ε, что $\mu C_\varepsilon < \varepsilon$ и $E \subset \operatorname{int} C_\varepsilon$.

▶

$$\mu^*(U_\delta(E)) \to \mu E = 0$$
$$\exists \delta > 0\colon \mu^*(U_\delta(E)) < \frac{\varepsilon}{2} \Rightarrow \exists C_\varepsilon \supset U_\delta(E)\colon \mu C_\varepsilon < \mu^* U_\delta(E) + \frac{\varepsilon}{2} < \varepsilon$$
$$\operatorname{int} C_\varepsilon \supset U_\delta(E) \text{ — открытое}$$

◀

$\mathfrak{Def}$ **3.2.3.** ∂E — граница множества:

$$\partial E = \{x \mid \forall r > 0, B_r(x) \cap E \neq \emptyset \wedge B_r(x) \setminus E \neq \emptyset\}$$

Замечание 3.2.2. Можно показать, что $\partial E = \operatorname{cl} E - \operatorname{int} E$.

Лемма 3.2.1. $x \in E$, $y \notin E$. Тогда на $[x, y]$ есть точка из границы E.

▶ $x_0 = x$, $y_0 = y$. Возьмём $m = \frac{x+y}{2}$. Если она в E, то перейдём к $x_1 = m$, $y_1 = y_0$, иначе к $x_1 = x_0$, $y_1 = m$. Получим сходящуюся последовательность вложенных отрезков. В их пересечении лежит одна точка c. Тогда для каждой окрестности этой точки есть отрезок $[x_i, y_i]$, начало которого лежит в E, а конец — не лежит. ◀

Лемма 3.2.2. Пусть D — клетчатое множество, а ограниченное множество E такое, что $\partial E \subset D$. Тогда $E \cup D$ тоже клетчатое.

▶ Q — ячейка, $Q \supset E \cup D$.

$$Q \setminus D = \bigcup_{i=1}^{m} P_k$$

$P_k \cap E \neq \emptyset$**:** Проверим, что $E \supset P_k$. Пусть $P_k \setminus E \neq \emptyset$. $x \in P_k \cap E$, $y \in P_k \setminus E$, значит на отрезке $[x, y]$ есть граничная точка E. Этот отрезок, а значит и точка, лежат в ячейке. Тогда $P_k \cap \partial E \neq \emptyset$, что неверно, так как $P_k \cap D = \emptyset$.

$P_k \cap E = \emptyset$**:** Ничего интересного.

Получили, что

$$E \cup D = D \cup \bigcup_{P_k \subset E} P_k$$

«$\supset$» очевидно, «$\subset$»: взяли точку $x \in E \setminus D$, тогда $x \in Q \setminus D$, значит $\exists k \colon x \in P_k$. ◀

Теорема 3.2.3 (Критерий измеримости)**.** E — ограниченное множество. Тогда E измеримо тогда и только тогда, когда граница имеет нулевую меру.

▶ $\Rightarrow$**:** E — измеримо, значит

$$\forall \varepsilon > 0, \exists A_\varepsilon, B_\varepsilon \text{ — клеточные} \colon A_\varepsilon \subset E \subset B_\varepsilon \land \mu B_\varepsilon - \mu A_\varepsilon < \varepsilon$$
$$\partial E = \operatorname{cl} E \setminus \operatorname{int} E \subset \operatorname{cl} B_\varepsilon \setminus \operatorname{int} A_\varepsilon$$
$$\tilde{A}_\varepsilon \text{ — клеточное, вложенное в } \operatorname{int} A_\varepsilon \quad \mu \tilde{A}_\varepsilon > \mu A_\varepsilon - \varepsilon$$
$$\tilde{B}_\varepsilon \text{ — клеточное, содержащее } \operatorname{cl} B_\varepsilon \quad \mu \tilde{B}_\varepsilon < \mu B_\varepsilon + \varepsilon$$
$$\partial E \subset \tilde{B}_\varepsilon \setminus \tilde{A}_\varepsilon$$
$$\mu(\tilde{B}_\varepsilon \setminus \tilde{A}_\varepsilon) \leqslant \mu \tilde{B}_\varepsilon - \mu \tilde{A}_\varepsilon < \mu B_\varepsilon + \varepsilon - \mu A_\varepsilon + \varepsilon < 3\varepsilon$$
$$\Rightarrow \mu^* \partial E = 0 \Rightarrow \mu(\partial E) = 0$$

$\Leftarrow$**:**

$$\mu(\partial E) = 0 \Rightarrow \exists C_\varepsilon \text{— клеточное} \colon \mu C_\varepsilon < \varepsilon \land C_\varepsilon \supset \partial E$$
$$\Rightarrow E \cup C_\varepsilon \text{ — клеточное}$$
$$B_\varepsilon = E \cup C_\varepsilon \quad A_\varepsilon = B_\varepsilon \setminus C_\varepsilon = E \setminus C_\varepsilon$$
$$B_\varepsilon \supset E \supset A_\varepsilon$$
$$\mu B_\varepsilon = \mu((B_\varepsilon \setminus C_\varepsilon) \sqcup C_\varepsilon) = \mu(A_\varepsilon \cup C_\varepsilon) \leqslant \mu A_\varepsilon + \mu C_\varepsilon < \mu A_\varepsilon + \varepsilon$$

◀

Лемма 3.2.3. $E, F \subset \mathbb{R}^n$. Тогда

$$\partial(E \cup F) \subset \partial E \cup \partial F$$
$$\partial(E \cap F) \subset \partial E \cup \partial F$$
$$\partial(E \setminus F) \subset \partial E \cup \partial F$$

▶ Возьмём $x \in \partial(E \cup F)$, но $x \notin \partial F$. Покажем, что она в ∂E. Хотим увидеть это:

$$B_r(x) \cap (E \cup F) \neq \emptyset \qquad (B_r(x) \cap E) \cup (B_r(x) \cap F) \neq \emptyset$$
$$B_r(x) \setminus (E \cup F) \neq \emptyset \qquad (B_r(x) \setminus E) \cap (B_r(x) \setminus F) \neq \emptyset$$

Покажем:

$$B_r(x) \setminus F \neq \emptyset \quad B_r(x) \setminus E \neq \emptyset$$
$$x \notin \partial F \Rightarrow B_r(x) \cap F = \emptyset \quad \text{для некоторого } r \Rightarrow B_r(X) \cap E \neq \emptyset$$

Остальное аналогично можно поупражняться. ◀

Следствие 3.2.3.1. E, F измеримы. Тогда измеримы $\operatorname{int} E$, $\operatorname{cl} F$, $E \cup F$, $E \cap F$, $E \setminus F$.

▶ Хотим понять, что

$$\partial(\operatorname{int} E) \subset \partial E$$
$$\partial(\operatorname{cl} E) \subset \partial E$$

Тогда победим из того, что $\mu(\partial E) = 0$.

$$\partial(\operatorname{int} E) = \operatorname{cl} \operatorname{int} E \setminus \operatorname{int} \operatorname{int} E = \operatorname{cl} \operatorname{int} E \setminus \operatorname{int} E \subset \operatorname{cl} E \setminus \operatorname{int} E = \partial E$$
$$\partial(\operatorname{cl} E) = \operatorname{cl} \operatorname{cl} E \setminus \operatorname{int} \operatorname{cl} E = \operatorname{cl} E \setminus \operatorname{int} \operatorname{cl} E \subset \operatorname{cl} E \setminus \operatorname{int} E = \partial E$$

Если E и F имзеримы, то мера объединения их границ нулевая, а тогда по лемме всё остальное тоже имеет нулевую меру границы. ◀

Теорема 3.2.4. E, F измеримы. Тогда

1. (монотонность меры) $E \subset F \Rightarrow 0 \leqslant \mu E \leqslant \mu F$
2. (полуаддитивность меры) $\mu(E \cup F) \leqslant \mu E + \mu F$
3. (аддитивность меры) $E \cap F = \emptyset \Rightarrow \mu(E \sqcup F) = \mu E + \mu F$

▶ 1. Знаем для μ^* и $\mu = \mu^*$.
2. Знаем для μ^* и $\mu = \mu^*$.
3. $A \subset E$, $B \subset F$ — клеточные, $A \cap F \neq \emptyset$. Тогда $A \sqcup B \subset E \sqcup F$.

$$\mu E + \mu F \geqslant \mu(E \cup F) \geqslant \mu(A \cup B) = \mu A + \mu B$$
$$\mu A + \mu B \leqslant \mu(E \cup F) \leqslant \mu E + \mu F$$

Перешли к супремумам по A и B.

$$\mu E + \mu F = \mu_* E + \mu_* F \leqslant \mu(E \cup F) \leqslant \mu E + \mu F$$

◀

Замечание 3.2.3. Аддитивности верхней и нижней меры нет.

$$E = [0,1] \cap \mathbb{Q} \quad [0,1] \setminus Q$$
$$\mu^* E = \mu^* F = 1 \quad \mu_* E = \mu_* F = 0$$
$$\mu_*(E \cup F) = \mu^*(E \cup F) = 1$$

Теорема 3.2.5. K — компакт, $f\colon K \to \mathbb{R}$ — непрерывная функция. Её график имеет меру ноль.

▶ f непрерывна на компакте, значит по теореме Кантора-Гейне f равномерна непрерывна:

$$\forall \varepsilon > 0, \exists \delta > 0\colon \forall x, y \in K, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

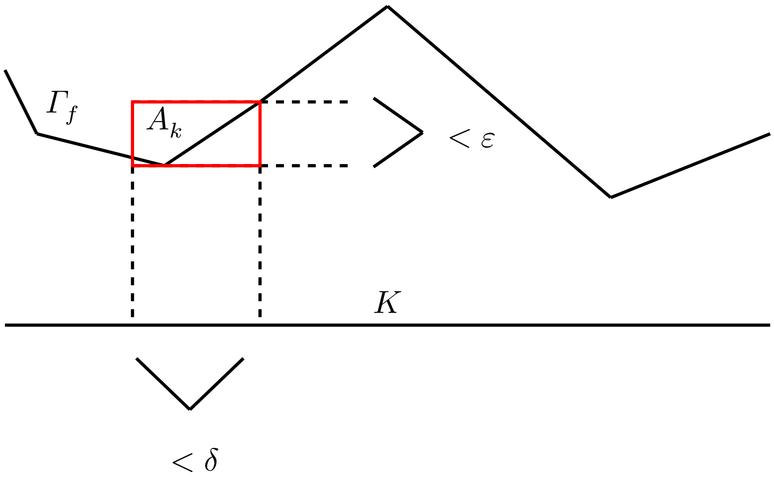
K компакт, значит ограничен и содержится в ячейке. Будем делить по всем осям пополам до тех пор, пока диаметр ячеек не станет меньше δ, получили какой-то набор ячеек P_k. Рассмотрим часть Γ_f, лежащую над ячейкой P_k:

$$\{(x, f(x)) \mid x \in K \cap P_k\}$$

Возьмём произвольные $x, y \in K \cap P_k$. Тогда $|x - y| \leqslant \operatorname{diam} P_k < \delta$, откуда $|f(x) - f(y)| < \varepsilon$. Зафиксируем $x_k \in K \cap P_k$.

$$A_k \leftrightharpoons \{(x, t)| \mid x \in P_k \wedge t \in (f(x_k) - \varepsilon, f(x_k) + \varepsilon]\}$$

Это ячейка, содержащая кусок графика.

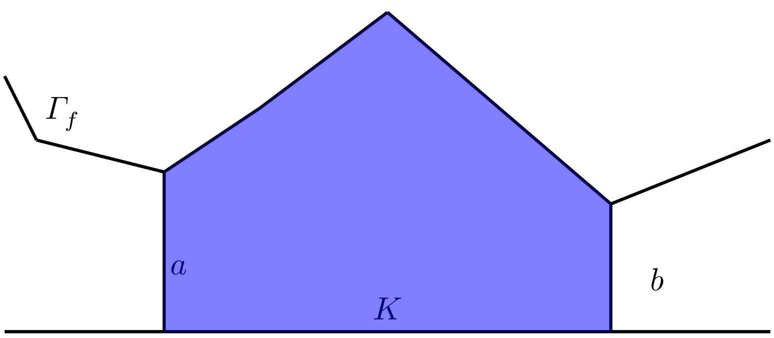


Тогда мы можем вписать наш график в $\cup_{k\colon P_k \cap K \neq \emptyset} A_k$. Давайте посчитаем меру этих множеств (здесь за μ_n обозначим меру подмножества K, а за μ_{n+1} — меру подмножества пространства, где нарисован график):

$$\mu_{n+1}(\cup A_k) \leqslant \sum \mu_{n+1} A_k \underbrace{\overset{\mu_{n+1} A_k = 2\varepsilon \mu_n P_k}{\leqslant}}_{\text{слева сумма не по всем } k\text{, а справа по всем}} 2\varepsilon \sum \mu_n P_k = 2\varepsilon \mu P$$

Так как P фиксировано, то μP тоже фиксировано, а ε мы выбираем любой. Значит, мера графика меньше любого положительного числа, то есть ноль. ◀

Следствие 3.2.5.1. Криволинейная трапеция измерима. Напоминание: криволинейная трапеция — это мы для непрерывной функции $f\colon [a, b] \to \mathbb{R}$ берём все точки вида $a \leqslant x \leqslant b$, $0 \leqslant y \leqslant f(x)$.



Более того, измеримы все множества вида

$$G_f = \{(x,t) \mid x \in K; 0 \leqslant t \leqslant f(x)\}$$

где $f\colon K \to \mathbb{R}$ (K — компакт, измеримо, f непрерывна и неотрицательна).

▶ Докажем, что если K измеримо, то такое множество тоже измеримо. Надо доказать, что граница G_f имеет меру ноль. Она состоит из следующих кусочков: график самой функции, компакт, вверхторчащие куски над границей K (назовём их цилиндрами над ∂K). Хотим показать, что все эти куски имеют меру ноль.

- $\partial G_f = 0$ по теореме.
- Поймём про K. Так как K компакт, то оно содержится в некоторой ячейке $P \subset \mathbb{R}^n$. А что происходит в $\mathbb{R}^{n+1}$? В ней $K \subset P \times (-\varepsilon, +\varepsilon]$ — ячейка в $\mathbb{R}^{n+1}$. Устремляем ε к нулю, получаем, что мера ячейки стремится к нулю, то есть мера K в $\mathbb{R}^{n+1}$ равна нулю.
- Так как f непрерывна на компакте, то она ограничена: $0 \leqslant f(x) \leqslant M$. Значит, оставшиеся цилиндры содержатся в $\partial K \times [0, M]$. Возьмём клеточное множество $C_\varepsilon \supset \partial K$ такое, что $\mu_n C_\varepsilon < \varepsilon$. Тогда $\partial K \times [0, M] \subset C_\varepsilon \times (-1, M]$ — клеточное множество.

$$\mu_{n+1}(C_\varepsilon \times (-1, M]) = (M+1)\mu_n C_\varepsilon < \varepsilon(M+1)$$

То есть мера цилиндров тоже ноль.

◀

Следствие 3.2.5.2. Если E имеет меру ноль, то ограниченный цилиндр над E тоже имеет меру ноль.

▶ В точности как в последнем пункте предыдущего доказательства. ◀

3.3. Кратный интеграл

𝔇𝔢𝔣 3.3.1. Пусть есть измеримое множество E. Тогда его разбиение τ есть набор множеств $\{E_k\}_{k=1}^{m(\tau)}$ ($m(\tau)$ — просто количество множеств в разбиении), где выполняются свойства:

1. E_k измеримо
2. E_k дизъюнктны
3. $E = \cup_{k=1}^{m(\tau)} E_k$.

𝔇𝔢𝔣 3.3.2. Мелкость (ранг) разбиения $|\tau|$ есть наибольший диаметр множеств E_k.

Замечание 3.3.1.

$$\mu E = \sum_{k=1}^{m(\tau)} \mu E_k$$

𝔇𝔢𝔣 3.3.3. Пусть τ и τ' — разбиения E. Мы скажем, что τ' *подчинено* τ (обозначается $\tau' \succ \tau$), если

$$\forall E'_k \in \tau' \colon \exists E_j \in \tau, E'_k \subset E_j$$

Смысл: мы сначала разбили E в τ, а потом еще доразбили до τ'.

Несколько свойств:

1. Подчинимость транзитивна: если $\tau_1 \succ \tau_2$ и $\tau_2 \succ \tau_3$, то $\tau_1 \succ \tau_3$.

2. $\forall \tau', \tau'' \colon \exists \tau, \tau', \tau'' \succ \tau$.

▶ Пусть $\tau' = \{E'_k\}$ и $\tau'' = \{E''_k\}$. Тогда положим $\tau = \{E'_k \cap E''_j\}$. ◀

𝔇𝔢𝔣 3.3.4. Оснащение разбиения (как в определении интеграла по Риману) — набор точек $\xi_k \in E_k$ (взяли в каждом множестве по точке).

𝔇𝔢𝔣 3.3.5. Сумма Римана: есть $f \colon E \to \mathbb{R}$ (E — измеримо), есть τ — разбиение E, ξ — его оснащение.

$$S(f, \tau, \xi) = \sum_{k=1}^{m(\tau)} f(\xi_k) \mu E_k$$

𝔇𝔢𝔣 3.3.6. Интеграл Римана: есть $f \colon E \to \mathbb{R}$ (E — измеримо). Обозначения:

$$I = \int_E f = \int_E f(x) \mathrm{d}x = \int \cdots \int_E f(x_1, \dots, x_n) \mathrm{d}x_1 \dots \mathrm{d}x_n$$

I называется интегралом Римана функции, если оно удовлетворяет следующему свойству:

$$\forall \varepsilon > 0 \colon \exists \delta, \forall \tau \text{— разбиение } E \text{ ранга} \leqslant \delta \colon \forall \xi \text{— оснащение разбиения} \colon \colon |I - S(f, \tau, \xi)| < \varepsilon$$

Замечание 3.3.2. $\int \cdots \int$ — это пока для нас просто обозначение. Но мы будем стремиться к тому, чтобы свести кратный интеграл к нескольким последовательным. Возможно, успеем даже в этом семестре.

Замечание 3.3.3. Если мы будет требовать от разбиения не дизъюнктность, а лишь что пересечение любых имеет меру ноль, то всё равно всё будет хорошо.

▶ В самом деле: $E = \cup E_k$ и $\operatorname{int} E_k \subset E$. А вот $\operatorname{int} E_k$ уже не пересекаются, иначе пересечение с каждой точкой содержит некоторую окрестность, то есть имеет ненулевую меру.

$$\sum_{k=1}^{m(\tau)} \mu E_k \geqslant \mu E \geqslant \sum_{k=1}^{m(\tau)} \mu \operatorname{int} E_k = \sum_{k=1}^{m(\tau)} \mu E_k$$

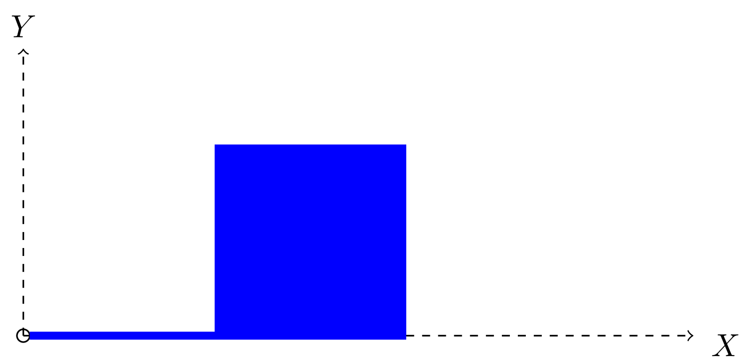
То есть замечание номер 3.3.1 всё еще в силе. В рассуждениях дальше тоже ничего не поломается: мы всего лишь добавляем какие-то граничные точки. ◀

Замечание 3.3.4. Мы когда-то уже говорили про интеграл Римана на прямой. Тогда упоминалось, что интегрируемы будут не только непрерывные функции. Это всё еще правда :). Правда, пока еще не было конкретики на тему «какие именно непрерывные будут интегрируемы». Ничего страшного, скоро будет.

Теорема 3.3.1. 1. $n = 1$. Если $f \colon [a, b] \to \mathbb{R}$ интегрируема, то f ограничена.

2. $n \geqslant 2$. Пусть для E существует последовательность разбиений τ_k такая, что $|\tau_k| \to 0$ и все множества всех разбиений имеют положительную меру. Тогда если f интегрируема на E, то f ограничена.

Замечание 3.3.5. При $n \geqslant 2$ простой ограниченности не будет. Возьмём такое множество E на плоскости:



Определим f как

$$f(x,y)=\begin{cases}\frac{1}{x}, & y=0\\ 0, & y\neq 0\end{cases}$$

Функция неограничена, да и множество меры не ноль. А интеграл ноль, так как кусочки разбиения, содержащие «хвост», имеют меру ноль. А мера кусочков разбиения, задевающих нижнюю сторону квадрата, ограничена мелкостью разбиения. То есть тоже стремится к нулю (а значения функции на нижней стороне квадрата ограничены). Значит, интеграл ноль. Очень. Жаль.

▶ Будем считать, что $n\geqslant 1$ и у нас есть последовательность разбиений τ_k из пункта два. Функция интегрируема, значит, $|I-S(f,\tau_k,\xi_k)|<1$ при достаточно больших k. Посмотрим на сумму Римана:

$$S=\sum_{j=1}^{m(\tau_k)} f(\xi_j^{(k)})\mu E_j^{(k)}$$

Предположим, что f неограничена. Значит, она неограничена на каком-то элементе разбиения $E_{j_0}^{(k)}$ (иначе она ограничена на каждом из конечного числа элементов, то есть ограничена вообще). Перепишем сумму:

$$S=f(\xi_{j_0}^{(k)})\underbrace{\mu E_{j_0}^{(k)}}_{>0}+\sum_{j\neq j_0} f(\xi_j^{(k)})\mu E_j^{(k)}$$

Второе слагаемое-сумму трогать не будем, зафиксируем. А в первом у нас фиксирована положительая мера. А вот точку мы можем менять произвольно (ведь оснащение можно менять как угодно, на мелкость это не влияет, а у нас должно быть верно для любого оснащения разбиения). Так как f неограничена, то мы можем сделать $f(\xi_{j_0}^{(k)}$ либо сколь угодно большим, либо сколь угодно маленьким. То есть S получаем сколь угодно большим или сколь угодно маленьким, что противоречит условию на сумму Римана.

Теперь поймём, что для невырожденного отрезка на прямой ($n=1$) условие из второго пункта заведомо выполнено. Это очевидно: взяли и нарезали отрезок на кусочки. А в вырожденном отрезке значение всего одно, значит, f ограничена. ◀

Упражнение: $E\subset \operatorname{cl}\operatorname{int}E$ и $f\colon E\to\mathbb{R}$ — интегрируема, то f ограничена.

Def 3.3.7. E — множество, $f\colon E\to\mathbb{R}$. $\omega(f,E)$ — колебание f на E.

$$\omega(f,E)=\sup_{x\in E}f(x)-\inf_{x\in E}f(x)=\sup_{x,y\in E}(f(x)-f(y))$$

Теорема 3.3.2 (критерий интегрируемости)**.** E — измерима, $f\colon E\to\mathbb{R}$. f интегрируема на E тогда и только тогда, когда

$$\forall\varepsilon>0,\exists\delta>0\colon\forall\tau\colon|\tau|<\delta,\sum_{k=1}^{m(\tau)}\omega(f,E_k)\mu E_k<\varepsilon$$

▶ **⇒:** f интегрируема, значит

$$\forall\varepsilon>0,\exists\delta>0\colon\forall\tau\colon|\tau|<\delta,\left|S(f,\tau,\xi)-\int\limits_E\right|<\varepsilon$$

Зафиксируем τ. Хотим найти два таких оснащения ξ_k и ξ_k', что $\omega(f,E_k)\leqslant 2(f(\xi_k)-f(\xi_k'))$.

f **постоянна на** E_k**:** очевидно, берём произвольные ξ_k и ξ_k'.

f **непостоянна на** E_k**:** $\gamma \leftrightharpoons \omega(f, E_k) = \sup_{E_k} f - \inf_{E_k} f$. Приблизимся к супремуму с точностью $\gamma/4$ в точке ξ_k, к инфимуму в ξ'_k с такой же точностью:

$$f(\xi_k) \geqslant \sup_{E_k} f - \gamma/4$$

$$f(\xi'_k) \leqslant \inf_{E_k} f + \gamma/4$$

$$f(\xi_k) - f(\xi'_k) \geqslant \sup f - \inf f - \gamma/2 = \gamma/2$$

$$\frac{1}{2}\omega(f, E_k) \leqslant f(\xi_k) - f(\xi'_k)$$

Тогда

$$\left|S(f, \tau, \xi) - \int_E f\right| < \varepsilon$$

$$\left|S(f, \tau, \xi') - \int_E f\right| < \varepsilon$$

$$|S(f, \tau, \xi) - S(f, \tau, \xi')| < 2\varepsilon$$

$$\sum_{k=1}^{m(\tau)} \left(f(\xi_k) - f(\xi'_k)\right) \mu E_k < 2\varepsilon$$

$$\frac{1}{2} \sum_{k=1}^{m(\tau)} \omega(f, E_k) \mu E_k \leqslant \sum_{k=1}^{m(\tau)} \left(f(\xi_k) - f(\xi'_k)\right) \mu E_k$$

$$\sum_{k=1}^{m(\tau)} \omega(f, E_k) \mu E_k \leqslant 2 \sum_{k=1}^{m(\tau)} \left(f(\xi_k) - f(\xi'_k)\right) \mu E_k < 4\varepsilon$$

$\Leftarrow$**:** Пусть

$$\forall \varepsilon > 0, \exists \delta > 0 \colon \forall \tau \colon |\tau| < \delta, \sum_{k=1}^{m(\tau)} \omega(f, E_k) \mu E_k \leqslant \varepsilon$$

Рассмотрим какие-нибудь $\tau' \succ \tau$ (τ' подчинено τ). Тогда каждое $E_k \in \tau$ представляется как дизюнктное объединение некоторых кусочков из разбиения τ'

$$E_k = \bigsqcup_{i=1}^{i_k} E'_{k_i}$$

Далее оценим

$$\left| f(\xi_k) \mu E_k - \sum_{i=1}^{i_k} f\left(\xi'_{k_i}\right) \mu E'_{k_i} \right| = \cdots \text{ (продолжение ниже)}$$

$\xi_k \in E_k$, $\xi'_{k_i} \in E'_{k_i} \subset E_k$.

$$|f(\xi_k) - f\left(\xi'_{k_i}\right)| \leqslant \omega(f, E_k)$$

$$\cdots = \left| \sum_{i=1}^{i_k} f(\xi_k) \mu E'_{k_i} - \sum_{i=1}^{i_k} f(\xi'_{k_i}) \mu E'_{k_i} \right| \leqslant$$

$$\leqslant \sum_{i=1}^{i_k} \left| f(\xi_k) - f\left(\xi'_{k_i}\right) \right| \mu E'_{k_i} \leqslant \sum_{i=1}^{i_k} \omega(f, E_k) \mu E'_{k_i} = \omega(f, E_k) \mu E_k$$

Суммируем по k
$$|S(f,\tau,\xi)-S(f,\tau,\xi')|\leqslant\sum_{k=1}^{m(\tau)}\omega(f,E_k)\mu E_k$$
Если при этом $|\tau|<\delta$, то $|S(f,\tau,\xi)-S(f,\tau,\xi')|<\varepsilon$ независимо от выбора τ'.
Пусть τ' и τ'' — разбиения мелкости меньше δ. Рассмотрим $\tau\succ\tau',\tau''$. Тогда
$$|S(f,\tau,\xi)-S(f,\tau',\xi')|<\varepsilon$$
$$|S(f,\tau,\xi)-S(f,\tau'',\xi'')|<\varepsilon$$
$$|S(f,\tau',\xi')-S(f,\tau'',\xi'')|<2\varepsilon$$
Рассмотрим последовательность разбиений τ_n такую, что $|\tau_n|\to 0$.
$$\forall\varepsilon>0,\exists N\colon \forall n,m>N,|\tau_n|,|\tau_m|<\delta\Rightarrow|S(f,\tau_n,\xi_n)-S(f,\tau_m,\xi_m)|<2\varepsilon$$
Это уже критерий Коши для последовательности $S(f,\tau_n,\xi_n)$. Значит, S_n имеет предел I.
Тогда взяв ряд неравенств (начиная с некоторого N)
$$|S(f,\tau',\xi')-S(f,\tau_n,\xi_n)|<2\varepsilon$$
в пределе получим
$$|S(f,\tau',\xi')-I|<2\varepsilon$$
◀

𝔇𝔢𝔣 3.3.8. Верхняя и нижняя сумма Дарбу $S_\tau(f)$ и $s_\tau(f)$:
$$S_\tau(f)=\sum_{k=1}^{m(\tau)}\sup_{x\in E_k}f(x)\cdot\mu E_k$$
$$s_\tau(f)=\sum_{k=1}^{m(\tau)}\inf_{x\in E_k}f(x)\cdot\mu E_k$$
Замечание 3.3.6.
$$S_\tau(f)\geqslant s_\tau(f)$$
Замечание 3.3.7.
$$S_\tau(f)-s_\tau(f)=\sum_{k=1}^{m(\tau)}\omega(f,E_k)\mu E_k$$
Теорема 3.3.3. $f\colon E\to\mathbb{R}$ — ограниченна. Тогда f интегрируема на E тогда и только тогда, когда
$$\forall\varepsilon>0,\exists\delta>0\colon\forall\tau\colon|\tau|<\delta,S_\tau-s_\tau<\varepsilon$$
▶ Очевидно из замечания 3.3.7 и теоремы 3.3.2. ◀

Следствие 3.3.3.1. $f\colon E\to\mathbb{R}$ — ограниченна и интегрируема. Тогда
$$s_\tau\leqslant\int\limits_E f\leqslant S_\tau$$

▶

$$s_\tau \leqslant S(f,\tau,\xi) \leqslant S_\tau$$

$$\sum_{k=1}^{m(\tau)} \inf_{x\in E_k} f(x)\mu E_k \leqslant \sum_{k=1}^{m(\tau)} f(\xi_k)\mu E_k \leqslant \sum_{k=1}^{m(\tau)} \sup_{x\in E_k} f(x)\mu E_k$$

$\tau' \succ \tau$

$$s_\tau \leqslant S(f,\tau',\xi') \leqslant S_\tau$$

$$\inf_{x\in E_k} f(x)\mu E_k \leqslant \sum_{i=1}^{i_k} f(\xi'_{k_i})\mu E'_{k_i} \leqslant \sup_{x\in E_k} f(x)\mu E_k$$

$\tau_n \succ \tau$, $|\tau_n| \to 0$, $S(f,\tau_n,\xi_n) \to \int_E f$

$$s_\tau \leqslant S(f,\tau',\xi') \leqslant S_\tau$$

$$s_\tau \leqslant \int\limits_E \leqslant S_\tau$$

◀

Теорема 3.3.4. $f\colon K \to \mathbb{R}$ непрерывна, K — измеримый компакт. Тогда f интегрируема на K.

▶ f непрерына на компакте, значит она ограничена и равномерно непрерывна.

$$\forall\varepsilon > 0, \exists\delta > 0\colon \forall x, y \in K, \|x - y\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

Если $\operatorname{diam} E < \delta$, то $\omega(f,E) < \varepsilon$. τ — произвольное разбиение мелкости меньше δ. Там все $\operatorname{diam} E_k < \delta$.

$$\sum_{k=1}^{m(\tau)} \omega(f,E_k)\mu E_k \leqslant \sum_{k=1}^{m(\tau)} \varepsilon\mu E_k = \varepsilon\mu K$$

◀

Упражнение. E — открытое измеримое множество. $f\colon E \to \mathbb{R}$ непрерывно и ограничено. Тогда f интегрируема.

3.4. Свойства кратного интеграла

1. E измеримо. Тогда

$$\int\limits_E 1 = \mu E$$

Замечание 3.4.1. 1 — «адын». ©Храбров.

2. $E \supset E'$ измеримо, f интегрируема на E. Тогда f интегрируема на E'.

▶ τ' — разбиение E'. Достроим, сохраняя мелкость, до разбиения E.

$$\forall\varepsilon > 0, \exists\delta > 0\colon \forall\tau\colon |\tau| < \delta, \sum_{k=1}^{m(\tau)} \omega(f,E_k)\mu E_k < \varepsilon$$

$$\sum_{k=1}^{m(\tau)} \omega(f,E'_k)\mu E'_k < \sum_{k=1}^{m(\tau)} \omega(f,E_k)\mu E_k < \varepsilon$$

◀

3. Аддитивность интеграла. $E = E' \sqcup E''$, E' и E'' измеримы, f ограничена. Тогда f интегрируема на E тогда и только тогда, когда f интегрируема на E' и E'', и в этом случае

$$\int\limits_E f = \int\limits_{E'} f + \int\limits_{E''} f$$

▶ Достаточно доказать влево. τ — разбиение E.

$$\tau' = \{E_k \cap E' \mid E_k \in \tau\}$$
$$\tau'' = \{E_k \cap E'' \mid E_k \in \tau\}$$
$$\tau_0 = \{E_k \mid E_k \cap E' \neq \emptyset \neq E_k \cap E''\}$$
$$\sum_{E_k \notin \tau_0} \omega(f, E_k)\mu E_k \leqslant \sum_{F \in \tau'} \omega(f, F)\mu F + \sum_{F \in \tau''} \omega(f, F)\mu F$$
$$\sum_{E_k \in \tau_0} \omega(f, E_k)\mu E_k \leqslant \sum_{E_k \in \tau_0} 2M\mu E_k = 2M\mu\left(\bigcup_{E_k \in \tau_0} E_k\right) \leqslant \cdots$$

$E_k \cap E' \neq \emptyset \neq E_k \cap E''$. Мы можем брать точки a_k и b_k из E_k, что на $[a_k, b_k]$ есть точка из $\delta E'$. $\operatorname{diam} E_k < \delta$, значит длина $[a_k, b_k]$ меньше δ. Тогда точка из границы $\delta E'$ лежит на рассоянии не более δ от концов.

Так как мы можем в качестве a перебрать любую точку, то $E_k \subset U_\delta(\partial E')$.

$$\ldots \leqslant 2M\mu(U_\delta(\partial E')) \to 2M\mu(\delta E') = 0$$

Осталось доказать, что

$$\int\limits_E f = \int\limits_{E'} f + \int\limits_{E''} f$$

Возьмём последовательность разбиений τ_l, $|\tau_l| \to 0$.

$$\tau_l = \left\{E_k^{(l)}\right\}$$
$$\tau_l' = \left\{E_k^{(l)} \cap E' \colon E_k^{(l)} \cap E' \neq \emptyset\right\}$$
$$\tau_l'' = \left\{E_k^{(l)} \cap E'' \colon E_k^{(l)} \cap E'' \neq \emptyset\right\}$$

Мы уже знаем, что

$$S(f, \tau_l, \xi_l) \to \int\limits_E f$$

и

$$S(f, \tau_l', \xi_l) + S(f, \tau_l'', \xi_l) = S(f, \tau_l, \xi_l)$$

ξ_l' и ξ_l'' — подправленные точки. Мы при разбиении некоторые множества разбили на 2. Точка могла оказаться не там, где надо. Подвинем её как-нибудь в нужное множество.

$$S(f, \tau_l', \xi_l') \to \int_{E'} f$$
$$S(f, \tau_l'', \xi_l'') \to \int_{E''} f$$

Осталось сравнить $S(f, \tau_l', \xi_l)$ и $S(f, \tau_l', \xi_l')$. Так как мы поправляли точки только у разбиваемых множеств, то

$$|S(f, \tau_l', \xi_l) - S(f, \tau_l', \xi_l') = \left|\sum_{E_k \in \tau_0} (f(\xi_{lk}) - f(\xi_{lk}'))\mu(E_k \cap E')\right| \leqslant$$
$$\leqslant 2M \sum_{E_k \in \tau_0} \mu(E_k \cap E) \leqslant 2M\mu(U_{2\tau l}(\partial E')) \to 0$$

◀

Упражнение: ограниченность f по делу.

4. E — множество нулевой меры, $f\colon E \to \mathbb{R}$. Тогда f интегрируема на E, и интергал равен нулю.

▶ E_k — разбиение E, $E_k \subset E \Rightarrow \mu E_k = 0$.

$$S(f, \tau, \xi) = \sum_{k=1}^{m(\tau)} f(\xi_k)\mu E_k = 0$$

◀

5. $f\colon E \to \mathbb{R}$ ограничена, $E \supset e$, $\mu e = 0$. Если, сохраняя ограниченость f, поменять значения f на e, интеграл не изменится.

▶ По п. 3:

$$\int\limits_E f = \int\limits_{E \setminus e} f + \int\limits_e f = \int\limits_{E \setminus e} f + 0$$

◀

6. $f\colon \operatorname{cl} E \to \mathbb{R}$ ограничена, E измеримо. Тогда интегрируемости f на E, $\operatorname{cl} E$ и $\operatorname{int} E$ равносильны и три интеграла равны.

▶

$$\operatorname{cl} E \setminus \operatorname{int} E = \partial E \quad \mu(\partial E) = 0$$

Изменим $f(x) = 0$ на ∂E, получили функцию $\tilde{f}$. Интеграл не поменялся.

$$\int\limits_E f = \int\limits_E \tilde{f} \quad \int\limits_{\operatorname{cl} E} f = \int\limits_{\operatorname{cl} E} \tilde{f} \quad \int\limits_{\operatorname{int} E} f = \int\limits_{\operatorname{int} E} \tilde{f}$$
$$\int\limits_E \tilde{f} = \int\limits_{\operatorname{cl} E} \tilde{f} = \int\limits_{\operatorname{int} E} \tilde{f}$$

◀

7. Интеграл линеен. $f, g\colon E \to \mathbb{R}$ интегрируемы, $\alpha, \beta \in \mathbb{R}$. Тогда $\alpha f + \beta g$ интегрируема и

$$\int\limits_E (\alpha f + \beta g) = \alpha \int\limits_E f + \beta \int\limits_E g$$

▶ f интегрируема, значит

$$\forall\varepsilon>0,\exists\delta>0\colon\forall\tau\colon|\tau|<\delta,\sum_{k=1}^{m(\tau)}\omega(f,E_k)\mu E_k<\varepsilon$$

Аналогично для g. Мы даже можем взять общий δ.

$$\omega(\alpha f+\beta g,E_k)=\sup_{E_k}|(\alpha f(x)+\beta g(x))-(\alpha f(y)+\beta g(y))|\leqslant$$
$$\leqslant\sup_{E_k}(|\alpha||f(x)-f(y)|+|\beta||g(x)-g(y)|)\leqslant$$
$$\leqslant|\alpha|\sup_{E_k}|f(x)-f(y)|+|\beta|\sup_{E_k}|g(x)-g(y)|=|\alpha|\omega(f,E_k)+|\beta|\omega(g,E_k)$$
$$\sum_{k=1}^{m(\tau)}\omega(\alpha f+\beta g,E_k)\mu E_k\leqslant|\alpha|\sum_{k=1}^{m(\tau)}\omega(f,E_k)\mu E_k+|\beta|\sum_{k=1}^{m(\tau)}\omega(g,E_k)\mu E_k\leqslant(|\alpha|+|\beta|)\varepsilon$$

Получили, что $\alpha f+\beta g$ интегрируема. Проверим равенство: возьмём последовательность τ_l, $|\tau_l|\to0$.

$$S(\alpha f+\beta g,\tau,\xi)=\alpha S(f,\tau,\xi)+\beta S(f,\tau,\xi)$$

В пределе

$$\int\limits_E(\alpha f+\beta g)=\alpha\int\limits_E f+\beta\int\limits_E g$$

◀

8. $f,g\colon E\to R$ интегрируемы и ограничены. Тогда fg интегрируема.

▶ $|f|,|g|\leqslant M$.

$$\omega(fg,E_k)=\sup_{E_k}|f(x)g(x)-f(y)g(y)|=\sup_{E_k}|f(x)g(x)-f(x)g(y)+f(x)g(y)-f(y)g(y)|\leqslant$$
$$\leqslant\sup_{E_k}(|f(x)||g(x)-g(y)|+|g(y)||f(x)-f(y)|)\leqslant|M|(\omega(f,E_k)+\omega(g,E_k))$$
$$\sum_{k=1}^{m(\tau)}\omega(fg,E_k)\mu E_k\leqslant2M\varepsilon$$

◀

9. $f,g\colon E\to\mathbb{R}$ интегрируемы, f ограничена, $\inf_E|g|>0$. Тогда $\frac{f}{g}$ интегрируема.

▶ Достаточно проверить ограниченость и интегрируемость $\frac{1}{g}$.

$$\sup\frac{1}{|g|}=\frac{1}{\inf|g|}<+\infty$$

Теперь интегрируемость:

$$\omega\left(\frac{1}{g},E_k\right)=\sup_{E_k}\left|\frac{1}{g(x)}-\frac{1}{g(y)}\right|=\sup_{E_k}\frac{|g(x)-g(y)|}{|g(x)g(y)|}\leqslant\frac{\omega(g,E_k)}{\alpha^2}$$
$$\sum_{k=1}^{m(\tau)}\omega\left(\frac{1}{g},E_k\right)\mu E_k\leqslant\frac{\varepsilon}{\alpha^2}$$

◀

10. $f\colon E \to \mathbb{R}$, интегрируема. Тогда $|f|$ интегрируема и

$$\left|\int_E f\right| \leqslant \int_E |f|$$

▶

$$||f(x)| - |f(y)|| \leqslant |f(x) - f(y)|$$
$$\omega(|f|, E_k) \leqslant \omega(f, E_k)$$
$$\sum_{k=1}^{m(\tau)} \omega(|f|, E_k)\mu E_k \leqslant \sum_{k=1}^{m(\tau)} \omega(f, E_k)\mu E_k$$

Берём τ_l:

$$S(f, \tau_l, \xi_l) = \sum_{k=1}^{m(\tau)} f\left(\xi_k^{(l)}\right)\mu E_k^{(l)} \leqslant \sum_{k=1}^{m(\tau)} |f\left(\xi_k^{(l)}\right)|\mu E_k^{(l)} = S(|f|, \tau_l, \xi_l)$$
$$S(f, \tau_l, \xi_l) = \sum_{k=1}^{m(\tau)} f\left(\xi_k^{(l)}\right)\mu E_k^{(l)} \geqslant \sum_{k=1}^{m(\tau)} -|f\left(\xi_k^{(l)}\right)|\mu E_k^{(l)} = -S(|f|, \tau_l, \xi_l)$$
$$-S(|f|, \tau_l, \xi_l) \leqslant S(f, \tau_l, \xi_l) \leqslant S(|f|, \tau_l, \xi_l)$$
$$-\int_E |f| \leqslant \int_E f \leqslant \int_E |f|$$

◀

11. $f, g\colon E \to \mathbb{R}$ интегрируемы, $f \geqslant g$. Тогда

$$\int_E f \geqslant \int_E g$$

▶

$$S(f, \tau_l, \xi_l) = \sum_{k=1}^{m(\tau)} f\left(\xi_k^{(l)}\right)\mu E_k^{(l)} \geqslant \sum_{k=1}^{m(\tau)} g\left(\xi_k^{(l)}\right)\mu E_k^{(l)} = S(g, \tau_l, \xi_l)$$
$$\int_E f \geqslant \int_E g$$

◀

12. $f\colon G \to \mathbb{R}$, G открыто, f интегрируема, $f \geqslant 0$, $a \in G$, f непрерывна в a, $f(a) > 0$. Тогда

$$\int_E f > 0$$

▶ По теореме о стабилизации знака существует окрестность V точки a, что

$$\forall x \in V, f(x) \geqslant \frac{f(a)}{2}$$

$P \subset V$ — ячейка с центром в a. Рассмотрим

$$g(x) = \begin{cases} f(a)/2 & x \in P \\ 0 & x \notin P \end{cases}$$
$$f \geqslant g$$
$$\int_G g = \frac{f(a)}{2}\mu P > 0$$

◀

13. Счётная аддитивность. $f\colon E \to \mathbb{R}$ интегрируема и ограничена, $E_k \subset E$ измеримы, $\mu E_k \to mu E$. Тогда
$$\int\limits_{E_k} f \to \int\limits_E f$$

▶

$$\int\limits_E f = \int\limits_{E_k} f + \int\limits_{E\setminus E_k} f$$
$$\left|\int\limits_E f - \int\limits_{E_k} f\right| = \left|\int\limits_{E\setminus E_k} f\right| \leqslant \int\limits_{E\setminus E_k} M = M\mu(E \setminus E_k) = M(\mu E - \mu E_k) \to 0$$

◀

14. *Объяснение* счётной аддитивности.
$$E = \bigsqcup_{k=1}^{\infty} E_k$$
E_k и E измеримы,
$$\sum_{k=1}^{n} \mu E_k \to \mu E$$
$f\colon E \to \mathbb{R}$ интегрируема. Тогда
$$\lim_{n\to\infty} \sum_{k=1}^{n} \int_{E_k} f = \int_E f$$

▶

$$\tilde{E}_n = \bigsqcup_{k=1}^{n} E_k \quad \mu\tilde{E}_n \to \mu E$$
$$\int\limits_{\tilde{E}_n} f = \sum_{k=1}^{n} \int\limits_{E_k} f$$
$$\int\limits_{\tilde{E}_n} f \to \int\limits_E f$$

◀

Теорема 3.4.1 (интегральная теорема о среднем)**.** Пусть $f, g\colon E \to \mathbb{R}$, причём f, g интегрируемы, $g \geqslant 0$ и f ограничена: $m \leqslant f \leqslant M$. Тогда существует $\lambda \in [m, M]$ такое, что $\lambda \int_E g = \int_E fg$

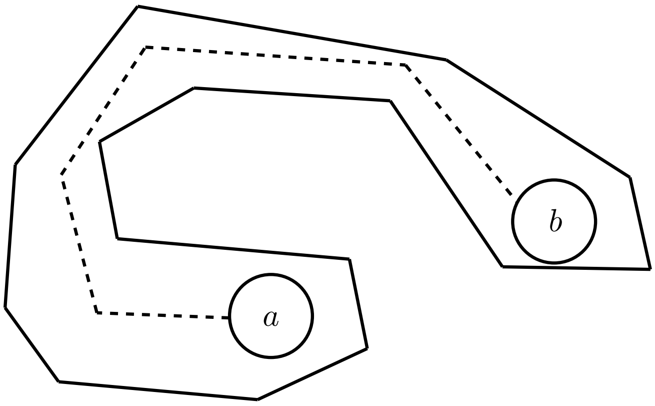
▶ Знаем, что $mg \leqslant fg \leqslant Mg$. Проинтегрируем, получим

$$m\int_E g = \int_E mg \leqslant \int_E fg \leqslant \int_E Mg = M\int_E g$$

- Если $\int_E g = 0$, то $\int_E fg = 0$ и нужное равенство тривиально.
- Если $\int_E > 0$, то $m \leqslant \frac{\int_E fg}{\int_E g} \leqslant M$, возьмём дробь за λ.

◀

Def 3.4.1. Множество E — линейно связное, если $\forall a, b \in E$ существует ломаная, их соединяющая, и целиком лежащая в E.



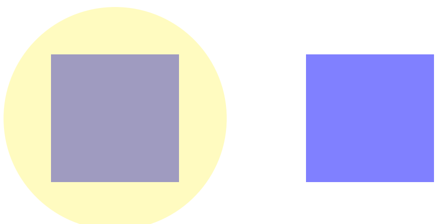
Следствие 3.4.1.1. Пусть K — линейно связный компакт (в одномерном случае — это только отрезки). Пусть также есть непрерывная $f\colon K \to \mathbb{R}$ и есть интегрируемая $g\colon K \to R$, $g \geqslant 0$. Тогда $\exists c \in K$ такая что $\int_K fg = f(c) \cdot \int_K g$.

▶ f непрерывна на компакте, значит, принимает наибольшее и наименьшее значение. Пусть $f(a)$ — минимум, $f(b)$ — максимум. Тогда $\forall x\colon f(a) \leqslant f(x) \leqslant f(b)$. Подставим $m = f(a)$ и $M = f(b)$ в теорему. Получим, что $\exists \lambda \in [f(a), f(b)]\colon \int_K fg = \lambda \int_K g$. Теперь надо показать, что λ — это значение функции в какой-то точке.

В самом деле, возьмём ломаную, соединяющую точки a и b. Давайте её как-нибудь запараметризуем. Пусть имеется m звеньев, разобьём отрезок $[0, 1]$ на m непустых кусков, каждый кусок отобразим в звено. Например, отрезок $[\frac{k-1}{m}, \frac{k}{m}]$ отобразили в звено номер k. Получили некоторое $\gamma\colon [0, 1] \to K$, отображающее на ломаную. Возьмём функцию $g(t) = f(\gamma(t))$. Она непрерывна, так как является композицией непрерывных функцией. То есть имеем непрерывную функцию $g\colon [0, 1] \to \mathbb{R}$, причём $g(0) = f(a)$, $g(1) = f(b)$, значит, по теореме Больцано-Коши, она принимает все промежуточные значения, в частности, λ. Значит $\lambda = f(c)$ для некоторого c на нашей ломаной. ◀

Def 3.4.2. Назовём E связным (просто связным), если E нельзя представить в виде дизъюнктного объединения двух множеств, каждое из которых одновременно и открыто, и замкнуто одновременно (открытость/замкнутость в E, не во всём $\mathbb{R}^n$).

Пример 3.4.1. Пусть E — это два замкнутых квадратика (обозначены синим).



Тогда каждый из квадратиков в E открыт (потому что можно взять его плюс небольшую открытую границу; обозначено жёлтым) и замкнут (потому что является дополнением другого квадратика).

Упражнение 3.4.1. Пусть K — связный компакт, есть непрерывная $f\colon K \to \mathbb{R}$, есть интегрируемая $g\colon K \to \mathbb{R}$, $g \geqslant 0$. Тогда $\exists c \in K$ такая, что $\int_K fg = f(c)\int_K g$. Указание: заменить последний кусок доказательства, посмотрев в прошлый семестр на доказательство теоремы Больцано-Коши. Утверждается, что оно проходит ровно для связных множеств.

Следствие 3.4.1.2. Пусть $f : K \to \mathbb{R}$ непрерывна, K — линейно связный компакт. Пусть K измеримо мерой μ. Тогда $\exists c \in K$ такое, что $\int_K f = f(c)\mu K$

▶ Возьмём $g \equiv 1$ и подставим в следствие 3.4.1.1. ◀

Пример 3.4.2. Связность существенна. Возьмём $K = [0,1] \cup [2,3]$ (компакт, измеримо) и такие f и g:

$$g(x) = 1$$
$$f(x) = \begin{cases} -1, & x \in [0,1] \\ 1, & x \in [2,3] \end{cases}$$

Заметим, что $\int_K fg = 0$, а $\int_K g = 2$. Но $f(c)\int_K g = 2f(c) \neq 0$ при любом c.

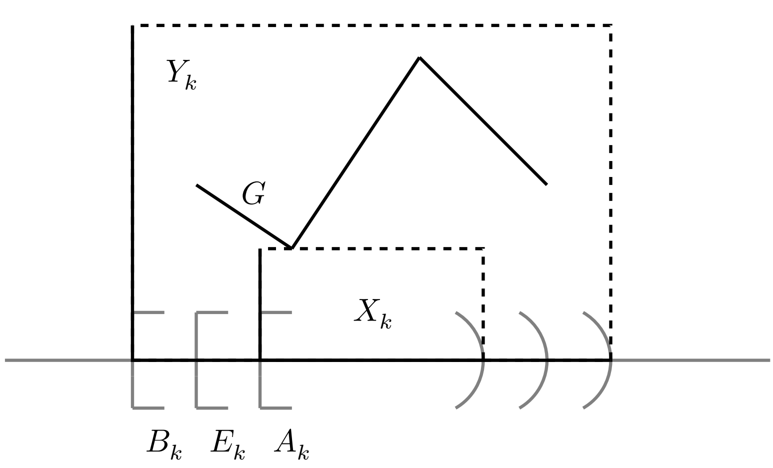
Теорема 3.4.2 (геометрический смысл многомерного интеграла)**.** Пусть есть $f\colon E \to \mathbb{R}$, интегрируемая и ограниченная, а также $f \geqslant 0$ на $E \subset \mathbb{R}^n$. Возьмём $G \subset \mathbb{R}^{n+1}$ следующего вида: $\{(x_1, \dots, x_{n+1}) \mid x = (x_1, \dots, x_n) \in E; 0 \leqslant x_{n+1} \leqslant f(x)\}$. Это, по сути, «подграфик» функции. Тогда G измеримо и $\mu_{n+1}G = \int_E f$ (то есть интеграл — объём под графиком).

▶ Пусть $I = \int_E f$, зафиксируем некоторое $\varepsilon > 0$. Возьмём $\delta > 0$ такое, что $|I - s_\tau(f)| < \varepsilon$ и $|I - S_\tau(f)| < \varepsilon$ для любого разбиения τ мелкости не больше δ. Здесь за s_τ и S_τ обозначены суммы Дарбу — нижняя и верхняя. Так как f интегрируема, то такое δ должно быть.

Возьмём некоторое $E_k \in \tau$. Приблизим его клеточными множествами A_k и B_k такими, что $A_k \subset E_k \subset B_k$, а также $\mu B_k - \mu A_k < \frac{\varepsilon}{m(\tau)}$ (здесь $m(\tau)$ — число множеств E_k в разбиении). Мы так можем выбрать, потому что E_k измеримо.

Теперь заводим множество $X_k = A_k \times [0, \inf_{E_k} f)$. И множество $Y_k = B_k \times [0, \sup_{E_k} f + \varepsilon)$. Это клеточные множества. Теперь хочется сказать, что

$$\underbrace{\bigcup_{k=1}^{m(\tau)} X_k}_{X} \subset G \subset \underbrace{\bigcup_{k=1}^{m(\tau)} Y_k}_{Y}$$



При этом по краям неравенства стоят клеточные множества. Мы зажали G между какими-то клеточными множествами $\mathcal{X}$ и $\mathcal{Y}$. Сейчас поймём, что их меры несильно отличаются друг от другая. Давайте сначала оценим меру $\mathcal{Y}$:

$$\mu_{n+1}\left(\bigcup_{k=1}^{m(\tau)} Y_k\right) \leqslant \sum_{k=1}^{m(\tau)} \mu_{n+1} Y_k = \sum_{k=1}^{m(\tau)} (\sup_{E_k} f + \varepsilon)\mu_n B_k \leqslant \sum_{k=1}^{m(\tau)} (\sup_{E_k} f + \varepsilon)(\mu_n A_k + \frac{\varepsilon}{m(\tau)}) =$$
$$= \sum_{k=1}^{m(\tau)} (\sup_{E_k} f \cdot \underbrace{\mu_n A_k}_{\leqslant \mu_n E_k} + \varepsilon\mu_n A_k + \frac{\varepsilon^2}{m(\tau)} + \sup_{E_k} f \cdot \frac{\varepsilon}{m(\tau)}) \leqslant S_\tau(f) + \varepsilon \sum_{k=1}^{m(\tau)} \mu A_k + \varepsilon^2 + \underbrace{\sup_E f}_{M} \cdot \varepsilon$$

Так как $A_k \subset E_k$, а E_k дизъюнктны, то $\sum_{k=1}^{m(\tau)} \mu A_k \leqslant \mu(\cup E_k) = \mu E$. Теперь оценим $\varepsilon < 1$ и заменим сумму Дарбу на интеграл (она не больше $I + \varepsilon \leqslant I + 1$):

$$\leqslant S_\tau(f) + \varepsilon(\mu E + M + 1) \leqslant I + \varepsilon(\mu E + M + 2)$$

Дальше смотрим на $\mathcal{X}$. Мы знаем, что X_k дизъюнктны, так как каждое из них есть «надстройка» над A_k, которые уже дизъюнктны. Поэтому:

$$\mu_{n+1}\left(\bigcup_{k=1}^{m(\tau)} X_k\right) = \sum_{k=1}^{m(\tau)} \mu_{n+1} X_k = \sum_{k=1}^{m(\tau)} \inf_{E_k} f \cdot \mu_n A_k \geqslant \sum_{k=1}^{m(\tau)} \inf_{E_k} f \cdot (\mu_n E_k - \frac{\varepsilon}{m(\tau)}) =$$
$$= \sum_{k=1}^{m(\tau)} \inf_{E_k} f \cdot \mu_n E_k - \sum_{k=1}^{m(\tau)} \underbrace{\inf_{E_k} f}_{\leqslant M} \cdot \frac{\varepsilon}{m(\tau)} \geqslant s_\tau(f) - M\varepsilon \geqslant I - (M+1)\varepsilon$$

Давайте теперь собирать. Мы знаем, что $\mathcal{X} = \cup_{k=1}^{m(\tau)} X_k$ — клеточное, $\mu\mathcal{X} \geqslant I - (M+1)\varepsilon$. Аналогично знаем, что $\mathcal{Y} = \cup_{k=1}^{m(\tau)} Y_k$ — клеточное, $\mu\mathcal{Y} \leqslant I + \varepsilon(M + 2 + \mu E)$. Меры $\mathcal{X}$ и $\mathcal{Y}$ отличаются на некоторую константу, зависящую только от f и E, умноженную на ε. Можем сделать зазор сколь угодно маленьким. Также знаем, что $X \subset G \subset Y$. Отсюда знаем, что G измеримо по критерию измеримости. Тогда можем написать неравенство на меры:

$$I - (M+1)\varepsilon \leqslant \mu_{n+1}\mathcal{X} \leqslant \mu_{n+1} G \leqslant \mu_{n+1}\mathcal{Y} \leqslant I + \varepsilon(M + 2 + \mu_n E)$$

Устремляем ε к нулю, получаем равенство $\mu_{n+1} G = I$. ◀

𝔇𝔢𝔣 3.4.3. Пусть $E \subset \mathbb{R}^n$ измеримо. Есть непрерывные функции $\varphi, \psi\colon \operatorname{cl} E \to \mathbb{R}$, причём $\varphi(x) \leqslant \psi(x) \forall x \in E$. Тогда определим $X \subset \mathbb{R}^{n+1}$ вот так:

$$X = \{(x_1, \ldots, x_n, x_{n+1} \mid x = (x_1, \ldots, x_n) \in E, \varphi(x) \leqslant x_{n+1} \leqslant \psi(x)\}$$

Назовём X элементарным множеством (обобщение криволинейной трапеции).

Следствие 3.4.2.1.

1. Элементарное множество измеримо
2. Конечное объединение элементарных множеств измеримо

▶ Второй пункт очевиден, потому что конечное объединение измеримых измеримо. Разберёмся с первым пунктом.

$\operatorname{cl} E$ — компакт, потому что оно, очевидно, замкнуто и ограничено (так как в мере Жордана измеримые всегда ограничены). Возьмём $m = \min_{e \in \operatorname{cl} E}\{\varphi(x), \psi(x)\}$. Введём две неотрицательных функции $f(x) = \varphi(x) - m$ и $g(x) = \psi(x) - m$, они еще и непрерывны. Значит, интегрируемы.

Возьмём из теоремы множества G_f и G_g. Они измеримы. Их разность $G_g \setminus G_f$ тогда тоже измерима. Теперь заметим, что $X = (G_g \setminus G_f) \cup \mathit{\Gamma}_f$ тоже измеримо (например, потому что $\mathit{\Gamma}_f$ — это граница измеримого G_f, то есть имеет меру ноль и заведомо измеримо). По смыслу мы вернули обратно график f, который выкинулся вместе с подграфиком.

Теперь заметим, то G_g — это G_ψ, сдвинутое вниз на m, аналогично для G_f и G_φ. Значит, X имеет такую же меру, как и $G_\psi - G_\varphi$ (потому что при сдвиге ячейки переходят в ячейки). ◀

Следствие 3.4.2.2. Получили формулу для меры элементарного множества:

$$\mu_{n+1} X = \int_E (\psi - \varphi)$$

▶ В процессе доказательства предыдущего следствия (напомниаем, что $\varphi \leqslant \psi$:

$$\mu_{n+1} G_g = \int_E g = \int_E \psi - m\mu E$$
$$\mu_{n+1} G_f = \int_E f = \int_E \varphi - m\mu E$$
$$\mu_{n+1}(G_g \setminus G_f) = \mu_{n+1} G_g - \mu_{n+1} G_f = \int_E \psi - \int_E \varphi = \int_E (\psi - \varphi)$$

◀

Лемма 3.4.1. Пусть γ — спрямляемая кривая в $\mathbb{R}^n$ $(n \geqslant 2)$. Тогда γ измерима, причём $\mu\gamma = 0$.

Замечание 3.4.2. При $n = 1$ мера равна длине отрезка, поэтому $n \geqslant 2$ существенно. Спрямляемость (конечность длины) тоже существенна, иначе есть безумная кривая Пеано, заметающая квадратик $[0, 1]^2$ и, соответственно, имеющая меру один.

▶ Знаем, что $L(s) < +\infty$. Нарежем на m кусочков длины $\frac{l(s)}{m}$. В каждой из $m + 1$ точек (границ кусочков) возьмём ячейку с центром в этой точке и стороной $\frac{2l(s)}{m}$. Тогда эти ячейки покроют нашу кривую (так как из каждой точки ячейка покрывает соседние кусочки). Считаем сумму мер ячеек: $(m + 1) \left(\frac{2l(s)}{m}\right)^n \to 0$. ◀