Содержание

0.1. Преобразование Абеля.

Пусть $\{a_n\}$, $\{b_n\}$ — (комплексные) последовательности, $m \in \mathbb{N}$ и пусть $A_k = \sum_{i=1}^{k} a_i$ для всех k. Тогда $a_k = A_k - A_{k-1}$ (считаем $A_0 = 0$) и

$$\sum_{k=m}^{n} a_k b_k = \sum_{k=m}^{n} (A_k - A_{k-1}) b_k = \sum_{k=m}^{n} A_k b_k - \sum_{k=m-1}^{n-1} A_k b_{k+1}.$$

Таким образом, справедливо тождество (*преобразование Абеля*)

$$\sum_{k=m}^{n} a_k b_k = A_n b_n - A_{m-1} b_m - \sum_{k=m}^{n-1} A_k (b_{k+1} - b_k). \tag{0.1}$$

Лемма 0.1 (Абель). *Пусть $\{a_n\}$ — (комплексная) последовательность, $\{b_n\}$ — монотонная последовательность. Если $\left|\sum_{i=m}^{k} a_i\right| \leqslant M$ для всех k, то справедлива оценка*

$$\left|\sum_{k=m}^{n} a_k b_k\right| \leqslant 2M(|b_m| + |b_n|).$$

□ Можно считать все a_k при $k < m$ равны нулю. Тогда из (0.1) с учетом $A_{m-1} = 0$ и сохранением знака разностей $b_{k+1} - b_k$ получаем

$$\left|\sum_{k=m}^{n} a_k b_k\right| \leqslant M \left(|b_n| + \left|\sum_{k=m}^{n-1} (b_k - b_{k+1})\right|\right) =$$
$$= M(|b_n| + |b_n - b_m|) \leqslant 2M(|b_m| + |b_n|).$$

Замечание. При $m = 1$ оценку можно уточнить: если $\{b_n\}$ нестрого убывает и неотрицательна, $a_k \in \mathbb{R}$ и $c \leqslant A_k \leqslant C$ для всех k, то $cb_1 \leqslant \sum_{k=1}^{n} a_k b_k \leqslant Cb_1$.

Лемма 0.2 (Абель). *Пусть функция f интегрируема, а функция g монотонна на $[a, b]$. Если $\left|\int_a^x f(t)dt\right| \leqslant M$ для всех $x \in [a, b]$, то справедлива оценка*

$$\left|\int_a^b f(x)g(x)dx\right| \leqslant 2M(|g(a)| + |g(b)|).$$

□ Применим предыдущую лемму к интегральным суммам произведения fg. Пусть $\varepsilon > 0$. Положим $I = \int_a^b f(t)dt$. Будем использовать принятые обозначения теории интеграла Римана.

Найдется такое $\delta > 0$, что для всякого отмеченного разбиения $(T,\, \xi)$, $|T| < \delta$, имеем $|\sigma_T(f,\, \xi) - I| < \frac{\varepsilon}{2}$. Рассмотрим такое разбиение $T = \{x_i\}_{i=0}^n$ Пусть $T_k = \{x_i\}_{i=0}^k$ – соответствующее разбиение отрезка $[x_0,\, x_k]$, $k = 1,\, \ldots,\, n$. Числа $\sigma_{T_k}(f,\, \xi_k)$ и $\int_{x_0}^{x_k} f(t)dt$ лежат на отрезке с концами $s_{T_k}(f)$ и $S_{T_k}(f)$. Учитывая, что $S_{T_k}(f) - s_{T_k}(f) \leqslant S_T(f) - s_T(f)$ и $I - \frac{\varepsilon}{2} \leqslant s_T(f) \leqslant S_T(f) \leqslant I + \frac{\varepsilon}{2}$, получаем неравенство

$$\left|\sigma_{T_k}(f,\, \xi_k) - \int_{x_0}^{x_k} f(t)dt\right| \leqslant \varepsilon.$$

Положим $A_k = \sum\limits_{i=1}^{k} f(c_i)\Delta x_i$. Так как $A_k = \sigma_{T_k}(f,\, \xi_k)$, то из последнего неравенства следует, что $|A_k| \leqslant M + \varepsilon$. Тогда по предыдущей лемме

$$\left|\sum_{k=1}^{n} f(c_k)g(c_k)\Delta x_k\right| \leqslant 2(M + \varepsilon)(|g(c_1)| + |g(c_n)|).$$

Оценка верна для любого набора отмеченных точек, а значит, и для набора с условием $c_1 = a$, $c_n = b$. Выбирая последовательность таких отмеченных разбиений, мелкость которых стремится к нулю, приходим к неравенству

$$\left|\int_a^b f(x)g(x)dx\right| \leqslant 2(M + \varepsilon)(|g(a)| + |g(b)|).$$

Для завершения доказательства осталось перейти к пределу при $\varepsilon \to +0$. ■

Задача (формула Бонне). Пусть $f \in \mathcal{R}[a,\, b]$, функция g нестрого убывает и неотрицательна на $[a,\, b]$. Покажите, что существует такая точка $c \in [a,\, b]$, что

$$\int_a^b f(x)g(x)dx = g(a)\int_a^c f(x)dx.$$

Указание. Используйте замечание после леммы 0.1.

1. Несобственный интеграл

В этом разделе интеграл Римана обобщается в двух направлениях: на случай, когда область интегрирования является лучом, и на случай, когда подынтегральная функция неограничена.

1.1. Определение и основные свойства.

Определение. Функция f называется *локально интегрируемой* (по Риману) на промежутке I, если $f \in \mathcal{R}[a, c]$ для каждого отрезка $[a, c]$, содержащегося в I.

Например, непрерывная на промежутке функция локально интегрируема.

Определение. Пусть $-\infty < a < b \leqslant +\infty$, и функция f локально интегрируема на $[a, b)$. Предел

$$\int_a^b f(x)dx := \lim_{c \to b-0} \int_a^c f(x)dx$$

называется *несобственным интегралом* (Римана) *от* f *по* $[a, b)$. Если указанный предел существует и конечен, то интеграл $\int_a^b f(x)dx$ называется *сходящимся*; иначе *расходящимся*.

Здесь и далее под $+\infty - 0$ подразумевается символ $+\infty$. Такое соглашение позволит нам рассматривать случаи $b \in \mathbb{R}$ и $b = +\infty$ одновременно.

Замечание. Если $b \in \mathbb{R}$, функция f локально интегрируема на $[a, b)$ и *ограничена*, то по теореме 1.6.9 $f \in \mathcal{R}[a, b]$ (при любом доопределении в точке b). В силу непрерывности интеграла с переменным верхним пределом

$$\lim_{c \to b-0} \int_a^c f(x)dx = \int_a^b f(x)dx,$$

т.е. несобственный интеграл совпадает с определенным интегралом. Поэтому новая ситуация возникает только в случае:

а) $b = +\infty$, или б) $b \in \mathbb{R}$ и f неограничена на $[a, b)$.

Аналогично определяется *несобственный интеграл от* f *по* $(a, b]$, $-\infty \leqslant a < b < +\infty$.

Ряд свойств определенного интеграла переносится на несобственные интегралы предельным переходом. Для определенности все утверждения будут формулироваться только для случая несобственного интеграла по $[a, b)$.

Свойство 1 (принцип локализации). Пусть f локально интегрируема на $[a, b)$, и $a^* \in (a, b)$. Тогда интегралы $\int_{a^*}^b f(x)dx$ и $\int_a^b f(x)dx$ сходятся или расходятся одновременно, а если сходятся, то

$$\int_a^b f(x)dx = \int_a^{a^*} f(x)dx + \int_{a^*}^b f(x)dx. \tag{1.1}$$

$\square$ Если $c \in (a, b)$, то по свойству аддитивности определенного интеграла

$$\int_a^c f(x)dx = \int_a^{a^*} f(x)dx + \int_{a^*}^c f(x)dx.$$

Поэтому пределы $\lim\limits_{c \to b-0} \int_a^c f(x)dx = \int_a^b f(x)dx$ и $\lim\limits_{c \to b-0} \int_{a^*}^c f(x)dx = \int_{a^*}^b f(x)dx$ существуют (конечны) одновременно. В случае существования равенство (1.1) следует из равенства определенных интегралов предельным переходом при $c \to b - 0$. $\blacksquare$

Следующие три свойства доказываются аналогично.

Свойство 2 (линейность). Пусть несобственные интегралы $\int_a^b f(x)dx$ и $\int_a^b g(x)dx$ сходятся, и α, $\beta \in \mathbb{R}$. Тогда сходится и интеграл $\int_a^b (\alpha f(x) + \beta g(x))dx$, причем верно равенство

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

Свойство 3. Пусть f локально интегрируема на $[a, b)$, и F — первообразная f на $[a, b)$. Тогда

$$\int_a^b f(x)dx = F(b-0) - F(a)$$

при существовании хотя бы одной из частей формулы.

Свойство 4 (интегрирование по частям). Пусть функции F и G дифференцируемы, а их производные f, g локально интегрируемы на $[a, b)$. Тогда

$$\int_a^b F(x)g(x)dx = F(x)G(x)\Big|_a^{b-0} - \int_a^b G(x)f(x)dx$$

(существование двух конечных пределов влечет существование третьего, и равенство выполняется).

Свойство 5 (замена переменной). Пусть f непрерывна на $[a,\, b)$, φ дифференцируема, строго монотонна на $[\alpha,\, \beta)$, причем φ' локально интегрируема на $[\alpha,\, \beta)$, $\varphi(\alpha) = a$ и $b = \lim\limits_{t\to\beta-0} \varphi(t)$. Тогда

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt$$

(существование одного из интегралов, влечет существование другого, и равенство выполняется).

□ Определим $F(c) = \int_a^c f(x)dx$ и $\Phi(\gamma) = \int_\alpha^\gamma f(\varphi(t))\varphi'(t)dt$. По формуле замены переменной в определенном интеграле $F(\varphi(\gamma)) = \Phi(\gamma)$ для всех $\gamma \in [\alpha,\, \beta)$.

Пусть (в $\overline{\mathbb{R}}$) определен несобственный интеграл $I = \int_a^b f(x)dx$. Тогда по свойству предела композиции существует

$$\lim_{\gamma\to\beta-0} \Phi(\gamma) = \lim_{c\to b-0} F(c) = I,$$

т.е. определен несобственный интеграл $\int_\alpha^\beta f(\varphi(t))\varphi'(t)dt = I$.

Поскольку определена обратная функция φ^{-1} и $\gamma = \varphi^{-1}(c) \to \to \beta$ при $c \to b-0$, то по свойству предела композиции существование $\lim\limits_{\gamma\to\beta-0} \Phi(\gamma)$ влечет существование равного $\lim\limits_{c\to b-0} F(c)$, т.е. правая часть равенства влечет существование левой. ■

Замечание. Для несобственных интегралов также примем соглашение $\int_a^b f(x)dx = -\int_b^a f(x)dx$, считая, что при $a > b$ выражение $[a,\, b)$ обозначает $(b,\, a]$.

Пример. Исследуем сходимость $\int\limits_1^{+\infty} \frac{dx}{x^\alpha}$ при всех $\alpha \in \mathbb{R}$.

□ Если $\alpha \neq 1$, то $\int\limits_1^{+\infty} \frac{dx}{x^\alpha} = \frac{x^{-\alpha+1}}{1-\alpha}\Big|_1^{+\infty} = \begin{cases} \frac{1}{\alpha-1}, & \alpha > 1 \\ +\infty, & \alpha < 1. \end{cases}$

Если $\alpha = 1$, то $\int\limits_1^{+\infty} \frac{dx}{x^\alpha} = \ln x\big|_1^{+\infty} = +\infty$. ■

Ответ: $\int_1^{+\infty} \frac{dx}{x^\alpha}$ сходится $\Longleftrightarrow \alpha > 1$.

Пример. Исследуем сходимость $\int\limits_0^1 \frac{dx}{x^\alpha}$ при всех $\alpha \in \mathbb{R}$.

$\square$ Интеграл рассматривается как несобственный по промежутку $(0,\, 1]$. Сделав в нем замену $x = \frac{1}{t}$, получим $\int_1^{+\infty} \frac{dt}{t^{2-\alpha}}$, который сходится $\iff \alpha < 1$. $\blacksquare$

Ответ: $\int_0^1 \frac{dx}{x^\alpha}$ сходится $\iff \alpha < 1$.

Теорема 1.1 (критерий Коши). *Пусть f локально интегрируема на $[a,\, b)$. Тогда для сходимости интеграла $\int_a^b f(x)dx$ необходимо и достаточно выполнения условия Коши:*

$$\forall \varepsilon > 0\ \exists b_\varepsilon \in [a,\, b)\ \forall \xi,\, \eta \in (b_\varepsilon,\, b)\ \ \left(\left|\int_\xi^\eta f(x)dx\right| < \varepsilon\right).$$

$\square$ Пусть $F(x) = \int_a^x f(t)dt$, $x \in [a,\, b)$. Так как $\int_\xi^\eta f(x)dx = F(\eta) - F(\xi)$, то доказываемое утверждение — переформулировка критерия Коши существования предела F при $x \to b - 0$. $\blacksquare$

Определение. Пусть f локально интегрируема на $[a,\, b)$. Несобственный интеграл $\int_a^b f(x)dx$ называется *абсолютно сходящимся*, если сходится интеграл $\int_a^b |f(x)|dx$. Если несобственный интеграл сходится, но не сходится абсолютно, то он называется *условно сходящимся*.

Замечание. Сходимость интеграла $\int_a^b |f(x)|dx$ не дает права использовать выражение $\int_a^b f(x)dx$, поскольку f может не быть интегрируемой на некотором отрезке $[a,\, c]$. Примером такой функции служит модификация функции Дирихле: $f(x) = \mathcal{D}(x) - \frac{1}{2}$ при $x \in [a,\, c]$ и $f(x) = 0$ при $x > c$.

Следствие. *Абсолютно сходящийся интеграл сходится.*

$\square$ Для любых ξ, η, таких что $a \leqslant \xi < \eta < b$, верно $\left|\int_\xi^\eta f(x)dx\right| \leqslant \int_\xi^\eta |f(x)|dx$. Отсюда следует, что если условие Коши выполняется для интеграла от $|f|$, то оно выполняется и для интеграла от f. Таким образом, если интеграл от $|f|$ сходится, то интеграл от f тоже сходится по критерию Коши. $\blacksquare$

Из последнего неравенства следует, что если интеграл $\int_a^b f(x)dx$ абсолютно сходится, то $\left|\int_a^b f(x)dx\right| \leqslant \int_a^b |f(x)|dx$.

1.2. Интегралы от неотрицательных функций.

Последнее следствие указывает на особую роль интегралов от неотрицательных функций. Изучим вопросы сходимости таких интегралов.

Лемма 1.1. *Пусть f локально интегрируема и неотрицательна на $[a, b)$. Тогда сходимость интеграла $\int_a^b f(x)dx$ равносильна ограниченности функции $F(x) = \int_a^x f(t)dt$ на $[a, b)$.*

□ Функция F нестрого возрастает на $[a, b)$, т.к. для любых x_1, $x_2 \in [a, b)$ из условия $x_1 < x_2$ следует $F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(t)dt \geqslant 0$. По теореме 1.4.3 о пределах монотонной функции существует $\lim\limits_{x \to b-0} F(x) = \sup_{[a, b)} F(x)$. Отсюда ввиду неотрицательности заключаем, что ограниченность F равносильна наличию конечного предела, т.е. сходимости интеграла. ■

Замечание. Несобственный интеграл $\int_a^b f(x)dx$ от неотрицательной функции либо сходится, либо расходится к $+\infty$. Сходимость равносильна ограниченности некоторой последовательности $I_n = \int_a^{c_n} f(x)dx$, где $c_n \in [a, b)$ и $c_n \to b$. Это вытекает из равенства $\lim\limits_{c \to b-0} \int_a^c f(x)dx = \lim\limits_{n \to \infty} I_n$.

Теорема 1.2 (признак сравнения). *Пусть f, g локально интегрируемы на $[a, b)$, и $0 \leqslant f \leqslant g$ на $[a, b)$.*

1) *Если $\int_a^b g(x)\,dx$ сходится, то и $\int_a^b f(x)\,dx$ сходится.*

2) *Если $\int_a^b f(x)\,dx$ расходится, то и $\int_a^b g(x)\,dx$ расходится.*

□ Для любого $x \in [a, b)$ выполнено $0 \leqslant \int_a^x f(t)\,dt \leqslant \int_a^x g(t)\,dt$. Если $\int_a^b g(x)\,dx$ сходится, то по лемме 1.1 функция $\int_a^x g(t)\,dt$ ограничена на $[a, b)$. Но тогда ограничена и функция $\int_a^x f(t)\,dt$, что по лемме 1.1 влечет сходимость $\int_a^b f(x)\,dx$.

Второе доказываемое утверждение вытекает из первого по правилу контрапозиции. ■

Следствие 1.1. *Пусть f, g локально интегрируемы и неотрицательны на $[a, b)$. Если $f(x) = O(g(x))$ при $x \to b-0$, то справедливо заключение предыдущей теоремы.*

□ Поскольку $f(x) = O(g(x))$ при $x \to b-0$ и f, g неотрицательны,

$$\exists C > 0\ \exists a^* \in [a, b)\ \forall x \in [a^*, b)\ (f(x) \leqslant C\,g(x)).$$

Если интеграл $\int_a^b g(x)dx$ сходится, то сходится и интеграл $\int_{a^*}^b Cg(x)dx$. Тогда по признаку сравнения сходится интеграл $\int_{a^*}^b f(x)dx$, а значит, сходится интеграл $\int_a^b f(x)dx$. ■

Следствие 1.2. *Пусть f, g локально интегрируемы и положительны на $[a, b)$. Если существует $\lim\limits_{x\to b-0} \dfrac{f(x)}{g(x)} \in (0, +\infty)$, то несобственные интегралы от функций f и g по $[a, b)$ сходятся или расходятся одновременно.*

□ В условиях следствия существует также $\lim\limits_{x\to b-0} \dfrac{g(x)}{f(x)} \in (0, +\infty)$. Поскольку существование конечного предела влечет ограниченность функции в некоторой окрестности предельной точки, то $f(x) = O(g(x))$ и $g(x) = O(f(x))$ при $x \to b-0$. Осталось воспользоваться следствием 1.1. ■

Пример. Исследуем сходимость интегралов

а) $\displaystyle\int\limits_1^{+\infty} x^{100}e^{-x}dx$, б) $\displaystyle\int\limits_0^1 \frac{dx}{\operatorname{tg} x}$.

□ а) Поскольку $e^{-x} = o\left(\dfrac{1}{x^{102}}\right)$, то $x^{100}e^{-x} = o\left(\dfrac{1}{x^2}\right)$ при $x \to$ $\to +\infty$. Интеграл $\int_1^{+\infty} \dfrac{dx}{x^2}$ сходится, поэтому по признаку сравнения (следствию 1.1) интеграл $\int_1^{+\infty} x^{100}e^{-x}dx$ также сходится.

б) Поскольку $\operatorname{tg} x \sim x$ при $x \to +0$ и интеграл $\int_0^1 \dfrac{dx}{x}$ расходится, то по признаку сравнения (следствию 1.2) интеграл $\int_0^1 \dfrac{dx}{\operatorname{tg} x}$ также расходится. ■

1.3. Интегралы от знакопеременных функций.

Изучим вопросы сходимости интегралов от функций, которые не сохраняют знак ни в какой окрестности точки b.

Лемма 1.2. *Пусть f, g локально интегрируемы на $[a, b)$.*

1) *Если $\int_a^b g(x)\,dx$ сходится, то интегралы $\int_a^b f(x)\,dx$ и $\int_a^b (f(x)+g(x))dx$ сходятся или расходятся одновременно.*
2) *Если $\int_a^b g(x)\,dx$ абсолютно сходится, то интегралы $\int_a^b f(x)\,dx$ и $\int_a^b (f(x)+g(x))dx$ либо одновременно расходятся, либо одновременно сходятся условно, либо одновременно сходятся абсолютно.*

□ Первое утверждение непосредственно вытекает из свойства линейности. Из неравенств

$$|f+g| \leqslant |f|+|g|, \quad |f| \leqslant |f+g|+|g|$$

по признаку сравнения следует, что интегралы от f и $f+g$ одновременно сходятся абсолютно. Теперь второе доказываемое утверждение следует из первого. ■

Теорема 1.3 (признак Дирихле). *Пусть функции f, g локально интегрируемы на $[a,\, b)$, причем*

1) *функция $F(x) = \int_a^x f(t)\,dt$ ограничена на $[a,\, b)$;*
2) *функция g монотонна на $[a,\, b)$;*
3) $\lim\limits_{x\to b-0} g(x) = 0$.

Тогда несобственный интеграл $\int_a^b f(x)g(x)dx$ сходится.

□ Пусть $|F| < M$ на $[a,\, b)$, тогда для всякого $\xi \in [a,\, b)$ справедлива оценка

$$\left|\int_\xi^x f(t)\,dt\right| = |F(x) - F(\xi)| < 2M.$$

Зафиксируем $\varepsilon > 0$. Выберем такое $b' \in [a,\, b)$, что $|g(x)| < \frac{\varepsilon}{8M}$ при всех $x \in (b',\, b)$. Воспользуемся леммой 0.2 Абеля. Тогда для любого $[\xi,\, \eta] \subset (b',\, b)$ выполнено

$$\left|\int_\xi^\eta f(t)g(t)\,dt\right| \leqslant 2\cdot 2M(|g(\xi)|+|g(\eta)|) < 4M\left(\frac{\varepsilon}{8M}+\frac{\varepsilon}{8M}\right) = \varepsilon.$$

Доказательство завершается ссылкой на критерий Коши. ■

Заметим, что первые два условия теоремы выполнены, если f непрерывна и имеет ограниченную на $[a,\, b)$ первообразную, а g дифференцируема и g' не меняет знака на этом промежутке.

Следующая теорема является следствием признака Дирихле.

Теорема 1.4 (признак Абеля). *Пусть функции f, g локально интегрируемы на $[a,\, b)$, причем*

1) *интеграл $\int_a^b f(x)dx$ сходится;*
2) *функция g монотонна на $[a,\, b)$;*
3) *функция g ограничена на $[a,\, b)$.*

Тогда несобственный интеграл $\int_a^b f(x)g(x)dx$ сходится.

$\square$ Так как g монотонна и ограничена, то существует $\lim\limits_{x\to b-0} g(x) = c \in \mathbb{R}$. Функции f и $g-c$ удовлетворяют признаку Дирихле, поэтому $\int_a^b f(x)(g(x)-c)dx$ сходится. Тогда

$$\int_a^b f(x)g(x)dx = \int_a^b f(x)(g(x)-c)dx + c\int_a^b f(x)dx$$

сходится как сумма сходящихся интегралов. $\blacksquare$

Следствие. *Пусть f локально интегрируема на $[a, b)$, g монотонна на $[a, b)$ и $\lim\limits_{x\to b-0} g(x) = c \in \mathbb{R}\setminus\{0\}$. Тогда интегралы $\int_a^b f(x)\,g(x)\,dx$ и $\int_a^b f(x)\,dx$ либо одновременно расходятся, либо одновременно сходятся условно, либо одновременно сходятся абсолютно.*

$\square$ Покажем, что интегралы $\int_a^b f(x)g(x)dx$ и $\int_a^b f(x)dx$ сходятся одновременно. Действительно, если интеграл $\int_a^b f(x)dx$ сходится, то $\int_a^b f(x)g(x)dx$ сходится по признаку Абеля. Так как $c \neq 0$, то на некотором промежутке $[a^*, b)$ функция g сохраняет знак и, значит, определена функция $h = \dfrac{1}{g}$, которая является монотонной на $[a^*, b)$. Поскольку $f = fgh$ на $[a^*, b)$, то из сходимости интеграла $\int_{a^*}^b f(x)g(x)dx$ по признаку Абеля следует сходимость $\int_{a^*}^b f(x)dx$, а следовательно, и сходимость $\int_a^b f(x)dx$.

Из существования конечных пределов g и h следует, что $|fg| = O(|f|)$ при $x\to b-0$ и $|f| = O(|fg|)$ при $x\to b-0$. По следствию 1.1 из признака сравнения заключаем, что интегралы $\int_a^b |f(x)g(x)|dx$ и $\int_a^b |f(x)|dx$ сходятся или расходятся одновременно. $\blacksquare$

Пример. Исследуем сходимость и абсолютную сходимость интеграла $I(\alpha) = \int\limits_1^{+\infty} \dfrac{\sin kx}{x^\alpha}\,dx$ $(k>0)$ при всех $\alpha\in\mathbb{R}$.

$\square$ Можно считать, что $k=1$ (иначе сделаем замену $t = kx$).

Пусть $\alpha > 1$. Поскольку $\dfrac{|\sin x|}{x^\alpha} \leqslant \dfrac{1}{x^\alpha}$ на луче $[1, +\infty)$, то по признаку сравнения интеграл $\int_1^{+\infty} \dfrac{|\sin x|}{x^\alpha}\,dx$ сходится, а значит, $I(\alpha)$ сходится абсолютно.

Пусть $\alpha \leqslant 0$. В этом случае интеграл $I(\alpha)$ расходится, т.к. выполняется отрицание условия Коши: для $\varepsilon = 2$ и всякого $\Delta \geqslant 1$

найдутся ξ, $\eta > \Delta$, $\xi = 2\pi n$, $\eta = \pi + 2\pi n$ $(n \in \mathbb{N})$, такие что

$$\left|\int_\xi^\eta \frac{\sin x}{x^\alpha}\,dx\right| = \int_\xi^\eta x^{-\alpha}\sin x\,dx \geqslant (2\pi n)^{-\alpha}\int_{2\pi n}^{\pi+2\pi n}\sin x\,dx \geqslant 2.$$

Пусть $0 < \alpha \leqslant 1$.

Функции $f(x) = \sin x$, $g(x) = \frac{1}{x^\alpha}$ удовлетворяют признаку Дирихле, поэтому интеграл $I(\alpha)$ сходится. Однако при $n \in \mathbb{N}$ имеем

$$\int_{\pi n}^{2\pi n} \frac{|\sin x|}{x^\alpha}\,dx \geqslant \frac{1}{(2\pi n)^\alpha}\int_{\pi n}^{2\pi n}|\sin x|\,dx =$$

$$= \frac{1}{(2\pi n)^\alpha}\sum_{m=0}^{n-1}\int_{\pi(n+m)}^{\pi(n+m+1)}|\sin x|\,dx = \frac{n}{(2\pi n)^\alpha}\int_0^\pi |\sin x|\,dx \geqslant \frac{1}{\pi^\alpha}.$$

Следовательно, по критерию Коши $\int_1^{+\infty}\frac{|\sin x|}{x^\alpha}\,dx$ расходится и, значит, $I(\alpha)$ сходится условно. ■

Задача. Пусть $f\colon [1, +\infty) \to \mathbb{R}$ непрерывна и $\int_1^{+\infty} f(x)dx$ сходится. Верно ли, что $\lim\limits_{x\to+\infty} f(x) = 0$? Тот же вопрос при дополнительном предположении, что а) f неотрицательна на $[1, +\infty)$; б) f равномерно непрерывна на $[1, +\infty)$.

Теперь определим интегралы с несколькими особенностями.

Определение. Пусть $a, b \in \overline{\mathbb{R}}$, $a < b$, и функция f определена на (a, b) кроме, быть может, конечного множества точек.

Точка $c \in (a, b)$ называется *особенностью* f, если $f \notin \mathcal{R}[\alpha, \beta]$ для любого $[\alpha, \beta] \subset (a, b)$, $\alpha < c < \beta$ (при доопределении в c).

Точка b называется *особенностью* f, если $b = +\infty$, или $b \in \mathbb{R}$ и $f \notin \mathcal{R}[\alpha, b]$ для любого $[\alpha, b]$, где $a < \alpha < b$. Аналогично определяется особенность в левом конце a.

Замечание. Если на интервале (c, d) нет особенностей функции f, то f локально интегрируема на (c, d).

□ Пусть $[u, v] \subset (c, d)$. По условию каждая точка $x \in [u, v]$ лежит внутри некоторого отрезка $[\alpha_x, \beta_x]$, на котором f интегрируема. По теореме Гейне-Бореля из покрытия $\{(\alpha_x, \beta_x)\}_{x\in[u, v]}$ отрезка $[u, v]$ можно выделить конечное подпокрытие. Объединяя элементы этого подпокрытия и пользуясь аддитивностью интеграла, заключаем, что f интегрируема на отрезке, содержащем $[u, v]$. ■

Пусть $c_1 < \ldots < c_{N-1}$ — все особенности f на (a, b), $c_0 = a$, $c_N = b$. Пусть ξ_k — произвольная точка (c_{k-1}, c_k), $k = 1, \ldots, N$. *Несобственным интегралом* функции f по (a, b) называется

$$\int_a^b f(x)\,dx := \sum_{k=1}^{N} \left(\int_{c_{k-1}}^{\xi_k} f(x)\,dx + \int_{\xi_k}^{c_k} f(x)\,dx \right),$$

если все интегралы в правой части (понимаемые как несобственные) и сумма имеют смысл в $\overline{\mathbb{R}}$. При этом $\int_a^b f(x)\,dx$ называется *сходящимся*, если сходятся все интегралы в правой части, иначе — *расходящимся*.

Корректность (независимость от ξ_k) следует из принципа локализации.

Таким образом, исследование интегралов с несколькими особенностями сводится к исследованию интегралов с одной особенностью.

2. Числовые ряды

Мы переходим к формализации понятия «суммы бесконечного числа слагаемых». Существенная часть теории в этом разделе будет направлена на установление условий, позволяющих работать с такими объектами также как с конечными суммами.

2.1. Сумма числового ряда.

Определение. Пусть $\{a_n\}$ — последовательность действительных (комплексных) чисел. Выражение

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \ldots \tag{2.1}$$

называется *числовым рядом* с n-м членом a_n. Число $S_N = \sum\limits_{n=1}^{N} a_n$ называется *N-й частичной суммой* ряда (2.1). Предел

$$\sum_{n=1}^{\infty} a_n = \lim_{N\to\infty} S_N$$

называется *суммой* ряда (2.1). Если указанный предел существует и конечен, то ряд (2.1) называется *сходящимся*; иначе *расходящимся*.

Замечание. С каждой последовательностью $\{s_n\}$ можно связать ряд (2.1), для которого $\{s_n\}$ является последовательностью частичных сумм: достаточно положить $a_1 = s_1$, $a_n = s_n - s_{n-1}$ при $n > 1$ (*телескопический ряд*).

Отметим, что нумерация членов ряда может начинаться с любого $m \in \mathbb{Z}$.

Примеры. Исследуем сходимость рядов

а) $\sum\limits_{n=0}^{\infty} z^n$, $z \in \mathbb{C}$ (*геометрический ряд*); б) $\sum\limits_{n=1}^{\infty} \dfrac{1}{n(n+1)}$.

□ а) При $z = 1$ ряд расходится, т.к. $S_N = N + 1 \to \infty$. При $z \neq$ $\neq 1$ частичная сумма $S_N = \dfrac{1 - z^{N+1}}{1 - z}$. Если $|z| < 1$, то $|z|^N \to 0$, а значит, $z^N \to 0$. Поэтому $S_N \to \dfrac{1}{1-z}$, ряд сходится, и его сумма равна $\dfrac{1}{1-z}$. Если $|z| \geqslant 1$, то $\{S_N\}$ не имеет (конечного) предела, т.к. в противном случае $z^N = S_N - S_{N-1} \to 0$ (см. свойство 3 ниже). Итак, $\sum\limits_{n=0}^{\infty} z^n$ сходится $\Longleftrightarrow |z| < 1$.

б) Так как

$$S_N = \sum_{n=1}^{N} \frac{1}{n(n+1)} = \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{N+1} \to 1,$$

то ряд сходится и его сумма равна единице. ■

Перечислим основные свойства числовых рядов.

Свойство 1 (принцип локализации). Для каждого $m \in \mathbb{N}$ ряды $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=m+1}^{\infty} a_n$ сходятся или расходятся одновременно, а если сходятся, то

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{m} a_n + \sum_{n=m+1}^{\infty} a_n.$$

□ Если $N > m$, то $\sum\limits_{n=1}^{N} a_n = \sum\limits_{n=1}^{m} a_n + \sum\limits_{m=m+1}^{N} a_n$. Поэтому пределы левой и правой частей при $N \to \infty$ существуют (конечны) одновременно. В случае существования заявленное равенство получается предельным переходом при $N \to \infty$. ■

Замечание. Ряд $r_N = \sum\limits_{n=N+1}^{\infty} a_n$ называется N*-м остатком* ряда (2.1). Принцип локализации можно переформулировать так: если ряд сходится, то сходится и любой его остаток; если некоторый остаток ряда сходится, то и весь ряд сходится.

Свойство 2 (линейность). Пусть ряды $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=1}^{\infty} b_n$ сходятся, и α, $\beta \in \mathbb{R}$ (или $\mathbb{C}$). Тогда сходится и ряд $\sum\limits_{n=1}^{\infty} (\alpha a_n + \beta b_n)$, причем верно равенство

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n.$$

□ Вытекает из свойства линейности предела последовательности. ■

Свойство 3 (необходимое условие сходимости ряда). Если ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится, то $\lim\limits_{n\to\infty} a_n = 0$.

$\square$ По определению $a_n = S_n - S_{n-1}$ (считаем $S_0 = 0$), поэтому $a_n \to S - S = 0$, где S — сумма ряда. $\blacksquare$

Как показывает следующий пример, стремление членов к нулю не является достаточным условием сходимости ряда.

Пример. *Гармонический ряд* $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ расходится: пусть $H_n = \sum\limits_{k=1}^{n} \frac{1}{k}$, тогда $H_{2n} - H_n = \sum\limits_{k=n+1}^{2n} \frac{1}{k} \geqslant \frac{1}{2n} n = \frac{1}{2} \not\to 0$; но $\frac{1}{n} \to 0$.

Рассмотрим вопрос расстановки скобок в рядах.

Определение. Пусть дана строго возрастающая последовательность целых чисел $0 = n_0 < n_1 < n_2 < \ldots$. Ряд $\sum\limits_{k=1}^{\infty} b_k$, где $b_k = a_{n_{k-1}+1} + \ldots + a_{n_k}$, называется *группировкой* ряда $\sum\limits_{n=1}^{\infty} a_n$.

Свойство 4. а) Если ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится, то любая его группировка сходится к той же сумме.

б) Пусть существует такое $L \in \mathbb{N}$, что $n_k - n_{k-1} \leqslant L$ для всех k. Если $a_n \to 0$ и группировка $\sum\limits_{k=1}^{\infty} b_k$, где $b_k = \sum\limits_{j=n_{k-1}+1}^{n_k} a_j$, сходится, то ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится к той же сумме.

$\square$ Обозначим через S_N частичную сумму исходного ряда, а через S_N^* частичную сумму его группировки.

а) Пусть $S_N \to S$. Так как $S_N^* = S_{n_N}$, то $S_N^* \to S$ как подпоследовательность.

б) В основе доказательства лежит наблюдение, что для всякого номера n найдется такое k, что $S_n = S_{n_k} + a_{n_k+1} + \ldots + a_n$ и $S_{n_k} = S_k^*$.

Пусть $\varepsilon > 0$. Выберем такие K, $M \in \mathbb{N}$, что $|S_k^* - S| < \frac{\varepsilon}{2}$ при $k \geqslant K$ и $|a_m| < \frac{\varepsilon}{2L}$ при $m \geqslant M$. Положим $N = \max\{n_K, M + L\}$.

Пусть $n \geqslant N$. Тогда $n_k \leqslant n < n_{k+1}$ для некоторого $k \geqslant K$ и, значит,

$$|S_n - S| \leqslant |S_{n_k} - S| + |a_{n_k+1}| + \ldots + |a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2L} \cdot L = \varepsilon. \quad \blacksquare$$

Применяя критерий Коши к последовательности частичных сумм $\{S_n\}$, получим *критерий Коши сходимости ряда.*

Теорема 2.1. *Для сходимости ряда* (2.1) *необходимо и достаточно выполнения условия Коши*

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall m,\, n \in \mathbb{N} \ \left(\left| \sum_{k=m}^{n} a_k \right| < \varepsilon \right).$$

Определение. Ряд $\sum\limits_{n=1}^{\infty} a_n$ называется *абсолютно сходящимся*, если сходится ряд $\sum\limits_{n=1}^{\infty} |a_n|$. Если ряд сходится, но не сходится абсолютно, то он называется *условно сходящимся.*

Следствие. *Абсолютно сходящийся ряд сходится.*

Рассмотренные свойства говорят о тесной связи между рядами и несобственными интегралами. Следующее важное утверждение позволяет напрямую переносить аналогичные результаты из теории интегралов в теорию рядов.

С *действительным* рядом (2.1) свяжем функцию $f_a \colon [1, +\infty) \to \mathbb{R}$, такую что $f_a(x) = a_n$ при всех $x \in [n, n+1)$, $n \in \mathbb{N}$.

Лемма 2.1. *Ряд* $\sum\limits_{n=1}^{\infty} a_n$ *и интеграл* $\int\limits_{1}^{+\infty} f_a(x)dx$ *сходятся или расходятся одновременно, причем если сходятся, то к одному значению.*

$\square$ Обозначим через S_n n-ю частичную сумму ряда (2.1). Поскольку $S_n = \int_1^{n+1} f_a(x)dx$, то сходимость интеграла влечет сходимость ряда. Обратное утверждение следует из оценки

$$\left| \int_1^x f_a(x)dx - S_n \right| \leqslant \int_x^{n+1} |f_a(x)|dx \leqslant |a_n|, \quad n = [x],$$

и необходимого условия сходимости ряда. $\blacksquare$

2.2. Ряды с неотрицательными членами.

Поскольку последовательность частичных сумм ряда с неотрицательными членами нестрого возрастает, то справедлива

Лемма 2.2. *Пусть $a_n \geqslant 0$ для всех $n \in \mathbb{N}$. Ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится тогда и только тогда, когда ограничена последовательность его частичных сумм $\{S_n\}$.*

Простым следствием леммы 2.1 и теоремы 1.2 является

Теорема 2.2 (Признак сравнения). *Пусть $0 \leqslant a_n \leqslant b_n$ для всех $n \in \mathbb{N}$.*

1) *Если ряд $\sum\limits_{n=1}^{\infty} b_n$ сходится, то и ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится.*

2) *Если ряд $\sum\limits_{n=1}^{\infty} a_n$ расходится, то и ряд $\sum\limits_{n=1}^{\infty} b_n$ расходится.*

Следующее два следствия выводятся из теоремы 2.2 также как аналогичные утверждения для интегралов выводятся из теоремы 1.2.

Следствие 2.1. *Пусть a_n, $b_n \geqslant 0$ для всех $n \in \mathbb{N}$ и $a_n = O(b_n)$ при $n \to \infty$. Тогда справедливо заключение предыдущей теоремы.*

Следствие 2.2. *Пусть a_n, $b_n > 0$ для всех $n \in \mathbb{N}$. Если существует $\lim\limits_{n\to\infty} \frac{a_n}{b_n} \in (0, +\infty)$, то ряды $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=1}^{\infty} b_n$ сходятся или расходятся одновременно.*

В лемме 2.1 с рядом связывался специальный интеграл. Для *монотонных* функций можно рассмотреть двойственный результат о связи интеграла с рядом из значений в целых точках.

Теорема 2.3 (интегральный признак). *Пусть функция f неотрицательна и нестрого убывает на $[1, +\infty)$. Тогда ряд $\sum\limits_{n=1}^{\infty} f(n)$ и интеграл $\int\limits_{1}^{+\infty} f(x)dx$ сходятся или расходятся одновременно.*

$\square$ Определим функции u, $v\colon [1, +\infty) \to \mathbb{R}$, так что $u|_{[n,\,n+1)} = f(n)$ и $v|_{[n,\,n+1)} = f(n+1)$ для каждого n. Ввиду монотонности функции f верно $v \leqslant f \leqslant u$ на $[1, +\infty)$.

Если ряд $\sum_{n=1}^{\infty} f(n)$ сходится, то по лемме 2.1 сходится $\int_1^{+\infty} u(x)dx$, а значит, по теореме 1.2 сходится $\int_1^{+\infty} f(x)dx$.

Если интеграл $\int_1^{+\infty} f(x)dx$ сходится, то сходится и $\int_1^{+\infty} v(x)dx$, что влечет сходимость ряда $\sum_{n=1}^{\infty} f(n+1)$. Этот ряд сходится одновременно с рядом $\sum_{n=1}^{\infty} f(n)$ как его остаток. ■

Пример. Ряд $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\alpha}}$ сходится $\Longleftrightarrow \alpha > 1$.

□ При $\alpha \leqslant 0$ ряд расходится, поскольку n-й член не стремится к нулю. При $\alpha > 0$ к функции $f(x) = \frac{1}{x^{\alpha}}$ применима теорема 2.3, поэтому $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\alpha}}$ и $\int\limits_{1}^{+\infty} \frac{dx}{x^{\alpha}}$ сходятся одновременно. Последний интеграл сходится только при $\alpha > 1$. ■

Задача. Покажите, что в условиях теоремы 2.3 последовательность $\alpha_n = \sum_{k=1}^{n} f(k) - \int_1^{n+1} f(x)dx$ сходится.

Следующие два признака основаны на сравнении с геометрическим рядом.

Теорема 2.4 (признак Коши). *Пусть $a_n \geqslant 0$ для всех $n \in \mathbb{N}$ и $q = \varlimsup\limits_{n\to\infty} \sqrt[n]{a_n}$.*

1) *Если $q < 1$, то ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится.*

2) *Если $q > 1$, то $a_n \nrightarrow 0$ и ряд $\sum\limits_{n=1}^{\infty} a_n$ расходится.*

□ 1) Пусть $q_0 \in (q,\, 1)$. Выберем N так, что $\sup\limits_{n\geqslant N} \sqrt[n]{a_n} < q_0$. Тогда $a_n < q_0^n$ для всех $n \geqslant N$ и ряд сходится по признаку сравнения с геометрическим рядом $\sum q_0^n$.

2) Так как q — частичный предел, то существует подпоследовательность $\sqrt[n_k]{a_{n_k}} \to q$. Тогда неравенство $a_{n_k} > 1$ выполняется для всех достаточно больших k. Следовательно, $a_n \nrightarrow 0$ и ряд расходится. ■

Теорема 2.5 (признак Даламбера). *Пусть $a_n > 0$ для всех $n \in \mathbb{N}$.*

1) *Если $\overline{r} = \varlimsup\limits_{n\to\infty} \frac{a_{n+1}}{a_n} < 1$, то ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится.*

2) *Если $\underline{r} = \varliminf\limits_{n\to\infty} \frac{a_{n+1}}{a_n} > 1$, то $a_n \nrightarrow 0$ и ряд $\sum\limits_{n=1}^{\infty} a_n$ расходится.*

□ 1) Пусть $r \in (\overline{r},\, 1)$. Выберем N так, что $\sup\limits_{n\geqslant N} \frac{a_{n+1}}{a_n} < r$. Тогда $a_{n+1} < ra_n$ при всех $n \geqslant N$ и, значит, $a_n < ra_{n-1} < \ldots < r^{n-N}a_N$

при $n > N$. Следовательно, ряд сходится по признаку сравнения с геометрическим рядом $\sum r^n$.

2) Пусть $\underline{r} > 1$. Тогда найдется такой номер N, что $\inf\limits_{n\geqslant N} \frac{a_{n+1}}{a_n} > 1$. Поэтому $a_n > a_{n-1} > \ldots > a_N > 0$ при $n > N$. Следовательно, $a_n \nrightarrow 0$ и ряд расходится. ■

Замечание. Если в теореме 2.4 $q = 1$ или в теореме 2.5 $\overline{r} \geqslant 1$, $\underline{r} \leqslant 1$, то в общем случае нельзя сделать вывод о сходимости ряда $\sum a_n$.

Пример. Пусть $a_n = 1$ и $b_n = \frac{1}{n^2}$. Тогда для каждого из этих рядов $q = \overline{r} = \underline{r} = 1$, но ряд $\sum a_n$ расходится, а ряд $\sum b_n$ сходится.

Задача. Пусть $a_n > 0$ для всех $n \in \mathbb{N}$. Докажите неравенства

$$\varliminf_{n\to\infty} \frac{a_{n+1}}{a_n} \leqslant \varliminf_{n\to\infty} \sqrt[n]{a_n} \leqslant \varlimsup_{n\to\infty} \sqrt[n]{a_n} \leqslant \varlimsup_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

Из последней цепочки неравенств следует, что если к ряду применим признак Даламбера, то применим и признак Коши. Обратное, неверно.

Пример. Пусть $0 < \alpha < \beta < 1$. Рассмотрим ряд $\sum\limits_{n=1}^{\infty} a_n$, где

$$a_n = \begin{cases} \alpha^n, & \text{если } n \text{ четно,} \\ \beta^n, & \text{если } n \text{ нечетно.} \end{cases}$$

Тогда отношения $\frac{a_{n+1}}{a_n}$ равны $\frac{\alpha^{n+1}}{\beta^n}$ или $\frac{\beta^{n+1}}{\alpha^n}$, и корни $\sqrt[n]{a_n}$ равны α или β в зависимости от четности n. Следовательно, $\underline{r} = 0$, $\overline{r} = +\infty$ и $q = \beta$. Поэтому ряд $\sum a_n$ сходится по признаку Коши, однако к нему не применим признак Даламбера.

Теорема 2.6 (признак Гаусса). *Пусть $a_n > 0$ для всех $n \in \mathbb{N}$. Пусть существуют такие $s > 1$ и ограниченная последовательность $\{\alpha_n\}$, что при всех n*

$$\frac{a_{n+1}}{a_n} = 1 - \frac{A}{n} + \frac{\alpha_n}{n^s}.$$

Тогда ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится при $A > 1$ и расходится при $A \leqslant 1$.

□ При $n > 1$ имеем

$$a_n = a_1 \prod_{k=1}^{n-1} \frac{a_{k+1}}{a_k} = a_1 \prod_{k=1}^{n-1} \left(1 - \frac{A}{k} + \frac{\alpha_k}{k^s}\right).$$

Так как $\ln(1+x) = x + O(x^2)$ при $x \to 0$, то

$$\ln\left(1 - \frac{A}{k} + \frac{\alpha_k}{k^s}\right) = -\frac{A}{k} + \frac{\alpha_k}{k^s} + O\left(\frac{1}{k^2}\right).$$

Поэтому

$$a_n = a_1 \exp\left(\sum_{k=1}^{n-1} \left(-\frac{A}{k} + \frac{\alpha_k}{k^s} + O\left(\frac{1}{k^2}\right)\right)\right).$$

Пользуясь равенством $\sum_{k=1}^{n-1} \frac{1}{k} = \ln n + O(1)$ при $n \to \infty$ и сходимостью рядов $\sum_{k=1}^{\infty} \frac{1}{k^p}$ при $p > 1$, получаем

$$a_n = a_1 \exp(-A \ln n + O(1)) = \frac{a_1 e^{O(1)}}{n^A}, \quad n \to \infty.$$

Теперь утверждение следует по признаку сравнения с рядом $\sum_{n=1}^{\infty} \frac{1}{n^A}$. ■

Пример. Исследуем сходимость ряда $\sum\limits_{n=1}^{\infty} \left(\frac{(2n-1)!!}{(2n)!!}\right)^p$.

□ Обозначим n-й член ряда через a_n. Тогда при $n > 1$ имеем

$$\frac{a_n}{a_{n-1}} = \left(\frac{2n-1}{2n}\right)^p = \left(1 - \frac{1}{2n}\right)^p = 1 - \frac{p}{2n} + O\left(\frac{1}{n^2}\right).$$

Применяя признак Гаусса, получаем, что при $p > 2$ ряд сходится, а при $p \leqslant 2$ расходится. ■

2.3. Ряды с произвольными членами.

Вернемся теперь к изучению рядов с произвольными (в общем случае комплексными) членами.

Лемма 2.3.

1) *Если ряд* $\sum\limits_{n=1}^{\infty} b_n$ *сходится, то ряды* $\sum\limits_{n=1}^{\infty} a_n$ *и* $\sum\limits_{n=1}^{\infty} (a_n + b_n)$ *сходятся или расходятся одновременно.*

2) *Если ряд* $\sum\limits_{n=1}^{\infty} b_n$ *абсолютно сходится, то ряды* $\sum\limits_{n=1}^{\infty} a_n$ *и* $\sum\limits_{n=1}^{\infty} (a_n + b_n)$ *либо одновременно расходятся, либо одновременно сходятся условно, либо одновременно сходятся абсолютно.*

□ Первое утверждение непосредственно вытекает из свойства линейности. Так как для всех $n \in \mathbb{N}$

$$|a_n + b_n| \leqslant |a_n| + |b_n|, \qquad |a_n| \leqslant |a_n + b_n| + |b_n|,$$

то по признаку сравнения следует, что ряды $\sum a_n$ и $\sum (a_n + b_n)$ одновременно сходятся абсолютно. Теперь второе доказываемое утверждение следует из первого. ■

Приведем некоторые признаки сходимости рядов вида $\sum a_n b_n$.

Теорема 2.7 (признак Дирихле). *Пусть* $\{a_n\}$ *— комплексная,* $\{b_n\}$ *— действительная последовательности. Пусть*

1) *последовательность* $A_N = \sum\limits_{n=1}^{N} a_n$ *ограничена;*

2) *последовательность* $\{b_n\}$ *монотонна;*

3) $\lim\limits_{n\to\infty} b_n = 0$.

Тогда ряд $\sum\limits_{n=1}^{\infty} a_n b_n$ *сходится.*

Задача. Покажите, что теорема 2.7 для $a_n \in \mathbb{R}$ следует из теоремы 1.3 и леммы 2.1. Общий случай получается применением к действительной и мнимой частям $a_n b_n$.

В следующем параграфе признак будет доказан в более общей ситуации. Это относится и к следующему утверждению.

Теорема 2.8 (признак Абеля). *Пусть* $\{a_n\}$ *— комплексная,* $\{b_n\}$ *— действительная последовательности. Пусть*

1) *ряд* $\sum\limits_{n=1}^{\infty} a_n$ *сходится;*

2) *последовательность* $\{b_n\}$ *монотонна;*

3) *последовательность* $\{b_n\}$ *ограничена.*

Тогда ряд $\sum\limits_{n=1}^{\infty} a_n b_n$ *сходится.*

Задача. Получите теорему 2.8 как следствие теоремы 2.7.

При применении признака Дирихле на практике особые затруднения вызывает проверка первого условия. Полезно выделить частные случаи, в которых условие ограниченности частичных сумм ряда выполняется автоматически.

Следствие (признак Лейбница). *Если последовательность $\{\alpha_n\}$ монотонна и $\alpha_n \to 0$, то ряд $\sum\limits_{n=1}^{\infty} (-1)^{n-1}\alpha_n$ сходится, причем $|S - S_n| \leqslant |\alpha_{n+1}|$.*

$\square$ Хотя сходимость ряда следует по признаку Дирихле, докажем ее напрямую, по пути получив заявленное неравенство.

Пусть для определенности $\{\alpha_n\}$ нестрого убывает и поэтому $\alpha_n \geqslant 0$. Тогда $S_{2n+2} - S_{2n} = \alpha_{2n+1} - \alpha_{2n+2} \geqslant 0$ для каждого n и, значит, $\{S_{2n}\}$ нестрого возрастает; $S_{2n+1} - S_{2n-1} = -\alpha_{2n} + \alpha_{2n+1} \leqslant 0$ для каждого n и, значит, $\{S_{2n-1}\}$ нестрого убывает. Кроме того, $S_{2n} - S_{2n-1} = -\alpha_{2n} \leqslant 0$. Поэтому для всех m, n выполнено

$$S_{2n} \leqslant S_{2k} \leqslant S_{2k-1} \leqslant S_{2m-1},$$

где $k = \max\{m, n\}$. В частности, монотонные последовательности $\{S_{2n}\}$, $\{S_{2n-1}\}$ ограничены и, значит, сходятся, $S_{2n} \to S'$ и $S_{2n-1} \to S''$. Тогда $S_{2n} \leqslant S' \leqslant S'' \leqslant S_{2n-1}$ и поскольку $S_{2n} - S_{2n-1} = -\alpha_{2n} \to 0$, то $S' = S'' =: S$. Кроме того, $|S_n - S| \leqslant |S_{n+1} - S_n| = |\alpha_{n+1}|$. $\blacksquare$

Пример. Пусть $a_n = \dfrac{(-1)^n}{\sqrt{n}}$ и $b_n = (-1)^n \left(\dfrac{1}{\sqrt{n}} + \dfrac{(-1)^n}{n}\right)$. Тогда $a_n \sim b_n$ при $n \to \infty$, однако ряд $\sum a_n$ сходится (по признаку Лейбница), а ряд $\sum b_n$ расходится (по п. 1 леммы 2.3).

Это доказывает, что признак сравнения для общих рядов не применим.

Следствие. *Если последовательность $\{\alpha_n\}$ монотонна и $\alpha_n \to 0$, $x \neq 2\pi m$ ($m \in \mathbb{Z}$), то ряды $\sum\limits_{n=1}^{\infty} \alpha_n \sin nx$ и $\sum\limits_{n=1}^{\infty} \alpha_n \cos nx$ сходятся.*

$\square$ Положим $s_N = \sum_{n=1}^{N} e^{inx}$. Рассматривая s_N как геометрическую прогрессию с коэффициентом $q = e^{ix}$, получим $s_N = \dfrac{e^{ix}(1 - e^{iNx})}{1 - e^{ix}}$. Так как $|e^{ix}| = 1$, то $|s_N| \leqslant \dfrac{2}{|1 - e^{ix}|} = \dfrac{1}{|\sin(x/2)|}$.

Для проверки ограниченности сумм $C_N = \sum_{n=1}^{N} \cos nx$ и $S_N = \sum_{n=1}^{N} \sin nx$ осталось заметить, что $C_N = \operatorname{Re} s_N$, а $S_N = \operatorname{Im} s_N$. Следовательно, данные ряды сходятся по признаку Дирихле. ■

Пример. Ряд $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ сходится по признаку Лейбница. Найдем его сумму.

Частичную сумму S_{2m} можно представить в виде разности частичных сумм гармонического ряда:

$$S_{2m} = 1 + \frac{1}{2} + \ldots + \frac{1}{2m-1} + \frac{1}{2m} - 2\left(\frac{1}{2} + \ldots + \frac{1}{2m}\right) = H_{2m} - H_m.$$

Осталось воспользоваться асимптотикой $H_n = \ln n + \gamma + o(1)$. В итоге $S_{2m} = (\ln 2m + \gamma + o(1)) - (\ln m + \gamma + o(1)) = \ln 2 + o(1)$. Следовательно, сумма ряда равна $\ln 2$.

2.4. Перестановки рядов.

Определение. Пусть дан ряд $\sum\limits_{n=1}^{\infty} a_n$ и биекция $\varphi\colon \mathbb{N} \to \mathbb{N}$, тогда ряд $\sum\limits_{n=1}^{\infty} a_{\varphi(n)}$ называется *перестановкой* ряда $\sum\limits_{n=1}^{\infty} a_n$.

Пример. Рассмотрим перестановку ряда $\sum\limits_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$:

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \ldots$$

(после одного положительного члена идут два отрицательных). Найдем ее сумму, сгруппировав положительный член с последующим отрицательным:

$$S_{\Pi} = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \cdots = \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots\right) = \frac{1}{2}\ln 2.$$

Такая ситуация невозможна для абсолютно сходящихся рядов.

Теорема 2.9. *Если ряд* $\sum\limits_{n=1}^{\infty} a_n$ *сходится абсолютно, то любая его перестановка* $\sum\limits_{n=1}^{\infty} a_{\varphi(n)}$ *сходится абсолютно к той же сумме.*

□ Абсолютная сходимость перестановки следует по лемме 2.2:

$$\sum_{n=1}^{N} |a_{\varphi(n)}| \leqslant \sum_{n=1}^{\max\limits_{1\leqslant j\leqslant N}\varphi(j)} |a_n| \leqslant \sum_{n=1}^{\infty} |a_n| < \infty.$$

Зафиксируем $\varepsilon > 0$ и найдем такой номер m, что $\sum\limits_{n=m+1}^{\infty} |a_n| < < \varepsilon$. Далее выберем номер M так, чтобы $\{1, \ldots, m\} \subset \{\varphi(1), \ldots, \varphi(M)\}$ (достаточно положить $M = \max\limits_{1\leqslant j\leqslant m} \varphi^{-1}(j)$).

Тогда для любого $N \geqslant M$ выполнено $\{1, \ldots, m\} \subset \{\varphi(1), \ldots, \varphi(N)\}$ и, значит,

$$\left|\sum_{n=1}^{\infty} a_n - \sum_{n=1}^{N} a_{\varphi(n)}\right| \leqslant \sum_{n=m+1}^{\infty} |a_n| < \varepsilon.$$

Это доказывает, что сумма ряда является пределом последовательности частичных сумм его перестановки. ■

Замечание. Также проверяется, что если $a_n \to 0$, то $a_{\varphi(n)} \to 0$.

Изучим перестановки условно сходящихся рядов. Для простоты ограничимся случаем действительных рядов.

Пусть ряд $\sum\limits_{n=1}^{\infty} a_n$ сходится условно. Пусть $\{n\colon a_n > 0\} = \{p_k\}$, $p_1 < p_2 < \ldots$ и $\{n\colon a_n \leqslant 0\} = \{q_k\}$, $q_1 < q_2 < \ldots$. Отметим, что эти множества бесконечны, иначе все члены ряда (начиная с некоторого номера) будут одного знака, а для таких рядов сходимость совпадает с абсолютной сходимостью.

Определим ряды $\sum_{k=1}^{\infty} b_k$, где $b_k = a_{p_k}$, и $\sum_{k=1}^{\infty} c_k$, где $c_k = a_{q_k}$.

Лемма 2.4. *Ряды* $\sum\limits_{k=1}^{\infty} b_k$ *и* $\sum\limits_{k=1}^{\infty} c_k$ *расходятся.*

□ Предположим, что хотя бы один из рядов сходится. Тогда в силу равенства $\sum\limits_{k=1}^{n} a_k = \sum\limits_{k=1}^{r_n} b_k + \sum\limits_{k=1}^{s_n} c_k$, $n = r_n + s_n$, сходится и другой ряд. Значит, в силу равенства $\sum\limits_{k=1}^{n} |a_k| = \sum\limits_{k=1}^{r_n} b_k + \sum\limits_{k=1}^{s_n} (-c_k)$ сходится ряд из $|a_k|$, что противоречит условной сходимости. ■

Теорема 2.10* (Риман). *Если ряд* $\sum\limits_{n=1}^{\infty} a_n$ *с действительны-*

ми членами сходится условно, то для любого $L \in \mathbb{R}$ существует такая перестановка $\sum\limits_{n=1}^{\infty} a_{\varphi(n)}$, что ее сумма равна L.

$\square$ По индукции определим числа a_k^* и суммы $S_n^* = \sum_{k=1}^n a_k^*$.

Пусть n_1 — наименьший из номеров N, что $b_1 + \ldots + b_N > L$. Номер n_1 найдется, т.к. ряд $\sum b_k$ расходится. Положим $a_k^* = b_k$ при $k \leqslant n_1$. Отметим, что если $n_1 > 1$, то $L < S_{n_1}^* \leqslant L + a_{n_1}^*$. Пусть m_1 — наименьший из номеров $M > n_1$, что $S_{n_1}^* + c_1 + \ldots + c_{M-n_1} < L$. Номер m_1 найдется, т.к. ряд $\sum c_k$ расходится. Положим $a_{n_1+k}^* = c_k$ при $k \leqslant m_1 - n_1$. Тогда $L + a_{m_1}^* \leqslant S_{m_1}^* < L$. Пусть n_2 — наименьший из номеров $N > m_1$, что $S_{m_1}^* + b_{n_1+1} + \ldots + b_N > L$. Определим $a_{m_1+k}^* = b_{n_1+k}$ при $k \leqslant n_2 - m_1$ и т.д. Остатки рядов $\sum b_k$ и $\sum c_k$ расходятся, поэтому по индукции будут построены последовательности $n_1 < m_1 < n_2 < m_2 < \ldots$ и $\{a_k^*\}$. Для $j \geqslant 2$ имеем $L < S_{n_j}^* \leqslant L + a_{n_j}^*$, и $L + a_{m_j}^* \leqslant S_{m_j}^* < L$.

По построению $\sum_{k=1}^{\infty} a_k^*$ является перестановкой исходного ряда. По замечанию $a_k^* \to 0$, так что $a_{n_j}^* \to 0$ и $a_{m_j}^* \to 0$ как подпоследовательности. Покажем, что сумма $\sum a_k^*$ равна L.

Пусть $n_j \leqslant n < m_j$ для некоторого j, тогда $S_{m_j}^* \leqslant S_n^* \leqslant S_{n_j}^*$ и, значит,

$$a_{m_j}^* \leqslant S_{m_j}^* - L \leqslant S_n^* - L \leqslant S_{n_j}^* - L \leqslant a_{n_j}^*.$$

Аналогично, если $m_j \leqslant n < n_{j+1}$, то $a_{m_j}^* \leqslant S_n^* - L < a_{n_{j+1}}^*$. Зафиксируем $\varepsilon > 0$ и найдем такой номер $j_0 \geqslant 2$, что $|a_{n_j}^*| < \varepsilon$ и $|a_{m_j}^*| < \varepsilon$ при всех $j \geqslant j_0$. Тогда $|S_n^* - L| < \varepsilon$ при всех $n \geqslant n_{j_0}$. $\blacksquare$

Задача. Докажите, что теорема 2.10 верна и при $L = \pm\infty$.

Замечание. Аналогичный результат в $\mathbb{C}$ (и даже в $\mathbb{R}^n$) утверждает, что если множество сумм ряда при различных перестановках содержит две точки, то оно содержит и проходящую через них прямую (теорема Леви–Штейница).

2.5. Произведение рядов.

Теорема 2.11 (Коши). *Если ряды $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=1}^{\infty} b_n$ сходятся абсолютно к A и B соответственно, то ряд $\sum\limits_{n=1}^{\infty} a_{i_n} b_{j_n}$ из всевозможных попарных произведений, занумерованных в произволь-*

ном порядке (т.е. $\varphi\colon \mathbb{N} \to \mathbb{N}^2$*,* $\varphi(n) = (i_n, j_n)$*, биекция), сходится абсолютно к* AB*.*

$\square$ Установим абсолютную сходимость ряда из произведений:

$$\sum_{n=1}^{N} |a_{i_n} b_{j_n}| \leqslant \sum_{i=1}^{\max\limits_{1\leqslant k\leqslant N} i_k} \sum_{j=1}^{\max\limits_{1\leqslant k\leqslant N} j_k} |a_i b_j| = \left(\sum_{i=1}^{\max\limits_{1\leqslant k\leqslant N} i_k} |a_i|\right)\left(\sum_{j=1}^{\max\limits_{1\leqslant k\leqslant N} j_k} |b_j|\right)$$

и, значит, $\sum\limits_{n=1}^{N} |a_{i_n} b_{j_n}| \leqslant \left(\sum\limits_{i=1}^{\infty} |a_i|\right)\left(\sum\limits_{j=1}^{\infty} |b_j|\right)$.

По теореме 2.9 любая перестановка $\sum a_{i_n} b_{j_n}$ сходится к той же сумме.

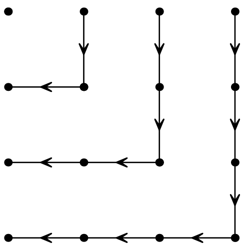

Рис. 2.1*а*

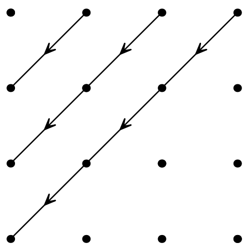

Рис. 2.1*б*

Рассмотрим перестановку «по квадратам» и ее частичные суммы $S_{N^2} = \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} a_i b_j$. Тогда $S_{N^2} = \Big(\sum\limits_{i=1}^{N} a_i\Big)\Big(\sum\limits_{j=1}^{N} b_j\Big) \to AB$. Так как в случае существования предел последовательности совпадает с пределом подпоследовательности, то сумма перестановки «по квадратам» (см. рис. 2.1*а*), а значит, и сумма $\sum\limits_{n=1}^{\infty} a_{i_n} b_{j_n}$, равна AB. $\blacksquare$

Чаще всего при определении произведения рядов используют

Определение. Ряд $\sum\limits_{n=1}^{\infty} c_n$, где $c_n = \sum\limits_{k=1}^{n} a_k b_{n+1-k}$, называется *произведением* рядов $\sum\limits_{n=1}^{\infty} a_n$ и $\sum\limits_{n=1}^{\infty} b_n$ *по Коши.*

В развернутом виде

$$(a_1 + a_2 + a_3 + \ldots)(b_1 + b_2 + b_3 + \ldots) = \\ = a_1b_1 + (a_1b_2 + a_2b_1) + (a_1b_3 + a_2b_2 + a_3b_1) + \ldots$$

т.е. в правой части рассматривается перестановка «по диагоналям» (см. рис. 2.1*б*) и группировка слагаемых, стоящих на одной диагонали. Следовательно, справедливо

Следствие. *Если ряды $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} b_n$ сходятся абсолютно, то их произведение по Коши сходится к произведению сумм этих рядов.*

Замечание. Произведение по Коши условно сходящихся рядов может дать расходящийся ряд.

Пример. Ряд $\sum\limits_{n=1}^{\infty} \dfrac{(-1)^{n-1}}{\sqrt{n}}$ сходится по признаку Лейбница, но не абсолютно. Рассмотрим его квадрат Коши, т.е. ряд с членами

$$c_n = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{\sqrt{k}} \frac{(-1)^{n-k}}{\sqrt{n+1-k}} = (-1)^{n-1} \sum_{k=1}^{n} \frac{1}{\sqrt{k(n+1-k)}}.$$

Поскольку $|c_n| \geqslant \sum\limits_{k=1}^{n} \dfrac{1}{\sqrt{n}\sqrt{n}} = 1$, ряд $\sum\limits_{n=1}^{\infty} c_n$ расходится.

Отметим, что если оба ряда сходятся и их произведение по Коши сходится, то сумма произведения совпадает с произведением сумм рядов. Это может быть доказано с использованием степенных рядов.

2.6. Суммируемые семейства.

Рассмотрим семейство $\{a_j\}_{j \in J}$ в $\mathbb{R}$, индексированное множеством J (возможно, несчетным), т.е. функцию $a\colon J \to \mathbb{R}$, $a_j = a(j)$. Определим сумму $\sum_{j \in J} a_j$. Обозначим через $\mathcal{F}(J)$ множество всех конечных подмножеств J.

Определение. Говорят, что семейство $\{a_j\}_{j \in J}$ *суммируемо* к числу s и пишут $\sum_{j \in J} a_j = s$, если для любого $\varepsilon > 0$ найдется такое $F_0 \in \mathcal{F}(J)$, что $\left|\sum\limits_{j \in F} a_j - s\right| < \varepsilon$ для всякого множества $F \in \mathcal{F}(J)$, $F \supset F_0$.

C1 (единственность). *Если семейство $\{a_j\}_{j\in J}$ суммируемо к числам s и t, то $t = s$.*

$\square$ Зафиксируем $\varepsilon > 0$. Существуют такие конечные F_1, F_2, что $\left|\sum_{j\in F} a_j - s\right| < \varepsilon$ для всех $F \in \mathcal{F}(J)$, $F \supset F_1$ и $\left|\sum_{j\in F} a_j - t\right| < \varepsilon$ для всех $F \in \mathcal{F}(J)$, $F \supset F_2$. Определим $F_0 = F_1 \cup F_2$. Если $F \in \mathcal{F}(J)$ содержит F_0, то оно содержит и F_1, и F_2. Поэтому

$$|t - s| \leqslant \left|\sum_{j\in F} a_j - t\right| + \left|\sum_{j\in F} a_j - s\right| < 2\varepsilon.$$

Так как $\varepsilon > 0$ — любое, то $t = s$. $\blacksquare$

Аналогично доказывается

C2 (линейность). *Если семейства $\{a_j\}_{j\in J}$, $\{b_j\}_{j\in J}$ суммируемы и α, $\beta \in \mathbb{R}$, то семейство $\{\alpha a_j + \beta b_j\}_{j\in J}$ также суммируемо, причем $\sum_{j\in J}(\alpha a_j + \beta b_j) = \alpha \sum_{j\in J} a_j + \beta \sum_{j\in J} b_j$.*

Для семеств с неотрицательными членами расширим понятие суммы: будем писать $\sum_{j\in J} a_j = +\infty$, если для любого $M \in \mathbb{R}$ найдется такое $F_0 \in \mathcal{F}(J)$, что $\sum_{j\in F} a_j > M$ для всякого множества $F \in \mathcal{F}(J)$, $F \supset F_0$.

Лемма 2.5. *Если $a_j \geqslant 0$ для всех $j \in J$, то*

$$\sum_{j\in J} a_j = \sup_{F\in\mathcal{F}(J)} \sum_{j\in F} a_j.$$

$\square$ Обозначим правую часть через s. Пусть $\varepsilon > 0$. По определению супремума найдется $F_0 \in \mathcal{F}(J)$, что $\sum_{j\in F_0} a_j > s - \varepsilon$. Тогда в силу неотрицательности a_j выполнено $s - \varepsilon < \sum_{j\in F_0} a_j \leqslant \sum_{j\in F} a_j \leqslant s$ для всякого конечного множества F, содержащего F_0. Это доказывает, что $\sum_{j\in J} a_j = s$.

Случай $s = +\infty$ рассматривается аналогично. $\blacksquare$

Пример. Исследуем суммируемость $\left\{\dfrac{1}{(m+n)^p}\right\}_{m,\,n\in\mathbb{N}}$.

Заметим, что если $T = \{(m,\,n)\colon 2 \leqslant m+n \leqslant N\}$, то

$$\sum_{T} \frac{1}{(m+n)^p} = \sum_{k=2}^{N}\sum_{n=1}^{k-1} \frac{1}{k^p} = \sum_{k=2}^{N} \frac{k-1}{k^p}.$$

При $p \leqslant 2$ частичные суммы ряда $\sum \frac{k-1}{k^p}$ неограничены и, значит, семейство несуммируемо.

Если $p > 2$ и $F \subset \mathbb{N} \times \mathbb{N}$ конечно, то существует $T \supset F$ и

$$\sum_F \frac{1}{(m+n)^p} \leqslant \sum_T \frac{1}{(m+n)^p} < \sum_{k=2}^{\infty} \frac{k-1}{k^p} < \infty.$$

Это означает, что семейство суммируемо.

Для $x \in \mathbb{R}$ положим $x^+ = \max\{x, 0\}$ и $x^- = \max\{-x, 0\}$. Тогда $x = x^+ - x^-$, $|x| = x^+ + x^-$.

Теорема 2.12. *Следующие утверждения эквивалентны:*

1) семейство $\{a_j\}_{j\in J}$ суммируемо;

2) частичные суммы ограничены, т.е. $\sup\limits_{F\in\mathcal{F}(J)} \left|\sum_{j\in F} a_j\right| < \infty$*;*

3) семейство $\{|a_j|\}_{j\in J}$ суммируемо.

$\square$ $(1 \Rightarrow 2)$ Пусть $\{a_j\}_{j\in J}$ суммируемо к s. По определению существует такое $F_0 \in \mathcal{F}(J)$, что $\left|\sum_{j\in E} a_j - s\right| < 1$ для всех $E \in \mathcal{F}(J)$, $E \supset F_0$ и, значит, $\left|\sum_{j\in E} a_j\right| < |s| + 1$. Для произвольного $F \in \mathcal{F}(J)$ имеем $\sum_{j\in F} a_j = \sum_{j\in F\cup F_0} a_j - \sum_{j\in F_0\setminus F} a_j$ и, значит,

$$\left|\sum_{j\in F} a_j\right| \leqslant \left|\sum_{j\in F\cup F_0} a_j\right| + \sum_{j\in F_0\setminus F} |a_j| \leqslant 1 + |s| + \sum_{j\in F_0} |a_j|.$$

$(2 \Rightarrow 1)$ Если $F \in \mathcal{F}(J)$, то

$$\sum_{j\in F} a_j^+ = \sum_{j\in F^+} a_j,$$

где $F^+ = \{j \in F\colon a_j \geqslant 0\}$. Следовательно, ограниченность частичных сумм $\{a_j\}_{j\in J}$ влечет ограниченность частичных сумм $\{a_j^+\}_{j\in J}$, а значит, по лемме 2.5 его суммируемость. Аналогично для $\{a_j^-\}_{j\in J}$. Так как $a_j = a_j^+ - a_j^-$, то по линейности получаем суммируемость $\{a_j\}_{j\in J}$.

Так как $|a_j| = a_j^+ + a_j^-$, то также установили, что ограниченность частичных сумм $\{a_j\}_{j\in J}$ влечет ограниченность частичных сумм $\{|a_j|\}_{j\in J}$. Очевидно верно и в обратную сторону. Следовательно, $(2 \Leftrightarrow 3)$. $\blacksquare$

Следствие. *Если семейство $\{a_j\}_{j\in J}$ суммируемо, то $S = \{j \in J\colon a_j \neq 0\}$ не более чем счетно.*

□ По теореме 2.12 $\sum_{j\in J}|a_j| = M < \infty$. Положим $S_n = \{j \in \in J\colon |a_j| > 1/n\}$. Тогда S_n конечно (количество $\leqslant nM$). Следовательно, $S = \bigcup_{n=1}^{\infty} S_n$ не более чем счетно. ■

Теорема 2.13. *Пусть* $\{J_n\}_{n=1}^{\infty}$ *— разбиение* J, *т.е.* $J = \bigcup_{n=1}^{\infty} J_n$ *и* $J_k \cap J_i = \emptyset$ *при* $i \neq k$. *Семейство* $\{a_j\}_{j\in J}$ *суммируемо тогда и только тогда, когда*

$$\sum_{n=1}^{\infty}\sum_{j\in J_n}|a_j| < \infty. \tag{2.2}$$

В этом случае

$$\sum_{j\in J} a_j = \sum_{n=1}^{\infty}\sum_{j\in J_n} a_j. \tag{2.3}$$

□ Пусть выполнено (2.2). Если $F \in \mathcal{F}(J)$, то существует такое N, что $F \subset \bigcup_{k=1}^{N} J_k$. Поэтому

$$\sum_{j\in F}|a_j| = \sum_{n=1}^{N}\sum_{j\in F\cap J_n}|a_j| \leqslant \sum_{n=1}^{N}\sum_{j\in J_n}|a_j| \leqslant \sum_{n=1}^{\infty}\sum_{j\in J_n}|a_j|,$$

так что $\{|a_j|\}_{j\in J}$ суммируемо.

Пусть $\{|a_j|\}_{j\in J}$ суммируемо. Тогда $\sum_{j\in J}|a_j| < \infty$, а значит, $\sum_{j\in J_n}|a_j| < \infty$ для всех n. Пусть $\varepsilon > 0$. Тогда по лемме 2.5 и определению супремума для всякого n найдется конечное множество $F_n \subset J_n$, что $\sum_{j\in F_n}|a_j| > \sum_{j\in J_n}|a_j| - \varepsilon 2^{-n}$. Зафиксируем N и пусть $F = \bigcup_{n=1}^{N} F_n$. Тогда

$$\sum_{n=1}^{N}\sum_{j\in J_n}|a_j| < \sum_{n=1}^{N}\left(\sum_{j\in F_n}|a_j| + \frac{\varepsilon}{2^n}\right) = \sum_{j\in F}|a_j| + \sum_{n=1}^{N}\frac{\varepsilon}{2^n} \leqslant \sum_{j\in J}|a_j| + \varepsilon.$$

Неравенство верно для всех N, поэтому $\sum_{n=1}^{\infty}\sum_{j\in J_n}|a_j| \leqslant \leqslant \sum_{j\in J}|a_j| + \varepsilon$ и формула (2.2) установлена. Так как $\varepsilon > 0$ — любое, то (2.3) выполняется при условии $a_j \geqslant 0$. Применяя (2.3) к $a_j^{\pm}$, заключаем, что (2.3) верно в общем случае. ■

Следствие. *Семейство* $\{a_n\}_{n\in\mathbb{N}}$ *суммируемо тогда и только тогда, когда* $\sum\limits_{n=1}^{\infty}|a_n|$ *сходится. В этом случае* $\sum\limits_{n\in\mathbb{N}} a_n = \sum\limits_{n=1}^{\infty} a_n$.

Следствие (Фубини для сумм). *Пусть* $\{a_{mn}\}_{m,\,n\in\mathbb{N}}$. *Тогда*

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} a_{mn} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn}$$

при условии, что хотя бы один из рядов сходится при замене a_{mn} *на* $|a_{mn}|$.

□ Из условия следует суммируемость $\{|a_{mn}|\}_{m,\,n\in\mathbb{N}}$. Применяя теорему 2.13 к разбиению $\mathbb{N}\times\mathbb{N}$ по строкам и по столбцам, получаем, что обе повторные суммы равны $\sum\limits_{(m,n)\in\mathbb{N}\times\mathbb{N}} a_{mn}$, а значит, равны между собой. ■

Замечание. Поскольку любое конечное $F\subset\mathbb{N}\times\mathbb{N}$ можно включить в квадрат, то в условиях следствия общее значение сумм вычисляется как $\lim\limits_{N\to\infty}\sum\limits_{n=1}^{N}\sum\limits_{m=1}^{N} a_{nm}$.

3. Функциональные ряды

В этом параграфе изучаются последовательности и ряды, членами которых являются функции. Будем предполагать, что функции определены на одном и том же *непустом* множестве E.

3.1. Равномерная сходимость.

Пусть заданы функции f, $f_n\colon E \to \mathbb{R}$ (или $\mathbb{C}$), $n \in \mathbb{N}$.

Определение. Последовательность $\{f_n\}$ *поточечно сходится* к функции f на E, если $f(x) = \lim\limits_{n\to\infty} f_n(x)$ для любого $x \in E$. Пишут $f_n \to f$ на E и f называют *предельной функцией* последовательности $\{f_n\}$.

Задача. Докажите, что если $f_n \to f$ на $E = \mathbb{R}$, и все функции f_n четные (нечетные, нестрого возрастают, нестрого убывают), то функция f также четная (нечетная, нестрого возрастает, нестрого убывает) на E.

Расшифруем определение предела:

$$f_n \to f \text{ на } E \Leftrightarrow \forall x \in E\ \forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geqslant N\ (|f_n(x) - f(x)| < \varepsilon).$$

При этом номер N может (при фиксированном ε) зависеть от x.

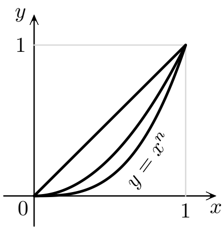

Рис. 3.1

Пример. Пусть $E = [0,\,1]$, $f_n(x) = x^n$. Тогда $f_n \to f$ на $[0,\,1]$, где

$$f(x) = \begin{cases} 0, & \text{если } x \in [0,\,1); \\ 1, & \text{если } x = 1. \end{cases}$$

При фиксированном $\varepsilon \in (0,\,1)$ выполнено $N(x) > \dfrac{\ln\varepsilon}{\ln x}$ и, значит, $N(x) \to \infty$ при $x \to 1$. (см. рис. 3.1)

Если N удается выбрать по ε одним для всех $x \in E$, то приходим к следующему понятию.

Определение. Последовательность $\{f_n\}$ *равномерно сходится* к функции f на множестве E, если

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geqslant N\ \forall x \in E\ (|f_n(x) - f(x)| < \varepsilon).$$

Пишут $f_n \rightrightarrows f$ на E или $f_n \underset{E}{\rightrightarrows} f$ при $n \to \infty$.

Замечание. Из определений очевидно, что равномерная сходимость влечет поточечную. В частности, если $f_n \rightrightarrows f$ на E, то функция f определена (на E) однозначно. Как показывает предыдущий пример, поточечная сходимость в общем случае не влечет равномерную.

Полезную переформулировку определения равномерной сходимости дает

Лемма 3.1 (супремум-критерий).
$f_n \rightrightarrows f$ на $E \Leftrightarrow \lim\limits_{n\to\infty} \rho_n = 0$, где $\rho_n = \sup\limits_{x\in E} |f_n(x) - f(x)|$.

$\square$ Вытекает из того, что высказывание $\forall x \in E\ (|f_n(x) - f(x)| \leqslant \varepsilon)$ равносильно высказыванию $\sup\limits_{x\in E}|f_n(x) - f(x)| \leqslant \varepsilon$. $\blacksquare$

Итак, равномерная сходимость $\{f_n\}$ означает, что графики функций f_n «прижимаются» с ростом n к графику f: уклонение ρ_n стремится к нулю при $n \to \infty$. (см. рис. 3.2)

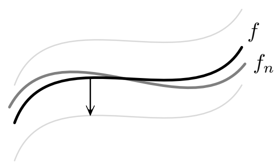

Рис. 3.2

Задача. Пусть $f_n \rightrightarrows f$ на E. Покажите, что $\lim\limits_{n\to\infty} |f_n(x_n) - f(x_n)| = 0$ для любой последовательности $\{x_n\}$ из E.

Определение. Функциональная последовательность *равномерно* (*поточечно*) *сходится* на E, если существует определенная на E функция, к которой последовательность равномерно (поточечно) сходится на E.

Рассмотрим *функциональный ряд* $\sum\limits_{n=1}^{\infty} f_n$, $f_n \colon E \to \mathbb{R}$ (или $\mathbb{C}$).

Определения. Функциональный ряд $\sum\limits_{n=1}^{\infty} f_n$ *сходится* на E, если для каждого $x \in E$ числовой ряд $\sum\limits_{n=1}^{\infty} f_n(x)$ сходится. При этом функция $S \colon x \mapsto \sum\limits_{n=1}^{\infty} f_n(x)$, называется *суммой* ряда.

Функциональный ряд $\sum\limits_{n=1}^{\infty} f_n$ *равномерно* (*поточечно*) *сходит-*

ся на E, если последовательность его частичных сумм $S_N = \sum_{n=1}^{N} f_n$ равномерно (поточечно) сходится на E.

Перечислим основные свойства равномерной сходимости.

Свойство 1. а) Если $f_n \rightrightarrows f$ на E, $g_n \rightrightarrows g$ на E, и α, $\beta \in \mathbb{R}$ (или $\mathbb{C}$), то $\alpha f_n + \beta g_n \rightrightarrows \alpha f + \beta g$ на E.

б) Если ряды $\sum_{n=1}^{\infty} f_n$ и $\sum_{n=1}^{\infty} g_n$ равномерно сходятся на E, то ряд $\sum_{n=1}^{\infty} (\alpha f_n + \beta g_n)$ также равномерно сходится на E, причем

$$\sum_{n=1}^{\infty} (\alpha f_n + \beta g_n) = \alpha \sum_{n=1}^{\infty} f_n + \beta \sum_{n=1}^{\infty} g_n.$$

$\square$ а) Пусть $x \in E$. Из неравенства треугольника следует, что

$$|\alpha f_n(x) + \beta g_n(x) - (\alpha f(x) + \beta g(x))| \leqslant \\ \leqslant |\alpha||f_n(x) - f(x)| + |\beta||g_n(x) - g(x)|.$$

Переходя в полученном неравенстве к супремуму по всем $x \in E$ сначала в правой части, затем в левой, имеем

$$\sup_{x\in E} |\alpha f_n(x) + \beta g_n(x) - (\alpha f(x) + \beta g(x))| \leqslant \\ \leqslant |\alpha| \sup_{x\in E} |f_n(x) - f(x)| + |\beta| \sup_{x\in E} |g_n(x) - g(x)|.$$

По супремум-критерию получаем требуемое утверждение.

б) Применим предыдущий пункт к последовательности частичных сумм ряда. $\blacksquare$

Следствие. *Если ряд* $\sum_{n=1}^{\infty} f_n$ *равномерно сходится на* E, *то* $f_n \rightrightarrows 0$ *на* E.

$\square$ Поскольку f_n можно записать в виде разности частичных сумм, $f_n = S_n - S_{n-1}$ (считаем $S_0 = 0$), то утверждение вытекает из определения равномерной сходимости ряда и свойства 1. $\blacksquare$

Свойство 2. Пусть функция $g\colon E \to \mathbb{R}$ (или $\mathbb{C}$) ограничена.

а) Если $f_n \rightrightarrows f$ на E, то $gf_n \rightrightarrows gf$ на E.

б) Если ряд $\sum_{n=1}^{\infty} f_n$ равномерно сходится на E, то ряд $\sum_{n=1}^{\infty} gf_n$ также равномерно сходится на E, причем $\sum_{n=1}^{\infty} gf_n = g \sum_{n=1}^{\infty} f_n$.

□ а) Пусть $|g| \leqslant M$ на E. Тогда справедливо неравенство:

$$\sup_{x\in E}|g(x)f_n(x) - g(x)f(x)| \leqslant M\sup_{x\in E}|f_n(x) - f(x)|.$$

По супремум-критерию получаем требуемое утверждение.

б) Применим предыдущий пункт к последовательности частичных сумм ряда. ■

Задача. Пусть $f_n \rightrightarrows f$ на E, и $g\colon D \to E$. Покажите, что $f_n \circ g \rightrightarrows f \circ g$ на D (замена переменной сохраняет равномерную сходимость).

Теорема 3.1 (критерий Коши равномерной сходимости). *Для равномерной сходимости последовательности $\{f_n\}$ на E необходимо и достаточно выполнения условия Коши:*

$$\forall\varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n,\, m \geqslant N\ \forall x \in E\ (|f_n(x) - f_m(x)| < \varepsilon). \tag{3.1}$$

□ ($\Rightarrow$) Пусть $\varepsilon > 0$. Найдем такой номер N, что $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$ для всех $n \geqslant N$ и $x \in E$. Тогда при $n,\, m \geqslant N$

$$|f_n(x) - f_m(x)| \leqslant |f_n(x) - f(x)| + |f_m(x) - f(x)| < \varepsilon.$$

($\Leftarrow$) Пусть $\{f_n\}$ удовлетворяет условию (3.1). Тогда для каждого $x \in E$ последовательность $\{f_n(x)\}$ является фундаментальной, а значит, сходится. Положим $f(x) = \lim\limits_{n\to\infty} f_n(x)$, $x \in E$. Выберем $\varepsilon > 0$ и найдем соответствующий номер N из условия (3.1). Переходя в неравенстве (3.1) к пределу при $m \to \infty$, получаем, что $|f_n(x) - f(x)| \leqslant \varepsilon$ при всех $n \geqslant N$ и $x \in E$. Это означает, что $f_n \rightrightarrows f$ на E. ■

Следствие (критерий Коши для рядов). *Для равномерной сходимости ряда $\sum\limits_{n=1}^{\infty} f_n$ необходимо и достаточно выполнения условия Коши:*

$$\forall\varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n,\, m \geqslant N\ \forall x \in E\ \left(\left|\sum_{k=m}^{n} f_k(x)\right| < \varepsilon\right).$$

Следствие. *Пусть $E \subset \mathbb{R}$ и все функции f_n непрерывны на замыкании $\overline{E}$. Если $\{f_n\}$ сходится равномерно на E, то $\{f_n\}$ сходится равномерно на $\overline{E}$.*

□ Выберем $\varepsilon > 0$ и найдем номер N, что $|f_n(x) - f_m(x)| \leqslant \varepsilon$ при всех $n, m \geqslant N$ и $x \in E$. Если $x_0 \in \overline{E}$, то найдется последовательность $\{x_k\}$ точек из E, сходящаяся к x_0. Тогда переходя к пределу при $k \to \infty$ в неравенстве $|f_n(x_k) - f_m(x_k)| \leqslant \varepsilon$ и пользуясь непрерывностью f_n, f_m, получаем $|f_n(x_0) - f_m(x_0)| \leqslant \varepsilon$. По критерию Коши $\{f_n\}$ сходится равномерно на $\overline{E}$. ■

Равномерная сходимость позволяет перебрасывать некоторые свойства приближающих функций на приближаемую (предельную). Приведем соответствующие теоремы о сохранении непрерывности, дифференцируемости и интегрируемости.

Теорема 3.2 (о непрерывности предельной функции). *Пусть $f_n\colon E \to \mathbb{R}$ непрерывна в точке $a \in E$ (на E) для каждого n. Если $f_n \rightrightarrows f$ на E, то f также непрерывна в точке a (на E).*

□ Зафиксируем $\varepsilon > 0$. Из условия равномерной сходимости найдем такой номер N, что $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$ для всех $x \in E$ и $n \geqslant N$. Тогда для $x \in E$ имеем

$$|f(x) - f(a)| \leqslant |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| < |f_N(x) - f_N(a)| + \frac{2\varepsilon}{3}.$$

Так как функция f_N непрерывна в точке a, то существует такое $\delta > 0$, что $|f_N(x) - f_N(a)| < \frac{\varepsilon}{3}$ для всех $x \in B_\delta(a) \cap E$. Но тогда $|f(x) - f(a)| < \varepsilon$ для всех $x \in B_\delta(a) \cap E$. Это означает, что функция f непрерывна в точке a. ■

Замечание. Если a — предельная точка E, то в условиях теоремы 3.2 $\lim\limits_{x\to a} \lim\limits_{n\to\infty} f_n(x) = \lim\limits_{n\to\infty} \lim\limits_{x\to a} f_n(x)$.

Следствие (о непрерывности суммы ряда). *Если ряд $\sum\limits_{n=1}^{\infty} f_n$ равномерно сходится на $E \subset \mathbb{R}$, и все f_n непрерывны в точке $a \in E$ (на E), то сумма ряда непрерывна в точке a (на E).*

В дальнейшем понятие непрерывности будет расширено на отображения метрических пространств. Теорема 3.2 и ее следствие справедливы в этой более общей ситуации.

Пример. Рассмотрим на $[0, 1]$ непрерывные функции $f_n(x) = n^\alpha x^n$. Выясним, при каких $\alpha \in \mathbb{R}$, последовательность $\{f_n\}$

сходится равномерно к функции $f_0 \equiv 0$:

$$\sup_{[0,\,1]} |f_n(x)| = n^\alpha \Rightarrow \big(f_n \rightrightarrows f_0 \text{ на } [0,\,1] \Leftrightarrow \alpha < 0\big).$$

Выясним, при каких условиях возможен предельный переход под знаком интеграла:

$$\int_0^1 n^\alpha x^n dx = \frac{n^\alpha}{n+1} \to 0 = \int_0^1 f_0(x)dx \Leftrightarrow \alpha < 1.$$

Это мотивирует следующее утверждение.

Теорема 3.3 (об интегрируемости предельной функции). *Если $f_n \rightrightarrows f$ на $[a,\,b]$ и $f_n \in \mathcal{R}[a,\,b]$ для всех $n \in \mathbb{N}$, то $f \in \mathcal{R}[a,\,b]$ и $\int_a^b f(x)dx = \lim\limits_{n\to\infty} \int_a^b f_n(x)dx$.*

□ Покажем, что f удовлетворяет критерию Дарбу. Зафиксируем $\varepsilon > 0$. Из равномерной сходимости найдем такой номер N, что $-\frac{\varepsilon}{4(b-a)} < f_n(x) - f(x) < \frac{\varepsilon}{4(b-a)}$ для всех $x \in [a,\,b]$ и $n \geqslant N$. Интегрируемость f_N влечет ее ограниченность на $[a,\,b]$, так что f также ограничена на $[a,\,b]$.

Пусть T — разбиение $[a,\,b]$. Для числовых функций $\sup_E(h + g) \leqslant \sup_E h + \sup_E g$. Поэтому для верхней суммы Дарбу верно

$$S_T(f) \leqslant S_T(f - f_N) + S_T(f_N) \leqslant \\ \leqslant S_T\left(\frac{\varepsilon}{4(b-a)}\right) + S_T(f_N) = \frac{\varepsilon}{4} + S_T(f_N).$$

Аналогично для нижней суммы Дарбу $s_T(f) \geqslant -\frac{\varepsilon}{4} + s_T(f_N)$. Поскольку $f_N \in \mathcal{R}[a,\,b]$, разбиение T можно выбрать так, что $S_T(f_N) - s_T(f_N) < \frac{\varepsilon}{2}$. Для такого разбиения $S_T(f) - s_T(f) < \varepsilon$. Следовательно, $f \in \mathcal{R}[a,\,b]$.

Поскольку при любом $n \geqslant N$ выполнено

$$\left|\int_a^b f_n(x)dx - \int_a^b f(x)dx\right| \leqslant \int_a^b |f_n(x) - f(x)|dx < \frac{\varepsilon}{b-a}(b-a) = \varepsilon,$$

то $\int_a^b f_n(x)dx \to \int_a^b f(x)dx$. ■

Замечание. В условиях теоремы 3.3 имеем $\lim\limits_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b (\lim\limits_{n\to\infty} f_n(x))dx$.

Задача. Выяснить, верно ли аналогичное утверждение для несобственных интегралов.

Следствие (о почленном интегрировании ряда). *Если ряд* $\sum\limits_{n=1}^{\infty} f_n$ *равномерно сходится на* $[a,\, b]$, *и все функции* $f_n \in \mathcal{R}[a,$ $b]$, *то сумма ряда интегрируема на* $[a,\, b]$ *и*

$$\int_a^b \left(\sum_{n=1}^{\infty} f_n(x)\right) dx = \sum_{n=1}^{\infty} \left(\int_a^b f_n(x)dx\right).$$

Условие равномерной сходимости для предельного перехода под знаком интеграла часто бывает обременительным (см. предыдущий пример). При изучении интеграла Лебега эти условия будут значительно ослаблены.

Теорема 3.4 (о дифференцируемости предельной функции). *Пусть последовательность функций* $f_n\colon [a,\, b] \to$ $\to \mathbb{R}$ $(a < b)$ *такова, что*

1) $f_n \to f$ *на* $[a,\, b]$;
2) *все* f_n *дифференцируемы на* $[a,\, b]$;
3) $f_n' \rightrightarrows g$ *на* $[a,\, b]$.

Тогда функция f *дифференцируема на* $[a,\, b]$, *причем* $f' = g$.

□ Докажем дифференцируемость функции f. Зафиксируем $x \in$ $\in [a,\, b]$ и рассмотрим на $[a,\, b]$ последовательность функций

$$\varphi_n(t) = \begin{cases} \dfrac{f_n(t) - f_n(x)}{t - x}, & \text{при } t \neq x, \\ f_n'(x), & \text{при } t = x. \end{cases}$$

Эта последовательность поточечно сходится к функции φ, где $\varphi(t) = \dfrac{f(t) - f(x)}{t - x}$ при $t \neq x$ и $\varphi(x) = g(x)$. Покажем, что сходимость равномерная. По теореме Лагранжа при $t \neq x$

$$\varphi_n(t) - \varphi_m(t) = \frac{(f_n(t) - f_m(t)) - (f_n(x) - f_m(x))}{t - x} = f_n'(c) - f_m'(c)$$

для некоторой точки c между t и x. Из того, что $\{f_n'\}$ удовлетворяет условию Коши равномерной сходимости на $[a,\, b]$ следует, что $\{\varphi_n\}$ также удовлетворяет этому условию, и значит, $\varphi_n \rightrightarrows \varphi$ на $[a,\, b]$ по критерию Коши (теорема 3.1). Поскольку f_n дифференцируема в точке x, функция φ_n непрерывна в точке x. Тогда по теореме 3.2 функция φ также непрерывна в точке x, т. е. $\lim\limits_{t\to x}\varphi(t)=\varphi(x)$ или $f'(x)=g(x)$. ■

Замечание 1. В условиях теоремы 3.4 $\frac{d}{dx}\lim\limits_{n\to\infty}f_n(x)=$ $=\lim\limits_{n\to\infty}\frac{d}{dx}f_n(x)$ для всех $x\in[a,\,b]$.

Замечание 2. Вместо условия 1 достаточно требовать сходимости $\{f_n(x_0)\}$ в *одной* точке. Сходимость в других точках вытекает из критерия Коши в силу равенства $(f_n(x)-f_m(x))-$ $-(f_n(x_0)-f_m(x_0))=(f_n'(c)-f_m'(c))(x-x_0)$, где c лежит между x и x_0. Кроме того, $|x-x_0|\leqslant b-a$, поэтому $f_n\rightrightarrows f$ на $[a,b]$.

Следствие (о почленном дифференцировании ряда). *Пусть последовательность функций $f_n\colon[a,\,b]\to\mathbb{R}$ такова, что*

1) *ряд $\sum\limits_{n=1}^{\infty}f_n$ сходится поточечно на $[a,\,b]$;*
2) *все f_n дифференцируемы на $[a,\,b]$;*
3) *ряд $\sum\limits_{n=1}^{\infty}f_n'$ равномерно сходится на $[a,\,b]$.*

Тогда сумма ряда дифференцируема и для всех $x\in[a,\,b]$

$$\left(\sum_{n=1}^{\infty}f_n(x)\right)'=\sum_{n=1}^{\infty}f_n'(x).$$

Отметим, что в теореме 3.4 условие равномерной сходимости производных нельзя заменить равномерной сходимостью самих функций. Например, последовательность $f_n\colon[-1,\,1]\to$ $\to\mathbb{R}$, $f_n(x)=\sqrt{x^2+\frac{1}{n}}$, сходится равномерно к недифференцируемой функции $f(x)=|x|$ (см. рис. 3.3).

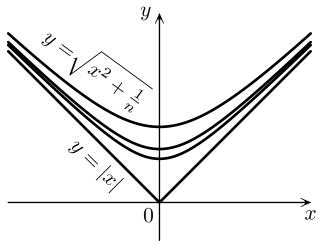

Рис. 3.3

3.2. Признаки равномерной сходимости.

Поскольку точная формула для суммы функционального ряда известна лишь в исключительных случаях, исследование равномерной сходимости по определению крайне затруднительно. На практике желательно иметь легко проверяемые признаки равномерной сходимости. Мы рассмотрим несколько таких утверждений, получающихся из аналогичных утверждений для числовых рядов.

Теорема 3.5 (признак Вейерштрасса). *Пусть $f_n\colon E \to \to \mathbb{C}$, и $a_n \in \mathbb{R}$ для любого $n \in \mathbb{N}$. Пусть*

1) $\forall n \in \mathbb{N}\ \forall x \in E\ (|f_n(x)| \leqslant a_n)$;

2) числовой ряд $\sum_{n=1}^{\infty} a_n$ сходится.

Тогда функциональный ряд $\sum\limits_{n=1}^{\infty} f_n$ сходится равномерно и абсолютно на множестве E.

$\square$ Пусть $\varepsilon > 0$. Так как ряд $\sum_{n=1}^{\infty} a_n$ сходится, то пользуясь критерием Коши как необходимым условием найдем такое N, что для всех $n \geqslant m \geqslant N$ верно $\sum_{k=m}^{n} a_k < \varepsilon$. Тогда при таких n, m и всех $x \in E$

$$\left|\sum_{k=m}^{n} f_k(x)\right| \leqslant \sum_{k=m}^{n} |f_k(x)| \leqslant \sum_{k=m}^{n} a_k < \varepsilon.$$

Пользуясь критерием Коши как достаточным условием, получаем, что ряды $\sum_{n=1}^{\infty} f_n$ и $\sum_{n=1}^{\infty} |f_n|$ сходятся равномерно на E. $\blacksquare$

Про ряд $\sum\limits_{n=1}^{\infty} a_n$ говорят, что он *мажорирует* функциональный ряд $\sum\limits_{n=1}^{\infty} |f_n|$. Существование мажорантного ряда достаточно, но не необходимо для равномерной сходимости даже в классе абсолютно сходящихся рядов.

Пример. Рассмотрим на луче $E = [1, +\infty)$ функциональный ряд $\sum\limits_{n=1}^{\infty} f_n$, где $f_n(x) = \frac{1}{x}$ при $x \in [n, n+1)$, и $f_n(x) = 0$ иначе. Ряд сходится на E к сумме $S(x) = \frac{1}{x}$, причем сходимость равномерная, т.к. $\sup_E |S(x) - S_n(x)| = \frac{1}{n}$. Поскольку $\sup_E |f_n(x)| = = f_n(n) = \frac{1}{n}$ и числовой ряд $\sum\limits_{n=1}^{\infty} f_n(n)$ расходится, то к данному функциональному ряду не применим признак Вейерштрасса.

Рассмотрим теперь более тонкие признаки равномерной сходимости Дирихле и Абеля.

Определение. Пусть функции g_n определены на множестве E. Последовательность $\{g_n\}$ называется *равномерно ограничен-*

ной на E, если найдется такое $C > 0$, что $|g_n(x)| \leqslant C$ для всех $n \in \mathbb{N}$ и $x \in E$.

Теорема 3.6 (признак Дирихле). *Пусть* $a_n\colon E \to \mathbb{R}$ *(или* $\mathbb{C}$*),* $b_n\colon E \to \mathbb{R}$*, такие что*

1) *последовательность* $A_N = \sum_{n=1}^N a_n$ *равномерно ограничена на* E*;*
2) *последовательность* $\{b_n(x)\}$ *монотонна при каждом* $x \in E$*;*
3) $b_n \rightrightarrows 0$ *на* E*.*

Тогда ряд $\sum\limits_{n=1}^{\infty} a_n b_n$ *равномерно сходится на* E*.*

$\square$ По условию существует такое $C > 0$, что $|A_n(x)| \leqslant C$ для всех $n \in \mathbb{N}$ и $x \in E$. Поэтому для всех $x \in E$ и $n \geqslant m$ имеем

$$\left|\sum_{k=m}^{n} a_n(x)\right| = |A_n(x) - A_{m-1}(x)| \leqslant 2C.$$

Зафиксируем $\varepsilon > 0$. Из равномерной сходимости $\{b_n\}$ найдем такой номер N, что $|b_n(x)| < \frac{\varepsilon}{8C}$ для всех $x \in E$ и $n \geqslant N$. Тогда по лемме 0.1 Абеля при всех $x \in E$ и $n \geqslant m \geqslant N$

$$\left|\sum_{k=m}^{n} a_k(x) b_k(x)\right| \leqslant 2 \cdot 2\left(|b_m(x)| + |b_n(x)|\right) < 4C\left(\frac{\varepsilon}{8C} + \frac{\varepsilon}{8C}\right) = \varepsilon.$$

Доказательство завершается ссылкой на критерий Коши. $\blacksquare$

Выделим частные случаи, в которых условие равномерной ограниченности частичных сумм ряда выполняется автоматически.

Следствие (признак Лейбница). *Если при каждом* $x \in E$ *последовательность* $\{\alpha_n(x)\}$ *монотонна, и* $\alpha_n \rightrightarrows 0$ *на* E*, то ряд* $\sum\limits_{n=1}^{\infty} (-1)^{n-1} \alpha_n$ *равномерно сходится на* E*.*

Следствие. *Пусть* I *— отрезок, не содержащий точек* $2\pi m$ $(m \in \mathbb{Z})$*. Если при каждом* $x \in I$ *последовательность* $\{\alpha_n(x)\}$ *монотонна, и* $\alpha_n \rightrightarrows 0$ *на* I*, то ряды* $\sum\limits_{n=1}^{\infty} \alpha_n(x) \sin nx$ *и* $\sum\limits_{n=1}^{\infty} \alpha_n(x) \cos nx$ *равномерно сходятся на* I*.*

□ Как было установлено ранее, $\left|\sum_{n=1}^{N}\sin(nx)\right| \leqslant \frac{1}{|\sin(x/2)|}$. Поскольку $\inf_I|\sin(x/2)| > 0$, суммы $\sum_{n=1}^{N}\sin(nx)$ равномерно ограничены на I. Аналогичная оценка справедлива и для сумм $\sum_{n=1}^{N}\cos(nx)$. По признаку Дирихле ряды равномерно сходятся на I. ■

Теорема 3.7 (признак Абеля). *Пусть* $a_n\colon E\to\mathbb{R}$ *(или* $\mathbb{C}$*),* $b_n\colon E\to\mathbb{R}$*, такие что*

1) *ряд* $\sum_{n=1}^{\infty} a_n$ *равномерно сходится на* E*;*

2) *последовательность* $\{b_n(x)\}$ *монотонна при каждом* $x\in E$*;*

3) *последовательность* $\{b_n\}$ *равномерно ограничена на* E*.*

Тогда ряд $\sum\limits_{n=1}^{\infty} a_nb_n$ *равномерно сходится на* E*.*

□ По условию существует такое $C>0$, что $|b_n(x)|\leqslant C$ для всех $n\in\mathbb{N}$ и $x\in E$. Зафиксируем $\varepsilon>0$. Пользуясь равномерной сходимостью ряда, найдем такое N, что $\left|\sum\limits_{k=m}^{n} a_k(x)\right| < \frac{\varepsilon}{4C}$ для всех $x\in E$ и $n\geqslant m\geqslant N$. Тогда по лемме 0.1 Абеля при таких n, m и $x\in E$

$$\left|\sum_{k=m}^{n} a_k(x)b_k(x)\right| \leqslant 2\cdot\frac{\varepsilon}{4C}(|b_m(x)|+|b_n(x)|) < \frac{\varepsilon}{2C}\cdot 2C=\varepsilon.$$

Доказательство завершается ссылкой на критерий Коши. ■

Характер монотонности в теоремах 3.6 и 3.7 в разных точках множества E может быть различным: в одних точках последовательность может возрастать, а в других — убывать.

Отметим, что признаки Абеля и Дирихле для числовых рядов (теоремы 2.7 и 2.8) являются частными случаями признаков для функциональных рядов.

Теорема 3.8 (признак Дини). *Пусть последовательность* $\{f_n\}$ *поточечно сходится к* f *на* $[a,\,b]$*, функция* f *и все* f_n *непрерывны на* $[a,\,b]$*, и последовательность* $\{|f_n(x)-f(x)|\}$ *нестрого убывает для всякого* $x\in[a,\,b]$*. Тогда* $f_n\rightrightarrows f$ *на* $[a,\,b]$*.*

□ Достаточно показать, что $g_n := |f_n-f|\rightrightarrows 0$ на $[a,\,b]$. Пусть $\varepsilon >$ > 0. Так как $f_n\to f$, то для всякой точки $x_0\in[a,\,b]$ найдется номер $N=N(x_0)$, что $0\leqslant g_N(x_0)<\varepsilon$. Функция g_N непрерывна, поэтому найдется такое $\delta = \delta(x_0) > 0$, что $0 \leqslant g_N(x) < \varepsilon$ для всех $x \in B_\delta(x_0) \cap [a,\, b]$. Но тогда в силу монотонности $0 \leqslant g_n(x) < \varepsilon$ для всех $n \geqslant N$ и $x \in B_\delta(x_0) \cap [a,\, b]$. Семейство $\{B_{\delta(x)}(x)\}_{x\in[a,\,b]}$ образует открытое покрытие $[a,\, b]$. По теореме Гейне–Бореля найдутся такие точки $x_1,\, \ldots,\, x_m \in [a,\, b]$, что $[a,\, b] \subset B_{\delta(x_1)}(x_1) \cup \cup \ldots \cup B_{\delta(x_m)}(x_m)$. Положим $N^* = \max_{1\leqslant i\leqslant m}\{N(x_i)\}$. Тогда при $n \geqslant N^*$ выполнено $0 \leqslant g_n(x) < \varepsilon$ для всех $x \in [a,\, b]$, что и требовалось доказать. ■

Следствие. *Пусть ряд* $\sum\limits_{n=1}^{\infty} f_n$ *поточечно сходится к* S *на* $[a,\, b]$, *функция* S *и все* f_n *непрерывны на* $[a,\, b]$, *причем* $f_n \geqslant 0$ *на* $[a,\, b]$. *Тогда ряд* $\sum\limits_{n=1}^{\infty} f_n$ *равномерно сходится на* $[a,\, b]$.

Часто равномерно сходящиеся ряды используют для построения функций с выделенными свойствами.

Пример (Ван-дер-Варден). Существует непрерывная функция $f\colon \mathbb{R} \to \mathbb{R}$, не дифференцируемая ни в одной точке.

□ Продолжим $|x|$ с отрезка $[-1,\, 1]$ на всю числовую прямую с периодом 2, т.е. рассмотрим функцию $\varphi\colon \mathbb{R} \to \mathbb{R}$, $\varphi(x \pm 2) = \varphi(x)$ и $\varphi(x) = |x|$ при $|x| \leqslant 1$. Заметим, что если на интервале $(x,\, y)$ нет целых чисел, то φ там кусочно линейна с угловым коэффициентом ± 1 и, значит,

$$|\varphi(x) - \varphi(y)| = |x - y|. \tag{3.2}$$

Определим функцию $f(x) = \sum\limits_{n=1}^{\infty} f_n(x)$, где $f_n(x) = 4^{-n}\varphi(4^n x)$. Функция f непрерывна как равномерно сходящийся по признаку Вейерштрасса ряд непрерывных функций.

Пусть $a \in \mathbb{R}$. Построим такую последовательность ненулевых чисел h_k, сходящуюся к 0, что конечного $\lim\limits_{k\to\infty} \dfrac{f(a+h_k) - f(a)}{h_k}$ не существует. Интервалы $\left(4^k a,\, 4^k a + \frac{1}{2}\right)$ и $\left(4^k a - \frac{1}{2},\, 4^k a\right)$ не могут одновременно содержать целых чисел. Поэтому найдется такое $h_k = \pm\frac{1}{2}4^{-k}$, что на интервале с концами $4^k a$ и $4^k(a + h_k)$ нет целых чисел. Фактически можно утверждать больше: при $n \leqslant \leqslant k$ на интервале с концами $4^n a$ и $4^n(a + h_k)$ нет целых чисел (если между $4^n a$ и $4^n(a + h_k)$ имеется целое число, то домножая соответствующее неравенство на 4^{k-n} получим целое число между $4^k a$ и $4^k(a+h_k)$, что противоречит выбору h_k). Поэтому по (3.2) имеем $|\varphi(4^n(a+h_k)) - \varphi(4^n a)| = 4^n|h_k|$, $1 \leqslant n \leqslant k$, а ввиду 2-периодичности φ имеем $|\varphi(4^n(a+h_k)) - \varphi(4^n a)| = 0$ при $n > k$. Следовательно,

$$|f_n(a+h_k) - f_n(a)| = 4^{-n}|\varphi(4^n(a+h_k)) - \varphi(4^n a)| = \begin{cases} |h_k|, & n \leqslant k, \\ 0, & n > k. \end{cases}$$

Тогда разностное отношение

$$\frac{f(a+h_k) - f(a)}{h_k} = \sum_{n=1}^{k} \frac{f_n(a+h_k) - f_n(a)}{h_k} = \sum_{n=1}^{k} \pm 1$$

есть четное число при четном k и нечетное — при нечетном k. Поэтому не существует предела разностных отношений при $k \to \infty$ и, значит, функция f не дифференцируема в точке a. ■

4. Степенные ряды

4.1. Свойства степенных рядов.

Определение. *Степенным рядом* называется функциональный ряд вида
$$\sum_{n=0}^{\infty} a_n(x-x_0)^n, \tag{4.1}$$
где a_n, $x_0 \in \mathbb{R}$ (или $\mathbb{C}$), x — действительная (комплексная) переменная.

В этом пункте будем рассматривать оба случая одновременно.

На множестве сходимости ряд (4.1) определяет функцию f — сумму этого ряда. По соглашению $(x-x_0)^0 = 1$ точка x_0 всегда лежит во множестве сходимости, причем $f(x_0) = a_0$.

Теорема 4.1 (Коши–Адамар). *Пусть* $R = \dfrac{1}{\varlimsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}$ (*считаем* $\frac{1}{+\infty} = 0$, $\frac{1}{0} = +\infty$). *Тогда*

1) *при* $|x-x_0| < R$ *ряд* (4.1) *сходится, причем абсолютно;*
2) *при* $|x-x_0| > R$ *ряд* (4.1) *расходится;*
3) *для* $r \in \mathbb{R}$, $0 < r < R$, *ряд* (4.1) *сходится равномерно на множестве* $\overline{B}_r(x_0) := \{x\colon\ |x-x_0| \leqslant r\}$.

□ Первые два пункта вытекают из признака Коши (теорема 2.4). Действительно, пусть $x \neq x_0$. Тогда
$$q := \varlimsup_{n\to\infty} \sqrt[n]{|a_n(x-x_0)^n|} = |x-x_0| \varlimsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|x-x_0|}{R}.$$
Если $|x-x_0| < R$, то $q < 1$ и по признаку Коши ряд (4.1) абсолютно сходится (а, следовательно, сходится). Если $|x-x_0| > R$, то $q > 1$ и по признаку Коши n-й член ряда (4.1) не стремится к нулю, поэтому ряд (4.1) расходится.

Пусть $r \in (0, R)$. По доказанному в точке $x = x_0 + r$ ряд абсолютно сходится, т. е. сходится числовой ряд $\sum\limits_{n=0}^{\infty} |a_n| r^n$. Если $|x-x_0| \leqslant r$, то $|a_n(x-x_0)^n| \leqslant |a_n| r^n$, и ряд (4.1) сходится равномерно на $\overline{B}_r(x_0)$ по признаку Вейерштрасса. ■

Определение. Величина R из теоремы 4.1 называется *радиусом сходимости* ряда (4.1). Множество $B_R(x_0) = \{x\colon |x-x_0| <$

$< R\}$ называется *интервалом сходимости* (для комплексного — *кругом сходимости*) ряда (4.1).

Замечание. Таким образом, множество сходимости действительного степенного ряда совпадает с интервалом сходимости, возможно, включающим одну или обе концевые точки. Для комплексного ряда — с кругом сходимости, возможно, включающим некоторое подмножество граничной окружности.

Радиус сходимости определяется условиями 1 и 2 теоремы 4.1. Оформим это в виде следующего наблюдения.

Следствие. *Если $R \in [0, +\infty]$ такое, что для всех x с условием $|x - x_0| < R$ ряд* (4.1) *абсолютно сходится, а для всех x с условием $|x - x_0| > R$ ряд* (4.1) *абсолютно расходится (т. е. расходится ряд из модулей членов), то R — радиус сходимости ряда.*

☐ Обозначим радиус сходимости ряда через $R_{\text{сх}}$. Пусть $R > R_{\text{сх}}$. Выберем x так, что $R_{\text{сх}} < |x - x_0| < R$. Из доказательства теоремы 4.1 следует, что $a_n(x - x_0)^n$ не стремится к нулю, что противоречит абсолютной сходимости ряда, которая имеется по условию. Аналогично устанавливается невозможность $R_{\text{сх}} > R$. ■

Пример. Найдем радиус сходимости ряда $\sum\limits_{n=1}^{\infty} \frac{n!}{n^n} x^{2n}$.

☐ Обозначим n-й член ряда как $u_n(x)$, тогда при $x \neq 0$ имеем

$$\frac{|u_{n+1}(x)|}{|u_n(x)|} = \frac{(n+1)n^n}{(n+1)^{n+1}}|x|^2 = \frac{|x|^2}{\left(1 + \frac{1}{n}\right)^n} \to \frac{|x|^2}{e}, \quad n \to \infty.$$

Если $\frac{|x|^2}{e} < 1 \Leftrightarrow |x| < \sqrt{e}$, то ряд абсолютно сходится по признаку Даламбера. Если $\frac{|x|^2}{e} > 1 \Leftrightarrow |x| > \sqrt{e}$, то ряд абсолютно расходится по признаку Даламбера. Следовательно, радиус сходимости $R = \sqrt{e}$. ■

Задача. Исследуйте сходимость ряда из предыдущего примера при $x = \pm\sqrt{e}$.

Теорема 4.2 (Абель). *Если степенной ряд* (4.1) *сходится в точке $x_1 \neq x_0$, то он сходится равномерно на отрезке с концами x_0, x_1.*

$\square$ Ряд $\sum\limits_{n=0}^{\infty} a_n(x_1-x_0)^n$ сходится по условию. Последовательность $\{t^n\}$ монотонна при любом $t \in [0,\,1]$, и $|t^n| \leqslant 1$ для всех $n \in \mathbb{N}$ и $t \in [0,\,1]$. Тогда по признаку Абеля ряд $\sum\limits_{n=0}^{\infty} a_n(x_1-x_0)^n t^n$ равномерно сходится на $[0,\,1]$. Сделав замену $t = \dfrac{x-x_0}{x_1-x_0}$, заключаем, что ряд (4.1) равномерно сходится на множестве $\{x = x_0 + t(x_1 - x_0),\ t \in [0,\,1]\}$ — отрезке с концами x_0, x_1. $\blacksquare$

Замечание. Если $x_1 \in B_R(x_0)$, то теорема 4.2 вытекает из п. 3 теоремы 4.1. Поэтому интерес представляет случай, когда x_1 лежит на границе интервала (круга) сходимости.

Пользуясь следствием теоремы 3.2, для действительных рядов получим

Следствие. *Сумма степенного ряда* (4.1) *на множестве сходимости является непрерывной функцией.*

Задача. Пусть ряды $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} b_n$ и $\sum_{n=0}^{\infty} c_n$, где $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$, сходятся к A, B и C соответственно. Покажите, что $C = AB$.

Одним из важнейших свойств степенных рядов является то, что их можно почленно дифференцировать. Ключевым здесь является следующий результат.

Лемма 4.1. *Если степенной ряд* (4.1) *имеет радиус сходимости R, то ряд $\sum\limits_{n=1}^{\infty} na_n(x-x_0)^{n-1}$, полученный почленным дифференцированием, также имеет радиус сходимости R.*

$\square$ Так как $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$, то множества частичных пределов последовательностей $\{\sqrt[n]{n|a_n|}\}$ и $\{\sqrt[n]{|a_n|}\}$ совпадают, а значит, $\varlimsup\limits_{n\to\infty} \sqrt[n]{n|a_n|} = \varlimsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$. По формуле Коши-Адамара заключаем, что радиус сходимости $\sum\limits_{n=1}^{\infty} na_n(x-x_0)^n$ равен R. Пусть $x \neq x_0$. Ряды $\sum\limits_{n=1}^{\infty} na_n(x-x_0)^n$ и $\sum\limits_{n=1}^{\infty} na_n(x-x_0)^{n-1}$ отличаются ненулевым множителем, а значит, сходятся одновременно (при

$x = x_0$ очевидно сходятся). Следовательно, радиус сходимости второго ряда также равен R. ■

Теорема 4.3. *Если* $f(x) = \sum\limits_{n=0}^{\infty} a_n(x-x_0)^n$ *— сумма степенного ряда с радиусом сходимости* $R > 0$, *то функция* f *бесконечно дифференцируема на интервале сходимости, и для каждого* $m \in \mathbb{N}$ *имеет место равенство*

$$f^{(m)}(x) = \sum_{n=m}^{\infty} n(n-1)\dots(n-m+1)\cdot a_n(x-x_0)^{n-m}. \qquad (4.2)$$

□ При дифференцировании радиус сходимости ряда не меняется. Поэтому нам достаточно доказать утверждение для $m = 1$, после чего применить индукцию.

Первый способ. Пусть $0 < r < R$. По п. 3 теоремы 4.1 исходный ряд и ряд $\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}$ равномерно сходятся на отрезке $[x_0 - r,\, x_0 + r]$. Обозначим через g сумму продифференцированного ряда. Тогда по следствию теоремы 3.4 функция f дифференцируема на $[x_0 - r,\, x_0 + r]$, причем $f' = g$. Так как $r \in (0,\, R)$ — любое, то равенство выполняется на $(x_0 - R,\, x_0 + R)$.

Второй способ. Без потери общности можно считать $x_0 = 0$. Пусть $t \in B_R(0)$. Покажем, что производная $f(x) = \sum_{n=0}^{\infty} a_n x^n$ в точке t равна числу $l = \sum_{n=1}^{\infty} na_n t^{n-1}$.

Зафиксируем такое r, что $|t| < r < R$. Для $x \neq t$, $|x| \leqslant r$ составим разность

$$\frac{f(x)-f(t)}{x-t} - l = \sum_{n=1}^{\infty} a_n\left(\frac{x^n - t^n}{x-t} - nt^{n-1}\right).$$

Выражение в скобках перепишем в виде:

$$x^{n-1} + tx^{n-2} + \ldots + xt^{n-2} + t^{n-1} - nt^{n-1} =$$
$$= (x^{n-1} - t^{n-1}) + t(x^{n-2} - t^{n-2}) + \ldots + t^{n-2}(x-t).$$

Каждая из разностей в скобках имеет множителем $(x - t)$. Частное при делении на $(x - t)$ имеет вид

$$(x^{n-2} + tx^{n-3} + \ldots + t^{n-2}) + t(x^{n-3} + tx^{n-4} + \ldots + t^{n-3}) + \ldots + t^{n-2}$$

(в первой сумме $(n-1)$ слагаемых, во второй — $(n-2)$ слагаемых, и т. д.). Поскольку $(n-1)+(n-2)+\ldots+1=\dfrac{n(n-1)}{2}$ и каждое слагаемое в частном по модулю не превосходит r^{n-2}, получаем оценку

$$\left|\frac{x^n-t^n}{x-t}-nt^{n-1}\right| \leqslant |x-t|\,|a_n|\cdot\frac{n(n-1)}{2}\,r^{n-2}.$$

По лемме 4.1 дважды продифференцированный ряд имеет тот же радиус сходимости R и, значит, сходится абсолютно при $x=$ $=r$, т.е. сходится ряд $\sum\limits_{n=2}^{\infty}|a_n|\dfrac{n(n-1)}{2}r^{n-2}$. Следовательно,

$$\left|\frac{f(x)-f(t)}{x-t}-l\right| \leqslant |x-t|\sum_{n=2}^{\infty}|a_n|\cdot\frac{n(n-1)}{2}\,r^{n-2},$$

что доказывает $\lim\limits_{x\to t}\dfrac{f(x)-f(t)}{x-t}=l.$ ■

Доказательство вторым способом справедливо и для комплексных степенных рядов (производная функции из $\mathbb{C}$ в $\mathbb{C}$ определяется также как и для действительной переменной).

Следствие (теорема единственности). *Если f — сумма степенного ряда* (4.1) *с радиусом сходимости $R>0$, то его коэффициенты однозначно определяются формулой $a_n=\dfrac{f^{(n)}(x_0)}{n!}$, $n=0,\,1,\,2,\,\ldots$*

□ В формуле (4.2) подставим $x=x_0$. Тогда слагаемые с номерами $n\geqslant m+1$ обнулятся, так что $f^{(m)}(x_0)=m!a_m$. ■

Следствие. *Сумма степенного ряда* (4.1) *с радиусом сходимости $R>0$ имеет первообразную $F(x)=C+\sum\limits_{n=0}^{\infty}\dfrac{a_n}{n+1}(x-x_0)^{n+1}$ при $|x-x_0|<R$.*

4.2. Ряды Тейлора.

В дальнейшем рассматриваются только действительные степенные ряды.

Определение. Если функция f определена в некоторой окрестности точки x_0 и имеет в точке x_0 производные всех порядков, то ряд $\sum\limits_{n=0}^{\infty}\dfrac{f^{(n)}(x_0)}{n!}(x-x_0)^n$ называется *рядом Тейлора* функции f в точке x_0. При $x_0 = 0$ ряд называют *рядом Маклорена*.

Предыдущее следствие можно переформулировать так: если функция f в окрестности точки x_0 является суммой степенного ряда с центром в x_0 (*разлагается* в ряд по степеням $x - x_0$), то этот ряд — ее ряд Тейлора в точке x_0. Следующий пример показывает, что ряд Тейлора бесконечно дифференцируемой функции f может сходиться к сумме, отличной от f.

Пример. Рассмотрим функцию $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0 \\ 0, & x \leqslant 0. \end{cases}$$

Существование производных любого порядка в точке $x \neq 0$ следует из теоремы о дифференцировании композиции. Более того, $f^{(n)}(x) = 0$ при $x < 0$ и $f^{(n)}(x) = p_n(1/x)e^{-1/x}$ при $x > 0$, где $p_n(t)$ — многочлен степени $2n$. Последнее утверждение можно установить по индукции: $p_0(t) = 1$ и дифференцирование $f^{(n)}$ дает соотношение $p_{n+1}(t) = t^2[p_n(t) - p_n'(t)]$.

Индукцией по n покажем, что $f^{(n)}(0) = 0$. Для $n = 0$ это верно по условию. Если предположить, что $f^{(n)}(0) = 0$, то $(f^{(n)})'_-(0) = 0$ и

$$(f^{(n)})'_+(0) = \lim_{h\to+0} \frac{f^{(n)}(h) - f^{(n)}(0)}{h} = \lim_{h\to+0} \frac{p_n(1/h)e^{-1/h}}{h} = \lim_{t\to+\infty} \frac{tp_n(t)}{e^t} = 0,$$

поскольку по правилу Лопиталя $\lim\limits_{t\to+\infty} \dfrac{t^m}{e^t} = 0$ для всех $m \in \mathbb{N}_0$. Это доказывает, что $f^{(n+1)}(0) = 0$.

Все коэффициенты ряда Маклорена функции f, а значит, и его сумма, равны нулю. Таким образом, функция f совпадает с суммой своего ряда Маклорена только при $x \leqslant 0$.

Другим препятствием разложимости функции f по степеням $x - x_0$ может являться расходимость ряда Тейлора при всех $x \neq x_0$.

Задача. Покажите, что функция $f(x) = \sum\limits_{n=0}^{\infty} \dfrac{\cos(n^2 x)}{2^n}$ бесконечно дифференцируема на $\mathbb{R}$, однако ее ряд Маклорена имеет нулевой радиус сходимости.

Приведем одно достаточное условие разложимости функции в степенной ряд.

Лемма 4.2. *Если функция f бесконечно дифференцируема на интервале $(x_0-\rho,\, x_0+\rho)$, и существует такое $C>0$, что $|f^{(n)}(x)| \leqslant \frac{Cn!}{\rho^n}$ для всех $n=0,\,1,\,\ldots$ и всех $x\in(x_0-\rho,\, x_0+\rho)$, то $f(x)=\sum\limits_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$ на этом интервале.*

☐ Поскольку $\left|\frac{f^{(n)}(x_0)}{n!}\right|^{1/n} \leqslant \frac{C^{1/n}}{\rho} \to \frac{1}{\rho}$, то по формуле Коши-Адамара радиус сходимости ряда Тейлора функции f в точке x_0 не меньше ρ. Покажем, что этот ряд сходится к f. По формуле Тейлора с остаточным членом в форме Лагранжа для всякого $x\in(x_0-\rho,\, x_0+\rho)$ и номера N найдется такая точка c, лежащая строго между x и x_0, что

$$\left|f(x)-\sum_{n=0}^{N}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\right| = \left|\frac{f^{(N+1)}(c)}{(N+1)!}(x-x_0)^{N+1}\right|.$$

Так как для $f^{(N+1)}(c)$ справедлива оценка из условия, то

$$\left|\frac{f^{(N+1)}(c)}{(N+1)!}(x-x_0)^{N+1}\right| \leqslant C\left|\frac{x-x_0}{\rho}\right|^{N+1} \to 0, \quad N\to\infty,$$

что завершает доказательство. ■

Следствие. *Если функция f бесконечно дифференцируема на $(x_0-\rho,\, x_0+\rho)$ и все ее производные равномерно ограничены (т.е. существует $M>0$, что $|f^{(n)}(x)|\leqslant M$ при $|x-x_0|<\rho$ и всех $n\in\mathbb{N}_0$), то на этом интервале f разлагается в ряд по степеням $x-x_0$.*

☐ Следует по лемме 4.2, поскольку последовательность $\{n!/\rho^n\}$ является бесконечно большой. ■

Следствие. *Ряды Маклорена функций e^x, $\sin x$, $\cos x$ сходят-*

ся к этим функциям в любой точке $x \in \mathbb{R}$, т.е.

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}.$$

$\square$ Указанные функции бесконечно дифференцируемы, причем

$$(e^x)^{(n)} = e^x, \ (\cos x)^{(n)} = \cos\left(x + \frac{\pi n}{2}\right), \ (\sin x)^{(n)} = \sin\left(x + \frac{\pi n}{2}\right).$$

Поэтому при $|x| < \delta$ $(\delta > 0)$ выполнено

$$|(e^x)^{(n)}| \leqslant e^{\delta}, \qquad |(\cos x)^{(n)}| \leqslant 1, \qquad |(\sin x)^{(n)}| \leqslant 1,$$

и значит, по лемме 4.2 все эти функции являются суммами своих рядов Маклорена на интервале $(-\delta, \delta)$. Так как $\delta > 0$ — любое, равенство имеет место для всех $x \in \mathbb{R}$. $\blacksquare$

Теорема 4.4 (биномиальный ряд). *Если $\alpha \notin \mathbb{N}_0$, $C_\alpha^n = \dfrac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}$, и $C_\alpha^0 = 1$, то справедливо представление*

$$(1+x)^\alpha = \sum_{n=0}^{\infty} C_\alpha^n x^n, \quad |x| < 1.$$

$\square$ Пусть $f(x) = (1+x)^\alpha$, тогда $f^{(n)}(x) = \alpha(\alpha-1)\cdot\ldots\cdot(\alpha-n+1)(1+x)^{\alpha-n}$ и, значит, $\dfrac{f^{(n)}(0)}{n!} = C_\alpha^n$. Так как при $x \neq 0$

$$\lim_{n\to\infty} \frac{|C_\alpha^{n+1} x^{n+1}|}{|C_\alpha^n x^n|} = \lim_{n\to\infty} \frac{|\alpha - n|}{n+1}|x| = |x|,$$

то по признаку Даламбера ряд абсолютно сходится при $|x| < 1$ и абсолютно расходится при $|x| > 1$. Следовательно, радиус сходимости исследуемого ряда равен 1.

Определим функцию $g(x) = \sum\limits_{n=0}^{\infty} C_\alpha^n x^n$, и покажем, что $g \equiv f$ на $(-1, 1)$, т. е. $(1+x)^{-\alpha} g(x) = 1$ при $x \in (-1, 1)$. Для этого найдем производную функции $(1+x)^{-\alpha} g(x)$. По теореме 4.3 имеем

$$\big((1+x)^{-\alpha} g(x)\big)' = (1+x)^{-\alpha} \sum_{n=1}^{\infty} n C_\alpha^n x^{n-1} - \alpha (1+x)^{-\alpha-1} \sum_{n=0}^{\infty} C_\alpha^n x^n =$$

$$= (1+x)^{-\alpha-1} \left[\sum_{n=1}^{\infty} n C_\alpha^n x^{n-1} + \sum_{n=0}^{\infty} n C_\alpha^n x^n - \alpha \sum_{n=0}^{\infty} C_\alpha^n x^n\right].$$

В первой сумме произведем замену индекса суммирования. После приведения подобных слагаемых получим

$$\big((1+x)^{-\alpha}g(x)\big)' = \\ = (1+x)^{-\alpha-1}\left[\sum_{n=0}^{\infty}(n+1)C_\alpha^{n+1}x^n - \sum_{n=0}^{\infty}(\alpha-n)C_\alpha^n x^n\right] = 0.$$

Отсюда следует, что $(1+x)^{-\alpha}g(x)$ постоянна на $(-1,\,1)$. Из условия $g(0)=1$ получаем, что $(1+x)^{-\alpha}g(x)=1$ для всех $x\in(-1,\,1)$. ■

Замечание. При $\alpha>0$ биномиальный ряд сходится к $(1+x)^\alpha$ равномерно на $[-1,\,1]$. Действительно, при $n>\alpha$

$$\frac{|C_\alpha^{n+1}|}{|C_\alpha^n|} = \frac{n-\alpha}{n+1} = 1-\frac{\alpha+1}{n}+O\left(\frac{1}{n^2}\right).$$

По признаку Гаусса ряд $\sum|C_\alpha^n|$ сходится. Следовательно, по признаку Вейерштрасса ряд $\sum C_\alpha^n x^n$ сходится равномерно на $[-1,\,1]$.

Пример. Так как $\frac{1}{1+x}=\sum\limits_{n=1}^{\infty}(-1)^{n-1}x^{n-1}$ при $|x|<1$, то по следствию 2 теоремы 4.3

$$\ln(1+x)=\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}x^n, \qquad |x|<1.$$

Ряд в правой части сходится при $x=1$, поэтому его сумма непрерывна на $(-1,\,1]$ и, значит, равенство имеет место при $x=1$. Получаем известный нам результат, что $\sum\limits_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}=\ln 2$.

Задача. Разложите функцию arctg в ряд по степеням x. С помощью полученного разложения найдите сумму ряда $\sum\limits_{n=0}^{\infty}\frac{(-1)^n}{2n+1}$.

5. Метрические пространства

5.1. Метрики и нормы.

Определение. Пусть $X \neq \varnothing$. Функция $\rho\colon X \times X \to \mathbb{R}$ называется *метрикой* на X, если для любых x, y, $z \in X$ выполнено:

1) $\rho(x, y) \geqslant 0, \quad \rho(x, y) = 0 \Leftrightarrow x = y$;
2) $\rho(x, y) = \rho(y, x)$;
3) $\rho(x, y) \leqslant \rho(x, z) + \rho(z, y)$ (*неравенство треугольника*).

Пара (X, ρ) называется *метрическим пространством.*

В дальнейшем часто под метрическим пространством будем понимать само множество X, предполагая наличие связанной с ним метрики.

Задача. Докажите, что $|\rho(x, y) - \rho(a, b)| \leqslant \rho(x, a) + \rho(y, b)$ для любых x, y, a, $b \in X$.

Пример. Пусть $X \neq \varnothing$. Положим $\rho(x, y) = 1$ при $x \neq y$ и $\rho(x, y) = 0$ при $x = y$. Тогда ρ — метрика на X (*дискретная* метрика). Проверка первых двух свойств тривиальна. Неравенство треугольника выполняется, т.к. левая часть всегда не больше 1, а правая часть не меньше 1.

Таким образом, любое (непустое) множество можно наделить метрикой.

Пример. Пусть G — граф (конечное множество вершин V, соединенных ребрами). Предположим, что граф G связен, т.е. существует путь, соединяющий любую пару различных вершин. Определим $\rho(v, v) = 0$ и $\rho(v, w)$ как длину кратчайшего пути из вершины v в вершину w. Нетрудно установить, что ρ — метрика.

Важные примеры метрических пространств возникают из других конструкций.

Определение. Пусть V — (вещественное) линейное пространство. Функция $\|\cdot\|\colon V \to \mathbb{R}$ называется *нормой*, если для любых x, $y \in V$ и $\alpha \in \mathbb{R}$ выполнено:

1) $\|x\| \geqslant 0, \quad \|x\| = 0 \Leftrightarrow x = 0$;
2) $\|\alpha x\| = |\alpha| \cdot \|x\|$;
3) $\|x + y\| \leqslant \|x\| + \|y\|$ (*неравенство треугольника*).

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

Лемма 5.1. *Всякое нормированное пространство является метрическим с* $\rho(x,\, y) = \|x - y\|$.

$\square$ Проверка свойств метрики тривиальна. Например, для установления неравенства треугольника достаточно применить свойство 3 нормы к векторам $x - y$ и $y - z$. $\blacksquare$

Еще раз подчеркнем, что в определении метрического простанства X — абстрактное множество, а в определении нормированного простанства V наделено линейной структурой.

Примеры. Для $X = \mathbb{R}^n$ и $x = (x_1,\, \ldots,\, x_n)^T$, $y = (y_1,\, \ldots,\, y_n)^T$ определим

1) $|x| = \sqrt{\sum\limits_{i=1}^{n} |x_i|^2}$ (*евклидова норма*), $\rho_2(x,\, y) = \sqrt{\sum\limits_{i=1}^{n} |x_i - y_i|^2}$ (*евклидова метрика*);

2) $\|x\|_p = \left(\sum\limits_{i=1}^{n} |x_i|^p\right)^{1/p}$, $\rho_p(x,\, y) = \left(\sum\limits_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}$, $p \geqslant 1$;

3) $\|x\|_\infty = \max\limits_{1 \leqslant i \leqslant n} |x_i|$, $\rho_\infty(x,\, y) = \max\limits_{1 \leqslant i \leqslant n} |x_i - y_i|$.

$\square$ Покажем, что $\|\cdot\|_p$ является нормой на $\mathbb{R}^n$. Достаточно проверить выполнения неравенства треугольника.

Сначала покажем, что если $\|x\|_p \leqslant 1$, $\|y\|_p \leqslant 1$ и $\alpha,\, \beta \in \mathbb{R}$ с $\alpha + \beta = 1$, то $\|\alpha x + \beta y\|_p \leqslant 1$.

В силу выпуклости функции $\varphi(s) = s^p$ на $[0,\, +\infty)$ имеем $|\alpha x_i + \beta y_i|^p \leqslant \alpha |x_i|^p + \beta |y_i|^p$. Просуммировав неравенства по всем $i = 1,\, \ldots,\, n$, получим $\|\alpha x + \beta y\|_p^p \leqslant \alpha + \beta = 1$.

Проверим, что $\|x + y\|_p \leqslant \|x\|_p + \|y\|_p$ (индекс p будем опускать). Если $x = 0$ или $y = 0$, то неравенство верно. Если $x \neq 0$ и $y \neq 0$, то положим $\alpha = \dfrac{\|x\|}{\|x\| + \|y\|}$, $\beta = \dfrac{\|y\|}{\|x\| + \|y\|}$, $\hat{x} = \dfrac{x}{\|x\|}$ и $\hat{y} = \dfrac{y}{\|y\|}$. Тогда, учитывая, что $\|\alpha\hat{x} + \beta\hat{y}\| \leqslant 1$, имеем

$$\|x + y\| = (\|x\| + \|y\|) \left\| \frac{x}{\|x\| + \|y\|} + \frac{y}{\|x\| + \|y\|} \right\| = \\ = (\|x\| + \|y\|)\|\alpha\hat{x} + \beta\hat{y}\| \leqslant \|x\| + \|y\|.$$

Проверка, что $\|\cdot\|_\infty$ является нормой, легко следует из свойств модуля числа. $\blacksquare$

Отметим, что не всякая метрика, заданная на линейном простанстве, порождается нормой.

Определение. Пусть $(X,\, \rho)$ — метрическое пространство и $a \in X$.

Множество $B_r(a) = \{x \in X \colon \rho(x,\, a) < r\}$ называется *открытым шаром* с центром a и радиусом r. Множество $\overline{B}_r(a) = \{x \in \in X \colon \rho(x,\, a) \leqslant r\}$ называется *замкнутым шаром* с центром a и радиусом r.

Задача. Изобразите шары $B_1((0,\, 0))$ в $\mathbb{R}^2$ с нормами $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_4$ и $\|\cdot\|_\infty$.

Определение. Множество $E \subset X$ называется *ограниченным*, если существуют такие $a \in X$ и $r > 0$, что $E \subset B_r(a)$.

Определение. Пусть $x_n \in X$ для всех $n \in \mathbb{N}$. Последовательность $\{x_n\}$ *сходится* к a (в X), если $\rho(x_n,\, a) \to 0$ при $n \to \infty$. Пишут $\lim\limits_{n\to\infty} x_n = a$, или $x_n \to a$.

Замечание. $x_n \to a \Leftrightarrow \forall \varepsilon > 0\ \exists N\ \forall n \geqslant N\ (x_n \in B_\varepsilon(a))$.

Свойство 1. Если $x_n \to a$ и $x_n \to b$, то $a = b$.

☐ Вытекает из неравенства $\rho(a,\, b) \leqslant \rho(a,\, x_n) + \rho(x_n,\, b)$. ■

Свойство 2. Если $x_n \to a$, то $\{x_n\}$ ограничена (т.е. ограничено множество значений $\{x_n\}$).

☐ По условию $\rho(x_n,\, a) \to 0$ и, значит, $\{\rho(x_n,\, a)\}$ ограничена в $\mathbb{R}$. Пусть $R > \sup \rho(x_n,\, a)$, тогда $x_n \in B_R(a)$ для всех n. ■

Свойство 3. Пусть $\{x_n\}$, $\{y_n\}$ — последовательности в нормированном пространстве $(V,\, \|\cdot\|)$, $\{\alpha_n\}$ — числовая последовательность. Если $x_n \to a$, $y_n \to b$ и $\alpha_n \to \alpha$, то $x_n + y_n \to a + b$, $\alpha_n x_n \to \alpha a$.

☐ Вытекает из неравенств

$$\|(x_n + y_n - (a + b)\| \leqslant \|x_n - a\| + \|y_n - b\|,$$

$$\|\alpha_n x_n - \alpha a\| \leqslant |\alpha_n - \alpha| \|x_n\| + |\alpha| \|x_n - a\|.$$

5.2. Топология метрических пространств.

Пусть $(X,\, \rho)$ — метрическое пространство, $E \subset X$ и $x \in X$.

Определение. 1) Точка x называется *внутренней точкой* множества E, если существует такое $\varepsilon > 0$, что $B_\varepsilon(x) \subset E$. Множество всех внутренних точек E называется *внутренностью* E и обозначается как $\operatorname{int} E$.

2) Точка x называется *внешней точкой* множества E, если существует такое $\varepsilon > 0$, что $B_\varepsilon(x) \subset X \setminus E$. Множество всех внешних точек E называется *внешностью* E и обозначается как $\operatorname{ext} E$. Очевидно, что $\operatorname{ext} E = \operatorname{int}(X \setminus E)$.

3) Точка x называется *граничной точкой* множества E, если для любого $\varepsilon > 0$ верно, что $B_\varepsilon(x) \cap E \neq \varnothing$ и $B_\varepsilon(x) \cap (X \setminus E) \neq$ $\neq \varnothing$. Множество всех граничных точек E называется *границей* E и обозначается как ∂E.

Замечание. Непосредственно из определения следует, что $X = \operatorname{int} E \cup \operatorname{ext} E \cup \partial E$, причем множества $\operatorname{int} E$, $\operatorname{ext} E$, ∂E попарно не пересекаются.

Определение. Множество $G \subset X$ называется *открытым*, если все его точки являются внутренними (т. е. $G = \operatorname{int} G$). Множество $F \subset X$ называется *замкнутым*, если $X \setminus F$ открыто.

Примеры. 1) Открытый шар — открытое множество.

Пусть $x \in B_r(a)$. Покажем, что x — внутренняя точка $B_r(a)$. Положим $\varepsilon = r - \rho(x, a)$. Тогда для любого $y \in B_\varepsilon(x)$ имеем $\rho(y, a) \leqslant \rho(y, x) + \rho(x, a) < \varepsilon + \rho(x, a) = r$. Следовательно, $B_\varepsilon(x) \subset$ $\subset B_r(a)$.

2) Замкнутый шар — замкнутое множество.

Пусть $x \in X \setminus \overline{B}_r(a)$. Положим $\varepsilon = \rho(x, a) - r$. Как и в предыдущем пункте устанавливается, что $B_\varepsilon(x) \subset X \setminus \overline{B}_r(a)$. Следовательно, $X \setminus \overline{B}_r(a)$ открыто.

3) $\operatorname{int} E$ — открытое множество.

Если $x \in \operatorname{int} E$, то существует $B_\varepsilon(x) \subset E$. Поскольку шар $B_\varepsilon(x)$ открыт, то любая его точка является внутренней для E, т.е. $B_\varepsilon(x) \subset \operatorname{int} E$.

Аналогично случаю $X = \mathbb{R}$ доказывается

Лемма 5.2. *Объединение произвольного семейства открытых множеств и пересечение конечного семейства открытых множеств являются открытыми множествами. Пересечение произвольного семейства замкнутых множеств и объединение конечного семейства замкнутых множеств являются замкнутыми множествами.*

Проверка замкнутости множества «по определению» не всегда удобна. Получим эквивалентные условия замкнутости.

Определение. Точка $x \in X$ называется *предельной точкой* множества E, если для любого $\varepsilon > 0$ выполнено $\mathring{B}_\varepsilon(x) \cap E \neq \varnothing$. (Здесь и далее $\mathring{B}_\varepsilon(x) = B_\varepsilon(x) \setminus \{x\}$.) Множество всех предельных точек E обозначим через E'.

Точка $x \in E$, не являющаяся предельной точкой E, называется *изолированной точкой* множества E.

Теорема 5.1 (критерии замкнутости). *Следующие утверждения эквивалентны:*

1) E замкнуто;

2) E содержит все свои граничные точки;

3) E содержит все свои предельные точки;

4) $\forall \{x_n\} \subset E \, (x_n \to x_0 \Rightarrow x_0 \in E)$.

□ (1 ⇒ 2) Пусть $x \in X \setminus E$. Так как дополнение $X \setminus E$ открыто, то существует $B_\varepsilon(x) \subset X \setminus E$. Отсюда следует, что точка x не является граничной для E. Поэтому $\partial E \subset E$.

(2 ⇒ 3) Любая предельная точка является граничной или внутренней. Поскольку внутренние точки всегда принадлежат E, получаем доказываемое утверждение.

(3 ⇒ 4) Пусть $\{x_n\} \subset E$ и $x_n \to x_0$. Если $x_0 \notin E$, то x_0 является предельной для E, что противоречит п. 3.

(4 ⇒ 1) Пусть $x \in X \setminus E$. Тогда $B_{1/n}(x) \subset X \setminus E$ для некоторого n. Иначе определена последовательность точек $x_n \in B_{1/n}(x) \cap E$, причем $x_n \to x$ и $x \notin E$, что противоречит п. 4. ■

Определение. Множество $\overline{E} = E \cup \partial E$ называется *замыканием* множества E.

Лемма 5.3. *Множество $\overline{E}$ замкнуто. Также $\overline{E} = E \cup E'$.*

□ Отметим, что поскольку $\operatorname{int} E \subset E \subset \operatorname{int} E \cup \partial E$, то $\overline{E} = \operatorname{int} E \cup \cup\, \partial E$ и, значит, $X \setminus \overline{E} = \operatorname{int}(X \setminus E)$ открыто.

Второе утверждение вытекает из двух очевидных наблюдений. Любая предельная точка является граничной или внутренней. Граничная точка, не лежащая во множестве, является предельной. ■

Следствие. *$x \in \overline{E} \Leftrightarrow$ существует последовательность $\{x_n\}$ точек из E, сходящаяся к x.*

Задача. Докажите, что $\overline{E} = \cap\{F\colon F$ замкнуто и $F \supset E\}$ и $\operatorname{int} E = \cup\{U\colon U$ открыто и $U \subset E\}$.

5.3. Подпространства метрического пространства.

Пусть $(X,\, \rho)$ — метрическое пространство, $E \subset X$ не пусто. Сужение $\rho|_{E\times E}$ очевидно задает метрику на E. Пара $(E,\, \rho|_{E\times E})$ называется *подпространством* $(X,\, \rho)$, а функция $\rho|_{E\times E}$ — *индуцированной метрикой.*

Пусть $B_r^E(x) = \{y \in E\colon \rho(y,\, x) < r\}$ — открытый шар в E. Тогда $B_r^E(x) = B_r^X(x) \cap E$, где $B_r^X(x)$ — открытый шар в X.

Оказывается, похожим образом описываются все открытые множества в E.

Лемма 5.4. *Множество $U \subset E$ открыто в $E \Leftrightarrow$ найдется такое открытое $V \subset X$, что $U = V \cap E$.*

□ Пусть множество U открыто в E. Тогда для каждой точки $x \in U$ найдется $B_{\varepsilon_x}^E(x) \subset U$ и, значит, $U = \bigcup_{x\in U} B_{\varepsilon_x}^E(x)$. Положим $V = \bigcup_{x\in U} B_{\varepsilon_x}^X(x)$. Тогда V открыто в X и $V \cap E = \bigcup_{x\in U}(B_{\varepsilon_x}^X(x) \cap E) = \bigcup_{x\in U} B_{\varepsilon_x}^E(x) = U$.

Обратно, пусть $U = V \cap E$, где V открыто в X. Если $x \in U$, то $x \in V$ и, значит, найдется $B_\varepsilon^X(x) \subset V$. Следовательно, $B_\varepsilon^E(x) = B_\varepsilon^X(x) \cap E \subset V \cap E = U$. Это доказывает, что U открыто в E. ■

Из леммы по законам де Моргана получаем

Следствие. *Множество $Z \subset E$ замкнуто в $E \Leftrightarrow$ найдется такое замкнутое $F \subset X$, что $Z = F \cap E$.*

Пример. Пусть $X = \mathbb{R}^1$ и $E = (-1,\, 3]$. Тогда множество $A = (-1,\, 1]$ замкнуто в E, множество $B = (2,\, 3]$ открыто в E, множество $C = (0,\, 2]$ не является ни открытым, ни замкнутым в E.

5.4. Компакты в метрических пространствах.

Пусть $(X,\, \rho)$ — метрическое пространство, и $K \subset X$.

Определение. Семейство $\{G_\lambda\}_{\lambda\in\Lambda}$ подмножеств X называется *покрытием* K, если $K \subset \bigcup_{\lambda\in\Lambda} G_\lambda$. Если все G_λ открыты, то покрытиеназыается *открытым.* Если $\{G_\lambda\}_{\lambda\in A}$ для некоторого

$A \subset \Lambda$ также является покрытием K, то оно называется *подпокрытием.*

Определение. Множество K называется *компактом* (в X), если из любого его открытого покрытия $\{G_\lambda\}_{\lambda\in\Lambda}$ можно выделить конечное подпокрытие, т. е. существуют $m \in \mathbb{N}$ и $\lambda_1, \ldots, \lambda_m \in \Lambda$, такие что $K \subset G_{\lambda_1} \cup \ldots \cup G_{\lambda_m}$.

Пример. Отрезок $[a, b]$ — компакт в $\mathbb{R}^1$ по теореме Гейне–Бореля.

Замечание. Множество K является компактом в (X, ρ), только если K является компактом «в себе», т. е. в подпространстве (K, ρ). Это следует из определения компактности и структуры открытых множеств подпространства.

Лемма 5.5. *Пусть (X, ρ) — метрическое пространство, и $K \subset X$. Если K — компакт, то K ограничено и замкнуто.*

□ Пусть $a \in X$. Так как $\bigcup\limits_{n=1}^{\infty} B_n(a) = X$, то $\{B_n(a)\}_{n\in\mathbb{N}}$ — открытое покрытие K. Поскольку K компакт, то существуют такие номера $n_1, \ldots, n_m$, что $K \subset B_{n_1}(a) \cup \ldots \cup B_{n_m}(a)$. Следовательно, $K \subset B_N(a)$, где $N = \max\limits_{1\leqslant k\leqslant m}\{n_k\}$. Это доказывает, что K ограничено.

Пусть $a \in X \setminus K$. Так как $\bigcup\limits_{n=1}^{\infty} \Big(X \setminus \overline{B}_{\frac{1}{n}}(a)\Big) = X \setminus \{a\}$, то $\big\{X \setminus \overline{B}_{\frac{1}{n}}(a)\big\}_{n\in\mathbb{N}}$ — открытое покрытие K. Поскольку K компакт, то

$$K \subset \left(X \setminus \overline{B}_{\frac{1}{n_1}}(a)\right) \cup \ldots \cup \left(X \setminus \overline{B}_{\frac{1}{n_m}}(a)\right) = X \setminus \overline{B}_{\frac{1}{N}}(a),$$

где $N = \max\limits_{1\leqslant k\leqslant m}\{n_k\}$. Поэтому $B_{\frac{1}{N}}(a) \subset X \setminus K$ и, значит, $X \setminus K$ открыто. ■

Лемма 5.6. *Замкнутое подмножество компакта является компактом.*

□ Пусть K — компакт и $F \subset K$ замкнуто в X. Рассмотрим открытое покрытие $\{G_\lambda\}_{\lambda\in\Lambda}$ множества F. Тогда $\{G_\lambda\}_{\lambda\in\Lambda} \cup \{X \setminus F\}$ образует открытое покрытие K, т. к. $\bigcup_{\lambda\in\Lambda} G_\lambda \cup (X \setminus F) = X$. Поскольку K — компакт, то $K \subset G_{\lambda_1} \cup \ldots \cup G_{\lambda_m} \cup (X \setminus F)$. Учитывая, что $F \subset K$, получаем $F \subset G_{\lambda_1} \cup \ldots \cup G_{\lambda_m}$. ■

Задача. Пусть $\{F_n\}$ — непустые компакты в X, и $F_1 \supset F_2 \supset \supset \ldots$. Покажите, что $\bigcap\limits_{n=1}^{\infty} F_n \neq \varnothing$ (обобщение теоремы Кантора).

Получим критерий компактности на языке последовательностей (*секвенциальная компактность*). Прежде докажем следующее полезное утверждение.

Лемма 5.7 (Лебега о покрытии). *Пусть $K \subset X$ такое, что каждая последовательность элементов из K имеет сходящуюся в K подпоследовательность. Пусть $\{G_\lambda\}_{\lambda\in\Lambda}$ — открытое покрытие множества K, тогда*

$$\exists \varepsilon > 0\ \forall x \in K\ \exists \lambda \in \Lambda\ (B_\varepsilon(x) \subset G_\lambda).$$

□ Если такого $\varepsilon > 0$ не найдется, то $\forall n \in \mathbb{N}\ \exists x_n \in K\ \forall \lambda \in \in \Lambda\ (B_{1/n}(x_n) \not\subset G_\lambda)$. Из последовательности $\{x_n\}$ выделим сходящуюся в K подпоследовательность, $x_{n_k} \to x$. Точка $x \in G_{\lambda_0}$ для некоторого $\lambda_0 \in \Lambda$. Множество G_{λ_0} открыто, поэтому существует такое $\alpha > 0$, что $B_\alpha(x) \subset G_{\lambda_0}$. Выберем k так, что $x_{n_k} \in \in B_{\frac{\alpha}{2}}(x)$ и $\frac{1}{n_k} < \frac{\alpha}{2}$. Тогда для всякого $z \in B_{\frac{1}{n_k}}(x_{n_k})$ имеем $\rho(z, x) \leqslant \rho(z, x_{n_k}) + \rho(x_{n_k}, x) < \alpha$, так что $z \in B_\alpha(x)$. Следовательно, $B_{\frac{1}{n_k}}(x_{n_k}) \subset G_{\lambda_0}$, что противоречит предположению. ■

Теорема 5.2. *Пусть (X, ρ) — метрическое пространство, и $K \subset X$. Множество K является компактом $\Leftrightarrow$ любая последовательность элементов из K имеет сходящуюся в K подпоследовательность.*

□ ($\Rightarrow$) Пусть $\{x_n\} \subset K$. Предположим, что $\{x_n\}$ не имеет сходящейся в K подпоследовательности. Другими словами, любая точка из K не является частичным пределом $\{x_n\}$. Это означает, что любая точка из K имеет окрестность, содержащую лишь конечное множество x_n, т.е.

$$\forall a \in K\, \exists \delta_a > 0\, \exists N_a\, \forall n \geqslant N_a\ (x_n \notin B_{\delta_a}(a)).$$

Семейство $\{B_{\delta_a}(a)\}_{a\in K}$ образует открытое покрытие K. Поскольку K компакт, то существуют такие $a_1, \ldots, a_m \in K$, что $K \subset B_{\delta_{a_1}}(a_1) \cup \ldots \cup B_{\delta_{a_m}}(a_m)$. Положим $N = \max\limits_{1\leqslant i\leqslant m} N_{a_i}$. Так как

$N \geqslant N_{a_i}$, то $x_N \notin B_{\delta_{a_i}}(a_i)$ для всех i. Следовательно, $x_N \notin \bigcup_{i=1}^m B_{\delta_{a_i}}(a_i)$. Значит, $x_N \notin K$, что противоречит условию.

($\Leftarrow$) Предположим, что каждая последовательность в K допускает сходящуюся подпоследовательность.

Отметим, что для любого $\varepsilon > 0$ существует конечное покрытие K открытыми шарами радиуса ε. В противном случае строим по индукции последовательность $\{x_n\}$: $x_1 \in K$, $x_n \in K \setminus \setminus \bigcup_{i=1}^{n-1} B_\varepsilon(x_i)$. По построению $\rho(x_i, x_j) \geqslant \varepsilon$ при $i \neq j$ и, значит, $\{x_n\}$ не имеет сходящейся подпоследовательности.

Пусть $\{G_\lambda\}_{\lambda \in \Lambda}$ — открытое покрытие K. По лемме 5.7

$$\exists \varepsilon > 0 \ \forall x \in K \ \exists \lambda \in \Lambda \ (B_\varepsilon(x) \subset G_\lambda).$$

Множество K покрывается конечным числом шаров радиуса ε, т.е. существуют $x_1, \ldots, x_m \in K$ такие, что $K \subset \bigcup_{i=1}^m B_\varepsilon(x_i)$. Следовательно, $K \subset \bigcup_{i=1}^m G_{\lambda_i}$, где λ_i удовлетворяет условию $B_\varepsilon(x_i) \subset G_{\lambda_i}$. Это доказывает, что K компакт. ■

В качестве следствия получим описание компактов в евклидовом пространстве $\mathbb{R}^n$. Начнем с важнейшего примера.

Пример. Брус $B = [a_1, b_1] \times \ldots \times [a_n, b_n]$ является компактом в $\mathbb{R}^n$.

□ По индукции: при $n = 1$ верно. Предполагая, что утверждение верно для n, докажем для $n+1$. Рассмотрим последовательность $\{x_k\}$ точек из $B = B_n \times [a_{n+1}, b_{n+1}]$ и пусть $x_k = (y_k, x_{n+1,k})$ (здесь $y_k \in B_n \subset \mathbb{R}^n$, а $x_{n+1,k} \in [a_{n+1}, b_{n+1}]$). Тогда по предположению существует $y_{k_i} \to y_0 \in B_n$ и ввиду компактности $[a_{n+1}, b_{n+1}]$ существует $x_{n+1, k_{i_j}} \to x_{n+1,0}$. Положим $x_0 = (y_0, x_{n+1,0})$. Так как $y_{k_{i_j}} \to y_0$, то $|x_k - x_0|^2 = |y_{k_{i_j}} - y_0|^2 + |x_{n+1, k_{i_j}} - - x_{n+1,0}|^2 \to 0$ и, значит, $x_{k_{i_j}} \to x_0 \in B$. По теореме B компакт. ■

Следствие. *Множество K в $\mathbb{R}^n$ является компактом $\Leftrightarrow$ K ограничено и замкнуто.*

□ ($\Rightarrow$) Вытекает из леммы 5.5.

($\Leftarrow$) Так как K ограничено, то $K \subset B_r(x)$ для некоторой точки $x = (x_1, \ldots, x_n)^T$ и $r > 0$. Рассмотрим замкнутый брус $[x_1 - r, x_1 + r] \times \ldots \times [x_n - r, x_n + r]$. Он содержит $B_r(x)$ и, следовательно, K. Тогда по лемме 5.6 K является компактом. ■

Следствие (теорема Больцано–Вейерштрасса). *. Любая ограниченная последовательность в $\mathbb{R}^n$ имеет сходящуюся подпоследовательность.*

□ Если последовательность ограничена, то она содержится в некотором замкнутом шаре, который (по следствию) является компактом. Осталось воспользоваться теоремой 5.2. ■

Замечание. В общих метрических пространствах ограниченность и замкнутость не влечет компактность.

Пример. $X = \mathbb{R}^1$ с дискретной метрикой, и $K = [0, 1]$. Тогда K ограничено и замкнуто, но из покрытия $\{B_{1/2}(x)\}_{x \in K}$ множества K нельзя выделить конечное подпокрытие, поскольку $B_{1/2}(x) = \{x\}$.

5.5. Полные метрические пространства.

Пусть (X, ρ) — метрическое пространство.

Определение. Последовательность $\{x_n\}$ называется *фундаментальной* в X, если

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n, m \geqslant N\ (\rho(x_n, x_m) < \varepsilon).$$

Лемма 5.8. *Всякая сходящаяся последовательность фундаментальна.*

□ Пусть $x \to a$ и $\varepsilon > 0$. Тогда найдется номер N, такой что $\rho(x_n, a) < \frac{\varepsilon}{2}$ для всех $n \geqslant N$ и, значит, $\rho(x_n, x_m) \leqslant \rho(x_n, a) + \rho(x_m, a) < \varepsilon$ для всех $n, m \geqslant N$. ■

Замечание. Обратное, вообще говоря, неверно. Например, в пространстве $X = (0, 1)$ с метрикой $\rho(x, y) = |x - y|$ последовательность $\{1/n\}$ фундаментальна, но не имеет предела в X.

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

Задача. Докажите, что подпространство полного метрического пространства является полным тогда и только тогда, когда оно замкнуто.

Теорема 5.3. *Пространство $\mathbb{R}^n$ с евклидовой метрикой полно.*

□ Пусть последовательность $\{x_k\}$, $x_k = (x_{1,k}, \ldots, x_{n,k})^T$, фундаментальна в $\mathbb{R}^n$, и $1 \leqslant i \leqslant n$. Тогда в силу неравенства $|x_{i,k} - x_{i,m}| \leqslant \rho_2(x_k, x_m)$, последовательность $\{x_{i,k}\}$ фундаментальна в $\mathbb{R}$. По критерию Коши она сходится к некоторому $a_i \in \mathbb{R}$. Положим $a = (a_1, \ldots, a_n)^T$. Тогда $\rho_2(x_k, a)^2 = \sum\limits_{i=1}^{n} |x_{i,k} - a_i|^2 \to 0$ при $k \to \infty$, т.е. $x_k \to a$. ■

Пример. Пусть $E \neq \varnothing$, и $B(E)$ — линейное пространство всех *ограниченных* функций $f\colon E \to \mathbb{R}$. Пространство $B(E)$ является нормированным относительно $\|f\| = \sup\limits_{x \in E} |f(x)|$.

Далее, рассмотрим сходимость:

$$f_n \to f \text{ в } B(E) \Leftrightarrow \|f_n - f\| \to 0$$
$$\Leftrightarrow \sup_{x \in E} |f_n(x) - f(x)| \to 0 \Leftrightarrow f_n \rightrightarrows f \text{ на } E.$$

Теорема 5.4. *Пространство $B(E)$ полно.*

□ Пусть последовательность $\{f_n\}$ фундаментальна, и $\varepsilon > 0$. Тогда найдется номер N, такой что $\sup\limits_{x \in E} |f_n(x) - f_m(x)| < \varepsilon$ для всех $n, m \geqslant N$. По критерию Коши равномерной сходимости найдется функция $f\colon E \to \mathbb{R}$, такая что $f_n \rightrightarrows f$, т. е. $\|f_n - f\| \to 0$.

Докажем ограниченность функции f. Возьмем, для определенности, $\varepsilon = 1$ и найдем такой номер N, что $|f_N(x) - f(x)| \leqslant 1$ при всех $x \in E$. Тогда $|f| \leqslant |f_N| + 1$ и, значит, $f \in B(E)$. ■

Следствие. *Пространство $C[a, b]$ непрерывных на $[a, b]$ функций является полным (как замкнутое подпространство $B[a, b]$).*

6. Непрерывные функции

6.1. Предел функции в точке.

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

Определение (Коши). Точка $b \in Y$ называется *пределом* функции f в точке a, если

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in X\ \big(0 < \rho_X(x, a) < \delta \Rightarrow \rho_Y(f(x), b) < \varepsilon\big)$$

или, что эквивалентно,

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in X\ \big(x \in \mathring{B}_\delta(a) \Rightarrow f(x) \in B_\varepsilon(b)\big).$$

Определение (Гейне). Точка $b \in Y$ называется *пределом* функции f в точке a, если

$$\forall \{x_n\},\ x_n \in X \setminus \{a\}\ \big(x_n \to a\ \Rightarrow\ f(x_n) \to b\big).$$

Как и в случае числовых функций, доказывается равносильность определений по Коши и по Гейне. В обоих случаях пишут $\lim\limits_{x\to a} f(x) = b$, или $f(x) \to b$ при $x \to a$.

Свойство 1 (единственность). Если $\lim\limits_{x\to a} f(x) = b$ и $\lim\limits_{x\to a} f(x) = c$, то $b = c$.

□ Пусть $x_n \to a$ и $x_n \neq a$. По определению Гейне $f(x_n) \to b$ и $f(x_n) \to c$. Так как последовательность в метрическом пространстве имеет не более одного предела, то $b = c$. ■

Пусть задана функция $f\colon X \setminus \{a\} \to \mathbb{R}^m$. Если $x \in X \setminus \{a\}$, то $f(x) = (y_1, \ldots, y_m)^T$, и значит, для каждого $i = 1, \ldots, m$ определена *i-я координатная функция* $f_i\colon X \setminus \{a\} \to \mathbb{R}$, $f_i(x) = y_i$. Пишут $f = (f_1, \ldots, f_m)^T$.

Лемма 6.1. *Пусть $f = (f_1, \ldots, f_m)^T$. Тогда $\lim\limits_{x\to a} f(x) = b \Leftrightarrow \lim\limits_{x\to a} f_i(x) = b_i$ для $i = 1, \ldots, m$.*

□ Вытекает из оценки $|x_i - b_i| \leqslant \rho_2(x, b) \leqslant \sum_{i=1}^m |x_i - b_i|$. ■

Примеры. 1) Пусть $f\colon \mathbb{R}^2 \setminus \{(0, 0)\} \to \mathbb{R}$, $f(x, y) = \dfrac{x^3 + y^3}{x^2 + y^2}$. Покажем, что $\lim\limits_{(x,y)\to(0,0)} f(x, y) = 0$. Действительно, т.к. $|x| \leqslant \sqrt{x^2 + y^2}$, $|y| \leqslant \sqrt{x^2 + y^2}$, то

$$|f(x, y)| \leqslant \frac{|x|^3 + |y|^3}{x^2 + y^2} \leqslant 2\sqrt{x^2 + y^2}.$$

Поэтому определение предела по Коши выполняется для $\delta = \frac{\varepsilon}{2}$.

2) Пусть $f\colon \mathbb{R}^2 \setminus \{(0, 0)\} \to \mathbb{R}$, $f(x, y) = \dfrac{xy + y^2}{x^2 + y^2}$. Предела $\lim\limits_{(x, y)\to(0, 0)} f(x, y) = 0$ не существует, т.к. $f\big(0, \frac{1}{k}\big) = 1$, $f\big(\frac{1}{k}, 0\big) = 0$ при всех $k \in \mathbb{N}$.

В дальнейшем, говоря о «пределе по подмножеству», всегда будем иметь в виду подпространство с индуцированной метрикой.

Свойство 2. Если a — предельная точка $E \subset X$, и $\lim\limits_{x\to a} f(x) = b$, то $\lim\limits_{x\to a} (f|_E)(x) = b$.

$\square$ Пусть $E \ni x_n \to a$ и $x_n \neq a$, тогда $(f|_E)(x_n) = f(x_n) \to b$. По определению Гейне $b = \lim\limits_{x\to a} (f|_E)(x)$. $\blacksquare$

Пример. Пусть $D \subset \mathbb{R}^n$ и $u \in \mathbb{R}^n$, $|u| = 1$. Пусть функция $f\colon D \to \mathbb{R}$ определена на $\{a + tu :\ 0 < t < \Delta\}$ для некоторого $\Delta > 0$. Тогда $\lim\limits_{t\to+0} f(a + tu)$ называется *пределом f в точке a по направлению u*.

По свойству 2 если $\lim\limits_{x\to a} f(x) = b$, то существует $\lim\limits_{t\to+0} f(a + tu) = b$. Следующий пример показывает, что обратное утверждение, вообще говоря, неверно.

Пример. Пусть $f\colon \mathbb{R}^2 \to \mathbb{R}$,

$$f(x, y) = \begin{cases} 1, & y = x^2,\ x > 0, \\ 0, & \text{иначе.} \end{cases}$$

Предел $\lim\limits_{t\to+0} f(t\alpha, t\beta) = 0$ для любого направления $u = (\alpha, \beta)$. Однако $\lim\limits_{(x, y)\to(0, 0)} f(x, y)$ не существует, т.к. $f(x, x^2) = 1$ при $x > 0$.

Свойство 3. Пусть a — предельная точка X и заданы функции $f, g\colon X \setminus \{a\} \to \mathbb{R}$. Если $\lim\limits_{x\to a} f(x) = b$ и $\lim\limits_{x\to a} g(x) = c$, то $\lim\limits_{x\to a} (f \pm g)(x) = b \pm c$ и $\lim\limits_{x\to a} (fg)(x) = bc$.

$\square$ Пусть $x_n \to a$ и $x_n \neq a$. По определению Гейне $f(x_n) \to b$ и $g(x_n) \to c$. Тогда по свойствам предела числовых последовательностей $f(x_n) \pm g(x_n) \to b \pm c$, $f(x_n)g(x_n) \to bc$. Осталось воспользоваться определением Гейне. $\blacksquare$

Свойство 4 (локальная ограниченность). Если существует $\lim\limits_{x \to a} f(x)$, то для некоторого $\delta > 0$ множество $f\big(\mathring{B}_\delta(a)\big)$ ограничено.

$\square$ Достаточно в определении Коши положить $\varepsilon = 1$. $\blacksquare$

Задача. Пусть X, Y — метрические пространства, причем Y полное, a — предельная точка X, и $f\colon X \setminus \{a\} \to Y$. Покажите, что $\lim\limits_{x \to a} f(x)$ существует тогда и только тогда, когда

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,\, x' \in X\ (x,\, x' \in \mathring{B}_\delta(a) \Rightarrow \rho_Y(f(x),\, f(x')) < \varepsilon).$$

6.2. Непрерывность функции.

Пусть $(X,\, \rho_X)$ и $(Y,\, \rho_Y)$ — метрические пространства, и задана функция $f\colon X \to Y$.

Определение. Функция f *непрерывна в точке* $a \in X$, если

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in X\ \big(\rho_X(x,\, a) < \delta \Rightarrow \rho_Y(f(x),\, f(a)) < \varepsilon\big)$$

или, что эквивалентно,

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in X\ \big(x \in B_\delta(a) \Rightarrow f(x) \in B_\varepsilon(f(a))\big).$$

Пример. Координатная функция $p_i\colon \mathbb{R}^n \to \mathbb{R}$, $p_i(x_1,\, \ldots,\, x_n) = x_i$, непрерывна в каждой точке $\mathbb{R}^n$ $(i = 1,\, \ldots,\, n)$. Это следует из неравенства $|x_i - a_i| \leqslant \rho_2(x,\, a)$.

Лемма 6.2. *Пусть $f\colon X \to Y$, $a \in X$. Следующие условия эквивалентны:*

(1) функция f непрерывна в точке a;

(2) $\forall \{x_n\}$, $x_n \in X$ $(x_n \to a \Rightarrow f(x_n) \to f(a))$;

(3) a — изолированная точка множества X или a — предельная точка X и $\lim\limits_{x \to a} f(x) = f(a)$.

$\square$ (1) $\Rightarrow$ (2) Выберем $\varepsilon > 0$ и соответствующее $\delta > 0$ из определения непрерывности. Если $x_n \to a$ (в X), то существует такой номер N, что $\rho_X(x_n,\, a) < \delta$ при всех $n \geqslant N$, но тогда $\rho_Y(f(x_n),\, f(a)) < \varepsilon$ при $n \geqslant N$. Это означает, что $f(x_n) \to f(a)$.

(2) $\Rightarrow$ (3) Если a — предельная точка X, то в силу (2) $\lim\limits_{x\to a} f(x) = f(a)$ по определению Гейне.

(3) $\Rightarrow$ (1) Если a изолирована, то $B_{\delta_0}(a)\cap X = \{a\}$ для некоторого $\delta_0 > 0$. Тогда для любого $\varepsilon > 0$ определение непрерывности в точке a выполняется при $\delta = \delta_0$. Пусть a предельная для X. По определению предела по Коши $\forall\varepsilon > 0\ \exists\delta > 0\ \forall x\ (0 < \rho_X(x, a) < \delta \Rightarrow \rho_Y(f(x), f(a)) < \varepsilon)$. Но последняя импликация верна и для $x = a$. Значит, функция f непрерывна в точке a. ■

Вместе со свойством 3 предела функции лемма дает

Следствие. *Если функции f, $g\colon X \to \mathbb{R}$ непрерывны в точке a, то в этой точке также непрерывны функции $f + g$, $fg\colon X \to \mathbb{R}$.*

Теорема 6.1 (о непрерывности композиции). *Пусть (X, ρ_X), (Y, ρ_Y) и (Z, ρ_Z) — метрические пространства. Если функция $f\colon X \to Y$ непрерывна в точке $a \in X$, и функция $g\colon Y \to Z$ непрерывна в точке $f(a) \in Y$, то их композиция $g\circ f\colon X \to Z$ непрерывна в точке a.*

□ Пусть $x_n \to a$. Тогда $f(x_n) \to f(a)$ и, значит, $g(f(x_n)) \to g(f(a))$. ■

Определение. Функция $f\colon X \to Y$ *непрерывна* (на X), если f непрерывна в каждой точке X.

Пример. *Многочленом* называется функция $P\colon \mathbb{R}^n \to \mathbb{R}$, $P(x) = \sum\limits_{(k_1,\ldots,k_n)} a_{k_1\ldots k_n} x_1^{k_1}\ldots x_n^{k_n}$, где суммирование ведется по конечному множеству наборов $(k_1, \ldots, k_n)$ целых неотрицательных чисел. Многочлен P непрерывен как линейная комбинация непрерывных функций $p_1^{i_1}\cdot\ldots\cdot p_n^{i_n}$, где $p_i(x) = x_i$.

Пример. Пусть $A \subset X$ непусто. Функция $d_A\colon X \to \mathbb{R}$, $d_A(x) = \inf\limits_{a\in A}\rho(x, a)$, непрерывна.

□ Покажем, что d_A непрерывна в точке $y \in X$. Для $x \in X$, $a \in A$ по неравенству треугольника имеем $\rho(y, a) \geqslant \rho(x, a) - \rho(x, y) \geqslant d_A(x) - \rho(x, y)$. Переходя к инфимуму по всем $a \in A$, получаем $d_A(y) \geqslant d_A(x) - \rho(x, y)$ или $d_A(x) - d_A(y) \leqslant \rho(x, y)$. Неравенство симметрично относительно x, y, поэтому $|d_A(x) - d_A(y)| \leqslant \rho(x, y)$. ■

Теорема 6.2 (критерий непрерывности). *Функция $f\colon X \to Y$ непрерывна на X $\Leftrightarrow$ для каждого открытого множества V из Y множество $f^{-1}(V)$ открыто в X.*

$\square$ ($\Rightarrow$) Пусть V открыто в Y. Если $x \in f^{-1}(V)$, то $f(x) \in V$ и, значит, существует такое $\varepsilon > 0$, что $B_\varepsilon(f(x)) \subset V$. Функция f непрерывна в точке x. Поэтому найдется такое $\delta > 0$, что $f(B_\delta(x)) \subset \subset B_\varepsilon(f(x))$. Отсюда следует, что $B_\delta(x) \subset f^{-1}(V)$. Это доказывает, что множество $f^{-1}(V)$ открыто.

($\Leftarrow$) Пусть $x \in X$ и $\varepsilon > 0$. Шар $B_\varepsilon(f(x))$ открыт в Y, поэтому множество $f^{-1}(B_\varepsilon(f(x)))$ открыто в X и $x \in f^{-1}(B_\varepsilon(f(x)))$. Значит, существует $\delta > 0$, что $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)))$, или $f(B_\delta(x)) \subset B_\varepsilon(f(x))$. Так как $\varepsilon > 0$ — любое, то f непрерывна в точке x. $\blacksquare$

Следствие. *Функция $f\colon X \to Y$ непрерывна на X $\Leftrightarrow$ для каждого замкнутого множества $F \subset Y$ множество $f^{-1}(F)$ замкнуто в X.*

$\square$ Следует из теоремы в силу равенства $X \setminus f^{-1}(F) = f^{-1}(Y \setminus F)$, верного для любого $F \subset Y$. $\blacksquare$

Задача. Покажите, что существует разрывная функция $f\colon X \to Y$, такая что $f(U)$ открыто для всякого открытого $U \subset \subset X$.

6.3. Непрерывные функции на компактах.

Теорема 6.3. *Если функция $f\colon K \to Y$ непрерывна, и K компакт, то $f(K)$ — компакт в Y.*

$\square$ Пусть $\{G_\lambda\}_{\lambda\in\Lambda}$ — открытое покрытие $f(K)$. Если $x \in K$, то существует такое $\lambda_0 \in \Lambda$, что $f(x) \in G_{\lambda_0}$ и, значит, $x \in f^{-1}(G_{\lambda_0})$. Следовательно, семейство $\{f^{-1}(G_\lambda)\}_{\lambda\in\Lambda}$ образует покрытие K. Это покрытие открыто по критерию непрерывности. Поскольку K компакт, то $K \subset f^{-1}(G_{\lambda_1}) \cup \ldots \cup f^{-1}(G_{\lambda_m})$.

Покажем, что $f(K) \subset G_{\lambda_1} \cup \ldots \cup G_{\lambda_m}$. Действительно, если $y \in \in f(K)$, то $y = f(x)$ для некоторого $x \in K$. Найдем такое k, что $x \in f^{-1}(G_{\lambda_k})$. Тогда, в свою очередь, $y = f(x) \in G_{\lambda_k}$. Следовательно, $f(K)$ — компакт. $\blacksquare$

Следствие (Теорема Вейерштрасса). *Если функция*

$f\colon K \to \mathbb{R}$ непрерывна, и K компакт, то существуют точки x_m, $x_M \in K$, такие что $f(x_M) = \sup\limits_{x\in K} f(x)$ и $f(x_m) = \inf\limits_{x\in K} f(x)$.

$\square$ Образ $f(K)$ является компактом в $\mathbb{R}$. Следовательно, $f(K)$ замкнуто и ограничено. Так как $f(K)$ ограничено, то $M = \sup_K f(x) \in \mathbb{R}$. Точка M является граничной для $f(K)$. Поэтому в силу замкнутости $M \in f(K)$, т.е. существует $x_M \in K$, что $f(x_M) = M$.

Для $\inf_K f(x)$ доказательство аналогично. $\blacksquare$

Определение. Пусть V — линейное пространство, $\|\cdot\|$, $\|\cdot\|^*$ нормы на V. Нормы $\|\cdot\|$ и $\|\cdot\|^*$ называются *эквивалентными*, если существуют такие $\alpha > 0$, $\beta > 0$, что $\alpha\|x\| \leqslant \|x\|^* \leqslant \beta\|x\|$ для всех $x \in V$.

Следствие. *Все нормы на конечномерном пространстве V эквивалентны.*

$\square$ Покажем сначала, что все нормы на $\mathbb{R}^n$ эквивалентны. Достаточно показать, что любая норма $\|\cdot\|$ эквивалентна евклидовой норме $\|\cdot\|_2$.

Если $x = x_1e_1 + \ldots + x_ne_n$ — разложение $x \in \mathbb{R}^n$ по стандартному базису, то по неравенствам треугольника и Коши-Буняковского-Шварца

$$\|x\| \leqslant \sum_{i=1}^{n} |x_i|\|e_i\| \leqslant \left(\sum_{i=1}^{n} \|e_i\|^2\right)^{1/2} \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} =: \beta\|x\|_2.$$

В частности, функция $\|\cdot\|$ непрерывна в $(\mathbb{R}^n, \|\cdot\|_2)$. Сфера $S = \{x \in \mathbb{R}^n\colon \|x\|_2 = 1\}$ является компактом, поэтому функция $\|\cdot\|$ по следствию 1 достигает инфимума $\alpha > 0$ на S. Тогда при $x \neq 0$ имеем $\left\|\dfrac{x}{\|x\|_2}\right\| \geqslant \alpha$ и, значит, $\|x\| \geqslant \alpha\|x\|_2$ (последнее неравенство верно и для $x = 0$). Следовательно, $\alpha\|x\|_2 \leqslant \|x\| \leqslant \beta\|x\|_2$ для всех $x \in \mathbb{R}^n$.

Пусть $x = \sum_{i=1}^{n} x_iv_i$ — разложение по базису (v_i) в V. Тогда отображение $\varphi(x) = (x_1, \ldots, x_n)^T$ задает изоморфизм между V и $\mathbb{R}^n$. Пусть $\|\cdot\|_V$, $\|\cdot\|_V^*$ — нормы на V. Определим на $\mathbb{R}^n$ нормы $\|\cdot\|$, $\|\cdot\|^*$ по формуле $\|y\| = \|\varphi^{-1}(y)\|_V$, $\|y\|^* = \|\varphi^{-1}(y)\|_V^*$. Поскольку $\|\cdot\|$, $\|\cdot\|^*$ эквивалентны, то эквивалентны и $\|\cdot\|_V$, $\|\cdot\|_V^*$. $\blacksquare$

Определение. Функция $f\colon X \to Y$ называется *равномерно непрерывной* (на X), если

$$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x,\, x' \in X\ (\rho_X(x',\, x) < \delta \Rightarrow \rho_Y(f(x'),\, f(x)) < \varepsilon).$$

Теорема 6.4 (Кантор). *Если функция $f\colon K \to Y$ непрерывна, и K компакт, то f равномерно непрерывна.*

□ Пусть $\varepsilon > 0$. По определению непрерывности $\forall a \in K\ \ \exists\delta_a > 0\ \ \forall x \in K\ \ \big(\rho_K(x,\, a) < \delta_a \Rightarrow\ \rho_Y(f(x),\, f(a)) < \frac{\varepsilon}{2}\big)$. Семейство $\{B_{\delta_a/2}(a)\}_{a\in K}$ образует открытое покрытие K. Поскольку K компакт, то $K \subset B_{\delta_{a_1}/2}(a_1) \cup \ldots \cup B_{\delta_{a_m}/2}(a_m)$.

Покажем, что $\delta = \min\limits_{1\leqslant i\leqslant m}\{\delta_{a_i}/2\}$ искомое. Пусть $\rho_K(x,\, x') < \delta$. Точка x лежит в некотором шаре $B_{\delta_{a_i}/2}(a_i)$. Поскольку $\rho_K(x',\, a_i) \leqslant \rho_K(x',\, x) + \rho_K(x,\, a_i) < \delta_{a_i}$, то $x,\, x' \in B_{\delta_{a_i}}(a_i)$, а значит,

$$\rho_Y(f(x'),\, f(x)) \leqslant \rho_Y(f(x'),\, f(a_i)) + \rho_Y(f(x),\, f(a_i)) < \varepsilon.$$

Определение. Пусть X, Y — метрические пространства. Функция $f\colon X \to Y$ называется *гомеоморфизмом*, если f биекция, а функции f и f^{-1} непрерывны.

Теорема 6.5. *Если $f\colon K \to Y$ непрерывная биекция, и K компакт, то f — гомеоморфизм.*

□ Покажем, что функция $f^{-1}\colon Y \to K$ непрерывна. Достаточно показать, что множество $(f^{-1})^{-1}(F)$ замкнуто для всякого замкнутого $F \subset K$. Это так, поскольку $(f^{-1})^{-1}(F) = f(F)$ является компактом, как образ компакта при непрерывном отображении. ■

6.4. Связные множества.

Определение. Метрическое пространство X называется *несвязным*, если существуют такие непустые открытые $U,\, V \subset X$, что $X = U \cup V$ и $U \cap V = \varnothing$.

Метрическое пространство X называется *связным*, если оно не является несвязным.

Множество $E \subset X$ называется *несвязным (связным)*, если оно несвязно (связно) как подпространство X.

Пример. Любое одноточечное множество является связным.

Замечание. Согласно устройству открытых множеств подпространства получаем, что E несвязно, если существуют открытые U, $V \subset X$, такие что $E \subset U \cup V$ и $E \cap U \neq \varnothing$, $E \cap V \neq \varnothing$, $U \cap V \cap E = \varnothing$. Отметим, что U и V можно всегда выбрать непересекающимися.

Задача. 1) Докажите, если $E \subset X$ связно, то $\overline{E}$ также связно.

2) Докажите, что если E_i связно для любого $i \in I$ и $\cap_{i \in I} E_i \neq \varnothing$, то $\cup_{i \in I} E_i$ также связно.

Покажем, что связными подмножествами $\mathbb{R}$ являются в точности промежутки. Напомним, что $I \subset \mathbb{R}$ является промежутком, если $\forall x,\ y \in I\ \forall z \in \mathbb{R}\ (x < z < y \Rightarrow z \in I)$.

Теорема 6.6. *Множество $I \subset \mathbb{R}$ связно $\Leftrightarrow$ I — промежуток.*

$\square$ ($\Rightarrow$) Если I не является промежутком, то существуют точки x, $y \in I$ и $z \in \mathbb{R}$, такие что $x < z < y$ и $z \notin I$. Рассмотрим $(-\infty,\ z) \cap I$ и $(z,\ +\infty) \cap I$. Это непустые (содержат соответственно точки x и y), непересекающиеся, открытые в I множества, объединение которых совпадает с I. Значит, множество I несвязно.

($\Leftarrow$) Предположим, что промежуток I не является связным множеством. Тогда найдутся открытые (в $\mathbb{R}$) множества U и V, такие что $I \subset U \cup V$, $I \cap U \neq \varnothing$, $I \cap V \neq \varnothing$ и $U \cap V \cap I = \varnothing$. Пусть $x \in I \cap U$ и $y \in I \cap V$. Без ограничения общности можно считать, что $x < y$ (тогда $[x,\ y] \subset I$).

Положим $S = \{z \in [x,\ y]\colon\ z \in U\}$. Так как S не пусто и ограничено, существует $c = \sup S$. В силу замкнутости отрезка $c \in [x,\ y]$. Имеем $[x,\ y] \subset I \subset U \cup V$. На отрезке $[x,\ y]$ у множеств U и V нет общих точек, поэтому $c \in U$ или $c \in V$.

Если $c \in U$, то $c \neq y$, и значит, найдется $\varepsilon > 0$, что полуинтервал $[c,\ c+\varepsilon)$ лежит одновременно в U и в $[x,\ y]$. Но тогда $[c,\ c+\varepsilon) \subset S$, что противоречит тому, что $c = \sup S$.

Если $c \in V$, то $c \neq x$, и значит, найдется $\varepsilon > 0$, что полуинтервал $(c-\varepsilon,\ c]$ лежит одновременно в V и в $[x,\ y]$. В частности, отрезок $[c-\varepsilon/2,\ c]$ не пересекается с S, что противоречит тому, что $c = \sup S$. $\blacksquare$

Теорема 6.7. *Если функция $f\colon S \to Y$ непрерывна, и множество S связно, то множество $f(S)$ связно в Y.*

$\square$ Предположим, что $f(S)$ несвязно. Тогда существуют открытые в Y множества U и V, такие что $f(S) \subset U \cup V$, $f(S) \cap U \neq \varnothing$, $f(S) \cap V \neq \varnothing$ и $U \cap V \cap f(S) = \varnothing$. Множества $f^{-1}(U)$ и $f^{-1}(V)$ не пусты, не пересекаются, открыты в S (по критерию непрерывности) и $S = f^{-1}(U) \cup f^{-1}(V)$ (т. к. U, V образуют покрытие $f(S)$). Это противоречит связности S. $\blacksquare$

Следствие (теорема о промежуточных значениях). *Если функция $f\colon S \to \mathbb{R}$ непрерывна, и множество S связно, то f принимает все промежуточные значения (т. е. если $u, v \in f(S)$ и $u < v$, то $[u, v] \subset f(S)$).*

$\square$ По теореме 6.7 множество $f(S)$ связно в $\mathbb{R}$ и, значит, по теореме 6.6 является промежутком. $\blacksquare$

Определение. Открытое связное множество в метрическом пространстве называется *областью.*

Пример. Выясним, является ли $E = \{(x, y, z) \in \mathbb{R}^3\colon\ e^{x^2+y^2} < 1 + z^2\}$ областью.

$\square$ Функция $f(x, y, z) = e^{x^2+y^2} - 1 - z^2$ непрерывна, поэтому множество $E = f^{-1}(-\infty, 0)$ открыто по критерию непрерывности. Однако E не является связным, т. к. $E \subset U \cup V$, где $U = \{(x, y, z) \in \mathbb{R}^3\colon\ z > 0\}$, $V = \{(x, y, z) \in \mathbb{R}^3\colon\ z < 0\}$, причем E пересекается и с U, и с V. $\blacksquare$

Имеется класс множеств, для которых проверка связности осуществляется несколько проще.

Определение. Метрическое пространство X называется *линейно связным*, если для любых точек $x, y \in X$ существует такая непрерывная функция $\gamma\colon [0, 1] \to X$, что $\gamma(0) = x$, $\gamma(1) = y$.

Пример. Покажем, что в нормированном пространстве шар $B_r(a)$ — линейно связное множество.

Пусть $x, y \in B_r(a)$. Определим на $[0, 1]$ функцию $\gamma\colon t \mapsto (1 - t)x + ty$. Покажем, что $\gamma(t)$ лежит в $B_r(a)$ для всех $t \in [0, 1]$.

Действительно,

$$\|\gamma(t) - a\| = \|(1-t)(x-a) + t(y-a)\| \leqslant \\ \leqslant (1-t)\|x-a\| + t\|y-a\| < (1-t)r + tr = r.$$

Теорема 6.8. *Линейно связное метрическое пространство связно.*

☐ Предположим, что линейно связное пространство X несвязно. Тогда найдутся непустые открытые множества U и V, такие что $X = U \cup V$ и $U \cap V = \varnothing$. Пусть $x \in U$ и $y \in V$. Так как X линейно связно, то существует непрерывная функция $\gamma\colon [0,\,1] \to X$, такая что $\gamma(0) = x$ и $\gamma(1) = y$. Тогда $\gamma^{-1}(U)$ и $\gamma^{-1}(V)$ не пусты, не пересекаются, открыты в $[0,\,1]$, и $[0,\,1] = \gamma^{-1}(U) \cup \gamma^{-1}(V)$, что невозможно, т. к. отрезок $[0,\,1]$ связен. ■

Лемма 6.3. *Связное открытое множество E в нормированном пространстве линейно связно.*

☐ Пусть $x \in E$. Рассмотрим множество U тех точек y, которые можно соединить с x кривой, т.е. существует непрерывная функция $\gamma\colon [0,\,1] \to E$, что $\gamma(0) = x$, $\gamma(1) = y$. Покажем, что U открыто. Для $y \in U$ в силу открытости E найдется такое $\varepsilon > 0$, что $B_\varepsilon(y) \subset E$. Любая пара точек в шаре может быть соединена отрезком: для $z \in B_\varepsilon(y)$ рассмотрим $\sigma\colon [0,\,1] \to B_\varepsilon(y)$, $\sigma(t) = (1 - t)y + tz$. Тогда кривая

$$\gamma \star \sigma(t) = \begin{cases} \gamma(2t), & 0 \leqslant t \leqslant 1/2, \\ \sigma(2t-1), & 1/2 \leqslant t \leqslant 1, \end{cases}$$

соединяет x и z и, значит, $B_\varepsilon(y) \subset U$. Также устанавливается, что $E \setminus U$ открыто. В силу связности $E \setminus U$ пусто, т.е. $E = U$. ■

Задача. Докажите, что множество $A = \{(0,\,y)\colon\ y \in [-1,\,1]\} \cup \{(x,\,\sin(1/x))\colon\ x \in (0,\,1]\}$ связно, но не линейно связно в $\mathbb{R}^2$.

6.5. Линейные отображения евклидовых пространств.

Определение. Отображение $L\colon \mathbb{R}^n \to \mathbb{R}^m$ называется *линейным*, если $\forall x_1,\, x_2 \in \mathbb{R}^n$ и $\forall \alpha_1,\, \alpha_2 \in \mathbb{R}$ выполнено $L(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 L(x_1) + \alpha_2 L(x_2)$. Множество всех линейных отображений из $\mathbb{R}^n$ в $\mathbb{R}^m$ образует линейное пространство, которое будем означать как $\mathcal{L}(\mathbb{R}^n,\, \mathbb{R}^m)$.

Пример. Пусть $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, $L(x) = Ax$ с $A = (a_{ij})$. Так как $|L(x)|^2 = \sum\limits_{i=1}^{m} (L_i, x)^2$, где $L_i = (a_{i1}, \ldots, a_{in})^T$, то по неравенству Коши–Буняковского–Шварца

$$|L(x)|^2 \leqslant \sum_{i=1}^{m} |L_i|^2 |x|^2 = |x|^2 \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2,$$

так что $|L(x)| \leqslant C|x|$ для $C = \left(\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} a_{ij}^2 \right)^{1/2}$.

Определение. Пусть $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Определим $\|L\| = \sup\limits_{x \neq 0} \dfrac{|L(x)|}{|x|}$ — *операторная норма* L.

Замечание. Согласно предыдущему примеру $\|L\| \in \mathbb{R}$. Кроме того, по определению супремума $|L(x)| \leqslant \|L\| |x|$ для всех $x \in \mathbb{R}^n$, и для всякого $\varepsilon > 0$ найдется такое $x_\varepsilon \in \mathbb{R}^n$, что $|L(x)| > (\|L\| - \varepsilon)|x_\varepsilon|$. Таким образом, $\|L\|$ — наименьшее из чисел $C > 0$, таких что $|L(x)| \leqslant C|x|$ для всех $x \in \mathbb{R}^n$. Такая переформулировка позволяет заключить, что для $\| \cdot \|$ выполнены все пункты определения нормы.

Итак, $(\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m), \| \cdot \|)$ является нормированным пространством. Нетрудно проверить, что $\|L_2 L_1\| \leqslant \|L_2\| \|L_1\|$.

7. Дифференциальное исчисление в $\mathbb{R}^n$

7.1. Дифференцируемость функции в точке.

Пусть $U \subset \mathbb{R}^n$ открыто, $a \in U$, и задана функция $f\colon U \to \mathbb{R}^m$.

Определение. Функция f называется *дифференцируемой в точке* a, если существует такое линейное отображение $L_a\colon \mathbb{R}^n \to \mathbb{R}^m$, что

$$f(a+h) = f(a) + L_a(h) + \alpha(h)|h|, \tag{7.1}$$

для некоторой функции α, такой что $\alpha(h) \to 0$ при $h \to 0$.

Линейное отображение L_a называется *дифференциалом* f в точке a и обозначается df_a.

Замечание. Так как a — внутренняя точка U, то $a + h \in U$ при всех малых по модулю $h \in \mathbb{R}^n$. Поэтому функция α определена в некоторой окрестности V точки $0 \in \mathbb{R}^n$, $\alpha\colon V \to \mathbb{R}^m$. Отметим, что формула (7.1) не определяет значение α в нуле. Будем считать, что $\alpha(0) = 0$, тем самым α непрерывна в нуле.

Формулу (7.1) можно записать в виде

$$f(a+h) = f(a) + df_a(h) + o(|h|), \quad h \to 0.$$

Замечание. Если функция f дифференцируема в точке a, то она непрерывна в a. Действительно, в силу непрерывности дифференциала из (7.1) следует, что $\lim\limits_{h\to 0} f(a+h) = f(a)$, что равносильно $\lim\limits_{x\to a} f(x) = f(a)$.

Определение. Пусть $v \in \mathbb{R}^n$, и функция f определена на множестве $\{a + tv\colon\ |t| < \Delta\}$ для некоторого $\Delta > 0$. Предел

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

называется *производной* f *по вектору* v в точке a и обозначается $\dfrac{\partial f}{\partial v}(a)$ (а также $f'_v(a)$ и $\partial_v f(a)$).

Замечание. По условию функция $\varphi(t) = f(a+tv)$ определена на интервале, содержащем 0, и $\dfrac{\partial f}{\partial v}(a) = \varphi'(0) \in \mathbb{R}^m$.

Пример. Пусть $f\colon \mathbb{R}^n \to \mathbb{R}$, $f(x) = |x|$. Пусть x, $v \in \mathbb{R}^n \setminus \{0\}$, тогда

$$\frac{\partial f}{\partial v}(x) = \frac{d}{dt}\Big|_{t=0}|x+tv| = \frac{d}{dt}\Big|_{t=0}\left(\sum_{i=1}^{n}(x_i+tv_i)^2\right)^{1/2} =$$

$$= \frac{1}{2|x|}\sum_{i=1}^{n}2x_iv_i = \left(\frac{x}{|x|},\, v\right).$$

Теорема 7.1. *Если функция $f\colon U \to \mathbb{R}^m$ дифференцируема в точке a, и $v \in \mathbb{R}^n$, то существует $\dfrac{\partial f}{\partial v}(a) = df_a(v)$.*

$\square$ Если $v = 0$, то равенство верно. Пусть $v \neq 0$. Выберем $\delta > 0$ так, что $B_\delta(a) \subset U$. Полагая в (1) $h = tv$ для всех t с $|t| < \dfrac{\delta}{|v|}$, получим

$$f(a+tv) = f(a) + df_a(tv) + \alpha(tv)|tv|.$$

В силу линейности $df_a(tv) = t df_a(v)$ и по непрерывности в нуле $\alpha(tv) \to 0$ при $t \to 0$. Поэтому

$$\frac{\partial f}{\partial v}(a) = \lim_{t\to 0}\frac{f(a+tv)-f(a)}{t} = \lim_{t\to 0}\left(df_a(v) \pm \alpha(tv)|v|\right) = df_a(v). \blacksquare$$

Следствие. *Если функция f дифференцируема в точке a, то ее дифференциал в точке a определен однозначно.*

Пример. Линейное отображение $L\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируемо в любой точке $a \in \mathbb{R}^n$ с $dL_a = L$. Это следует из равенства $L(a+h) = L(a) + L(h)$.

Непосредственно из определения вытекает *линейность* дифференцирования: если функции f, $g\colon U \to \mathbb{R}^m$ дифференцируемы в точке a и λ, $\mu \in \mathbb{R}$, то функция $\lambda f + \mu g$ также дифференцируема в точке a, и $d(\lambda f + \mu g)_a = \lambda df_a + \mu dg_a$.

7.2. Дифференцируемость в различных размерностях.
Случай функций из $\mathbb{R}$ в $\mathbb{R}^m$.

Дифференцируемость функции $\gamma\colon (\alpha, \beta) \to \mathbb{R}^m$ в точке $a \in (\alpha, \beta)$ определялась ранее как существование производной $\gamma'(a) = \lim\limits_{t\to 0}\dfrac{\gamma(a+t)-\gamma(a)}{t}$. Это согласуется с определением дифференцируемости, поскольку наличие конечного предела равносильно равенству

$$\gamma(a+t) - \gamma(a) = t\gamma'(a) + t\sigma(t),$$

где $\sigma(t) \to 0$ при $t \to 0$. Таким образом, $d\gamma_a(t) = t\gamma'(a)$.

Случай функций из $\mathbb{R}^n$ в $\mathbb{R}$.

Пусть $U \subset \mathbb{R}^n$ открыто, и задана функция $f\colon U \to \mathbb{R}$. Пусть $e_1, \ldots, e_n$ — стандартный базис в $\mathbb{R}^n$.

Определение. Производная по вектору e_k в точке a, т.е. $\dfrac{\partial f}{\partial e_k}(a) = \lim\limits_{t \to 0} \dfrac{f(a + te_k) - f(a)}{t}$, называется *частной производной* функции f по переменной x_k в точке a и обозначается $\dfrac{\partial f}{\partial x_k}(a)$ (а также $f'_{x_k}(a)$ и $\partial_k f(a)$).

Согласно определению $\dfrac{\partial f}{\partial x_k}(a)$ есть обычная производная функции $t \mapsto f(a_1, \ldots, a_{k-1}, t, a_{k+1}, \ldots, a_n)$ в точке a_k. В этом и состоит правило вычисления частной производной.

Следующее утверждение дает *необходимые условия дифференцируемости.*

Следствие. *Если $f\colon U \to \mathbb{R}$ дифференцируема в точке a, то она непрерывна в этой точке и имеет частные производные $\dfrac{\partial f}{\partial x_k}(a)$, $k = 1, \ldots, n$. Кроме того, $df_a(h) = \sum\limits_{k=1}^{n} \dfrac{\partial f}{\partial x_k}(a)\, h_k$ для всех $h \in \mathbb{R}^n$.*

□ Как отмечалось, из дифференцируемости следует непрерывность в точке и по теореме 1 существуют $\dfrac{\partial f}{\partial x_k}(a) = df_a(e_k)$. Тогда из линейности дифференциала получаем

$$df_a(h) = df_a\left(\sum_{k=1}^{n} h_k e_k\right) = \sum_{k=1}^{n} h_k df_a(e_k) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a)\, h_k. \blacksquare$$

Замечание. Дифференциал коорд-ной функции $p_k(x_1, \ldots, x_n) = x_k$ не зависит от точки. Обозначим его через dx_k. Тогда $dx_k(h) = h_k$ для любого $h \in \mathbb{R}^n$. Линейные функции $dx_1, \ldots, dx_n$ образуют в пространстве $(\mathbb{R}^n)^*$ базис, двойственный к $e_1, \ldots, e_n$. Это приводит к функциональной записи дифференциала:

$$df_a = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a)\, dx_k.$$

На значение $df_a(h)$ можно смотреть как на скалярное произведение вектора h и вектора, составленного из частных производных.

Определение. Вектор $\left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)^T$ называется *градиентом* функции f в точке a и обозначается $\operatorname{grad} f(a)$ или $\nabla f(a)$ (последний символ читается «набла f»).

С помощью градиента условие дифференцируемости (при $m = 1$) можно переписать в виде

$$f(a+h) = f(a) + (\operatorname{grad} f(a),\, h) + o(|h|).$$

Следствие. *Если f дифференцируема в точке a и $\operatorname{grad} f(a) \neq 0$, то для любого $v \in \mathbb{R}^n$ с $|v| = 1$ выполнено*

$$\left|\frac{\partial f}{\partial v}(a)\right| \leqslant |\operatorname{grad} f(a)|,$$

причем равенство имеет место лишь при $v = \pm\frac{\operatorname{grad} f(a)}{|\operatorname{grad} f(a)|}$.

$\square$ Поскольку

$$\frac{\partial f}{\partial v}(a) = df_a(v) = (\operatorname{grad} f(a),\, v),$$

то по неравенству КБШ $\left|\frac{\partial f}{\partial v}(a)\right| \leqslant |\operatorname{grad} f(a)| \cdot |v| = |\operatorname{grad} f(a)|$, причем равенство достигается лишь в случае коллинеарности $\operatorname{grad} f(a)$ и v, т. е. при $v = \pm\frac{\operatorname{grad} f(a)}{|\operatorname{grad} f(a)|}$. $\blacksquare$

Замечание. Так как $\frac{\partial f}{\partial v}(a) = \frac{d}{dt}\big|_{t=0} f(a+tv)$, то следствие дает характеристику градиента: ненулевой градиент направлен в сторону наибольшего роста функции, а его длина равна максимальной скорости роста. Это свойство градиента лежит в основе ряды вычислительных методов.

Утверждение, обратное к теореме 7.1, вообще говоря, неверно. Следующий пример показывает, что существование производных по всем векторам при $n > 1$ не гарантирует даже непрерывность функции.

Пример. Пусть $f\colon \mathbb{R}^2 \to \mathbb{R}$, где

$$f(x,\, y) = \begin{cases} 1, & y = x^2,\ x > 0 \\ 0, & \text{иначе.} \end{cases}$$

Тогда $\frac{\partial f}{\partial v}(0,\, 0) = 0$ для любого $v \in \mathbb{R}^2$, но функция f разрывна в точке $(0,\, 0)$.

Тем не менее, в терминах частных производных можно получить довольно простой признак дифференцируемости.

Теорема 7.2. *Пусть* $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$. *Если все частные производные* $\dfrac{\partial f}{\partial x_k}$ *определены в окрестности точки* $a \in U$ *и непрерывны в точке* a, *то* f *дифференцируема в точке* a.

$\square$ Пусть все $\dfrac{\partial f}{\partial x_k}$ определены в $B_r(a) \subset U$. Рассмотрим $h = (h_1, \ldots, h_n)^T$ с $|h| < r$ и определим точки $x_0 = a$, $x_k = x_{k-1} + h_k e_k$. Тогда $x_n = a + h$ и

$$f(a+h) - f(a) = \sum_{k=1}^{n} \big(f(x_k) - f(x_{k-1})\big) = \\ = \sum_{k=1}^{n} \big(f(x_{k-1} + h_k e_k) - f(x_{k-1})\big).$$

Функция $g(t) = f(x_{k-1} + te_k) - f(x_{k-1})$ на отрезке с концами 0 и h_k (при $h_k \neq 0$) имеет производную $g'(t) = \dfrac{\partial f}{\partial x_k}(x_{k-1} + te_k)$. По теореме Лагранжа о среднем $g(h_k) - g(0) = g'(\xi_k)h_k$ для некоторого ξ_k между 0 и h_k. Положим $c_k(h) = x_{k-1} + \xi_k e_k$ и отметим, что $c_k \to a$ при $h \to 0$. Теперь $f(x_k) - f(x_{k-1}) = \dfrac{\partial f}{\partial x_k}(c_k)h_k$ и, значит,

$$f(a+h) - f(a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(c_k)h_k.$$

Вычитая из обеих частей равенства сумму $\sum_{k=1}^{n} \dfrac{\partial f}{\partial x_k}(a)h_k$, получим

$$f(a+h) - f(a) - \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a)h_k = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(c_k) - \frac{\partial f}{\partial x_k}(a)\right) h_k = \\ = \sum_{k=1}^{n} \left(\frac{\partial f}{\partial x_k}(c_k) - \frac{\partial f}{\partial x_k}(a)\right) \frac{h_k}{|h|}|h| =: \alpha(h)|h|.$$

В силу непрерывности $\dfrac{\partial f}{\partial x_k}$ в точке a и неравенства $|h_k| \leqslant |h|$ имеем $\alpha(h) \to 0$ при $h \to 0$. Следовательно, функция f дифференцируема в точке a. $\blacksquare$

Вернемся к общему *случаю функций из* $\mathbb{R}^n$ *в* $\mathbb{R}^m$.

Пусть $U \subset \mathbb{R}^n$ открыто, и функция $f\colon U \to \mathbb{R}^m$, $f = (f_1, \ldots, f_m)^T$.

Лемма 7.1. *Функция f дифференцируема в точке $a \Leftrightarrow$ все ее координатные функции f_i дифференцируемы в точке a.*

$\square$ Пусть f дифференцируема в точке a. Запишем равенство (7.1) покоординатно:

$$f_i(a+h) = f_i(a) + L_i(h) + \alpha_i(h)|h|. \tag{7.2}$$

Координатные функции L_i дифференциала L_a линейны, а условие "$\alpha(h) \to 0$ при $h \to 0$" равносильно условиям "$\alpha_i(h) \to 0$ при $h \to 0$", $i = 1, \ldots, m$. Поэтому функция f_i дифференцируема в точке a и $d(f_i)_a = L_i$.

Обратно, если выполнены условия (7.2)) с линейными функциями L_i и $\alpha_i(h) \to 0$, то выполнено равенство (7.1) с отображением $L_a = (L_1, \ldots, L_m)^T$ и $\alpha = (\alpha_1, \ldots, \alpha_m)^T$. $\blacksquare$

Для дифференциала df_a, как линейного отображения, можно рассмотреть его матрицу $Df(a)$ в стандартных базисах.

Определение. Пусть $f\colon U \to \mathbb{R}^m$ дифференцируема в точке a. Матрица $Df(a)$ размера $m \times n$, такая что $df_a(h) = Df(a)h$ для всех $h \in \mathbb{R}^n$, называется *матрицей Яко́би* функции f в точке a.

Из леммы 7.1 следует, что $df = (df_1, \ldots, df_m)^T$. Тогда по следствию 1 ij-й элемент матрицы Якоби в точке a равен значению $d(f_i)_a(e_j)$, т.е. $\dfrac{\partial f_i}{\partial x_j}(a)$. Поэтому

$$Df(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \ldots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \ldots & \dfrac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

Таким образом, строками матрицы Якоби f в точке a в точности являются градиенты ее координатных функций в этой точке.

Пример. Пусть $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x+y, xy)^T$. Тогда $Df(x, y) = \begin{pmatrix} 1 & 1 \\ y & x \end{pmatrix}$.

7.3. Дифференцирование композиции.

Теорема 7.3. *Пусть U открыто в $\mathbb{R}^n$, а V открыто в $\mathbb{R}^m$. Если функция $f\colon U \to \mathbb{R}^m$ дифференцируема в точке a, функция $g\colon V \to \mathbb{R}^k$ дифференцируема в точке $f(a)$, и $f(U) \subset V$, то композиция $g \circ f\colon U \to \mathbb{R}^k$ дифференцируема в точке a, причем $d(g \circ f)_a = dg_{f(a)} \circ df_a$.*

□ Положим $b = f(a)$. По определению дифференцируемости

$$f(a+h) = f(a) + df_a(h) + \alpha(h)|h|, \quad \alpha(h) \to 0 \text{ при } h \to 0,$$
$$g(b+u) = g(b) + dg_b(u) + \beta(u)|u|, \quad \beta(u) \to 0 \text{ при } u \to 0.$$

Подставляя во второе равенство вместо u выражение $\varkappa(h) = df_a(h) + \alpha(h)|h|$, получаем

$$\begin{aligned} g\big(f(a+h)\big) = g\big(b + \varkappa(h)\big) &= \\ &= g(b) + dg_b\Big(df_a(h) + \alpha(h)|h|\Big) + \beta(\varkappa(h))|\varkappa(h)| = \\ &= g(b) + dg_b\big(df_a(h)\big) + |h| dg_b\big(\alpha(h)\big) + \beta(\varkappa(h))|\varkappa(h)| = \\ &= g(f(a)) + dg_b\, df_a(h) + \gamma(h)|h|, \end{aligned}$$

где $\gamma(h) = dg_b(\alpha(h)) + \beta(\varkappa(h))\dfrac{|\varkappa(h)|}{|h|}$. Для завершения доказательства нужно показать, что функция γ является бесконечно малой при $h \to 0$.

По теореме о непрерывности композиции функции $dg_b(\alpha(h))$ и $\beta(\varkappa(h))$ непрерывны в нуле и их значения там равны 0. Существует такое $C > 0$, что $|df_a(h)| \leqslant C|h|$, откуда заключаем, что дробь $\dfrac{|\varkappa(h)|}{|h|}$ ограничена в некоторой проколотой окрестности 0. Поэтому $\gamma(h) \to 0$ при $h \to 0$ как сумма двух бесконечно малых. ■

Следствие. *Если $f,\, g\colon U \to \mathbb{R}$ дифференцируемы в точке a, то функции $f \cdot g$, $\dfrac{f}{g}$ (при условии $g \neq 0$ на U) также дифференцируемы в точке a, причем $d(f \cdot g)_a = g(a)\, df_a + f(a)\, dg_a$ и $d\big(\dfrac{f}{g}\big)_a = \dfrac{g(a)\, df_a - f(a)\, dg_a}{g^2(a)}$.*

□ Функция $h\colon U \to \mathbb{R}^2$, $h = (f,\, g)^T$, дифференцируема в точке a и $dh_a = (df_a,\, dg_a)^T$, функция $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$, $\varphi(x,\, y) = xy$, дифференцируема в каждой точке $\mathbb{R}^2$ и $d(xy) = xdy + ydx$. Тогда композиция $\varphi\circ h$ дифференцируема в точке a и $d(\varphi\circ h)_a = d\varphi_{h(a)}\circ dh_a$, т.е. $d(f\cdot g)_a = g(a)\,df_a + f(a)\,dg_a$.

Доказательство для $\dfrac{f}{g}$ аналогично. ∎

Поскольку матрица композиции линейных отображений равна произведению матриц, то из доказанной теоремы получаем, что матрица Якоби $g\circ f$ является произведением матриц Якоби g и f. Рассмотрим подробнее случай $k=1$, т.е случай числовой функции g:

$$\left(\frac{\partial(g\circ f)}{\partial x_1}(a)\dots\frac{\partial(g\circ f)}{\partial x_n}(a)\right) =$$

$$= \left(\frac{\partial g}{\partial y_1}(b)\dots\frac{\partial g}{\partial y_m}(b)\right)\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \dots & \dots & \dots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a)\end{pmatrix}.$$

Откуда следует, что $\dfrac{\partial(g\circ f)}{\partial x_j}(a) = \sum\limits_{i=1}^{m}\dfrac{\partial g}{\partial y_i}(b)\cdot\dfrac{\partial f_i}{\partial x_j}(a)$.

Запишем дифференциал композиции:

$$d(g\circ f)_a = \frac{\partial(g\circ f)}{\partial x_1}(a)\,dx_1 + \dots + \frac{\partial(g\circ f)}{\partial x_n}(a)\,dx_n =$$

$$= \left(\sum_{i=1}^{m}\frac{\partial g}{\partial y_i}(b)\frac{\partial f_i}{\partial x_1}(a)\right)dx_1 + \dots + \left(\sum_{i=1}^{m}\frac{\partial g}{\partial y_i}(b)\frac{\partial f_i}{\partial x_n}(a)\right)dx_n$$

$$= \sum_{i=1}^{m}\frac{\partial g}{\partial y_i}(b)\left(\sum_{j=1}^{n}\frac{\partial f_i}{\partial x_j}(a)\,dx_j\right) = \sum_{i=1}^{m}\frac{\partial g}{\partial y_i}(b)df_{i,\,a}.$$

Это свойство носит название *инвариантности формы первого дифференциала* относительно замен переменных.

7.4. Частные производные и дифференциалы высших порядков.

Пусть $U\subset\mathbb{R}^n$ открыто и задана функция $f\colon U\to\mathbb{R}$ и $a\in U$, $k\in\mathbb{N}$.

Определение. Частной производной 0-го порядка в точке a считаем $f(a)$. Пусть частная производная $\dfrac{\partial^{k-1}f}{\partial x_{i_{k-1}}\dots\partial x_{i_1}}$ порядка

$(k-1)$ определена в некоторой окрестности точки a и имеет в точке a производную по переменной x_{i_k}. Тогда

$$\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a) = \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \dots \partial x_{i_1}}(x) \right) \bigg|_{x=a}$$

называется *частной производной порядка* k функции f в точке a.

Если $i_k = \dots = i_1$, т.е. $x_{i_k}, \dots, x_{i_1}$ — одна и та же переменная, то такую производную называют *чистой* и обозначают $\dfrac{\partial^k f}{\partial x_{i_1}^k}(a)$. Если среди переменных $x_{i_k}, \dots, x_{i_1}$ встречаются различные, то частную производную $\dfrac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a)$ называют *смешанной*.

Вопрос о равенстве смешанных производных изучим для случая двух переменных, x и y.

Теорема 7.4 (Юнг). *Пусть $U \subset \mathbb{R}^2$ открыто, и функция $f\colon U \to \mathbb{R}$. Если частные производные $\dfrac{\partial f}{\partial x}$ и $\dfrac{\partial f}{\partial y}$ определены в окрестности точки (a, b) и дифференцируемы в (a, b), то*

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

$\square$ Выберем окрестность $B_\delta(a, b)$, в которой определены $\dfrac{\partial f}{\partial x}$ и $\dfrac{\partial f}{\partial y}$. Рассмотрим выражение

$$\Delta(t) = f(a+t, b+t) - f(a+t, b) - f(a, b+t) + f(a, b), \;\; 0 < |t| < \delta.$$

Отметим, что $\Delta(t) = g(t) - g(0)$, где $g(s) = f(a+s, b+t) - f(a+s, b)$. Функция g на отрезке с концами 0 и t имеет производную $g'(s) = \dfrac{\partial f}{\partial x}(a+s, b+t) - \dfrac{\partial f}{\partial x}(a+s, b)$. По теореме Лагранжа $g(t) - g(0) = g'(\xi)t$ для некоторого ξ между 0 и t. Следовательно, $\Delta(t) = \frac{\partial f}{\partial x}(a+\xi, b+t)t - \frac{\partial f}{\partial x}(a+\xi, b)t$. Воспользуемся дифференцируемостью $\dfrac{\partial f}{\partial x}$ в точке (a, b):

$$\begin{aligned}\Delta(t) &= \left[\tfrac{\partial f}{\partial x}(a, b) + \tfrac{\partial^2 f}{\partial^2 x}(a, b)\xi + \tfrac{\partial^2 f}{\partial y \partial x}(a, b)t + \alpha(t)\sqrt{\xi^2 + t^2} \right] t - \\ &\quad - \left[\tfrac{\partial f}{\partial x}(a, b) + \tfrac{\partial^2 f}{\partial^2 x}(a, b)\xi + \beta(t)|\xi| \right] t = \\ &= \left[\tfrac{\partial^2 f}{\partial y \partial x}(a, b) \pm \alpha(t)\sqrt{1 + \tfrac{\xi^2}{t^2}} \pm \beta(t)\tfrac{|\xi|}{|t|} \right] t^2,\end{aligned}$$

где $\alpha(t) \to 0$, $\beta(t) \to 0$ при $t \to 0$. Следовательно, существует $\lim\limits_{t\to 0} \dfrac{\Delta(t)}{t^2} = \dfrac{\partial^2 f}{\partial y \partial x}(a,\, b)$.

Так как $\Delta(t) = h(t) - h(0)$, где $h(s) = f(a+t,\, b+s) - f(a,\, b+ + s)$, то аналогичные выкладки приводят к равенству $\lim\limits_{t\to 0} \dfrac{\Delta(t)}{t^2} = = \dfrac{\partial^2 f}{\partial x \partial y}(a,\, b)$. Теперь утверждение следует из единственности предела. ■

Распространим теорему на случай n переменных.

Следствие. *Пусть $k \in \mathbb{N}$, $k \geqslant 2$. Если все частные производные до порядка $k-2$ дифференцируемы в некоторой окрестности точки a, а все частные производные порядка $k-1$ дифференцируемы в точке a, то*

$$\frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(a) = \frac{\partial^k f}{\partial x_{j_k} \ldots \partial x_{j_1}}(a)$$

при условии, что списки $(i_1,\, \ldots,\, i_k)$ и $(j_1,\, \ldots,\, j_k)$ отличаются лишь порядком.

□ Индукция по k. При $k = 2$ зафиксируем все переменных, кроме двух, положив $x_r = a_r$, $r \neq i_1, i_2$. Тогда имеем функцию переменных x_{i_1} и x_{i_2}, и равенство вытекает из теоремы 4.

Пусть $k > 2$. Рассмотрим сначала случай, когда список $(j_1,$ $\ldots,\, j_k)$ получен из $(i_1,\, \ldots,\, i_k)$ с помощью одной транспозиции, т. е. перестановкой i_r и i_{r-1}.

Определим функцию $g = \dfrac{\partial^{r-2} f}{\partial x_{i_{r-2}} \ldots \partial x_{i_1}}$. По теореме 4 в окрестности точки a имеет место равенство $\dfrac{\partial^2 g}{\partial x_{i_r}\, \partial x_{i_{r-1}}} = \dfrac{\partial^2 g}{\partial x_{i_{r-1}}\, \partial x_{i_r}}$. При $r = k$ имеем $\dfrac{\partial^2 g}{\partial x_{i_r}\, \partial x_{i_{r-1}}}(a) = \dfrac{\partial^2 g}{\partial x_{i_{r-1}}\, \partial x_{i_r}}(a)$, что лишь формой записи отличается от требуемого равенства; при $r < k$ еще надо продифференцировать по переменным $x_{i_{r+1}}, \ldots,$ x_{i_k} и подставить $x = a$. Поскольку любая перестановка порождается транспозициями соседних элементов, доказательства следует и в общем случае. ■

При определении дифференциалов высших порядков нам потребуется следующее понятие.

Определение. Функция $\varphi\colon (\mathbb{R}^n)^k \to \mathbb{R}$ называется *k-линейной*, если φ линейна по каждому аргументу, т.е. $\varphi(\ldots,\, \lambda v + \mu w,\, \ldots) = \lambda\varphi(\ldots,\, v,\, \ldots) + \mu\varphi(\ldots,\, w,\, \ldots)$.

Функция φ через компоненты векторов-аргументов записывается многочленом степени $\leqslant k$. В частности, любая k-линейная функция непрерывна.

Пусть $U \subset \mathbb{R}^n$ открыто, и задана функция $f\colon U \to \mathbb{R}$.

Определение. Положим $d^1 f_a = df_a$. Пусть дифференциал $d^{k-1}f$ определен в некоторой окрестности точки a. Тогда *дифференциалом k-го порядка* функции f в точке a называется k-линейная функция, определяемая формулой

$$d^k f_a(v_1,\, \ldots,\, v_k) = d\big(d^{k-1} f(v_1,\, \ldots,\, v_{k-1})\big)_a(v_k), \quad v_i \in \mathbb{R}^n.$$

Если $d^k f_a$ существует, то функция f называется *k раз дифференцируемой* в точке a.

Таким образом, чтобы получить значение $d^k f_a$ на векторах v_1, $\ldots$, v_k, нужно рассмотреть функцию $x \mapsto \varphi(x) = df_x^{k-1}(v_1,\, \ldots,\, v_{k-1})$, определенную в окрестности a, и взять ее дифференциал в точке a на векторе v_k, т.е. $d^k f_a(v_1,\, \ldots,\, v_k) = d\varphi_a(v_k)$.

Вопрос о порядке векторов в определении снимается следующим фактом.

Лемма 7.2. *Дифференциал $d^k f$ симметричен, т.е. на наборах k векторов, отличающихся лишь порядком, принимает одинаковые значения.*

$\square$ Достаточно установить совпадение на наборах векторов стандартного базиса и воспользоваться линейностью.

Покажем по индукции, что $d^k f_a(e_{i_1},\, \ldots,\, e_{i_k}) = \dfrac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(a)$. При $k = 1$ это следует из теоремы 1 и определения частной производной. Пусть утверждение верно для $k-1$, т.е. $d^{k-1} f(e_{i_1},\, \ldots,\, e_{i_{k-1}}) = \dfrac{\partial^{k-1} f}{\partial x_{i_{k-1}} \ldots \partial x_{i_1}}$. Тогда, дифференцируя полученную функцию в точке a, имеем

$$d^k f_a(e_{i_1},\, \ldots,\, e_{i_k}) = d\left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \ldots \partial x_{i_1}}\right)_a (e_{i_k}) =$$

$$= \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \ldots \partial x_{i_1}} \right) \Bigg|_{x=a} = \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(a).$$

Симметричность $d^k f$ на наборах базисных векторов теперь вытекает по следствию теоремы 7.4. ■

Из доказательства получаем

Следствие. *Функция f дифференцируема k раз в точке a тогда и только тогда, когда частные производные до порядка $k-2$ дифференцируемы в некоторой окрестности точки a, и все частные производные порядка $k-1$ дифференцируемы в точке a.*

Лемма 7.2 наряду с k-линейным отображением $d^k f_a$ позволяет рассматривать соответствующую k-форму $h \mapsto d^k f_a(h, \ldots, h) =: d^k f_a(h)$. Форма $d^k f_a(h)$ является однородным многочленом степени k от компонент вектора $h = (h_1, \ldots, h_n)^T$:

$$d^k f_a(h) = \sum_{i_k=1}^{n} \ldots \sum_{i_1=1}^{n} \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(a)\, h_{i_1} \ldots h_{i_k}.$$

Часто под дифференциалом $d^k f$ понимают именно k-форму.

Теперь мы можем доказать многомерный вариант формулы Тейлора.

Теорема 7.5 (Лагранж). *Пусть функция $f\colon U \to \mathbb{R}$ $(p+1)$ раз дифференцируема на открытом U в $\mathbb{R}^n$. Если $a \in U$, $h \in \mathbb{R}^n$, такие что отрезок $[a,\, a+h] \subset U$, то найдется $\theta \in (0,\, 1)$, что*

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{1}{k!}\, d^k f_a(h) + \frac{1}{(p+1)!}\, d^{p+1} f_{a+\theta h}(h).$$

□ Рассмотрим функцию $t \mapsto g(t) = f(a+th)$. Так как отрезок $[a,\, a+h] = \{a+th\colon t \in [0,\, 1]\}$ лежит в U, то g определена на интервале, содержащем $[0,\, 1]$. Учитывая, что функция f на U является $p+1$ раз дифференцируемой, по следствию леммы 7.2 для любого $t \in [0,\, 1]$ имеем

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+th) h_i,$$
$$g''(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(a+th) h_i h_j$$

и так далее. Получаем

$$g^{(k)}(t) = \sum_{i_k=1}^{n} \ldots \sum_{i_1=1}^{n} \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(a+th)h_{i_1} \ldots h_{i_k}$$

при $k = 1, \ldots, p+1$. Сравнивая с выражения для $d^k f$, приходим к равенствам

$$g^{(k)}(t) = d^k f_{a+th}(h), \quad k = 1, \ldots, p+1.$$

Запишем для g формулу Тейлора с остаточным членом в форме Лагранжа:

$$g(t) = g(0) + \sum_{k=1}^{p} \frac{g^{(k)}(0)}{k!} t^k + \frac{g^{(p+1)}(\theta_t)}{(p+1)!} t^{p+1}, \quad \theta_t \in (0, 1).$$

Теперь искомая формула получается при $t = 1$ и $\theta = \theta_1$. ■

Перейдем к обсуждению формулы Тейлора с остаточным членом в форме Пеано.

Прежде отметим, что k-линейные функции образуют линейное пространство и $\|\varphi\| = \sup\limits_{|x_1|=\ldots=|x_k|=1} |\varphi(x_1, \ldots, x_k)|$ задает на нем норму (sup можно заменить на max, поскольку φ непрерывна и множество, определяемое $|x_1| = \ldots = |x_k| = 1$, является компактом). Так как для ненулевых векторов $|\varphi(x_1, \ldots, x_k)| = |\varphi(\frac{x_1}{|x_1|}, \ldots, \frac{x_k}{|x_k|})||x_1| \cdot \ldots \cdot |x_k|$, то справедлива оценка

$$|\varphi(x_1, \ldots, x_k)| \leqslant \|\varphi\| \, |x_1| \cdot \ldots \cdot |x_k|.$$

Такая оценка очевидно верна и для случая, когда хотя бы один из векторов нулевой.

Замечание. Разлагая аргументы x_i по стандартному базису, по линейности получаем
$\varphi(x_1, \ldots, x_k) = \sum\limits_{(i_1, \ldots, i_k)} \varphi(e_{i_1}, \ldots, e_{i_k}) x_{1i_1} \ldots x_{ki_k}$. Если длины векторов равны 1, то их компоненты $\leqslant 1$ по модулю. Следовательно, справедлива оценка

$$|\varphi(e_{j_1}, \ldots, e_{j_k})| \leqslant \|\varphi\| \leqslant \sum_{(i_1, \ldots, i_k)} |\varphi(e_{i_1}, \ldots, e_{i_k})|. \tag{7.3}$$

Лемма 7.3. *Пусть φ — k-линейная симметрическая функция, и $\Phi\colon \mathbb{R}^n \to \mathbb{R}$, $\Phi(x) = \varphi(x, \ldots, x)$. Тогда функция Φ дифференцируема и $d\Phi_x(h) = k\varphi(x, \ldots, x, h)$.*

□ Имеем

$$\Phi(x+h) - \Phi(x) = \varphi(x+h, \ldots, x+h) - \varphi(x, \ldots, x) = \\ = k\varphi(x^{k-1}, h) + \text{слагаемые } \varphi(x^p, h^q),$$

где $p + q = k$, $q \geqslant 2$ ((x^p, h^q) – сокращенная запись p аргументов x, q аргументов h).

Так как $|\varphi(x^p, h^q)| \leqslant C|x|^p|h|^q$ для некоторой константы $C \geqslant 0$ и $q \geqslant 2$, то $\varphi(x^p, h^q) = o(|h|)$ при $h \to 0$, что доказывает утверждение. ■

Теорема 7.6. *Если функция $f\colon U \to \mathbb{R}$ p раз дифференцируема в точке a, то*

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{1}{k!}\, d^k f_a(h) + o\big(|h|^p\big), \;\; h \to 0.$$

□ Индукция по p. При $p = 1$ равенство верно по определению дифференцируемости. Предположим, утверждение верно при $p - 1$. Рассмотрим функцию

$$g(x) = f(a+x) - f(a) - df_a(x) - \frac{1}{2}d^2 f_a(x) - \ldots - \frac{1}{p!}d^p f_a(x).$$

Зафиксируем $v \in \mathbb{R}^n$. По лемме 7.3 $d(d_a^k f(x))_x(v) = kd^k f_a(x, \ldots, x, v)$. Поэтому

$$dg_x(v) = df_{a+x}(v) - df_a(v) - d^2 f_a(x, v) - \ldots - \frac{1}{(p-1)!}d^p f_a(x, \ldots, x, v).$$

Применим предположение индукции к функции $y \mapsto df_y(v)$:

$$df_{a+x}(v) = df_a(v) + d^2 f_a(x, v) + \ldots + \\ + \frac{1}{(p-1)!}d^p f_a\big(x, \ldots, x, v\big) + \alpha(x, v)|x|^{p-1},$$

где $\alpha(x, v) \to 0$ при $x \to 0$. Сравнивая последние равенства, получаем, что $dg_x(v) = \alpha(x, v)|x|^{p-1}$. Функция $|dg_x|$ при фиксированном x непрерывна, поэтому по теореме Вейерштрасса существует $|dg_x(v_0)| = \max_{|v|=1} |dg_x(v)|$.

Зафиксируем $\varepsilon > 0$. Найдем такое $\delta > 0$, что $|dg_x(v_0)| \leqslant \varepsilon |x|^{p-1}$ при всех $x \in \mathbb{R}^n$, $|x| < \delta$. Применим предыдущую теорему при $p = 0$ в шаре $B_\delta(0)$:

$$|g(h)| = |g(h) - g(0)| = |dg_{\theta h}(h)|, \;\; \theta \in (0,\, 1).$$

Так как $|dg_{\theta h}(h)| = \left|dg_{\theta h}\left(\frac{h}{|h|}\right)\right||h| \leqslant |dg_{\theta h}(v_0)||h|$, то $|g(h)| \leqslant \leqslant \varepsilon |h|^{p-1}|h|$ и, значит, $g(h) = o(|h|^p)$ при $h \to 0$. ■

Определение. Функция f называется *k раз непрерывно дифференцируемой* на U, если все ее частные производные k-го порядка непрерывны на U. Множество всех k раз непрерывно дифференцируемых на U функций обозначим как $C^k(U)$.

Лемма 7.4. *Функция $f \in C^k(U) \Leftrightarrow f$ k раз дифференцируема в каждой точке U и отображение $d^k f$ из U в пространство k-линейных функций непрерывно.*

□ Отметим, что непрерывность частных производных k-го порядка влечет по теореме 7.2 дифферецируемость частных производных $(k-1)$-го порядка, т.е. k раз дифференцируемость. Так как $d^k f_a(e_{i_1},\, \ldots,\, e_{i_k}) = \dfrac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}$, то из оценки (7.3), примененной к $\varphi = d^k f_x - d^k f_a$, получаем

$$\lim_{x \to a} \|df_x - df_a\| = 0 \Leftrightarrow \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(x) \to \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(a),$$

что завершает доказательство. ■

8. Мера Лебега в $\mathbb{R}^n$

8.1. Объем бруса.

Определение. *Брусом* в $\mathbb{R}^n$ называется множество вида

$$B = I_1 \times \ldots \times I_n,$$

где все I_k — ограниченные промежутки. Если $a_k \leqslant b_k$ — концы I_k, то *объемом* бруса B называется число $|B| = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$.

Если все I_k — интервалы (отрезки), то брус B называется *открытым* (*замкнутым*).

Если хотя бы один из промежутков I_k вырожден, то брус B называется *вырожденным.* В частности, пустое множество $\varnothing$ является вырожденным брусом. Объем вырожденного бруса равен нулю.

Следующее свойство бруса интуитивно ясно, но его обоснование требует некоторой работы. При доказательстве следуем фон Нейману.

С1. *Если* $B_1, \ldots, B_k$, B *— брусы и* $B \subset \bigcup_{i=1}^k B_i$, *то* $|B| \leqslant \sum_{i=1}^k |B_i|$.

□ Для ограниченного промежутка $I \subset \mathbb{R}$ верно $|I| - 1 \leqslant \#(I \cap \mathbb{Z}) \leqslant |I| + 1$, где через $\#A$ обозначено число элементов конечного множества A. Тогда $N|I| - 1 \leqslant \#(NI \cap \mathbb{Z}) \leqslant N|I| + 1$ и, значит, $|I| - \frac{1}{N} \leqslant \frac{1}{N}\#\left(I \cap \frac{1}{N}\mathbb{Z}\right) \leqslant |I| + \frac{1}{N}$ для всех $N \in \mathbb{N}$. Следовательно,

$$|I| = \lim_{N\to\infty} \frac{1}{N}\#\left(I \cap \frac{1}{N}\mathbb{Z}\right).$$

Пусть $B = I_1 \times \ldots \times I_n$. Тогда

$$|B| = \lim_{N\to\infty} \prod_{j=1}^n \frac{1}{N}\#\left(I_j \cap \frac{1}{N}\mathbb{Z}\right) = \lim_{N\to\infty} \frac{1}{N^n}\#\left(B \cap \frac{1}{N}\mathbb{Z}^n\right).$$

Если $B \subset \bigcup_{i=1}^k B_i$, то

$$\frac{1}{N^n}\#\left(B \cap \frac{1}{N}\mathbb{Z}^n\right) \leqslant \sum_{i=1}^k \frac{1}{N^n}\#\left(B \cap \frac{1}{N}\mathbb{Z}^n\right).$$

Переходя в этом неравенстве к пределу при $N \to \infty$, получаем искомую оценку. ■

C2. *Для любого бруса $B \subset \mathbb{R}^n$ и любого $\varepsilon > 0$ найдутся замкнутый брус B' и открытый брус B^o, такие что $B' \subset B \subset B^o$ и $|B'| > |B| - \varepsilon$, $|B^o| < |B| + \varepsilon$.*

□ Пусть $B = \varnothing$, тогда $B' = B^0 = \varnothing$ подходят. Пусть $B = I_1 \times \ldots \times I_n$, где I_k — промежуток с концами $a_k \leqslant b_k$. Если $|B| > 0$, то положим

$$B'_\delta = [a_1 + \delta, b_1 - \delta] \times \ldots \times [a_n + \delta, b_n - \delta],$$

$$B^0_\delta = (a_1 - \delta, b_1 + \delta) \times \ldots \times (a_n - \delta, b_n + \delta).$$

Поскольку $|B^0_\delta|$, $|B'_\delta| \to |B|$ при $\delta \to +0$, искомые брусы B^0 и B' в этом случае существуют. Если $|B| = 0$, т.е. $a_k = b_k$ для некоторого k, то положим $B' = \varnothing$, а B^0 — как выше. ■

Роль брусов отчасти проясняется следующим утверждением.

Лемма 8.1. *Каждое непустое открытое множество U в $\mathbb{R}^n$ представимо в виде счетного объединения непересекающихся кубов.*

□ Куб $\left[\frac{k_1}{2^m}; \frac{k_1+1}{2^m}\right) \times \ldots \times \left[\frac{k_n}{2^m}; \frac{k_n+1}{2^m}\right)$, где $k_i \in \mathbb{Z}$, $m \geqslant 0$, будем называть двоичным m-го ранга.

Обозначим через A_0 множество всех кубов ранга 0, содержащихся в U. Если множества $A_0, \ldots, A_{m-1}$ уже определены, то обозначим через A_m множество всех кубов ранга m, содержащихся в U и не лежащих ни в одном кубе из $A_0, \ldots, A_{m-1}$. Положим $A = \bigcup_{m=0}^{\infty} A_m$. Тогда A — счетное множество непересекающихся кубов. Покажем, что $U = \bigcup_{Q \in A} Q$. Пусть $x \in U$. Ввиду открытости U существует шар $\overline{B}_r(x) \subset U$. Если m таково, что $\frac{\sqrt{n}}{2^m} \leqslant r$, то содержащий точку x куб $Q_m(x)$ ранга m удовлетворяет включению $Q_m(x) \subset \overline{B}_r(x)$ и, значит, множество $\{m \in \mathbb{N}_0 \colon\ Q_m(x) \subset U\}$ непусто. Обозначим через m_0 его минимум. Тогда $Q_m(x) \not\subset U$ при $m < m_0$, а $Q_{m_0}(x) \subset U$. Следовательно, $Q_{m_0}(x) \in A_{m_0}$ и поэтому $x \in \bigcup_{Q \in A} Q$. Учитывая, что обратное включение очевидно, равенство установлено. ■

8.2. Алгебры множеств.

Для целей анализа достаточно работать с семействами подмножеств $\mathbb{R}^n$, содержащими все открытые множества, и замкнутыми относительно стандартных теоретико-множественных операций.

Определение. Семейство $\mathcal{A} \subset \mathcal{P}(\mathbb{R}^n)$ называется *алгеброй*, если

(1) $\varnothing \in \mathcal{A}$;

(2) если $E \in \mathcal{A}$, то $E^c := \mathbb{R}^n \setminus E \in \mathcal{A}$;

(3) если E, $F \in \mathcal{A}$, то $E \cup F \in \mathcal{A}$.

Будем говорить, что $\mathcal{A}$ образует *σ-алгебру*, если (3) выполняется для счетных объединений, т.е. если $E_k \in \mathcal{A}$ для $k = 1, 2, \ldots$, то $\bigcup\limits_{k=1}^{\infty} E_k \in \mathcal{A}$.

Пример. 1) Пусть $A \subset \mathbb{R}^n$. Семейство $\{\varnothing,\ A,\ \mathbb{R}^n \setminus A,\ \mathbb{R}^n\}$ является σ-алгебра в $\mathbb{R}^n$.

2) Совокупность всех конечных объединений промежутков в $\mathbb{R}$ является алгеброй, но не σ-алгеброй.

3) Рассмотрим σ-алгебру $\mathcal{A}$, содержащую все одноэлементные подмножества. Тогда $\mathcal{A}$ будет включать все не более чем счетные множества, а также множества, дополнения к которым не более чем счетно.

Замечание. Из определения следует, что если $\mathcal{A}_i$ — σ-алгебры ($i \in I$), то $\bigcap\limits_{i \in I} \mathcal{A}_i$ — σ-алгебра.

Определение. *Борелевская σ-алгебра* $\mathcal{B}(\mathbb{R}^n)$ — минимальная σ-алгебра, содержащая все открытые множества $\mathbb{R}^n$. Множества, принадлежащие $\mathcal{B}(\mathbb{R}^n)$, называются *борелевскими*.

Чтобы понять, что такая алгебра существует, нужно взять пересечение всех σ-алгебр, содержащих все открытые множества $\mathbb{R}^n$. Это свойство является характеристическим для $\mathcal{B}(\mathbb{R}^n)$.

Пример. Открытые и замкнутые множества являются борелевскими. Счетные пересечения открытых множеств (G_δ-множества) и счетные объединения замкнутых (F_σ-множества) также являются борелевскими. Процесс можно продолжить, получая $G_{\delta\sigma}$, $F_{\sigma\delta}$-множества и т.д. Однако потребуются некоторые технические усилия (трансфинитная индукция), чтобы исчепать $\mathcal{B}(\mathbb{R}^n)$ таким способом. Это накладывает определенные трудности в работе с $\mathcal{B}(\mathbb{R}^n)$. Типичный пример работы с борелевской алгеброй иллюстрирует

Лемма 8.2. *Пусть $\sigma(C)$ — пересечение всех σ-алгебр, содер-*

жащих совокупность C лучей вида $(-\infty,\, b)$, $b \in \mathbb{R}$. Тогда $\sigma(C) = \mathcal{B}(\mathbb{R})$.

$\square$ По определению $\sigma(C)$ включает лучи $[a,\, +\infty)$ как дополнения, а значит, и лучи $(a,\, +\infty) = \cup_{k=1}^{\infty}[a - \frac{1}{k},\, +\infty)$. Интервалы $(a,\, b) = (-\infty,\, b) \cap (a,\, +\infty)$ также лежат в $\sigma(C)$. Любое открытое множество на $\mathbb{R}$ есть объединение не более чем счетного набора открытых промежутков (проверить!). Поэтому $\sigma(C)$ включает также и все открытые множества, а значит, $\mathcal{B}(\mathbb{R}) \subset \sigma(C)$. Обратное включение очевидно. $\blacksquare$

Задача. Докажите, что $\mathcal{B}(\mathbb{R}^n)$ есть пересечение всех σ-алгебр, содержащих двоичные кубы (см. лемма 8.1).

Наша цель — построить на $\mathbb{R}^n$ σ-алгебру $\mathcal{M}$, содержащую все борелевские множества, и меру $\mu\colon \mathcal{M} \to [0,\, +\infty]$, такую что

а) $\mu(R) = |R|$ для любого бруса R;

б) если $E_k \in \mathcal{M}$ попарно не пересекаются, то $\mu\big(\bigcup_{k=1}^{\infty} E_k\big) = \sum_{k=1}^{\infty} \mu(E_k)$ (*счетная аддитивность*);

в) $\mu(E + t) = \mu(E)$ (*инвариантность относительно сдвигов*).

8.3. Внешняя мера Лебега и ее свойства.

Начнем построение с того, что *каждому* $E \subset \mathbb{R}^n$ сопоставим внешнюю меру $\mu^*(E)$.

Определение. *Внешней мерой Лебега* множества $E \subset \mathbb{R}^n$ называется величина

$$\mu^*(E) = \inf\left\{\sum_{i=1}^{\infty} |B_i|\colon\ E \subset \bigcup_{i=1}^{\infty} B_i\right\},$$

где точная грань берется по всем счетным наборам брусов, объединение которых содержит E.

Очевидно, что $0 \leqslant \mu^*(E) \leqslant \infty$.

Теорема 8.1. *Внешняя мера Лебега обладает следующими свойствами:*

1) если $E \subset F$, то $\mu^*(E) \leqslant \mu^*(F)$ (монотонность)*;*

2) если $E = \bigcup\limits_{k=1}^{\infty} E_k$, то $\mu^*(E) \leqslant \sum\limits_{k=1}^{\infty} \mu^*(E_k)$ (счетная полуаддитивность)*;*

3) $\mu^*(R) = |R|$ для любого бруса R (нормировка).

$\square$ Так как любое покрытие F является покрытием $E \subset F$, то п. 1 непосредственно вытекает из определения.

Докажем п. 2. Считаем, что $\sum\limits_{k=1}^{\infty} \mu^*(E_k) < \infty$ (иначе утверждение очевидно верно). Возьмем произвольное $\varepsilon > 0$ и для каждого множества E_k рассмотрим счетный набор брусов $\{B_{i,\,k} \colon i \in \mathbb{N}\}$, образующих покрытие E_k, и такой что

$$\sum_{i=1}^{\infty} |B_{i,\,k}| \leqslant \mu^*(E_k) + \frac{\varepsilon}{2^k}.$$

Тогда счетный набор брусов $\{B_{i,\,k} \colon i,\, k \in \mathbb{N}\}$ образует покрытие $E = \bigcup_{k=1}^{\infty} E_k$ и

$$\mu^*(E) \leqslant \sum_{i,\,k=1}^{\infty} |B_{i,\,k}| = \sum_{k=1}^{\infty} \left(\sum_{i=1}^{\infty} |B_{i,\,k}| \right) \leqslant$$

$$\leqslant \sum_{k=1}^{\infty} (\mu^*(E_k) + 2^{-k}\varepsilon) = \sum_{k=1}^{\infty} \mu^*(E_k) + \varepsilon.$$

Поскольку $\varepsilon > 0$ — любое, требуемое неравенство установлено.

Докажем п. 3. Поскольку $\{R\}$ — покрытие R, то $\mu^*(R) \leqslant |R|$. Докажем, что $\mu^*(R) \geqslant |R|$. Рассмотрим сначала случай, когда брус R замкнут.

Пусть счетный набор брусов $\{B_k\}$ образует покрытие R. Достаточно доказать, что $|R| \leqslant \sum\limits_{k=1}^{\infty} |B_k|$.

Пусть $\varepsilon > 0$. По свойству 2 для каждого k подберем открытый брус $B_k^o \supset B_k$ с $|B_k^o| \leqslant |B_k| + \frac{\varepsilon}{2^k}$. Поскольку $R \subset \bigcup_{k=1}^{\infty} B_k^o$ и R — компакт, можно выделить конечное подпокрытие $R \subset \bigcup_{k=1}^{N} B_k^o$. Тогда по свойству 1 $|R| \leqslant \sum\limits_{k=1}^{N} |B_k^o|$ и, значит, $|R| \leqslant \sum\limits_{k=1}^{N} (|B_k| + \frac{\varepsilon}{2^k}) \leqslant \sum\limits_{k=1}^{\infty} |B_k| + \varepsilon$. Так как $\varepsilon > 0$ — любое, то требуемое неравенство установлено.

Пусть теперь R произвольный брус. По свойству 2 для $\varepsilon > 0$ найдется замкнутый брус $R' \subset R$ с $|R'| > |R| - \varepsilon$. Откуда по монотонности $\mu^*(R) \geqslant \mu^*(R') = |R'| > |R| - \varepsilon$. Так как $\varepsilon > 0$ — любое, то неравенство установлено и в этом случае. $\blacksquare$

8.4. Измеримые множества.

Определение. Множество $E \subset \mathbb{R}^n$ называется *измеримым*, если для любого $A \subset \mathbb{R}^n$ выполнено

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

Это означает, что E разрезает каждое множество A на части, внешние меры которых при сложении дают внешнюю меру A.

Замечание. Для проверки измеримости достаточно устанавливать, что $\mu^*(A) \geqslant \mu^*(A \cap E) + \mu^*(A \cap E^c)$, т. к. противоположное неравенство выполняется по свойству полуаддитивности.

Пример 1. Если $\mu^*(E) = 0$, то множество E измеримо. Действительно, т. к. $\mu^*(A) \geqslant \mu^*(A \cap E^c)$ и $\mu^*(A \cap E) = 0$, то $\mu^*(A) \geqslant \geqslant \mu^*(A \cap E) + \mu^*(A \cap E^c)$.

Пример 2. Покажем, что для каждого $a \in \mathbb{R}$ и номера $k \in \{1, \ldots, n\}$ полупространство

$$H = H_{a,k} = \{(x_1, \ldots, x_k, \ldots, x_n)^T \colon x_k > a\}$$

измеримо. Для этого рассмотрим множество $A \subset \mathbb{R}^n$ и произвольное его покрытие брусами $\{B_i\}_{i=1}^{\infty}$. Положим $B_i^1 = \{x \in B_i \colon x_k > a\}$ и $B_i^2 = \{x \in B_i \colon x_k \leqslant a\}$. Тогда наборы брусов $\{B_i^1\}$ и $\{B_i^2\}$ образуют покрытия $A \cap H$ и $A \cap H^c$ соответственно, причем $|B_i| = |B_i^1| + |B_i^2|$. Имеем

$$\sum_{i=1}^{\infty} |B_i| = \sum_{i=1}^{\infty} |B_i^1| + \sum_{i=1}^{\infty} |B_i^2| \geqslant \mu^*(A \cap H) + \mu^*(A \cap H^c).$$

Переходя к инфимуму по всем покрытиям $\{B_i\}$, получаем $\mu^*(A) \geqslant \mu^*(A \cap H) + \mu^*(A \cap H^c)$. Следовательно, H измеримо. Аналогично доказывается измеримость полупространств по неравенствам $<$, $\leqslant$ и $\geqslant$.

Следующая теорема является важнейшей в этом разделе.

Теорема 8.2 (Каратеодори). *Совокупность $\mathcal{M}$ всех измеримых множеств является σ-алгеброй. Функция μ^* на $\mathcal{M}$ счетно аддитивна.*

$\square$ Из определения сразу вытекает, что $\varnothing$ измеримо, и если $E \in \in \mathcal{M}$, то $E^c \in \mathcal{M}$. Дальнейшее доказательство проведем в три этапа.

I. Покажем, что если E, $F \in \mathcal{M}$, то $E \cup F \in \mathcal{M}$.

Пусть $A \subset \mathbb{R}^n$. Так как E, F измеримы, то

$$\mu^*\big(A \cap (E \cup F)\big) + \mu^*\big(A \cap (E \cup F)^c\big) = \\ = \mu^*\big(A \cap (E \cup F) \cap E\big) + \mu^*\big(A \cap (E \cup F) \cap E^c\big) + \\ + \mu^*\big(A \cap (E \cup F)^c\big) = \mu^*\big(A \cap E\big) + \mu^*\big(A \cap E^c \cap F\big) + \\ + \mu^*\big(A \cap E^c \cap F^c\big) = \mu^*\big(A \cap E\big) + \mu^*\big(A \cap E^c\big) = \mu^*(A).$$

II. Пусть $E_k \in \mathcal{M}$, причем $E_i \cap E_j = \varnothing$ при $i \neq j$. Покажем, что $F = \bigcup\limits_{k=1}^{\infty} E_k \in \mathcal{M}$.

Положим $F_n = \bigcup_{k=1}^n E_k$. Если $A \subset \mathbb{R}^n$, то, пользуясь измеримостью E_k,

$$\mu^*(A \cap F_n) = \mu^*(A \cap F_n \cap E_n) + \mu^*(A \cap F_n \cap E_n^c) = \mu^*(A \cap E_n) + \mu^*(A \cap F_{n-1}).$$

Продолжая процесс, получим $\mu^*(A \cap F_n) = \sum_{k=1}^n \mu^*(A \cap E_k)$.

Поскольку $F_n \in \mathcal{M}$, то

$$\mu^*(A) = \mu^*(A \cap F_n) + \mu^*(A \cap F_n^c) \geqslant \sum_{k=1}^{n} \mu^*(A \cap E_k) + \mu^*(A \cap F^c).$$

Переходя к пределу при $n \to \infty$, получим $\mu^*(A) \geqslant \sum_{k=1}^{\infty} \mu^*(A \cap E_k) + \mu^*(A \cap F^c)$. Откуда по свойству счетной полуаддитивности

$$\mu^*(A) \geqslant \sum_{k=1}^{\infty} \mu^*(A \cap E_k) + \mu^*(A \cap F^c) \geqslant \mu^*(A \cap F) + \mu^*(A \cap F^c) \geqslant \mu^*(A).$$

Это доказывает, что $F \in \mathcal{A}$. Если еще положить $A = F$, то $\mu^*(F) = \sum_{k=1}^{\infty} \mu^*(E_k)$.

III. Пусть $A_k \in \mathcal{M}$. Покажем, что $A = \bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{M}$.

Положим $E_1 = A_1$, $E_k = A_k \setminus \bigcup\limits_{i<k} E_i$. Тогда E_k попарно не пересекаются, и $A = \bigcup\limits_{k=1}^{\infty} E_k \in \mathcal{M}$ по предыдущему пункту. ■

Следствие. *Всякий брус измерим.*

□ Брус можно записать в виде конечного пересечения полупространств. Теперь результат следует из измеримости полупространств по теореме 8.2. ■

Следствие. *Все борелевские множества измеримы.*

□ По лемме 8.1 всякое открытое множество можно записать в виде счетного объединения кубов (брусов) и, значит, открытые множества измеримы. Тогда $\mathcal{M}$ также включает минимальную σ-алгебру, содержащую все открытые множества, т.е. $\mathcal{B} \subset \mathcal{M}$. ■

Определение. Сужение $\mu = \mu^*|_{\mathcal{M}}$ называется *мерой Лебега.* Как установлено в теореме 8.2, мера μ счетно аддитивна: если $E_k \in \mathcal{M}$ и $E_i \cap E_j = \varnothing$ при $i \neq j$, то $\mu\big(\bigcup_{k=1}^{\infty} E_k\big) = \sum_{k=1}^{\infty} \mu(E_k)$.

Полезно думать о мере как о «непрерывной» функции в следующем смысле.

Теорема 8.3. *Пусть $A_i \in \mathcal{M}$ для всех $i \in \mathbb{N}$.*

1) Если $A_1 \subset A_2 \subset \ldots,$ и $A = \bigcup\limits_{i=1}^{\infty} A_i$, то $\mu(A) = \lim\limits_{i\to\infty} \mu(A_i)$ (непрерывность снизу)*;*

2) Если $A_1 \supset A_2 \supset \ldots,$ $\mu(A_1) < \infty$, и $A = \bigcap\limits_{i=1}^{\infty} A_i$, то $\mu(A) = \lim\limits_{i\to\infty} \mu(A_i)$ (непрерывность сверху).

□ Докажем п. 1. Положим $B_1 = A_1$, $B_i = A_i \setminus A_{i-1}$, $i \geqslant 2$. Тогда $B_i \in \mathcal{M}$, $B_i \cap B_j = \varnothing$ при $i \neq j$, и $\bigcup\limits_{i=1}^{m} B_i = \bigcup\limits_{i=1}^{m} A_i$ для $m \in \mathbb{N} \cup \infty$. Поэтому

$$\mu(A) = \mu\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{m\to\infty} \sum_{i=1}^{m} \mu(B_i) = \lim_{m\to\infty} A_m.$$

Докажем п. 2. Применим п. 1 к $A_1 \setminus A_i$ ($i \in \mathbb{N}$). Тогда $A_1 \setminus A = \bigcup_{i=1}^{\infty} (A_1 \setminus A_i)$ и

$$\mu(A_1) - \mu(A) = \mu(A_1 \setminus A) = \lim_{m\to\infty} \mu(A_1 \setminus A_m) = \mu(A_1) - \lim_{m\to\infty} \mu(A_m).$$

Осталось из обеих частей вычесть $\mu(A_1)$ и изменить знак. ■

Задача. Покажите, что условие $\mu(A_1) < \infty$ в п. 2 важно.

Получим эквивалентные определения измеримости множеств, с которыми легче работать на практике.

Лемма 8.3. *Если множество $E \subset \mathbb{R}^n$ измеримо, то для любого $\varepsilon > 0$ найдется такое открытое множество $G \supset E$, что $\mu(G \setminus E) < \varepsilon$.*

□ Рассмотрим сначала случай, когда E ограничено, а значит, $\mu^*(E) < \infty$. Для $\varepsilon > 0$ рассмотрим покрытие E счетным набором брусов $\{B_i\}$ с $\sum_{i=1}^{\infty} |B_i| < \mu(E) + \frac{\varepsilon}{2}$. Для каждого $i \in \mathbb{N}$ найдем такой открытый брус B_i^o, содержащий B_i, что $|B_i^o| < |B_i| + \frac{\varepsilon}{2^{i+1}}$. Положим $G = \bigcup_{i=1}^{\infty} B_i^o$. Тогда множество G открыто, содержит E и

$$\mu(G) \leqslant \sum_{i=1}^{\infty} |B_i^0| < \sum_{i=1}^{\infty} \Big(|B_i| + \frac{\varepsilon}{2^{i+1}} \Big) \leqslant \sum_{i=1}^{\infty} |B_i| + \frac{\varepsilon}{2} < \mu(E) + \varepsilon.$$

По аддитивности меры $\mu(G \setminus E) = \mu(G) - \mu(E) < \varepsilon$.

Перейдем к общему случаю. Поскольку $\mathbb{R}^n = \bigcup_{k=1}^{\infty} A_k$, где $A_k = \{x \in \mathbb{R}^n\colon\ k - 1 \leqslant |x| < k\}$, то E есть счетное объединение непересекающихся ограниченных измеримых множеств $E_k = E \cap \cap A_k$. По доказанному существует такое открытое множество $G_k \supset E_k$, что $\mu(G_k \setminus E_k) < \frac{\varepsilon}{2^k}$. Тогда множество $G = \bigcup_{k=1}^{\infty} G_k$ открыто, содержит E, $G \setminus E \subset \bigcup_{k=1}^{\infty} (G_k \setminus E_k)$ и

$$\mu(G \setminus E) \leqslant \sum_{k=1}^{\infty} \mu(G_k \setminus E_k) < \varepsilon. \blacksquare$$

Следствие. *Если множество $E \subset \mathbb{R}^n$ измеримо, то для любого $\varepsilon > 0$ найдется такое замкнутое множество $F \subset E$, что $\mu(E \setminus F) < \varepsilon$.*

□ Применив лемму 8.3 к множеству E^c, найдем такое открытое множество $G \supset E^c$, что $\mu(G \setminus E^c) < \varepsilon$. Тогда множество $F = G^c$ замкнуто, $F \subset E$, $E \setminus F = G \setminus E^c$ и, значит, $\mu(E \setminus F) < \varepsilon$. ■

Напомним, что счетные пересечения открытых множеств называются множествами типа G_δ, а счетные объединения замкнутых — множествами типа F_σ.

Теорема 8.4. *Множество $E \subset \mathbb{R}^n$ измеримо тогда и только тогда, когда найдется такое множество Ω типа G_δ, что $E \subset \subset \Omega$ и $\mu(\Omega \setminus E) = 0$.*

□ ($\Rightarrow$) Найдем $\forall k \in \mathbb{N}$ открытое $G_k \supset E$ с $\mu(G_k \setminus E) < \frac{1}{k}$. Положим $\Omega = \bigcap\limits_{k=1}^{\infty} G_k$. Тогда $E \subset \Omega$ и $\mu(\Omega \setminus E) \leqslant \mu(G_k \setminus E) < \frac{1}{k}$. Следовательно, $\mu(\Omega \setminus E) = 0$.

($\Leftarrow$) Поскольку $E = \Omega \setminus (\Omega \setminus E)$, то E измеримо как разность двух измеримых множеств. ■

Следствие. *Множество $E \subset \mathbb{R}^n$ измеримо тогда и только тогда, когда найдется такое множество Δ типа F_σ, что $\Delta \subset$ $\subset E$ и $\mu(E \setminus \Delta) = 0$.*

Таким образом, всякое измеримое множество является борелевским «с точностью» до множества меры нуль. Теперь покажем, что (ограниченное) множество измеримо, если оно «хорошо приближается» брусами.

Теорема 8.5. *Пусть $\mu^*(E) < \infty$. Множество E измеримо тогда и только тогда, когда для любого $\varepsilon > 0$ найдутся брусы $B_1, \ldots, B_N$, такие что $\mu^*(E \Delta \bigcup_{k=1}^N B_k) \leqslant \varepsilon$.*

□ Если E измеримо, то $E \subset \bigcup_{k=1}^\infty B_k$ с $\sum_{k=1}^\infty |B_k| \leqslant \mu(E) + \frac{\varepsilon}{2}$. Ряд сходится, поэтому существует такой номер N, что $\sum_{k=N+1}^\infty |B_k| <$ $< \frac{\varepsilon}{2}$. Положим $C = \bigcup_{k=1}^N B_k$. Тогда

$$\mu(E \Delta C) = \mu(E \setminus C) + \mu(C \setminus E) \leqslant \mu\Big(\bigcup_{k \geqslant N+1} B_k\Big) + \mu\Big(\bigcup_{k \geqslant 1} B_k \setminus E\Big) \leqslant$$

$$\leqslant \sum_{k=N+1}^\infty |B_k| + \sum_{k=1}^\infty |B_k| - \mu(E) < \varepsilon.$$

Обратно, пусть $\mu^*(E \Delta C) < \varepsilon$. Тогда тем более $\mu^*(C \setminus E) < \varepsilon$ и $\mu^*(E \setminus C) < \varepsilon$. Пусть $A \subset \mathbb{R}^n$. Так как $E \subset C \cup (E \setminus C)$ и $E^c \subset$ $\subset C^c \cup (C \setminus E)$, то в силу измеримости C имеем

$$\mu^*(A \cap E) + \mu^*(A \cap E^c) \leqslant \mu^*(A \cap C) + \mu^*(E \setminus C) +$$
$$+ \mu^*(A \cap C^c) + \mu^*(C \setminus E) < \mu^*(A) + 2\varepsilon.$$

Следовательно, $\mu^*(A \cap E) + \mu^*(A \cap E^c) \leqslant \mu^*(A)$ и, значит, E измеримо. ■

Покажем, что мера Лебега инвариантна относительно сдвигов.

Лемма 8.4. *Если E измеримо и $y \in \mathbb{R}^n$, то множество $E +$ $+ y = \{x + y \colon x \in E\}$ измеримо и $\mu(E + y) = \mu(E)$.*

□ Очевидно, что если B брус, то $B + y$ брус того же объема. Пусть $A \subset \bigcup_{k=1}^\infty B_k$. Тогда $A + y \subset \bigcup_{k=1}^\infty (B_k + y)$, откуда следует, что $\mu^*(A + y) \leqslant \sum_{k=1}^\infty |B_k|$ и, значит, $\mu^*(A + y) \leqslant \mu^*(A)$. Поскольку $A = (A+y) - y$, то верно противоположное неравенство, т.е. $\mu^*(A+y) = \mu^*(A)$. Пусть E измеримо и $A \subset \mathbb{R}^n$. Тогда

$$\begin{aligned} &\mu^*(A \cap (E+y)) + \mu^*(A \cap (E+y)^c) = \\ &\qquad = \mu^*\big(((A-y) \cap E) + y)\big) + \mu^*\big(((A-y) \cap E^c) + y)\big) = \\ &\qquad = \mu^*\big((A-y) \cap E)\big) + \mu^*\big((A-y) \cap E^c)\big) = \mu^*(A-y) = \mu^*(A), \end{aligned}$$

так что $E + y$ также измеримо. ■

Задача. Докажите, что если $E \subset \mathbb{R}^n$ измеримо и $\alpha > 0$, то множество $\alpha E = \{\alpha x \colon x \in E\}$ измеримо и $\mu(\alpha E) = \alpha^n \mu(E)$.

Пример. Пусть $E \subset \mathbb{R}$ измеримо и имеет конечную меру. Покажем, что $\mu(E \setminus (E+t)) \to 0$ при $t \to 0$.

□ Для произвольного $J \subset \mathbb{R}$ справедливо включение

$$E \setminus (E+t) \subset (J \setminus (J+t)) \cup (E \setminus J) \cup ((J+t) \setminus (E+t)).$$

Действительно, т.к. $E \subset J \cup (E \setminus J)$, то $E \cap (E+t)^c \subset (J \cap (E+ +t)^c) \cup (E \setminus J)$. Для завершения заменим $(E+t)^c$ на $(J+t)^c \cup \cup ((E+t)^c \setminus (J+t)^c)$.

Зафиксируем $\varepsilon > 0$ и выберем в качестве J объединение промежутков (брусов) $I_1, \ldots, I_m$, так что $\mu(E \Delta J) < \varepsilon$. Поскольку $J \setminus (J+t) \subset \cup_j (I_j \setminus (I_j+t))$, $\mu(I_j \setminus (I_j+t)) \leqslant |t|$ и $(J+t) \setminus (E+ +t) = (J \setminus E) + t$, то

$$\mu(E \setminus (E+t)) \leqslant \sum_j \mu(I_j \setminus (I_j+t)) + \mu(E \setminus J) + \mu(J \setminus E) < m|t| + 2\varepsilon.$$

Следовательно, $\mu(E \setminus (E+t)) < 3\varepsilon$ при достаточно малых t. ■

Пример (неизмеримого множества). На $[0,\,1]$ введем отношение эквивалентности $x \sim y \Leftrightarrow x - y \in \mathbb{Q}$. Тогда справедливо представление $[0,\,1] = \bigsqcup_\alpha H_\alpha$, где H_α — различные классы эквивалентности.

□ Рассмотрим множество V, содержащее ровно по одну элементу из каждого класса H_α и только такие элементы. Множество V существует по аксиоме выбора.

Пусть $\{r_n\}_{n=0}^{\infty}$ — некоторая нумерация $\mathbb{Q} \cap [-1,\,1]$ $(r_0 = 0)$. Определим $V_n = V + r_n$.

1) V_n попарно не пересекаются.

Если $x \in V_i \cap V_j$, то $x = x_i + r_i$ и $x = x_j + r_j$ для некоторых $x_i \in \in V$, $x_j \in V$. Тогда $x_i + r_i = x_j + r_j$, $x_i - x_j \in \mathbb{Q}$. По определению V $x_i = x_j$, но тогда $r_i = r_j$. Следовательно, $V_i = V_j$.

2) $[0,\, 1] \subset \bigsqcup_{n=0}^{\infty} V_n \subset [-1,\, 2]$.

Если $y \in [0,\, 1]$, то y лежит в некотором классе H_α, т.е. $y = = x + r$, где $x \in V$, $r \in \mathbb{Q}$. Так как $r = y - x \in \mathbb{Q} \cap [-1,\, 1]$, то $r = r_n$ для некоторого n. Это доказывает первое включение. Второе включение следует из того, что $V \subset [0,\, 1]$ и $r_n \in [-1,\, 1]$.

3) Множество V неизмеримо.

В противном случае все V_n измеримы и $\mu(V_n) = \mu(V)$. Следовательно, измеримо $S = \bigsqcup_{n=0}^{\infty} V_n$ и по счетной аддитивности $\mu(S) = \sum_{n=0}^{\infty} \mu(V_n)$. Так как $S \subset [-1,\, 2]$, то $\mu(S) \leqslant 3$. Сходимость ряда $\sum_{n=0}^{\infty} \mu(V)$ возможна лишь в случае $\mu(V) = 0$. Тогда $\mu(S) = = 0$, что противоречит включению $[0,\, 1] \subset S$. ■

9. Интеграл Лебега

9.1. Измеримые функции.

Пусть E измеримо и задана функция $f\colon E \to \overline{\mathbb{R}}$.

Определение. Функция f называется *измеримой*, если $f^{-1}([-\infty, a)) = \{x \in E\colon f(x) < a\}$ измеримо для всех $a \in \mathbb{R}$.

Смысл данного определения отчасти объясняет

Лемма 9.1.* *Пусть E измеримо и задана функция $f\colon E \to \overline{\mathbb{R}}$. Тогда следующие утверждения эквивалентны:*

1) f измерима;

2) $f^{-1}(U)$ измеримо для всех открытых U в $\overline{\mathbb{R}}$;

2) $f^{-1}(\Omega)$ измеримо для всех множеств Ω из $\mathcal{B}(\overline{\mathbb{R}})$ (т.е. таких, что $\Omega \cap \mathbb{R}$ борелевское).

$\square$ Теоретико-множественные операции сохраняются при взятии прообразов:

$f^{-1}\big(\overline{\mathbb{R}} \setminus A\big) = E \setminus f^{-1}(A)$, $f^{-1}\big(\cup_{i=1}^{\infty} A_i\big) = \cup_{i=1}^{\infty} f^{-1}(A_i)$.

Отсюда следует, что совокупность $\mathcal{A} = \{A \subset \overline{\mathbb{R}}\colon\ f^{-1}(A)$ измеримо$\}$ образует σ-алгебру. Если функция f измерима, то $\mathcal{A}$ включает все лучи $[-\infty, a)$. Тогда по лемме 8.3 совокупность $\mathcal{A}$ включает также все борелевские множества. Поскольку $f^{-1}(-\infty) = \cap_{i=1}^{\infty} f^{-1}([-\infty, -i))$, то $-\infty \in \mathcal{A}$, а значит, и $+\infty \in \mathcal{A}$. Следовательно, $\mathcal{B}(\overline{\mathbb{R}}) \subset \mathcal{A}$. Это доказывает импликацию $(1 \Rightarrow 3)$. Импликации $(3 \Rightarrow 2 \Rightarrow 1)$ очевидны. $\blacksquare$

Таким образом, функция измерима в точности тогда, когда она сохраняет измеримость прообразов множеств из $\mathcal{B}(\overline{\mathbb{R}})$.

Пример. 1) Пусть $A \subset \mathbb{R}^n$. Определим *индикатор* (*характеристическую функцию*) A:

$$I_A\colon \mathbb{R}^n \to \mathbb{R}, \qquad I_A(x) = \begin{cases} 1, & x \in A; \\ 0, & x \notin A. \end{cases}$$

Поскольку $\{x\colon I_A(x) < a\}$ пусто при $a \leqslant 0$, совпадает с A^c при $a \in (0, 1]$ и совпадает с $\mathbb{R}^n$ при $a > 1$, то функция I_A измерима $\Leftrightarrow$ A измеримо.

2) Если функция $f\colon E \to \mathbb{R}$ непрерывна на измеримом E, то она измерима. Действительно, по критерию непрерывности множество $\{x\colon f(x) < a\} = f^{-1}(-\infty, a)$ открыто в E. Это означает, что существует такое открытое в $\mathbb{R}^n$ множество G, что $\{x\colon f(x) <$

$< a\} = E \cap G$. Тогда $\{x\colon f(x) < a\}$ измеримо, как пересечение измеримых множеств.

Задача. Пусть $f\colon E \to \mathbb{R}$ монотонна на измеримом E, то f измерима.

Замечание. Если в определении знак $<$ заменить на $\leqslant$, $>$ или $\geqslant$, то получим равносильное определение измеримости. Это вытекает из равенств

$$\{x\colon f(x) \leqslant a\} = \bigcap_{k=1}^{\infty} \{x\colon f(x) < a + 1/k\},$$
$$\{x\colon f(x) > a\} = E \setminus \{x\colon f(x) \leqslant a\},$$
$$\{x\colon f(x) \geqslant a\} = \bigcap_{k=1}^{\infty} \{x\colon f(x) > a - 1/k\},$$
$$\{x\colon f(x) < a\} = E \setminus \{x\colon f(x) \geqslant a\}.$$

Измеримость функций согласована с алгебраическими операциями.

Теорема 9.1. *Если функции f, $g\colon E \to \mathbb{R}$ измеримы, и $\lambda \in \mathbb{R}$, то функции $f+g$, λf, $|f|$ и fg также измеримы.*

$\square$ Пусть $\{r_k\}$ — некоторая нумерация $\mathbb{Q}$. Ввиду того, что $\alpha < \beta \Leftrightarrow \alpha < r < \beta$ для некоторого $r \in \mathbb{Q}$, имеем

$$\{x\colon f(x) + g(x) < a\} = \{x\colon f(x) < a - g(x)\} =$$
$$= \bigcup_{k=1}^{\infty} \{x\colon f(x) < r_k < a - g(x)\} =$$
$$= \bigcup_{k=1}^{\infty} \{x\colon f(x) < r_k\} \cap \{x\colon g(x) < a - r_k\}.$$

Следовательно, $\{x\colon f(x) + g(x) < a\}$ измеримо.

Функция λf измерима при $\lambda > 0$, т.к. $\{x\colon \lambda f(x) < a\} = \{x\colon f(x) < \frac{a}{\lambda}\}$. Случай $\lambda = 0$ очевиден, случай $\lambda < 0$ рассматривается аналогично.

Функция f^2 измерима, т.к. $\{x\colon f^2(x) < a\} = \{x\colon f(x) > -\sqrt{a}\} \cap \{x\colon f(x) < \sqrt{a}\}$ при $a > 0$ и $\{x\colon f^2(x) < a\}$ пусто при $a \leqslant 0$. Аналогичные рассуждения применимы и для $|f|$. Но тогда измерима и функция $fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$. ■

Задача. Докажите, что если $g\colon E \to \mathbb{R}$ измерима, и $g \neq 0$, то функция $\frac{1}{g}$ измерима на E.

Замечание. Теорема 9.1 остается справедливой для функций со значениями в $\overline{\mathbb{R}}$, если операции допустимы. Например, для $f+g$ необходимы условия $f^{-1}(+\infty)\cap g^{-1}(-\infty)=\varnothing$, $f^{-1}(-\infty)\cap \cap g^{-1}(+\infty)=\varnothing$.

Определение. Функции $f^+=\max\{f,\,0\}$ и $f^-=\max\{-f,\,0\}$ называются *положительной* и *отрицательной* частями функции f.

Из определения следует, что $f=f^+-f^-$, $|f|=f^++f^-$ и $0\leqslant \leqslant f^\pm\leqslant |f|$.

Следствие. *Измеримость f равносильна одновременной измеримости f^+ и f^-.*

□ Если f измерима, то функции $f^\pm=\frac{1}{2}(|f|\pm f)$ измеримы. Если же $f^\pm$ измеримы, то $f=f^+-f^-$ измерима. ■

Задача. Пусть функции $f_1,\,\ldots,\,f_m\colon E\to\mathbb{R}$ измеримы, $f= =(f_1,\,\ldots,\,f_m)\colon E\to\mathbb{R}^m$, и функция $g\colon\mathbb{R}^m\to\mathbb{R}$ непрерывна. Покажите, что $g\circ f\colon E\to\mathbb{R}$ измерима.

Теорема 9.2. *Если функции $f_k\colon E\to\overline{\mathbb{R}}$ измеримы, то функции $\sup\limits_k f_k$, $\inf\limits_k f_k$, $\varlimsup\limits_{k\to\infty} f_k$ и $\varliminf\limits_{k\to\infty} f_k$ также измеримы.*

□ Измеримость $g=\sup\limits_k f_k$ следует из равенства $\{x\colon g(x)\leqslant a\}= =\bigcap\limits_{k=1}^{\infty}\{x\colon f_k(x)\leqslant a\}$, измеримость $h=\inf\limits_k f_k$ следует из равенства $\inf\limits_k f_k=-\sup\limits_k(-f_k)$.

Далее, поскольку $\varlimsup\limits_{k\to\infty} f_k=\inf\limits_k\sup\limits_{m\geqslant k} f_m$, $\varliminf\limits_{k\to\infty} f_k=\sup\limits_k\inf\limits_{m\geqslant k} f_m$, то оба предела измеримы. ■

Поточечная сходимость сохраняет измеримость (в отличие, например, от непрерывности). Этот факт объясняет, почему измеримые функции образуют достаточно широкий класс для нужд анализа.

Следствие. *Если функции $f_k\colon E\to\overline{\mathbb{R}}$ измеримы, и $f_k\to f$ на E, то $f\colon E\to\overline{\mathbb{R}}$ измерима.*

Пусть $E\subset\mathbb{R}^n$ и $Q(x)$ — формула на E. Говорят, что Q верна *почти всюду* (п.в.) на E, если $\mu\{x\in E\colon Q(x)\ \text{ложно}\}=0$.

Лемма 9.2. *Пусть f, $g\colon E \to \overline{\mathbb{R}}$. Если $f = g$ п.в. на E, и f измерима, то g также измерима.*

$\square$ По условию мера множества $Z = \{x \in E\colon f(x) \neq g(x)\}$ равна 0. Для любого $a \in \mathbb{R}$ справедливо равенство

$$\{x \in E\colon g(x) < a\} = (\{x \in E\colon f(x) < a\} \cap Z^c) \cup (\{x \in E\colon g(x) < a\} \cap Z).$$

Поскольку функция f измерима, множество $\{x\colon f(x) < a\}$ измеримо, а значит, измеримо и $\{x\colon f(x) < a\} \cap Z^c$; множество $\{x\colon g(x) < a\} \cap Z$ измеримо как подмножество меры нуль. Следовательно, $\{x\colon g(x) < a\}$ измеримо. $\blacksquare$

Следствие. *Пусть f, $f_k\colon E \to \overline{\mathbb{R}}$. Если функции f_k измеримы и $f_k \to f$ п.в. на E, то f измерима на E.*

$\square$ Поскольку все f_k измеримы, то $g = \varlimsup\limits_{k\to\infty} f_k$ измерима и $f = g$ п.в. Следовательно, f также измерима. $\blacksquare$

Определение. Функция $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ называется *простой*, если φ измерима и множество ее значений конечно.

Любая линейная комбинация индикаторов измеримых множеств, очевидно, является простой функцией. С другой стороны, для простой функции φ всегда существует разбиение $\mathbb{R}^n$ конечным числом измеримых множеств, на которых φ постоянна (*допустимое* разбиение). Такое разбиение можно получить, например, так. Пусть $\varphi(\mathbb{R}^n) = \{a_1, \ldots, a_m\}$, где все a_i попарно различны, и $A_i = \varphi^{-1}(a_i)$. Тогда набор $\{A_i\}_{i=1}^n$ образует измеримое разбиение $\mathbb{R}^n$ и $\varphi = \sum\limits_{i=1}^{n} a_i I_{A_i}$.

Допустимое разбиение, вообще говоря, не единственно: каждое из A_i можно разбивать на «более мелкие» измеримые части. Кроме того, не исключается случай, когда некоторые элементы пусты.

Изучим вопрос приближения измеримой функции простыми.

Теорема 9.3. *Если функция $f\colon E \to [0, +\infty]$ измерима, то существует такая последовательность простых неотрицательных функций $\{\varphi_k\}$, что для каждого $x \in E$*

(1) $0 \leqslant \varphi_1(x) \leqslant \varphi_2(x) \leqslant \ldots$,

(2) $\lim\limits_{k\to\infty} \varphi_k(x) = f(x)$.

□ Для $k \in \mathbb{N}$ и $j = 1, 2, \ldots, k2^k$ положим

$$E_{k,j} = \left\{x \in E\colon \frac{j-1}{2^k} \leqslant f(x) < \frac{j}{2^k}\right\},$$

$$F_k = \{x \in E\colon f(x) \geqslant k\}.$$

Тогда все $E_{k,j}$ и F_k измеримы, попарно не пересекаются и в объединении дают E. Определим

$$\varphi_k = \sum_{j=1}^{k2^k} \frac{j-1}{2^k} I_{E_{k,j}} + kI_{F_k}.$$

Зафиксируем $x \in E$ и покажем, что $\{\varphi_k(x)\}$, возрастая, стремится к $f(x)$.

Начнем с проверки возрастания. Пусть $k \in \mathbb{N}$. Если $f(x) \geqslant k$, то $\varphi_{k+1}(x) \geqslant k = \varphi_k(x)$. Если $f(x) < k$, то $\frac{j-1}{2^k} \leqslant f(x) < \frac{j}{2^k}$ для некоторого j, $1 \leqslant j \leqslant k2^k$. Возможны два варианта: $\frac{2j-2}{2^{k+1}} \leqslant f(x) < \frac{2j-1}{2^{k+1}}$ или $\frac{2j-1}{2^{k+1}} \leqslant f(x) < \frac{2j}{2^{k+1}}$. В обоих случаях $\varphi_{k+1}(x) \geqslant \frac{2j-2}{2^{k+1}} = \frac{j-1}{2^k} = \varphi_k(x)$.

Если $f(x) \in \mathbb{R}$, то $0 \leqslant f(x) - \varphi_k(x) < 2^{-k}$ при всех $k \geqslant [f(x)] + 1$. Откуда следует, что $\varphi_k(x) \to f(x)$. Если $f(x) = +\infty$, то $\varphi_k(x) = k$ для всех k и, значит, также $\varphi_k(x) \to f(x)$. ■

Замечание. Если в условиях теоремы функция f ограничена, то $\varphi_k \rightrightarrows f$ на E: неравенство $0 \leqslant f(x) - \varphi_k(x) < 2^{-k}$ выполняется равномерно по x.

9.2. Интеграл Лебега от неотрицательных функций.

Определение. Пусть φ — неотрицательная простая функция, $\{A_i\}_{i=1}^m$ — ее допустимое разбиение, $\varphi = \sum\limits_{i=1}^m a_i I_{A_i}$. Интегралом от φ по измеримому множеству E называется сумма

$$\int_E \varphi\, d\mu := \sum_{i=1}^m a_i \mu(A_i \cap E).$$

Считаем, что $0 \cdot \infty = 0$.

Лемма 9.3. *Пусть φ, ψ — неотрицательные простые функции, и $\lambda \in [0, +\infty)$.*

1) Если $\varphi \leqslant \psi$ на E, то $\int\limits_E \varphi d\mu \leqslant \int\limits_E \psi d\mu$ (монотонность);

2) $\int\limits_E \lambda\varphi d\mu = \lambda \int\limits_E \varphi d\mu$ (однородность);

3) $\int\limits_E (\varphi + \psi) d\mu = \int\limits_E \varphi d\mu + \int\limits_E \psi d\mu$ (аддитивность).

□ Пусть $\{A_i\}_{i=1}^m$ — допустимое разбиение для φ, $\varphi|_{A_i} = a_i$, и $\{B_j\}_{j=1}^k$ — допустимое разбиение для ψ, $\psi|_{B_j} = b_j$. Тогда множества $C_{ij} = A_i \cap B_j$ образуют общее допустимое разбиение и для φ и для ψ. Поскольку $A_i = A_i \cap \mathbb{R}^n = A_i \cap \bigcup_{j=1}^k B_j = \bigcup_{j=1}^k C_{ij}$, то по свойству аддитивности меры

$$\int_E \varphi d\mu = \sum_{i=1}^m a_i \mu(A_i \cap E) = \sum_{i=1}^n a_i \mu\Big(\bigcup_{j=1}^k C_{ij} \cap E\Big) = \sum_{i=1}^m \sum_{j=1}^k a_i \mu(C_{ij} \cap E).$$

Аналогично $\int\limits_E \psi d\mu = \sum\limits_{j=1}^k \sum\limits_{i=1}^m b_j \mu(C_{ij} \cap E)$. Если $C_{ij} \cap E$ непусто, то для любого $x \in C_{ij} \cap E$ выполнено $a_i = \varphi(x) \leqslant \psi(x) = b_j$. Следовательно, $\int\limits_E \varphi d\mu \leqslant \int\limits_E \psi d\mu$.

Свойства 3 устанавливается аналогично. Свойство 2 очевидно верно. ■

Попутно при доказательстве монотонности установлена корректность определения, т.е. независимость от выбора допустимого разбиения.

Определение. Если $f\colon E \to [0, +\infty]$ — неотрицательная измеримая функция, то

$$\int_E f d\mu := \sup\left\{\int_E \varphi d\mu\colon\ 0 \leqslant \varphi \leqslant f \text{ на } E,\ \varphi \text{ – простая}\right\}.$$

Замечание. Покажем, что данное определение согласуется с определением интеграла от простой функции. Чтобы их различить, перед знаком введенного ранее интеграла поставим (s). Пусть f — простая неотрицательная функция. Если $0 \leqslant \varphi \leqslant f$ на E и φ — простая, то по свойству монотонности $(s)\int\limits_E \varphi d\mu \leqslant$

$\leqslant (s)\int\limits_E f d\mu$. Переходя к супремуму по φ, получаем $\int\limits_E f d\mu \leqslant$ $\leqslant (s)\int\limits_E f d\mu$. Противоположное неравенство очевидно, т. к. f сама является простой функцией.

Сразу из определения вытекают следующие свойства.

Пусть f, $g\colon E \to [0, +\infty]$ — неотрицательные измеримые функции.

C1 (монотонность). Если $f \leqslant g$ на E, то $\int\limits_E f d\mu \leqslant \int\limits_E g d\mu$.

C2 (однородность). Если $\lambda \in [0, +\infty)$, то $\int\limits_E \lambda f d\mu = \lambda \int\limits_E f d\mu$.

C3. Если множество $E_0 \subset E$ измеримо, то $\int\limits_{E_0} f d\mu = \int\limits_E f I_{E_0} d\mu$.

□ Для простой функции утверждение верно, т.к. в интеграле учитываются только ненулевые значения. Пусть φ простая и $0 \leqslant$ $\leqslant \varphi \leqslant f$ на E_0. Тогда $\int\limits_{E_0} \varphi d\mu = \int\limits_E \varphi I_{E_0} d\mu \leqslant \int\limits_E f I_{E_0} d\mu$. Переходя к супремуму по всем таким φ, получаем $\int\limits_{E_0} f d\mu \leqslant \int\limits_E f I_{E_0} d\mu$.

Обратно, пусть ψ простая и $0 \leqslant \psi \leqslant f I_{E_0}$ на E. Тогда $\psi = 0$ на $E \setminus E_0$ и, значит, $\psi = \psi I_{E_0}$. Поэтому

$$\int_E \psi d\mu = \int_E \psi I_{E_0} d\mu = \int_{E_0} \psi d\mu \leqslant \int_{E_0} f d\mu.$$

Следовательно, $\int\limits_E f I_{E_0} d\mu \leqslant \int\limits_{E_0} f d\mu$. ■

C4. Если $E_0 \subset E$ измеримо, то $\int\limits_{E_0} f d\mu \leqslant \int\limits_E f d\mu$.

□ Так как $f I_{E_0} \leqslant f$ на E, то по свойствам 1 и 3 имеем

$$\int_{E_0} f\, d\mu = \int_E f I_{E_0}\, d\mu \leqslant \int_E f\, d\mu. \blacksquare$$

Перечисленных свойств достаточно, чтобы доказать следующий фундаментальный факт.

Теорема 9.4 (Б. Леви). *Пусть функции $f_k\colon E \to [0, +\infty]$ измеримы, и $f_k \to f$ на E. Если $0 \leqslant f_k(x) \leqslant f_{k+1}(x)$ для всех $x \in$ $\in E$ и $k \in \mathbb{N}$, то*

$$\lim_{k\to\infty} \int_E f_k d\mu = \int_E f d\mu.$$

$\square$ Функция f измерима как предел измеримых функций. Интегрируя $f_k \leqslant f_{k+1} \leqslant f$, получаем $\int_E f_k d\mu \leqslant \int_E f_{k+1} d\mu \leqslant \int_E f d\mu$. Следовательно, последовательность $\left\{\int_E f_k d\mu\right\}$ возрастает (в $\overline{\mathbb{R}}$) и, значит, существует $\lim\limits_{k\to\infty} \int\limits_E f_k d\mu \leqslant \int\limits_E f d\mu$.

Докажем противоположное неравенство. Для этого достаточно показать, что $\lim\limits_{k\to\infty} \int\limits_E f_k d\mu \geqslant \int\limits_E \varphi d\mu$ для всякой простой функции φ с $0 \leqslant \varphi \leqslant f$. Рассмотрим такую функцию φ. Фиксируем $t \in \in (0, 1)$ и рассмотрим $E_k = \{x\colon f_k(x) \geqslant t\varphi(x)\}$. В силу возрастания $\{f_k(x)\}$ имеем $E_k \subset E_{k+1}$. Покажем, что $\bigcup_{k=1}^{\infty} E_k = E$. Пусть $x \in E$. Если $\varphi(x) = 0$, то $x \in E_k$ для всех k. Если же $\varphi(x) > 0$, то $f(x) \geqslant \varphi(x) > t\varphi(x)$. По определению предела найдется номер m, такой что $f_m(x) > t\varphi(x)$, т.е. $x \in E_m$. Это доказывает прямое включение. Обратное включение следует по построению.

По монотонности и определению E_k имеем

$$\int_E f_k d\mu \geqslant \int_{E_k} f_k d\mu \geqslant t \int_{E_k} \varphi d\mu. \tag{$*$}$$

Пусть $\varphi = \sum\limits_{i=1}^{m} a_i I_{A_i}$, где $\{A_i\}$ — допустимое разбиение. Тогда по свойству непрерывности меры

$$\int_{E_k} \varphi d\mu = \sum_{i=1}^{m} a_i \mu(A_i \cap E_k) \to \sum_{i=1}^{m} a_i \mu(A_i \cap E) = \int_E \varphi d\mu.$$

Переходя в (*) к пределу при $k \to \infty$, получаем $\lim\limits_{k\to\infty} \int\limits_E f_k d\mu \geqslant \geqslant t \int\limits_E \varphi d\mu$. Для завершения доказательства осталось t устремить к 1. $\blacksquare$

Задача. Пусть $\{f_k\}$ – последовательность неотрицательных измеримых на E функций, и $f_k \to f$ п.в. на E. Докажите, что если существует такое $C > 0$, что $\int\limits_E f_k d\mu \leqslant C$ для всех k, то $\int\limits_E f d\mu \leqslant \leqslant C$ (лемма Фату).

Теорема Леви в сочетании с теоремой о приближении измеримой функции простыми позволяет переносить свойства интеграла с простых функций на неотрицательные измеримые.

С5 (аддитивность по функциям). Если f, $g \geqslant 0$ измеримы на E, то $\int\limits_E (f+g)d\mu = \int\limits_E f d\mu + \int\limits_E g d\mu$.

$\square$ Пусть $\{\varphi_k\}$ и $\{\psi_k\}$ — возрастающие (по k) последовательности простых неотрицательных функций, такие что $\varphi_k \to f$ и $\psi_k \to g$ на E. Тогда $\{\varphi_k + \psi_k\}$ — возрастающая последовательность простых неотрицательных функций и $\varphi_k + \psi_k \to f + g$ на E. По теореме Леви и свойству аддитивности интеграла от простых функций

$$\int_E (f+g)d\mu = \lim_{k\to\infty} \int_E (\varphi_k + \psi_k)d\mu =$$
$$= \lim_{k\to\infty} \int_E \varphi_k d\mu + \lim_{k\to\infty} \int_E \psi_k d\mu = \int_E f d\mu + \int_E g d\mu. \blacksquare$$

Следствие (теорема Леви для рядов). *Если функции f_k измеримы на E и $f_k \geqslant 0$, то*

$$\int_E \sum_{k=1}^{\infty} f_k d\mu = \sum_{k=1}^{\infty} \int_E f_k d\mu.$$

$\square$ По свойству 5 для каждого n выполнено $\int\limits_E \sum\limits_{k=1}^{n} f_k d\mu = \sum\limits_{k=1}^{n} \int\limits_E f_k d\mu$. Правая часть при $n \to \infty$ стремится к $\sum\limits_{k=1}^{\infty} \int\limits_E f_k d\mu$. Ввиду неотрицательности функций f_k частичные суммы ряда нестрого возрастают. Поэтому по теореме Леви при $n \to \infty$ левая часть стремится к $\int\limits_E \sum\limits_{k=1}^{\infty} f_k d\mu$. $\blacksquare$

Теорема 9.5 (счетная аддитивность интеграла). *Пусть E_k измеримы при всех $k \in \mathbb{N}$, попарно не пересекаются и $E = \bigcup\limits_{k=1}^{\infty} E_k$. Если $f \geqslant 0$ измерима на E, то*

$$\int_E f d\mu = \sum_{k=1}^{\infty} \int_{E_k} f d\mu.$$

$\square$ Поскольку E_k образуют разбиение E, то $I_E = \sum\limits_k I_{E_k}$ и $f = f I_E = \sum\limits_k f I_{E_k}$. Тогда по теореме Леви для рядов и свойству 3 имеем

$$\int_E f d\mu = \sum_k \int_E f I_{E_k} d\mu = \sum_k \int_{E_k} f d\mu. \blacksquare$$

Теорема 9.6 (неравенство Чебышева). *Если $f \geqslant 0$ измерима на E, то для любого $t \in (0, +\infty)$ выполнено*

$$\mu\{x \in E\colon\ f(x) \geqslant t\} \leqslant \frac{1}{t} \int_E f d\mu.$$

$\square$ Положим $E_t = \{x \in E\colon\ f(x) \geqslant t\}$. Тогда в силу монотонности интеграла

$$\int_E f d\mu \geqslant \int_{E_t} f d\mu \geqslant \int_{E_t} t d\mu = t\mu(E_t). \ \blacksquare$$

9.3. Интеграл Лебега в общем случае.

Определение. Пусть $f\colon E \to \overline{\mathbb{R}}$ — произвольная измеримая функция. Тогда

$$\int_E f d\mu := \int_E f^+ d\mu - \int_E f^- d\mu,$$

при условии, что хотя бы один из интегралов $\int_E f^\pm d\mu$ конечен. Функция f называется *интегрируемой* (*по Лебегу*) на E, если оба интеграла $\int_E f^\pm d\mu$ конечны.

Замечание. Данное определение согласуется с определением интеграла от неотрицательной функции, т.к. в этом случае $f^+ = f$, $f^- = 0$ и $\int_E 0 d\mu = 0$.

Замечание. Если f измерима на E, то интегрируемость f и $|f|$ на E равносильны. В случае интегрируемости

$$\left| \int_E f d\mu \right| \leqslant \int_E |f| d\mu$$

.

Действительно, если f интегрируема, то оба интеграла $\int_E f^\pm d\mu$ конечны. Тогда в силу $|f| = f^+ + f^-$ получаем, что конечен $\int\limits_E |f|\, d\mu = \int\limits_E f^+\, d\mu + \int\limits_E f^-\, d\mu$. Если же $|f|$ интегрируем, то в силу неравенства $0 \leqslant f^\pm \leqslant |f|$ получаем, что $\int_E f^\pm d\mu$ конечны, а значит, f интегрируема.

Имеем

$$\left|\int_E f d\mu\right| = \left|\int_E f^+ d\mu - \int_E f^- d\mu\right| \leqslant \int_E f^+ d\mu + \int_E f^- d\mu \leqslant \int_E |f| d\mu.$$

Замечание. Если f интегрируема на E, то f конечна п.в. на E.

Пусть $A = \{x\colon |f(x)| = +\infty\}$. Тогда по неравенству Чебышева для любого $t \in (0, +\infty)$ выполнено $\mu(A) \leqslant \mu\{x\colon |f(x)| \geqslant t\} \leqslant \leqslant \frac{1}{t}\int_E |f| d\mu$. Переходя к пределу при $t \to +\infty$ и учитывая конечность интеграла в правой части, заключаем, что $\mu(A) = 0$.

Лемма 9.4. *Если $E_0 \subset E$ и $\mu(E \setminus E_0) = 0$, то интегралы $\int\limits_E f d\mu$ и $\int\limits_{E_0} f d\mu$ существуют одновременно и в случае существования равны.*

□ Отметим, что сужения f на E и f на E_0 измеримы одновременно. По аддитивности интеграла по множествам имеем

$$\int_E f^\pm d\mu = \int_{E_0} f^\pm d\mu + \int_{E \setminus E_0} f^\pm d\mu = \int_{E_0} f^\pm d\mu.$$

В последнем равенстве учли, что интеграл от измеримой функции по множеству меры нуль равен 0. Это следует из определений интеграла последовательно для простых, неотрицательных измеримых, произвольных измеримых функций (для простой функции надо учесть ее ограниченность). ■

Следствие. *Пусть $f,\, g\colon E \to \overline{\mathbb{R}}$ и $f = g$ п.в. на E. Если f интегрируема на E, то g интегрируема на E и $\int_E g d\mu = \int_E f d\mu$.*

В качестве другого следствия получим признак интегрируемости.

Следствие. *Пусть f измерима на E и существует такая интегрируемая на E функция g, что $|f| \leqslant g$ п.в. на E. Тогда f интегрируема на E.*

□ Ввиду монотонности интеграла от неотрицательных функций

$$\int_E |f| d\mu = \int_{E \cap \{|f| \leqslant g\}} |f| d\mu \leqslant \int_{E \cap \{|f| \leqslant g\}} g d\mu = \int_E g d\mu < +\infty. \ ■$$

Теорема 9.7. *Пусть функции $f,\, g\colon E \to \overline{\mathbb{R}}$ интегрируемы на E, и $\lambda \in \mathbb{R}$, тогда*

1) если $f \leqslant g$, *то* $\int\limits_E f d\mu \leqslant \int\limits_E g d\mu$ (монотонность);

2) $\int\limits_E \lambda f d\mu = \lambda \int\limits_E f d\mu$ (однородность);

3) $\int\limits_E (f+g) d\mu = \int\limits_E f d\mu + \int\limits_E g d\mu$ (аддитивность).

□ 1) Пусть $f \leqslant g$, тогда $f^+ \leqslant g^+$, $f^- \geqslant g^-$ и, значит, $\int_E f^+ d\mu \leqslant \int_E g^+ d\mu$, $\int_E f^- d\mu \geqslant \int_E g^- d\mu$. Вычитая второе неравенство из первого, получаем требуемое.

2) Пусть $\lambda \geqslant 0$, тогда $(\lambda f)^+ = \lambda f^+$ и $(\lambda f)^- = \lambda f^-$ и, значит,

$$\int_E \lambda f d\mu = \int_E \lambda f^+ d\mu - \int_E \lambda f^- d\mu = \lambda \int_E f^+ d\mu - \lambda \int_E f^- d\mu = \lambda \int_E f d\mu.$$

Далее, т.к. $(-f)^+ = \max\{-f, 0\} = f^-$, $(-f)^- = \max\{f, 0\} = f^+$, то

$$\int_E (-f) d\mu = \int_E f^- d\mu - \int_E f^+ d\mu = -\int_E f d\mu.$$

Случай $\lambda < 0$ сводится к рассмотренным, т.к. $\lambda = (-1)|\lambda|$.

3) Пусть $h = f + g$. Так как f и g конечны п.в., то найдется такое $E_0 \subset E$ с $\mu(E \setminus E_0) = 0$, что h определена на E_0. На E_0 функция h интегрируема (т.к. $|h| \leqslant |f| + |g|$) и выполнено $h^+ - h^- = h = f + g = (f^+ - f^-) + (g^+ - g^-)$ или $h^+ + f^- + g^- = h^- + f^+ + g^+$. Тогда

$$\int_{E_0} f^- d\mu + \int_{E_0} g^- d\mu + \int_{E_0} h^+ d\mu = \int_{E_0} f^+ d\mu + \int_{E_0} g^+ d\mu + \int_{E_0} h^- d\mu.$$

Все интегралы конечны, их перегруппировка дает $\int_{E_0} h d\mu = \int_{E_0} f d\mu + \int_{E_0} g d\mu$. Определяя h на $E \setminus E_0$ произвольно, заключаем, что последнее равенство выполняется и на E. ■

Важнейшей в теории интеграла является следующая *теорема Лебега о мажорированной сходимости.*

Теорема 9.8 (Лебег). *Пусть функции* f_k, $f\colon E \to \overline{\mathbb{R}}$ *измеримы, и* $f_k \to f$ *п.в. на* E. *Если существует такая интегрируемая функция* $g\colon E \to [0, +\infty]$, *что* $|f_k| \leqslant g$ *п.в. на* E, *то*

$$\lim_{k\to\infty} \int_E f_k d\mu = \int_E f d\mu.$$

$\square$ Так как при интегрировании можно пренебречь множеством меры нуль, считаем, что $f_k \to f$ на E, и g конечна на E.

Переходя к пределу в $|f_k| \leqslant g$, получаем, что $|f| \leqslant g$ на E. Следовательно, по лемме функции f_k, f интегрируемы на E.

Положим $h_k = \sup\limits_{m \geqslant k} |f_m - f|$. Ясно, что $0 \leqslant h_{k+1}(x) \leqslant h_k(x)$ для всех $x \in E$, и

$$\lim_{k\to\infty} h_k(x) = \inf_k \sup_{m \geqslant k} |f_m(x) - f(x)| = \varlimsup_{k\to\infty} |f_k(x) - f(x)| = 0.$$

Кроме того, $0 \leqslant h_k \leqslant 2g$ для всех $k \in \mathbb{N}$. Применим теорему Леви к последовательности $\{2g - h_k\}$. Тогда

$$\lim_{k\to\infty} \int_E (2g - h_k) d\mu = \int_E 2g d\mu.$$

Поэтому $\int\limits_E h_k d\mu \to 0$ и, значит, $\int\limits_E |f_k - f| d\mu \leqslant \int\limits_E h_k d\mu \to 0$. Поскольку $\left| \int\limits_E f_k d\mu - \int\limits_E f d\mu \right| \leqslant \int\limits_E |f_k - f| d\mu$, то требуемое равенство установлено. $\blacksquare$

Теорема 9.9. *Пусть f ограничена на $[a, b]$. Функция f интегрируема по Риману на $[a, b] \Leftrightarrow f$ непрерывна п.в. на $[a, b]$. В этом случае f интегрируема по Лебегу, и оба интеграла совпадают.*

$\square$ I. Пусть $f \in \mathcal{R}[a, b]$ и $J = \int_a^b f(x)dx$. Покажем, что f непрерывна п.в. и $\int\limits_{[a,b]} f d\mu = J$.

Для разбиения $T = \{x_i\}_{i=0}^n$ отрезка $[a, b]$ положим $m_i = \inf\limits_{[x_{i-1}, x_i]} f(x)$, $M_i = \sup\limits_{[x_{i-1}, x_i]} f(x)$ и определим простые функции

$$\varphi_T = \sum_{i=1}^n m_i I_{[x_{i-1}, x_i)}, \qquad \psi_T = \sum_{i=1}^n M_i I_{[x_{i-1}, x_i)}$$

(в последний промежуток включим точку $b = x_n$). Очевидно, что $\int\limits_{[a,b]} \varphi_T d\mu = s_T$, $\int\limits_{[a,b]} \psi_T d\mu = S_T$.

Рассмотрим последовательность T_k разбиений $[a, b]$ на 2^k равных отрезка. Тогда $T_{k+1} \subset T_k$ и мелкость $|T_k| \to 0$. Положим $\varphi_k = \varphi_{T_k}$ и $\psi_k = \psi_{T_k}$. Поскольку при измельчении верхние суммы Дарбу убывают, а нижние возрастают, то $\varphi_k(x) \leqslant \varphi_{k+1}(x) \leqslant f(x) \leqslant \psi_{k+1}(x) \leqslant \psi_k(x)$ для всех $x \in [a, b]$. Поэтому существуют пределы $\varphi(x) = \lim\limits_{k\to\infty} \varphi_k(x)$ и $\psi(x) = \lim\limits_{k\to\infty} \psi_k(x)$, причем $\varphi(x) \leqslant f(x) \leqslant \psi(x)$. Функции φ и ψ измеримы, и если $|f| \leqslant M$, то $|\varphi_k|, |\psi_k| \leqslant M$. Поэтому по теореме Лебега

$$\int_{[a,\,b]} (\psi - \varphi) d\mu = \lim_{k\to\infty} (S_{T_k} - s_{T_k}) = 0.$$

Откуда следует, что $\psi = \varphi = f$ п.в. на $[a, b]$.

Положим $Z = \{x\colon\ \psi(x) \neq \varphi(x)\}$. Пусть $x \notin Z \cup (\bigcup_{k=1}^{\infty} T_k)$ и $\varepsilon > 0$. Выберем разбиение T_k, так что $\psi_k(x) - \varphi_k(x) < \varepsilon$ и рассмотрим $(x - \delta, x + \delta)$, лежащий в отрезке T_k. Так как ψ_k, φ_k постоянны на каждом отрезке из T_k, то $|f(t) - f(x)| < \varepsilon$ для всех $t \in (x - \delta, x + \delta)$. Это означает, что f непрерывна в точке x. Следовательно, функция f непрерывна п.в. на $[a, b]$. Кроме того, по теореме Лебега

$$J = \lim_{k\to\infty} S_{T_k} = \lim_{k\to\infty} \int_{[a,\,b]} \psi_k d\mu = \int_{[a,\,b]} f d\mu.$$

II. Пусть f непрерывна п.в. на $[a, b]$, и $\varepsilon > 0$. Рассмотрим разбиения T_k как в п. I. Пусть x не является точкой разрыва и не лежит в $\bigcup_{k=1}^{\infty} T_k$. Тогда в силу непрерывности в точке x имеем $\varphi_k(x) \to f(x)$ и $\psi_k(x) \to f(x)$. По теореме Лебега $S_{T_k} = \int\limits_{[a,\,b]} \psi_k d\mu \to \int\limits_{[a,\,b]} f d\mu$ и $s_{T_k} = \int\limits_{[a,\,b]} \varphi_k d\mu \to \int\limits_{[a,\,b]} f d\mu$. Откуда $S_{T_k} - s_{T_k} \to 0$ и по критерию Дарбу $f \in \mathcal{R}[a, b]$. ■

Задача. Пусть f локально интегрируема (по Риману) на $[a, b)$ $(-\infty < a < b \leqslant +\infty)$. Докажите, что

1) если $f \geqslant 0$, то f интегрируема на $[a, b) \Leftrightarrow \int_a^b f(x)dx$ сх;

2) в общем случае если f интегрируема на $[a, b)$, то $\int_a^b f(x)dx$ сходится, но обратная импликация неверна.

9.4. Формула суммирования Эйлера.

Теорема 9.10 (Эйлер). *Пусть* $f\colon [1, +\infty) \to \mathbb{R}$ *дифференцируема, а ее производная локально интегрируема. Тогда для любого* $n \in \mathbb{N}$ *выполнено*

$$\sum_{k=1}^{n} f(k) = \int_1^n f(t)dt + \frac{f(1)+f(n)}{2} + \int_1^n \left(\{t\} - \frac{1}{2}\right) f'(t)dt.$$

$\square$ Интегрирование по частям дает

$$\int_k^{k+1} f(t)dt = (t-k)f(t)\Big|_k^{k+1} - \int_k^{k+1} (t-k)f'(t)dt = \\ = f(k+1) - \int_k^{k+1} \{t\}f'(t)dt.$$

Суммируя полученные равенства от 1 до $n-1$, имеем

$$\int_1^n f(t)dt = \sum_{k=2}^{n} f(k) - \int_1^n \{t\}f'(t)dt.$$

Складывая это выражение с $\frac{f(n)}{2} - \frac{f(1)}{2} = \int_1^n \frac{f'(t)}{2}dt$, после элементарных преобразований получаем искомое равенство. $\blacksquare$

При дополнительных ограничениях формула Эйлера принимает наиболее интересный вид.

Следствие. *Пусть функция $f\colon [1, +\infty) \to \mathbb{R}$ дифференцируема и производная $f'(t)$ монотонно стремится к нулю при $t \to \to +\infty$. Тогда для любого $n \in \mathbb{N}$ выполнено*

$$\sum_{k=1}^{n} f(k) = \int_1^n f(t)dt + C(f) + \frac{f(n)}{2} + \varepsilon_n,$$

где

$$C(f) = \frac{f(1)}{2} + \int_1^{+\infty} \left(\{t\} - \frac{1}{2}\right) f'(t)dt,$$

$$\varepsilon_n = -\int_n^{+\infty} \left(\{t\} - \frac{1}{2}\right) f'(t)dt,$$

и все интегралы в правых частях этих формул существуют.

$\square$ Интеграл от $\{t\} - \frac{1}{2}$ по отрезку $[k,\, k+1]$ равен нулю для любого целого k. Поэтому

$$\int_1^x \left(\{t\} - \frac{1}{2}\right) dt = \int_{[x]}^x \left(t - [x] - \frac{1}{2}\right) dt = \frac{1}{2}(\{x\}^2 - \{x\}) \qquad (9.1)$$

и, значит, функция $F(x) = \int_1^x (\{t\} - \frac{1}{2})dt$ ограничена. Так как функция f' монотонно стремится к нулю, то несобственный интеграл в $C(f)$ сходится по признаку Дирихле. Поэтому данное утверждение есть лишь переформулировка формулы Эйлера. ■

Отметим, что $\varepsilon_n \to 0$ как «хвост» сходящегося интеграла.

Пример (формула Стирлинга). $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, $n \to \infty$.

□ Применим следствие теоремы 9.10 к функции $f(t) = \ln t$:

$$\sum_{k=1}^{n} \ln k = n \ln n - n + \frac{1}{2} \ln n + C_f + 1 + \varepsilon_n.$$

Поскольку $\sum\limits_{k=1}^{n} \ln k = \ln n!$, то избавляясь от логарифмов и полагая $c = e^{C_f+1}$, получаем

$$n! = c \exp\left(n \ln n - n + \frac{1}{2} \ln n + \varepsilon_n\right) = c\sqrt{n}\left(\frac{n}{e}\right)^n (1 + o(1)).$$

Найдем константу c при помощи формулы Валлиса. Имеем

$$\frac{(2n)!!}{(2n-1)!!} = \frac{2^{2n}(n!)^2}{(2n)!} = \frac{a\sqrt{n}(1+o(1))^2}{\sqrt{2}(1+o(1))}.$$

Тогда по формуле Валлиса

$$\pi = \lim_{n\to\infty} \frac{1}{n} \frac{a^2 n (1+o(1))^4}{2(1+o(1))^2} = \frac{c^2}{2}.$$

Поэтому $c = \sqrt{2\pi}$, и значит, $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + o(1))$. ■